LATE ASSIGNMENT OF SYNTAX THEORY: EVIDENCE FROM CHINESE AND ENGLISH

Submitted by Xingjia Shen to the University of Exeter as a thesis for the degree of Doctor of Philosophy in the School of Psychology in October 2006.

I certify that all material in this thesis which is not my own work has been identified and that no material is included for which a degree has previously been conferred upon me.

é é é é é é é é é .(Signature)

Abstract

The attraction of the well-structured arguments of the mental syntactic processing device (parser) in Chomskyôs theory has led to an overemphasis on syntactic processing to the exclusion of semantic and other processing in the initial sentence processing stage (Frazier & Clifton, 1996; Gibson & Hickok, 1993; Pickering & van Gompel, 2006). The current thesis joins some others (Green & Mitchell, 2006; MacDonald et al., 1994; Townsend & Bever, 2001, etc.), investigating the timecourse of the information processing of sentences.

The first interest centres on ambiguous sentence resolution. Crosslinguistic studies have shown different resolutions in processing the relative clause (RC) attachment as in ñ*the servant of the actress who was on the balcony*ò (Cuetos & Mitchell, 1988). Three studies confirmed that there is an NP-low preference in Chinese; however, this effect was delayed in comparison to its English counterparts. The NP-low preference can be explained by syntax-first, syntax parallel, and syntax later theories. However, the delay effect questions the traditional syntax-first theories. This leads to the second investigation of direct comparison of the timecourse of syntactic and semantic processing using anomalous materials in English and Chinese. Two experiments have confirmed that the syntactic anomaly is recognised later than semantic anomaly in both languages.

The empirical investigation in the current thesis used various methodologies, including self-paced reading, a questionnaire, and eye-tracking studies, where the design of materials strictly followed linguistic principles. All the results support the late assignment of syntax theory (LAST) (Townsend & Bever, 2001). In fact, LAST can explain most of the evidence for syntax-first and syntax-parallel theories, and it is in line with the latest development of the linguistic UG theories (the Minimalist Programme).

Table of Contents

List of Figures

List of Tables

Acknowledgement

First of all, I shall thank all my friends and colleagues at the School of Psychology, University of Exeter, for providing all kinds of help and support during the whole period of my PhD. They are the academic fellows, technicians, secretaries and many postgraduate and undergraduate students who participated in my experiments, exchanged their opinions with me on my research topic, and even helped me with my English. Particularly, I need to thank Prof. Don Mitchell, my supervisor, for his continuous encouragement and guidance throughout my research. He has constantly informed me of the latest development in the field as well as providing insightful suggestions to my study. Moreover, I was very lucky to have worked as his research assistant, which not only funded almost two years of my PhD, but also strengthened my skills that will be extremely valuable for my future development. I would also like to thank Dr. Tim Hodgson, who has generously taught me eye-tracking technology, and wrote some of the software used in my experiments. Thanks also to Dr. Aureliu Lavric who taught me the basic knowledge of EEG/ERPs technology, and has provided valuable discussion and suggestions to my research. Aureliuôs handout for the undergraduate course ñLanguage and Thoughtò has helped me to have an in-depth understanding of different language processing models.

Discussions at conferences such as AMLaP in Glasgow and Ghent, PCOEAL in Hong Kong and Camling in Cambridge helped to shape this thesis. I would specially thank Prof. Janet Fodor, Dr. Barbara Hemforth, Dr. Yuki Kamide and Dr. Edson Miyamoto. They had not only shown interest in my presentations, providing valuable questions and suggestions, but also provided me with adequate follow-up information about their own research work.

The eye-tracking data analyses could never start without the help from my dear husband, Lari, who deserves special thanks for writing Matlab scripts and various small programmes to ñcleanò the massive dataset. My most useful programming knowledge was taught by him. He also spent much of his spare time discussing my research topic, even though he might not be particularly interested in it. His suggestions helped to clear up my ideas. Without his support, I could not have finished this work.

I would like to express my thanks to two special friends who have always been lovely and supportive: Mr. Mikihiro Tanaka who always updated me with important information, especially from conferences that I could not go; and Miss Paula McDonald, who helped to proof-read this thesis.

In the end, I would like to dedicate this work to my dear parents, who always believed that I could finish this work even when I doubted it, and were supportive both spiritually and financially throughout the years. This thesis is written in memory of my dear grandmother, Ms. Zhenghua Zhang, who left me forever just before I started the PhD, and my grandfather, Bangxing Shen, who loved me the most.

Chapter 1

Introduction

Natural language is like a device that encodes oneøs thoughts and is decoded by another so that people can communicate. However in terms of using a language, it is quite often that the procedure of encoding to decoding can be ñlost in translationò. Suppose a person enters the office Monday morning and heard his colleague chatting with another across the room:

1) *"…and I saw an iceberg on the shoulder of a giant. …"*

He did not catch the previous conversation. One may find it an impossible task to make sense of this sentence. This sentence has two interpretations: I was standing on the shoulder of a giant, and saw an iceberg; or, I saw a giant carrying an iceberg and the iceberg was on the giantøs shoulder.

In the past three decades, psychologists and psycholinguists have done much work attempting to explain the phenomena of human language processing. It has been argued that the sentence is a natural level of linguistic representation (e.g., Chomsky, 1959; Townsend & Bever, 2001). When we understand a sentence, we are hardly conscious of the structures; rather, we directly absorb their ñmeaningsò before moving on to the next sentence. Therefore, it is unsurprising that it took some years for psychologists to realize and admit that knowledge and usage of language involve sentence-level syntactic structure (Cairns, 1999 ; also mentioned in Townsend & Bever, 2001). This finding raised the question: what does syntax have to do with sentence processing; and when and how do structure and meaning interact? Is syntax really assigned earlier than other information in achieving the meaning of a sentence?

Before discussing the function of syntax, it is necessary to have a clear definition of human language. Hockett (1959; 1960; 1961; cf. Pinker, 1999; Also see Whitney, 1998) suggested six fundamental characters of human languages: Semanticity, Arbitrariness, Discreteness, Duality of Patterning, Productivity and Displacement.

Semanticity means that the language symbols have to be able to carry meanings. **Arbitrariness** means the symbols and their representees are forced to link to each other for no particular reasons. Even when the pictographic forms of some languages (such as Ancient Egyptian hieroglyphs, or Ancient Chinese pictograms), were formed to represent the shapes of physical objects, there was no particular reason for the way the system of phonemes was linked to the glyphs. Crosslinguistically, there are as many different names for one object as there are languages in the world, whereas animal signals have only one type of symbol for one predator across one species (Burling, 1993).

Discreteness means that each symbol can only represent a specific meaning. If one wants to modify this meaning, he/she will have to use other symbols to combine with this one, or change symbols. These characteristics do not yet separate human languages from all animal signal systems, since some of the latter also bear these features, but the following features are seen in human languages only.

In human language, meaningful symbols are made up of meaningless components. For example, phonemes are meaningless, but they are the basic units that combine to form morphemes. Morphemes carry basic meanings and are the basic component of meaningful words, and then words combine into sentences. The **duality of patterning** means that there are two levels in a language: meaningful symbols and its meaningless component. Moreover, the combination is not random. It is rather governed by a set of rules (Chomsky, 1965; Pinker, 1999). Phonology is the governor of the sound combination based on phonological rules, whereas syntax is the governor of the combination of words based on syntactic rules. With these rules, the meanings can be passed on persistently. This feature also makes it possible for language to be more flexible and creative (hence the feature of **productivity**), so that it can be used to describe things that are ñphysically and temporally displacedò (Whitney, 1998) (hence the feature of **displacement**).

From the above, we can see that the **duality of patterning** is the most crucial characteristic separating human language from other animal signals. Moreover, the rules that are used to conduct the combination are the crucial link between meaning and symbols, as well as the link between smaller meaning units to a complicated expression of thoughts.

Thus, it has become clear that human language is some kind of combination of symbols based on rules, and the goal of this combination is to express and transfer meanings. It follows three steps, encoding meaning → language form → decoding meaning, to complete one cycle of the meaning transfer. While linguists have been trying hard to describe the rules in the language form (the middle of the cycle), and psycholinguists have tried hard to determine how the meaning is encoded into and decoded from the form (Altmann, 1997), philosophers are trying to use their results to answer such questions as ñwhat is meaning?ò, ñDoes language determine our thinking?ò, ñWhat is the *innateness* of language all about?ò (Cooper, 1973; 1975).

One of the most intriguing topics in language phenomena is the question of how we understand other peopleôs production of language (speech or written forms). Chomsky (1959; 1965; also see Townsend & Bever, 2001 for a discussion) has suggested that the natural level of language representation is the sentence level. This is because sentences are generated by syntactic rules, whereas phonological rules determine the phonetic forms of a sentence; and morphological rules are embedded in the syntactic rules in such a way that the changes of lexical forms are assigned when the words are part of a sentence (i.e., inflection involves computational operations within a broader syntactic scope (Chomsky & Lasnik, 1995)).

Now have a look at the conversation between the two crocodiles in the comic strip shown below:

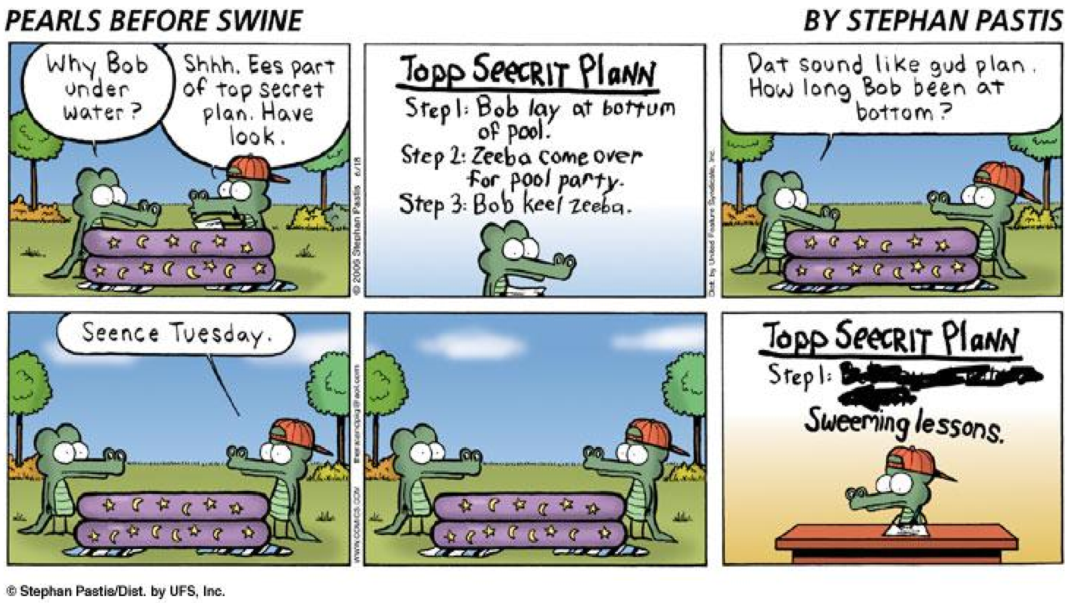


Figure 1 The dialogue between two bogan-crocodiles

The crocodiles represent a kind of misbehaving social class in the neighbourhood depicted in Stephan Pastisô comic strip. Their use of language mimics the ñboganò[1] language, which is full of mistakes in terms of ñstandardò English: some words are phonetically wrong, (for example, the phoneme /ð/ in the word *that* has been mispronounced as /d/* (*dat*)); other words lack inflection, (such as *sound**s in *dat sound like gud plan.*); moreover, the combination of words do not strictly follow English grammar (as in the first sentence, *why* *is *Bob under water?*, there is no main verb in the whole sentence). However, it does not stop one from understanding the meaning of this conversation: one planned to murder Zeeba, but unwisely drowned Bob.

Figure 1 and sentence 1) above characterise some typical language problems, namely, anomaly and ambiguity. *Anomaly* is when sentences do not follow rules and simply go wrong, and *ambiguity* is when the interpretation of sentences has more than one solution (several possible meanings). Sentence 1) exhibits what is called *Global Ambiguity*, because there is no way to resolve the ambiguous meaning just from the structure of the sentence itself. There is also another type of ambiguity, as in 2):

---

[1] Bogan is an Australian and New Zealand English slang term, at times derogatory, for a person who is, or is perceived to be, unsophisticated or of a lower class background. The stereotype includes having speech and mannerisms that are considered to denote poor education and uncultured upbringing.

2) The horse raced past the barn *fell*.

When the processing reaches *fell*, the sentence becomes unambiguous and the meaning of the sentence can be resolved by re-analysing the structure. This type of ambiguity is called *local ambiguity*.

Although the purpose for humans to use a language is to communicate, to express ideas, and to get to know others better, from the example above, we can see that sometimes language does not entirely serve this purpose. Most normal sentences[2], as well as anomalous or ambiguous sentences can be understood, at some point or another, by a human language structure processing mechanism ï the parser[3]. According to Chomsky (1965; Chomsky & Lasnik, 1995), there is a fundamental distinction between human language *competence* (i.e., fluent native speakersô tacit knowledge of their language) and *performance* (i.e., what people actually say or understand when receiving speech or reading). As generative grammar (i.e. transformational grammar) serves the purpose of providing a structural description to an infinite range of sentences, logically, the performance is seen as an ñimperfect reflection of competenceò (Radford, 1997)[4]. Sentence processing uses all the information available in the incoming message. However, it is unknown how the parser recovers from the loss or mismatch of information in a situation where the speakers commit as many performance errors in sentence production as the crocodiles in Figure 1 do. Things get more complicated when speakers produce sentences like 1), which are fully grammatical: how does the parser help to achieve the goal of understanding the meaning of the sentence from so many interpretations?

The key dispute in the answers to the questions above is in the time-course of the processing of different information. At some stage the processing of the sentence structure will be performed, just like processing of other information such as context,

[2] Here *normal* means grammatically correct. It is not in the same sense of *acceptable* (Chomsky, 1965).
[3] Parsing refers to the processing of syntax. Sentence processing involves parsing as well as other aspects of processing, e.g., semantic processing and discourse. Therefore *parsing* and *sentence processing* are NOT interchangeable in this thesis.
[4] Note that misproductions and misinterpretations do not always mean lack of competence, but rather *performance errors*. Native speakers making those errors can be attributed to various reasons such as tiredness, external distraction and so on. A grammar tells people the *normal* way the native speakers do. With language competence, one can speak as fluently as a native speaker.

prosody, and even plausibility derived from ñworld knowledgeò (i.e., common sense). The Chomskyan linguists and psycholinguists (e.g., Frank, 2004; Frazier, 1978; Frazier & Clifton, 1996; Frazier & Rayner, 1982; Petersson et al., 2004), hold the belief that the processing of syntax is logically and also behaviourally prior to other information, but the language data collected to date actually varies (i.e., syntax can be processed simultaneously with or even later than other information) (e.g., Bornkessel et al., 2004; Connell & Keane, 2006; McDonald et al., 2005; McRae et al., 1998; Runner et al., 2006; Townsend & Bever, 2001, etc.).

The current thesis also attempts to investigate this issue. First of all the development of syntactic theories will be introduced, primarily in Chapter 2, so that the argument of sentence processing can have a clear linguistic foundation. This is because simple ñstructureò cannot cover the connotation of syntax, as in some languages (such as Chinese), the linear structure and the syntax structure are not identical. A clear understanding of the concept of ñsyntaxò makes it a good foundation for arguments of syntax processing.

Chapter 3 reviews contemporary psycholinguistic theories. Those theories have been categorised into three groups according to their proposal of the timecourse of syntactic processing (i.e., syntax processing prior to other processing, syntax processing simultaneously to other processing and syntax processing posterior to other processing). While Chapter 4 introduces the commonly used psycholinguistic methodologies, Chapter 5 provides evidence, using methodologies introduced in Chapter 4, to argue for and against the theories introduced in Chapter 3.

Chapters 2-5 form the literature review of the current thesis. As the old metaphor goes ñdwarves standing on the shoulders of giants see further than the giants themselvesò (Latin: ñPigmaei gigantum humeris impositi plusquam ipsi gigantes videntò)[5], this part can be seen as mounting to the shoulder of the giants who provided adequate evidence and theories from previous studies.

[5] This was first recorded in the twelfth century and attributed to Bernard of Chartres by John of Salisbury.

Chapter 6 provides some Chinese and English data for a global ambiguous sentence processing of relative clause (RC) attachment (as in ñ*the servant* (NP-low) *of the actress* (NP-high) *who was on the balcony* (RC)ò). The reason of choosing this type of anomaly is that according to crosslinguistic data, the preference of the attachment is not unique, and some of the preferences are against the universal principles of syntax assignment. If the universality and the principles have both been questioned, the foundation of dominant syntax-first theories has been weakened.

Study 1 used self-paced reading time measures; Study 2 used a crosslinguistic questionnaire and Study 3 used eye movement measures. Whereas the preference of attachment undoubtedly prefers the lower NP in both languages, the mystery of the delayed effects in Chinese has led to the original question of the timecourse of syntactic processing. This is because the Chinese language features have made it impossible to directly use syntactic manipulation as other studies of this kind, so semantic manipulation was applied. However, to test whether the delay was caused by semantic manipulation, direct comparison of the timecourse of syntactic and semantic processing are considered.

Chapter 7 reports two experiments using eye-tracking technology, directly comparing syntactic and semantic processing timecourse by manipulating the anomaly of these kinds. The results from both studies show clear later awareness of syntactic anomaly. Thus the delayed effects in Chinese were not caused by semantic manipulation, and syntactic processing operates later than semantic processing. This supports the rarely-accepted late assignment of syntax theory (LAST). Moreover, it is proposed that the delayed effects may originate from the specific language feature.

Chapters 6 & 7 form the second part of the current thesis. The evidence is like the tip of a huge iceberg, emerging from the sea of the psycholinguistic field.

Chapters 8 & 9 focus on LAST, explaining the linguistic foundation for this theory, as well as the explicability of this theory to the evidence provided in Part II of the current thesis, and to other evidence ñsupporting ò syntax early (or simultaneous) theories. It also proposes a pre-syntactic processing (PSP) stage for simple structured sentences.

Part I.

On the Shoulder of A Giant

Chapter 2:

The Development of Chomskyan Syntax Theories

As the argument of processing time-course mainly departs from the stage of syntax processing, before moving on to the sentence parsing models, it is essential to know what syntax exactly involves. The Chomskyan theories of Universal Grammar (UG) have been developed for at least three generations since the 1960s: the *Transformational Grammar* (TG) (e.g., Chomsky, 1959; 1965), *Government and Binding* (GB) (See Black, 1999, for a detailed introduction; Chomsky, 1986) and the newly developed *Minimalist Programme* (MP) (e.g., Chomsky, 1995; 1999; 2001; and see Radford, 1997, for a detailed introduction).

The consistency shared by the three generations is *Universality* of the generative grammar despite the existing language forms. However, after 50 years of development, the current UG is to some extent a different one to its 1960sôand 1980sôcounterparts.

2.1 Transformational Grammar (1950s-1970s)

In the 1960sôTG theory, the system of rules was analysed into three major components of the generative grammar: syntactic, phonological and semantic components. The relationship between each component is shown in Figure 2 below.

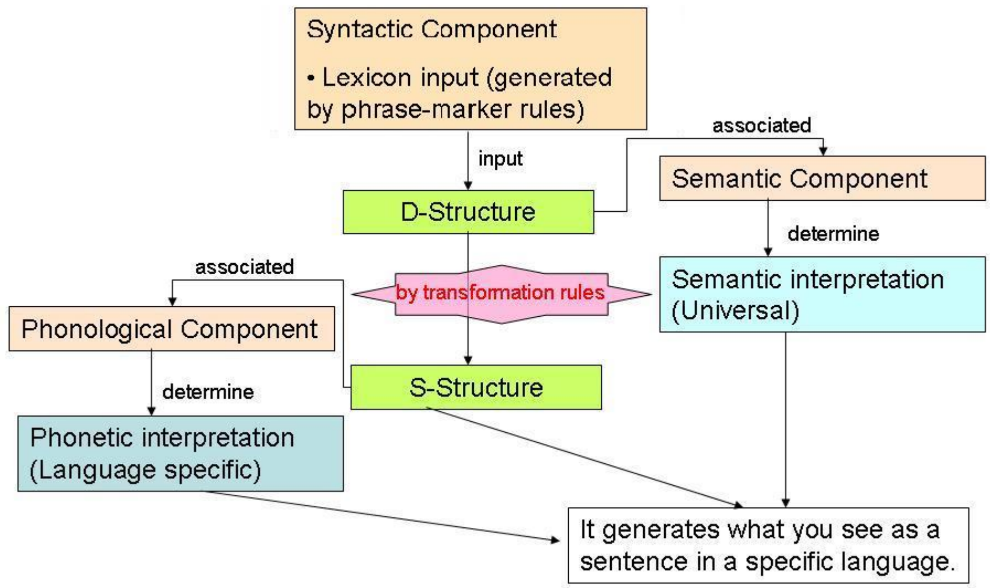


Figure 2 The UG explained by TG

It shows that the syntactic component forms a Deep-structure (D-structure), which is associated with the semantic component. As the semantic components determine the semantic interpretation ï the meaning ï it should be identical throughout different languages. The D-structure undergoes an operation called *grammatical transformation* to form a *Surface-structure* (S-structure), which is associated with the phonological component. As the phonological component determines phonetic interpretation, it is language-specific. Moreover, TG assumes that ñthe semantic component is a system of rules that assign readings to constituents of Phrase-markers[6] ï a system that has no intrinsic structure beyond this.ò (Chomsky, 1965, p.160). This means, although briefly, that Chomsky did not deny the preliminary function of the semantic factor. It is just not yet been entered into the computational system of a sentence (cf. Chomsky & Lasnik, 1995). It is understandable that the assignment of the syntactic category cannot be an empty entity without lexical entries. In fact, not only semantic factors, but also phonetic structures and functioning of transformation rules should be presented at the lexicon entry as well. The rules (phrase marker rules, transformation rules, etc.) can be generalised into abstract forms and seemingly can be ñcontext freeò, but the application

6 ñA phrase-marker maybe based on a vocabulary of symbols that includes both *formatives* (the, boy, etc.) and *category symbols* (S, NP, V, etc.). The former is further divided into *lexical items* (sincerity, boy) and *grammatical items* (Perfect, Possessive, etc.)ò (Chomsky, 1965)

of rules cannot utterly abandon other components of language.

As Chomsky admitted, the boundary between semantics and syntax is impossible to determine (Chomsky, 1965). TG has been criticised as too ñpowerfulò in terms of the transformational movements (See Black, 1999 for a review). It also has problems in solving the structural difficulties by assigning ñillogical categoriesò. For example:

3) I read that the economy is poor.

Under TG, the *head* of the complement clause is the sentence ñthe economy is poorò, where the complementiser ñthatò is the *specifier* of the sentence. However, using a sentence to be the head violates an alleged universal *Phrase-marker Rule* as in 4).

4) Phrase-marker Rule[7]:

   For any lexical or functional category X, X_0 = head (head cannot be a sentence),

   XP $\rightarrow$ Specifier Xô

   Xô$\rightarrow$ X_0 Complements (=YP)

Rule 4) requires that the head of a phrase cannot be a sentence, but the simplicity of this theory does not specify in what category the sentence ñthe economy is poorò should be, and therefore causes some conflicts in categorisation.

2.2 Government and Binding (1980s-1990s)

The GB theory introduced in the 1980s is based on TG. It is very important to the current psycholinguistic research because a large amount of the development of the psycholinguistic parsing models (especially structure-based models (see Carroll, 1994, for an introduction; or, e.g., Frazier, 1987, etc.) can be traced back to it.

According to GB, the universality is still embedded in the D-Structure which is combined by lexicon items. It is then mapped into the S-Structure, where, however, it is then factored into the *Phonological Form* (PF) and the *Logical Form* (LF). The

7 This ruleô s written form and also the tree-structure in Figure 4 below are based on head-initial languages. Other type of languages and forms will be discussed later when parameter setting is introduced.

relationship and transformation between each level is represented in Figure 3 below.

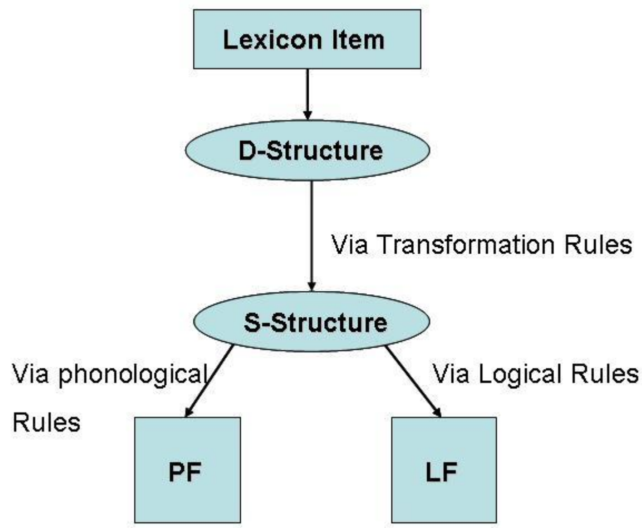


Figure 3 The UG explained by GB

The PF is the boundary between syntax and phonology. The sounds and grouping of the lexicon items are directly represented by the PF. Moreover, the border between the semantics and syntax is at the LF, where ñpredication relationships and the scope of quantifiers and operators of various kinds are explicitly represented in the phrase structureò (Black, 1999, p. 2).

The most important development by GB is the X-bar theory[8]. This has enabled stricter control of the transformations. It follows the universal Phrase-marker Rule 4). The tree structure can be seen below:

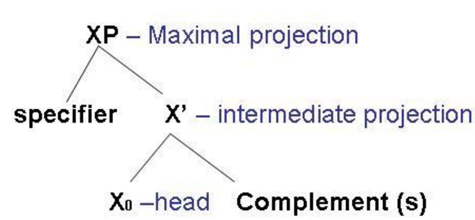


Figure 4 The Tree-diagram of the X-bar Theory in head-initial languages

Differences between languages are parameterised by each languageôs capacity at the

[8] Although the original X-bar theory has already been raised in 1960s, itôs been developed to a more mature stage from GB.

S-Structure. Table 1 gives a list of the possible combinations of some parameterisation of word order existing in the current human languages, derived from Black (1999, p. 19).

Table 1 X-bar's Parameterisation

SVO	XP → Specifier Xô	VOS	XP → XôSpecifier
e.g., English	Xô→ X0 Complement(s)	e.g., Tzotzil[9]	Xô→ X0 Complement(s)
SOV	XP → XôSpecifier	OVS	XP → XôSpecifier
e.g., Japanese	Xô→ Complement(s) X0	e.g., Hixkaryana[10]	Xô→ Complement(s) X0

Word order is very important in GB, as it assigns the initial parameter. However, it is not difficult to see the theoretical deficiency from the parameter table above: There are six combinations of possible word order, whereas the parameter for X-bars can only allow four. The remaining two orders, namely VSO and OSV, are left with no free parameters. Unless these two types of word order do not exist in the world, the deficiency can be veiled. Unfortunately, among the 5,000 known natural languages (Ruhlen, 1987), the chance of missing those two types of order is small. Indeed, although few, there are still records of these two types: Hawaiian (VSO) and Xavante[11] (OSV). The parameter settings of those two types of language are somewhat arbitrary.

Nevertheless, following the X-bar theory, more constraints are given to the movement so that the movement cannot be as free as in the TG. Two basic principles are:

5) a. Principle of No Loss of Information (PNLI)

-- nothing can move to a position that is already phonetically filled;

b. Principle of Structure-preserving;

-- movement must fit into the existing position generated at D-structure.

These principles suggest that in the D-structure, there must be some empty categories to allow movements to fit in. In addition, the movements always happen in the ñheadò or ñspecifiedò position.

[9] A Mayan language spoken in Mexico.

[10] A southern Guiana Carib language spoken in Northen Brazil.

[11] A type of Ge language spoken by the Xavante people in about 60 villages in the area surrounding Eastern Mato Grosso, Brazil.

Another development by GB is that it has added *semantic roles* into the lexical entry. General semantic roles include AGENT, THEME, RECIPIENT, GOAL, LOCATIVE, etc., which can be linked to arguments assigned by predicates.

Now it is possible to apply those two theories to real sentence examples. Taking sentence like 6) as an example, the differences between TG and GB are rather substantial.

6) a. His friend carried him to the hospital. (Active)
   b. He was carried by his friend to the hospital. (Passive)

The TG theory treats the passive sentence 6b) as the surface structure of a), i.e., Sentence 6a) is the D-structure of b). Sentence 6b) is realised by movements of shifting the subject to the *by*-phrase (or omitting it), and the objects to the subject position, then add the passive AUX *be*. The rule for this type of transformation in TG is

7) a. Most transitive verbs have passive alternates;
   b. No intransitive verbs have passive alternates;
   c. The ósubjectô of a passive verb corresponds to the object of its transitive alternates.

This transformation is illustrated by the tree-diagrams below[12].

[12] We only focus on the main information transformation here. The change of the case of the pronouns (e.g., the change of the accusative case *him* to the nominative case *he*) will not be described here.

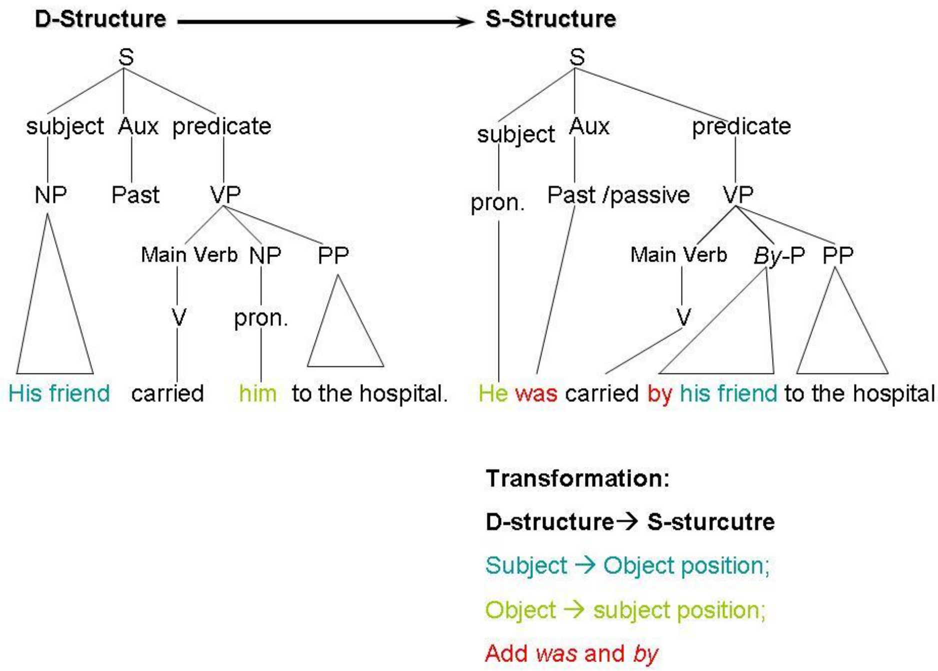


Figure 5 The transformation of passive sentences under TG.

However, this transformation violates the new rules 5) introduced in GB: the subject position has been taken in D-structure, so simply transforming the subject position will cause *Loss of Information*, which violates rule 5a); and the position for *by*-phrase was not preserved in D-structure, which violates rule 5b). Fortunately, under GB, the semantic roles are assigned at the lexical entry, which directly applies at D-structure[13]. The developers of the GB have added other rules (or hypothesis) to constrain the semantic role assignment:

8) Uniformity of Theta Assignment[14] Hypothesis (UTAH)
   The THEME role is always assigned to the direct object, since it follows the main verb and is in the position of a normal ñobjectò; the RECIPENT role is assigned to the indirect object.
9) Movement into a position linked with an assigned semantic role is not

---

[13] For example, the lexical entry for the verb *ask* is: ask, V, [__ NP (PP[to])], where *ask*Õs semantic role at the entry is <AGENT, THEME, GOAL>. ñ__ò means the prime component that *ask* follows. No ñ__ò means what *ask* can take. But the lexical entry for its passive form *asked* is: asked, V[passive], [__ NP (PP[to]) (PP[by])], and the semantic role assigned to each entry is <THEME, GOAL, AGENT>.

[14] Theta-assignment (ɗ-assignment) and semantic role assignment is inter-changeable in this thesis.

allowed[15].

This actually invalidates the TG rules in 7) because the D-structure of a passive sentence is no longer a transitive one (e.g., 7a)), but a structure with passive semantic assignment directly. The transformation has become two movements: the RECEIPTANCE moves up to the specifier position of an IP, and the AUX at the head of the Vômoves up to the head of an Iô.

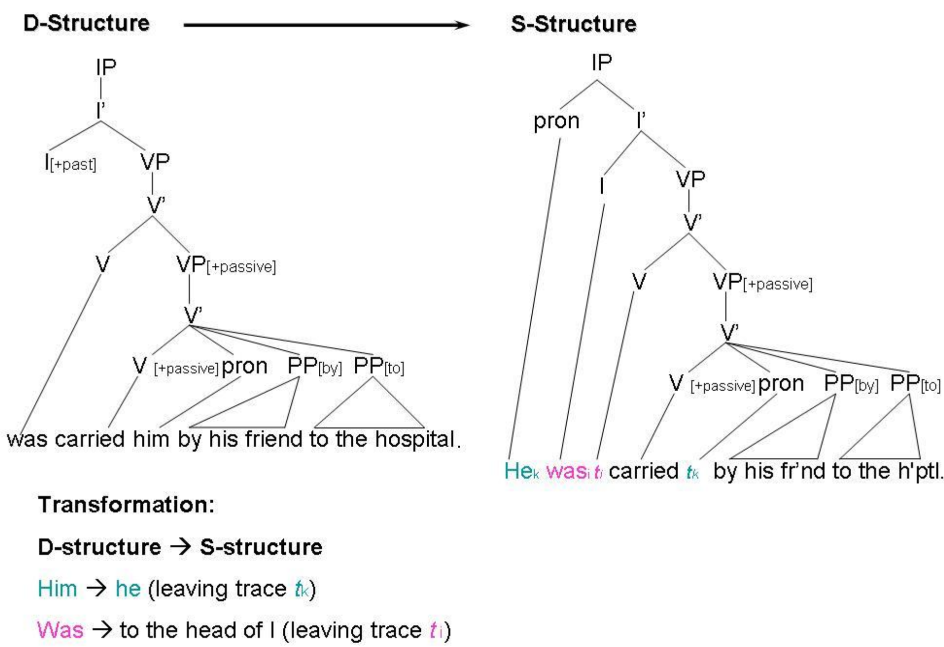


Figure 6 The transformation of passive sentences under GB

Figures 5 & 6 show the fundamental differences between TG and GB. In GB theory, movements do not simply happen without trace. In fact, the argument structure of lexical items must be preserved throughout the derivation. The movement traces are bound by their antecedent, and the empty slots (empty category) left by the movements (as well as the controlled null pronominal (PRO)) share the semantic properties of its antecedent.

However, GB also meets substantial problems. In addition to the parameter-setting problem discussed above, GB also assumes that word order similar to an SVO language

[15] The semantic role stays with the original position. It is not part of the structure, but part of the lexical subcategorisation, which goes with the D-structure. D-structure does not change.

is ñdefaultò (e.g., the Internal Subject Hypothesis (ISH)[16] (Please see Black, 1999 for a detailed introduction; also, Diesing, 1990; Kitakawa, 1986; Koopman & Sportiche, 1991; Kuroda, 1988). Nonetheless, there is no convincing reason for this hypothesis or a hypothesis of any type of word order to be the default one.

One further point on GB concerns the question of the system it describes (as illustrated in Figure 3 above). Chomsky and colleagues (See Chomsky & Lasnik, 1995, for a introduction) assume that the lexical entry is not a computational system, but rather *appears in* this system. The other levels (the D-Structure, S-Structure, and LF / PF) form a computational system generated from the structure of a certain language. Although it is not certain whether the relationships between the three levels are directional, the school of Chomskyan syntactic theory proposed a positive suggestion, which has become the dominant assumption in the psycholinguistic research field for sentence processing:

> ñWe will tentatively proceed on the assumption that the relations are, in fact, directional: D-Structure is mapped to S-Structure, which is (independently) mapped to PF and LF.ò (Chomsky & Lasnik, 1995, p. 23)

However, the assumption of the directional mapping is illogical when directly used in the psycholinguistic field. Psycholinguistic research studies language *performance*, which involves both language production and interpretation. This means the performance of transforming between meaning and structures has to be bi-directional, or circular. As the Chomskyan School *never* claimed the relationship is bi-directional, nor circular, the fundamental question is how the parser attains the input of the lexical entry. Essentially, any sentence is presented by certain form of a natural language, and the parser receives the form as some percept of a signal, say sound. This means that the lexical entry achieved by a parser should be some type of PF, which already violates the system illustrated by Figure 3 (as the entry to the system should be lexical items). Furthermore, the sentence that a parser received has to be at least phrase by phrase (if

[16] ISH proposes that the subject begins in the specifier of VP in all languages. The order of VSO is obtained by moving the verb up to the I_0-head position (Black, 1999). However, as the detail of this hypothesis is out of the scope of the current thesis, it will not be discussed here.

not word-by-word (allow skipping in reading text) or character-by-character, such as the Chinese language). This means it should first appear in the parser as its S-Structure. If the meaning cannot be obtained directly from this structure before the parser goes to check with the D-Structure, then the information that is input into the D-Structure is not from the lexical entry (i.e., the S-Structure is only a *medium* between the D-Structure and meaning; without parsing the D-Structure, the role of a medium cannot be realised). The whole system will then collapse. If the parser can directly parse from the S-Structure, then what is the role of D-Structure?

2.3 The Minimalist Program (1990s - now)

The MP theory takes a different turn in terms of the D- / S- Structures. Although early work on MP still involves the separation of D-Structure and S-Structure (Chomsky, 1995), the MP minimises this distinction and eliminate them entirely in later developments. The very recent development of this theory even has a tendency to eliminate the LF and PF levels, but suggests that derivations proceed in *phases* (Chomsky, 1999; 2001; Radford, 2001), which will be discussed in detail below.

The MP theory is based on Chomskyôs idea that language is *perfect* in the sense that it only uses information that is *conceptually necessary*. Thus, the movement in a sentence structure has become more restricted, that the movement will ONLY be triggered by the necessary NEED / REQUIREMENT under certain principles. The reason is that the so-called innate ñLanguage Facultyò[17] can only hold limited amounts of structure in its ñactive memoryò (Chomsky, 1999, p. 9), the computational burden of structure has to be reduced to a minimum. Chomsky proposes that syntactic structures are built up one phase at a time (Chomsky, 1999; 2001). A phase can be a complete clausal complex (i.e., CP), or a complete thematic complex (i.e., v*P). At the end of each phase, the *domain* (i.e. complement of the phase head) undergoes transfer to the phonological and semantic components. Chomsky suggests that *no future syntactic operations on that domain* from that point are allowed. This means that the syntactic operation is actually *bottom-up* and one-directional only.

[17] Will be discussed in more detail in the next section 2.4

The fundamental principle for UG is **Locality** (i.e., the nearest expression of a relevant type should be placed locally in the structure).

10) **Locality Principle:**

All grammatical operations are local. (see Radford, 2001 p. 9)

This principle is maintained under MP. This can be understood from the example 11) (example taken from Radford, 1997) below (* indicating problematic sentences):

11) a. *Who had* he said would do what? (cf. He *had* (1st AUX) said *who* (1st Wh-word) would do what?)

b. * Who would he had said do what? (cf. He had said *who* (1st Wh-word) *would* (2nd AUX) do what?)

c. * What had he said who would do? (cf. He *had* (1st AUX) said who would do *what* (2nd Wh-word)?)

d. * What would he had said who do (cf. He had said who *would* (2nd AUX) do *what* (2nd Wh-word)?)

Phases also follow the rule 10) and it is ensured via another extended principle 12) called **Phase Impenetrability Condition (PIC)**:

12) **PIC**

The domain (i.e. complement) of a phase head is impenetrable to an external probe[18] (i.e. to a probe which lies outside the phase). (Radford, 2001, p. 195)

Another universal principle that has been maintained in the UG (GB & MP) is known as **Extended Projection Principle** (EPP).

13) **EPP:**

A finite tense constituent T requires an extended projection into a TP containing a subject. (Chomsky, 1982, p. 10)

[18] Chomsky has suggested that agreement involves a relation between a *probe* and a *goal.* A Probe is a highest head TP or VP in a phase, which is searching for a nominal Goal to agree with. Under the LP rule 9), a Probe searches for the closest goal within the immediate structure containing the probe. (Chomsky, 1999; 2001)

Examples are shown in 14) below:

14) We are trying to finish the work.

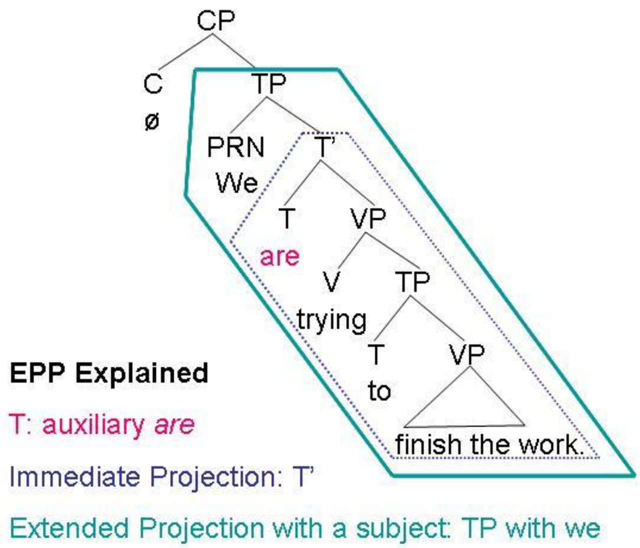


Figure 7 Tree-structure illustrating EPP

Furthermore, some of the rules actually invalidate the previous theories on movement, such as 14) and 15), (for both rules, see Radford, 2001, p.83):

15) **Headedness Principle**

Every projection is headed (i.e., has a head).

16) **Binarity Principle**

All structure is binary (i.e., all non-terminal nodes are binary-branching).

This means the structure illustrated in Figure 6 above does not follow the new rules. The correct one for sentence 6b under MP has become:

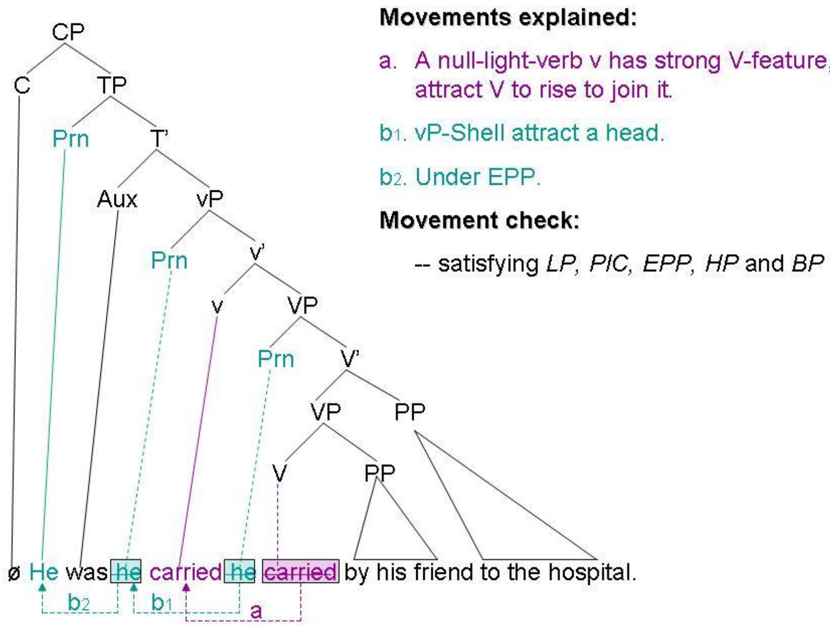


Figure 8 The transformation of passive sentences under MP

Figure 8 has some crucial differences from Figures 5 & 6: a. the transformation is not from a D-Structure to an S-Structure; b. the movements are more restricted; c. each movement has a logical reason; d. the movements are limited to only those necessary.

The problem is if the structure is built bottom-up, it means the parser has to wait until all the lexical inputs are given. This means, initially, other information processings than parsing (the syntactic processing) would happen. This will be further discussed when processing models are reviewed.

2.4 The Innateness of Natural Language (A Summary)

The UG theory is based on an assumption that there is specific biologically endowed human language ability ï the *language faculty*. It can provide children with a genetically transmitted algorithm for developing a grammar on the basis of their experience. It is also used in transforming the articulatory-perceptual system into a conceptual-intentional system in the adult language system.

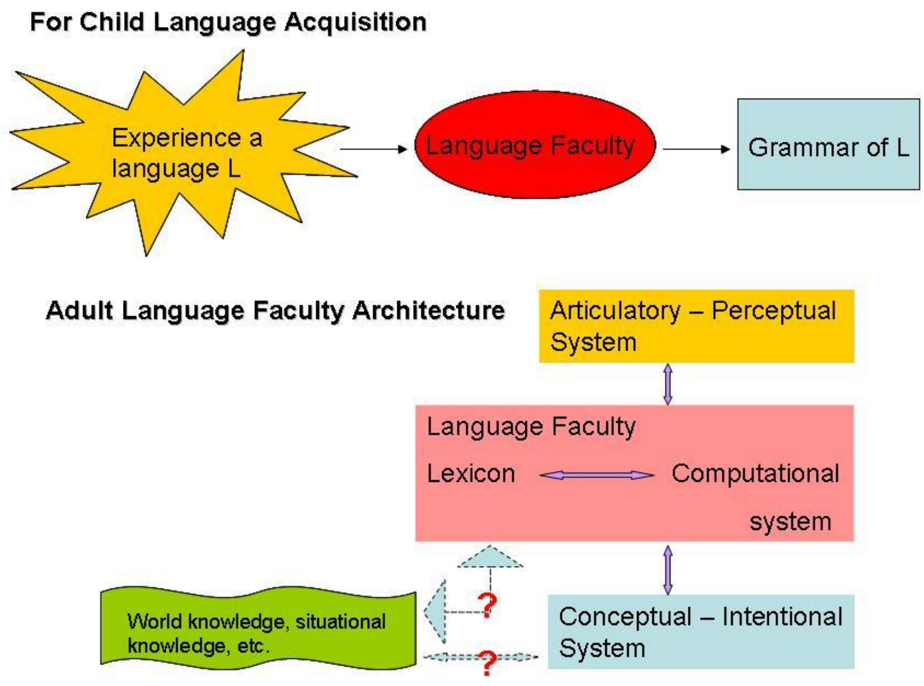


Figure 9 The Language Faculty

It is unarguable that each language has its own grammatical system that is learnable, but UG does not describe any of those specific language systems. In fact, the UG is a set of rules that allows the individual language system to exist and provides a rationale for the learnability of those rules. Chomsky (1965) argues that a child is exposed to a limited amount of language input, but this does not stop him from acquiring the competence of his first language, and in fact, a child acquires his first language at a speed that an adult learner of a second language will never be able to catch up with. He suggests that the language faculty is innate or there is no explanation to where and how a child acquires the languageôs grammar.

This view has been challenged by many philosophers (e.g., Cooper, 1975; Steinberg et al., 2001) who state that the innate component should not be described in any ñepistemic termsò. Steinberg et al. have questioned the implication of the statement that a child has the benefit of UG in acquiring a language ï does it mean that an adultôs UG dies out with age? The answer is false. Moreover, he argues that if a child was born in a bilingual family of two totally different language systems (say, English (SVO) and Japanese (SOV), it is difficult to see what the initial parameter setting for this child would be. Empirical evidence shows that a child can acquire two languages fluently by the age of four. Early bilinguals do not necessary have a dominant first language.

Moreover, the red question-marks in Figure 9 question the interference of the world knowledge, situational knowledge, etc. with the architecture. Chomskyan scholars suggest that this interference does not affect the language faculty, but rather in the level of the conceptual-intentional system. However, considering the use of the sentences below:

17) A: ñ*Nice to meet you!*ò
    B: ñ*You too!*ò

When A says the rather-incomplete sentence (i.e., the subject is missing), his intention is not ambiguous within an English speaking environment. When B answers ñYou too!ò, no one would interpret it as ***You*** *also think it was nice to meet me*. Now how does B know that ñyou tooò should be the lexicon to be picked up, but not ñme tooò or anything else? Without world and situational knowledge, the conceptual-intention of Bôs conversation is impossible to reach. If non-linguistic information has to be used to assist the transformation between the articulatory system and the conceptual system, then the *language faculty* is not really pure-linguistic.

This is the key motivation of contemporary psycholinguistic research. If this innate faculty is pure-linguistic, then logically the most important information used in sentence processing, especially in the early stage, should be pure-linguistic; moreover, there should also be some universal parsing strategy. If the faculty is not only linguistic-related, but also mixed with other information, then the existence of the UG could be questioned, and there is no guarantee that the linguistic information should be processed prior to others. In the next section, some of the most important sentence processing models will be introduced.

Chapter 3:
Some Psycholinguistic Models

In a sense, development in psycholinguistic research trails theoretical development in linguistics. Psycholinguistic studies do not attempt to focus on any specific language system (but more general language features), however, each individual study is based on some specific language(s). It should be encouraged to follow those applied linguistic rules on each specific language. With the reviewed rules in Chapter 2, the current thesis attempts to report studies applying principles in the linguistic system.

The development of contemporary syntax theory had already begun in the 1950s, but early psycholinguistic studies concerning this matter only started in late 1960s (e.g., J. A. Fodor & Garrett, 1967; J. A. Fodor et al., 1968). In the 1970s, Holmes (1979) criticised the delay in the awareness of the importance of syntactic processing in the psycholinguistic field:

> ñGiven the initial interest in syntax triggered by Chomskyøs (1957, 1965) theories of language, it might seem somewhat surprising that psychologists have come up with so few lasting generalizations about this issue.ö (Holmes, 1979, p. 227)

This is understandable because psycholinguistic research is based on linguistic theories. For example, early parsing models should specify whether the assignment of structure is to the D-Structure or S-Structure, as this is the key distinction in the syntactic theories before the 1980s. However, few researchers besides Holmes have made this point very clear. Holmesô model follows Figure 2 above. An experiment was conducted where ambiguity was manipulated to determine which structure(s) has(ve) effects on sentence comprehension:

18) a. We heard about the finding of the geologists.

b. The spokesman confirmed the killing of the terrorists.

19) a. The councillors discussed the difficulties with the mayor.

b. The teachers arranged the discussion with the migrant parents.

20) a. The judge who was presiding patiently questioned the witness.

b. The girl who had watched the traffic nervously crossed the street.

Sentences 18) are D-Structurally ambiguous, with 18a) biased to the subject structure and 18b biased to the object structure. Sentences 19) are S-structurally ambiguous: in 19a, the PP *with + Noun* is considered to modify the verb, but in 19b, the PP *with + Noun* is considered more as the proceeding nominal. Sentences 20) are mixed-ambiguous, as being both D- and S- Structurally ambiguous. On-line reading time of a complete sentence was measured.

The finding was that the ambiguity in D-Structure had no (significant) bias effects, but the bias effects of the S-Structure were reliable. Moreover, sentences with non-standard bias were more time-consuming than biased ambiguous sentences.

Further studies carried out by Holmes and colleagues (Langford & Holmes, 1978) investigated context effects on the surface structure. Their results indicate that context information makes the interpretation faster.

Trying to make use of the existence of *D-Structure* in interpreting their results, Holmes postulates that a model of comprehension has to include several stages:

21) HolmesôModel:

Stage 1. Assigning a surface structure;

Stage 2. Semantic analyses;

Stage 3. Integrating word meanings and the results of the S-Structure analysis, using D-Structure to check the comprehension accuracy.

Holmesômodel is not the first or the most influential model in psycholinguistic history. The assignment of D-Structure remains mysterious since none of the stages really process D-Structure. However, this model attempts to distinguish the processing of S-and D- structures, which clearly defines the term of ñstructure-processingò under the linguistic theory at that time. Although there is no longer a distinction between D- and S- Structure in current linguistic theory, early psycholinguists would not anticipate this, so clarifying two structures should be essential when modelling parsing strategies. Nevertheless, there appears to be a consensus that initial parsing begins with S-Sructure (Annotated to S-Structure: Gorrell, 1995; Marcus, 1980; Prichett, 1992; see Townsend & Bever, 2001, for a detailed introduction) from Holmesô results and the fact that humans approach sentential information as its surface structure. However, syntax theories at that time (as in Figures 2 and 3) assume S-Structure cannot stand alone without D-Structure, as D-Structure is mapped to S-Structure (or S-Structure is the medium between D-Structure and meaning) and the relationship is not bi-directional. Thus without dealing with D-Structure, the theories were incomplete. If one suggests that the linguistic theory is wrong (which seems to be the case by later development of the Universal Grammar theories), then there is no theoretical support that syntactic processing has to be the initial stage. Although some (e.g., Marslen-Wilson, 1976) have suggested that sentence processing is ñonline and interactiveò, structure-first models have been widely developed and become a dominant view in the psycholinguistic field since the late 1970s.

This thesis examines the time-course of parsing in sentence processing. As linguistic theory has abandoned the distinction between D-Structure and S-Structure, the current thesis will not focus on this issue and Holmesô model any further. Next, contemporary models of sentence processing will be introduced in an order according to the stages where parsing occurs[19].

3.1 Parsing Prior to Other Processing Types

The best phrase to describe structural models is ñsyntax proposes, semantics disposesò (Crain & Steedman, 1985). Following Chomsky's UG theory, a group of linguists and psycholinguists (Atkinson, 1992; 1982; Chomsky, 1986; 1995; Cook, 1993; Pinker,

[19] In this chapter, only the model will be introduced. The empirical evidence will be introduced in Chapter 3 after the psychological methodologies are explained.

1984; Radford, 1997, etc.) believe that certain structural principles of language are universal and innate, which indicates that it is possible for some of the human parsing procedures to be universal and innate as well. There arguments follow the famous Chomskyan ñsentencesò (Chomsky, 1957):

22) a. Colorless green ideas slept furiously.
    b. Ideas furiously green slept colorless.

22a makes people feel more comfortable than 22b because 22a follows a set of rules that is used in a real natural language (i.e. English). The Chomskyan School proposes that this implies that the structural processor is essential or that at least it plays an important role in sentence processing (Cairns, 1999). In agreement with this point of view, the following structure-based models have been proposed.

3.1.1 The Garden-Path Theory Group and Serial Models

The *Garden-Path Theory Group* (GPTG) includes the traditional *Garden Path* (GPT) (Frazier, 1978; 1987), *Relativized Relevance* (RR) (Frazier, 1990a; 1990b) and *Construal* (Frazier & Clifton, 1996) theories. This group is one of the most cited and discussed universal structural theories in sentence processing. It is considered to be a purely structural universal hypothesis. It assumes that processing resources are limited and the human language processor uses only syntactic information.

◆ The Garden-Path Theory (GPT)

GPT proposes that when a word is received, the parser selects a single structure that can absorb the word into the previous structure, while minimizing changes to the structure being constructed; when the choice is proved to be wrong, it will activate the parser to re-analyse the structure. Moreover, only a single syntactic analysis is initially pursued during the processing and any semantic or contextual information is not involved at all in the selection of the initial syntactic analysis.

This theory is composed of two key principles:

23) Garden Path Theory (GPT):

a. *Minimal Attachment* (MA) ï ñDo not postulate any potentially unnecessary nodesò (Frazier, 1987, p. 562),

b. *Late Closure* (LC) ï ñIf grammatically permissible, attach new items into the clause or phrase currently being processedò (Frazier, 1978; 1987, p.562; Frazier & Clifton, 1996; Frazier & Rayner, 1982; cf. Right Association, Kimball, 1973).

The GPT is very successful in explaining temporally ambiguous sentences such as sentence 2) in Chapter 1 (P.10): It is the reduced relative clause (*that was raced*) that causes a misanalysis because the parser takes it for granted that the VP *raced* could attach to the first NP *the horse* as soon as *raced* is received (i.e., according to MA, the parser attaches *raced* to *the horse* because it is the simplest grammatically possible analysis; according to LC, the most recently processed grammatically possible phrase is the TP headed by the NP *the horse*.) Thus, sentence parsing is completed at the stage of ñ*The horse raced past the barn”*. However, when *fell* is received, the original analysis is proved to be wrong. It costs time and energy to re-analyse and assign *raced past the barn* to be the reduced relative clause that modifies the NP *the horse*. This is consistent with empirical results indicating that people's reading time is longer for 2) than for the full relative sentence as 24) below:

24) The horse *that was* raced past the barn fell.

When explaining how parser works, the model proposes another two stages called *Sausage-Machine* (Frazier & Fodor, 1978):

25) Stage I (bottom-up): A *preliminary phrase packager* (PPP, or ñsausage machineò) generate words into phrases, under MA and LC, etc.;

Stage II (top-down): A *sentence structure supervisor* (SSS) adds higher nodes to connect the phrases the PPP made into a complete phrase marker.

This means not only the parser is initially used in human sentence processing, but parsing itself can also be divided into two serial stages. The parser initially generates only one possible structure, but in later stages, other possibilities will be taken into concern. In this sense, a sentence like sentence 1) in Chapter 1 should be interpreted as *the iceberg* is *on the shoulder of a giant*, in Stage I under PPP, but the ambiguity happens at Stage II, where the SSS has found two possible notes for the PP at this time.

GPT is a pure structure model, as the processing of information other than structures is not included primarily at all.

Linguistically, *MA* cannot always be right under the binarity law in MP introduced in 16) above. Moreover, GPT cannot explain other language phenomenon such as the sentences in 26):

26) a. Centre-embedded sentences: The rat the cat the dog chased ate died.

b. Early context effects in identical structures: cf. i & ii (sentences used by Taraban & McClelland, 1988, Experiment 2)

i. The janitor cleaned the storage area with the broom because of many complaints. (Faster)

ii. The janitor cleaned the storage area with the solvent because of many complaints. (Slower)

c. Crosslinguistic difference (sentences used by Cuetos & Mitchell, 1988):

i. The journalist interviewed *the daughter* of *the colonel* who had had the accident. (English ï *the colonel* had the accident)

ii. El periodista entrevisto a *la hija* del *coronel* que tuvo el accidente. (Spanish ï *la hija* (the daughter) had the accident)

The current thesis focuses on the issue similar to 26c).

◆ The Revised Garden-Path Theory (RR)

To defend the *Garden-Path Theory*, Frazier refined her theory with adding a new principle, *Relativized Relevance* (Frazier, 1990b) which suggests that all other information (e.g., grammatical, informative and appropriate information in the discourse) being equal, the parser will prefer construing a phrase ñas being relevant to the main assertion of the current sentenceò (Frazier, 1990b, p. 321). In other words, parsers process a sentence initially based on syntactic principles such as *Late Closure*; then consider other information at a post-syntactic phase, using principles such as *Relativized Relevance.*

One essential problem of this revision of GPT is that RR has changed the grammatical assumption in *MA* (as in 23a), as criticised by Kamide (1998). In the traditional GPT (Frazier, 1978; 1987; Frazier & Fodor, 1978), MA is based on the TG (Transformational Grammar), in which the phrase marker is ñstraightforwardò. However, in RR (Frazier, 1990b), MA is based on GB, specifically, the X-bar theory, where a distinction between arguments and adjuncts has been specified. If the argument phrase happens to be a verb phrase, the predictions from GPT and RR remain consistent. However, if the argument and the verb phrase are not identical, the two versions of the theory will contradict each other.

Consider sentence 1) in the introduction: ñ*I saw an iceberg on the shoulder of a giant*ò again. Let us only focus on the MAôs prediction (ignoring the LC for the moment). Traditional GPT predicts that the VP attachment only needs two nodes (S and VP), whereas the NP attachment needs three nodes (S, VP, and NP); so VP attachment is preferred. However, GB makes a crucial distinction between adjuncts and complements. Thus, the node numbers of the two conditions between the PP and the modifiee are the same. As a result, MA in RR does not differentiate between the two alternatives in sentence 1), leaving only LC to determine the attachment. The structures constructed by TG and GB are illustrated below:

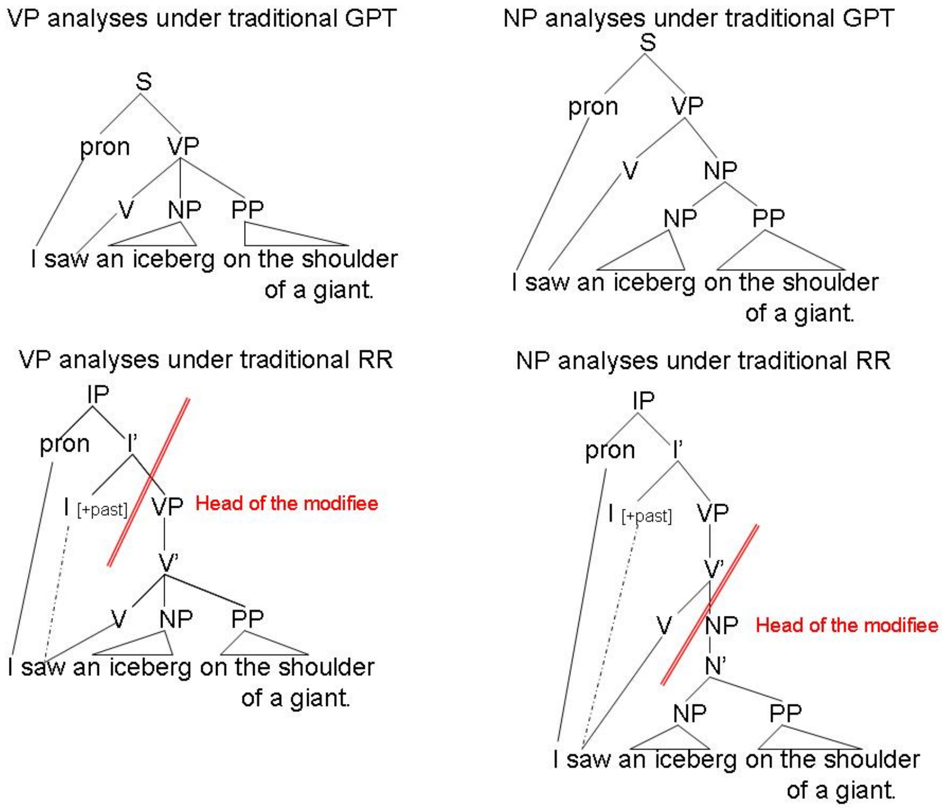


Figure 10 Comparison between the predictions of GPT and RR

GPT and RR do not allow any under-specification in the syntactic relations encoded in the phrase marker. Adding a new node in the tree will change the syntactic relations between nodes that were already there. Revision for the syntactic structure will lead to longer parsing time.

However, in order to explain the phenomenon in 26c), further development to the minimal attachment theory is required.

◆ Construal

Construal (Frazier & Clifton, 1996) is a top-down, principle-based, single-analysis module. It divides syntactic structure into two categories, *primary relations* and *non-primary relations*. The primary phrases include subject, predicate, sentential complements, and syntactic positions that occupy argument positions such as agent, goal, instrument and theme. The non-primary phrases are elaborations of arguments (e.g., relative clauses, prepositional phrases and adjunct phrases). Primary relations are parsed as under GPT (essentially MA); those and only those non-primary relations which are neither the subject and main predicate of any clause, nor the complements and obligatory constituents of primary phrases, are described by the theory in 27) (cf. Frazier & Clifton, 1996, pp. 41-42; 1997, p. 279):

27) Construal Principle:

a. Analyse an input, X, as instantiating a primary phrase if possible.

b. Otherwise *associate*[20] X into the current thematic processing domain (the extended projection of the last actual theta-assigner).

To account for the *attachment* preferences in 26c), Construal first categorises this type of ambiguity as belonging to the non-primary relationship. Thus, 27a) is not satisfied so 27b) should be applied. This means that there should be no initial preference in any language to attachment bias; the relative clause is *construed* / *associated* (not *attached*) to the current processing domain, while the parser would use all the other information available to interpret the meaning of the sentence. It is said that this could avoid having to stipulate the reanalysis of a parsing decision that would launch in the absence of conflict (Fernández, 2002).

◆ Summary

The development of the most pure-structure sentence processing theory has shown that in practice, sentence parsing cannot be simply based on structural cues. Although using words and syntactic rules, the primary construction of sentences is determined by a constituent structure module, a thematic processing module has to be operating simultaneously to indicate the most plausible host for the association of the non-primary phrases.

Figure 11 attempts to display the contemporary GPTG, which was originally designed by Townsend and Bever (2001):

[20] According to Construal, only primary relationships can be ñattachedò, whereas the non-primary relationships are ñassociatedò.

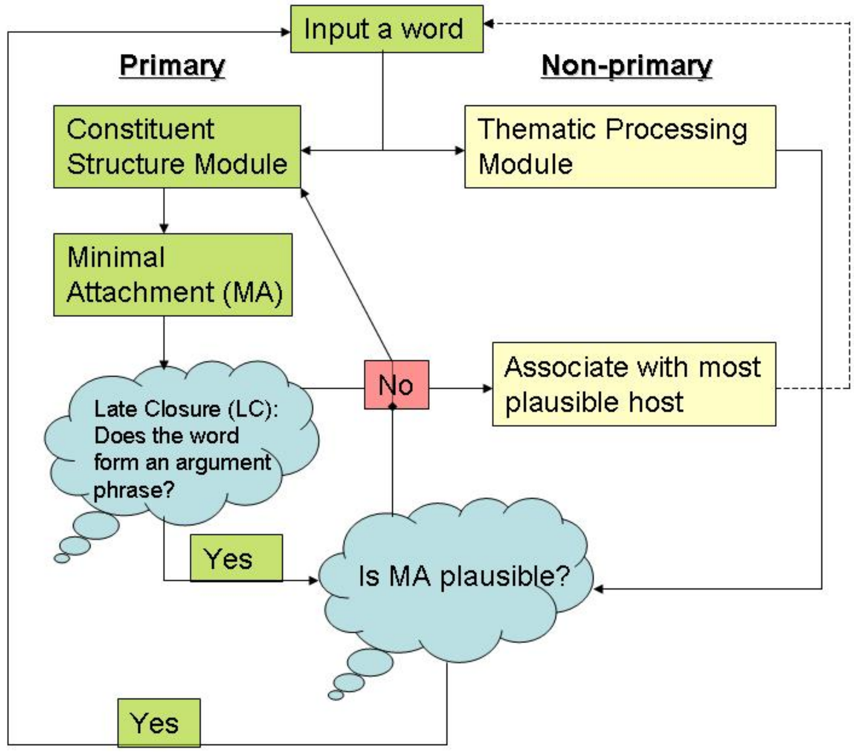


Figure 11 A Model of the Garden Path Theory Group

3.1.2 The Structural-competing Theories

Like the GPTG, structural-competing theory groups propose that parsing is the initial stage of sentence processing. Unlike the GPTG, they are not based on the same theoretical framework, and the initial construction of structure can have several possibilities. The structures then compete with each other based on parameter settings in each language. This thesis will focus on two models here.

◆ Recency / Predicate Preference Theory (*RP / PP*[21])

The *RP / PP* theory can be seen as a proposal resembling Frazier et al.ô theory but also considers parameter-setting, proposed by Gibson and colleagues (Gibson, 1991; Gibson, Pearlmutter et al., 1996; Gibson & Schütze, 1999; Gibson, Schütze et al., 1996; Miyamoto et al., 1999). *RP / PP* proposes that preference ranking is affected by two

[21] In this thesis, when referring to the Predicate Preference (*PP*), *PP* will be italicised to differentiate it from the abbreviation of prepositional phrase (PP).

factors, *Recency Preference* (*RP*) and *Predicate Proximity* (*PP*).

28) a. Recency Preference (*RP*) (Gibson, Schütze et al., 1996, p. 26):

Preferentially attach structures for incoming lexical items to structures built more recently;

b. Predicate Proximity (*PP*) (Gibson, Schütze et al., 1996, p. 41):

Attach as close as possible to the head of a predicate phrase[22].

The former is a variant of Late Closure, which postulates preferential attachment of incoming lexical items to the more recent sites; the latter is an extension of *Relativized Relevance*[23], which postulates preferential attachment to the head of a predicate phrase. Thus, for sentences with structure similar to that of 26c, *RP* predicts an attachment to the NP *the colonel*, while *PP* predicts the contrary, i.e. an attachment to the NP *the daughter*. The two factors compete with each other, and the final choice of preferred host falls to the one that entails the lowest processing-load cost, which is supposed to be measured by the formulae below:

29) a. Recency Preference (*RP*):

$C(X) = f(Y)\text{ŏ}X_{RP}$

b. Predicate Proximity (*PP*):

$C'(X) = g(Z)\text{ŏ}X_{PP}$

Here C(X) or C'(X) = the cost of the processing load of *RP or PP* from an attachment of structure X at one site; Y = the number of more recent words that would also allow an attachment of X; Z = the number of sites that is structurally closer to the head of a predicate phrase that allow an attachment of X; X_{RPorPP} = the initial cost increment (Gibson, Schütze et al., 1996, p.43; p.45). Thus, if two sites, $C_{NP1}(X) + C'_{NP1}(X) > C_{NP2}(X) + C'_{NP2}(X)$, NP_2 attachment is preferred, as in English; and vice versa, as in Spanish. The theory assumes the *RP* to be the universal factor since it matches the general properties of the working memory, while assuming *PP* to be the parameterised

[22] Note that this means it is not necessary to attach directly to the VP, but close to it.
[23] The *Predicate Proximity* allows attachment to *any* S node, while *Relativized Relevanc* restricts to the root S node. Thus, the *Predicate Proximity* has the advantage of the possibility in explaining the VP- and NP-site attachment preferences. (Gibson, Schütze et al., 1996)

factor, which differs from one language to another. Gibson et al. propose that *f* is a decay function and *g* is a single-step function[24].

This *RP / PP* theory satisfactorily explains the U-shape preference (NP_3, NP_1, NP_2) of the three-site ambiguities (as in (16) below) both in Spanish and English, as well as languages studied later such as German (Hemforth et al., 1998) and Japanese (Miyamoto et al., 1999):

30) ... the lamp near the painting of the house *that was damaged in the flood*
NP_1 NP_2 NP_3 RC

RP predicts NP_3 attachment of the RC, meaning that NP_3 *the house* is favoured, while *PP* predicts NP_1 so that NP_1 *the lamp* is favoured. Since neither of these factors will predict a preference for NP_2, the middle site is the least favoured one. And since C(X) decreases as the number of sites grows while C'(X) remains a one-step function; the cost of low attachment is smaller than the highest attachment in both languages[25].

RP / PP assumes that the parameters of *PP* depend on (or are affected by) “word order” freedom. For example, a language with a freer word order, such as Spanish, is more affected by the predicate, while a language with a less free word order, such as English, is more likely to be local. One theoretical criticism to *RP / PP* is that the theory is “somewhat speculative”(Gibson, Pearlmutter et al., 1996, p.47). The functions in 29) are based on assumptions with little independent evidence. Moreover, it has been criticised for “making no commitment to a specific definition of verb / argument distance” (Mitchell & Brysbaert, 1998). This means that in a free-order language, the “position” of an argument is uncertain. If this is the case, then the formula should concern a parameter setting that addresses this “distance” issue.

---

[24] Since *Recency* follows working memory, *f* is a decay function. Since if *g* increases more rapidly than *f*, then C'(X) will add more cost to the lower sites than C(X) to the higher one, so the highest sites will become incrementally favored. The only remaining possibility is that *g* changes more slowly than *f*, and an ideal for *g* is to be a single-step function. (Gibson et al. 1996)

[25] As this chapter focuses on the theoretical aspects, how Gibson et al. proved this empirically will be introduced in the Chapter 5.

◆ The Semantics-Oriented Unification-based Language (SOUL) System and Attachment Binding (AB)

The SOUL system was originally proposed by Konieczny and colleagues (1997). It is based on the head-driven phrase structure grammar (HPSG) (Pollard & Sag, 1994; cf. Pritchett, 1992). The original SOUL system does not generate ñcompeting interpretationsò, and it has two key assumptions about sentence processing: a. the processing system relies on the rich information in the lexical level (cf. Pritchett, 1992); b. the parser generates many possible interpretations at the same time so that the interpretations can compete with each other (cf. the *RP* / *PP* theory (Gibson, Pearlmutter et al., 1996; Gibson & Schütze, 1999) introduced above).

At the sentence level, the attachment of a phrase follows rules below chronologically (adopted from Kamide, 1998, p. 71):

31) SOUL Principles:

a. Head Attachment: If possible, attach a constituent *g* to a phrasal unit whose lexical head has already been read (cf. *PP* (Gibson, Pearlmutter et al., 1996) or *RR* (Frazier, 1990b));

b. Preferred Role Attachment: If more than one potential unit remains as a potential attachment site, attach *g* to a phrasal unit whose head provides a required (obligatory) or expected (optional) thematic or time / place role for *g*;

c. Recent Head Attachment: If further attachment possibilities for *g* remain, attach *g* to the phrase whose lexical head has been read more recently (cf. *RP* (Gibson, Pearlmutter et al., 1996), and LC (Frazier, 1978, , 1987)).

Under SOUL Principles the preference of sentence 1) would be ñthe iceberg is on the shoulder of a giantò. This is because initially the principle 31a) leads the reader to choose attachment for the PP *on the shoulder of a giant*, but found both the VP *see…* and the NP *a giant* are the potential heads that have been read; then principle 31b) cannot be used to decide because both the VP and NP can expect an attachment of PP; in the end, 31c) rules over and the more recently read head, the NP wins.

In the case of sentence 26c above, similar predictions can be given to account for the English attachment choice. But the Spanish data throws doubt on this account.

Thus, a revised SOUL system, the *Attachment-binding* (AB) theory, was proposed by Hemforth and her colleagues (Hemforth et al., 1998) in an attempt to account for the differences in RC-attachment preference (as in sentence 26c) cross-linguistically. It proposes that two types of possible analysis, NP_1 and NP_2-attachment preferences, have been generated at the same time. As in the original SOUL system, AB draws attention to the lexicon (category) entries. The thematic and pragmatic habits, (forms of anaphoric reference, topicalisation of the preverbal position, etc.) have been considered as language parameters that can account for the language differences. For example, in languages where the RC is introduced by a relative pronoun, attachment preferences have a tendency toward the prominent discourse entities, which constrain anaphoric binding of the pronouns. Languages such as German and Spanish, which have a strict usage of the relative pronoun (it is rarely omitted in those languages), are more likely to bias for NP_1, since this site is the salient discourse entity; yet in languages such as English, where the omission of the relative pronoun is freer, the anaphoric binding is weaker, and NP_2 is favoured. This account therefore has another name: *Relative Pronoun Drop Model*.

◆ Summary

In structural-competition models, sentence structures following universal linguistic properties are generated in the initial stages of the sentence processing. The selection of the final attachment can be determined by cognitive processing capacities, or factors specific to individual languages.

For local ambiguity (i.e., garden-path sentences, as sentence 2) above), if the new constituent cannot be attached (i.e., indicating the analysis has gone wrong), re-analysis via competition[26] should be allowed. However, for the global ambiguity (as 1) and 26c), once the final decision has been made, there should not be any change in the structure.

[26] This is because for the purpose of re-analyses, the newly constructed structures should compete with the old mistaken one, although, of course, the new one will win.

This is illustrated by Figure 12 below:

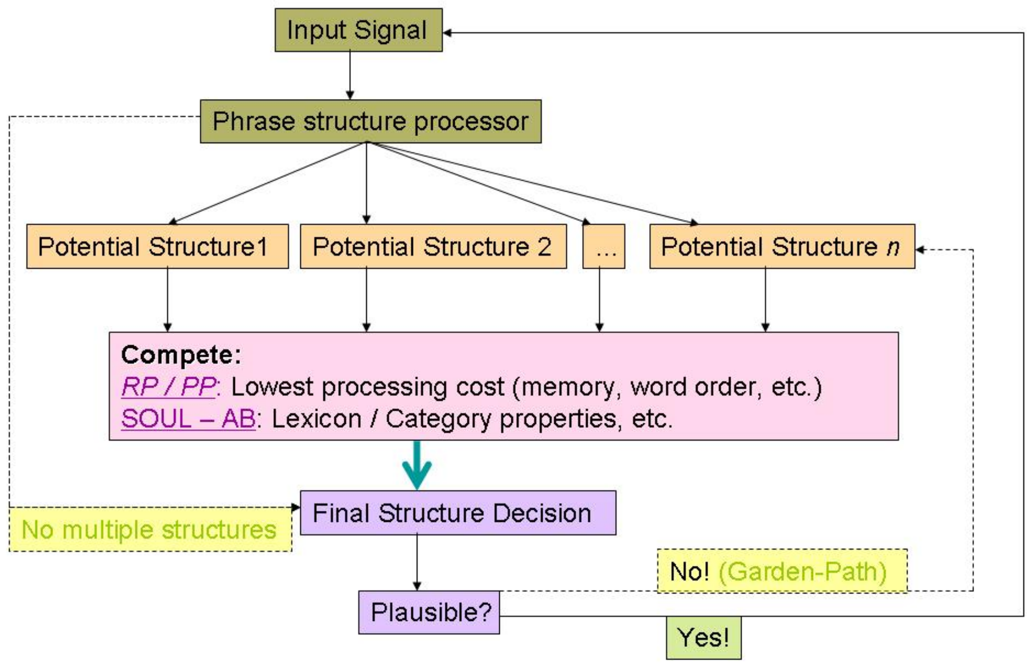


Figure 12 Illustration of Structural-competition Theories

3.2 Parsing Parallel with Other Processing Types

Unlike the structural-first models, the *Parallel Models* claim that structural cues are used at the same time as other cues, such as semantics, context, frequencies, thematic fit, animacy and many others.

The parallel models have been strongly inspired by the development of the connectionist models in cognitive science and artificial intelligence. Some earlier development in the parallel models (e.g., the connectionist network by Elman, 1993; the interactive activation model by McClelland & Rumelhart, 1981; and the parallel distributed processing model by Seidenberg & McClelland, 1989) are mostly used in simulations and predictions in language learning.

Models of this type attempt to explain the rule-based performance from a statistical point of view ï ñrule-governed performance without any actual ruleò (McClelland & Rumelhart, 1981, p. 375) ï which is not really specific to language usage, but rather applies to general human learning. They propose that people learn by gathering all relevant information. In terms of language learning, it ñrequires tracking correlations and covariation across multiple types of linguistic information within and across modalities (e.g., a speech signal and the context in which it is uttered)ò (MacDonald & Seidenberg, in press).

The basic assumptions are: i). Statistical ways of learning are the basic means to general learning; ii). Only relevant information is used in the computation; the model is highly constrained to perceptual capacities (i.e., what people can hear, read, or say), but the combination of this information is not a linear relationship; iii). The information in a network is encoded by ñweightsò, which determines the output in performing tasks (i.e., input weights are computed into output, and the output should satisfy the constrained information in i)). Between input and output, the weights can be tuned by feedbacks.

There are many distinct statistical models (e.g., Bates & MacWhinney, 1989; Elman, 1993; MacDonald et al., 1994; McClelland & Rumelhart, 1981; McRae et al., 1998; Mitchell, 1994; Seidenberg & McClelland, 1989). They differ in terms of the input information weight, feedback type (positive or negative), activations in training, etc.. Here two accounts are introduced that are relevant to the main topic of the current thesis: the frequency account (Cuetos & Mitchell, 1988; Cuetos et al., 1996; Mitchell et al., 1995)[27] and constraint-based competition-integration model (McRae et al., 1998; Spivey-Knowlton, 1996).

3.2.1 The Exposure-based Tuning Hypothesis (ETH)

Since the crosslinguistic differences in the RC attachment preferences break the syntactical-universality (as in sentence 26c) above, the difference between Spanish and English speakers), a pure statistical hypothesis, Exposure-based Tuning Hypothesis (ETH) has been suggested. The ETH (Corley, 1995; Cuetos & Mitchell, 1988; Cuetos et

[27] In fact, ETH was not proposed as a model against syntax-first theories. It postulates that the goal of the frequency information processing is to access the syntax structure, instead of the meaning of the sentence directly. However, the current thesis does not focus on the ñgoalò of processing, but rather the ñfactò of processing. Moreover, it is arguable that if the frequency were to be used only for syntax selection, it is unclear why it is difficult to process a simple structure of very low-frequency semantic (but high-frequency word form) lexical inputs, such as *The man ducks many parties*. Therefore, the current thesis categorises the ETH into the Parallel model.

al., 1996; Mitchell & Brysbaert, 1998; Mitchell et al., 1995) assumes that the choice of sentential analysis is based on the reader/listenerôs experience of resolution of previously encountered similar ambiguous structures. The full description is in 32) below:

32) ña. The HSPM (human sentence processing mechanism) is a probabilistic serial, or weighted parallel, device, in which an initial commitment is specified as the only, or favoured, analysis at the point at which an ambiguity is encountered.
b. The *only* determinant of an initial commitment for any ambiguity is that, given that any analyses are syntactically allowable, the analysis chosen will be the one to which an individual has previously most frequently been exposed.ò (Corley, 1995, p. 72)

In other words, a parser with processing tendencies (e.g., attachment preference) is entirely determined by the environment of a speaker. Thus, the syntactic rules and principles prevalent in the input material may lead the parser to imitate those rules/principles, but this does not imply that the parser uses the rules/principles themselves. Moreover, a cumulative frequency input change will cause a short-term bias change individually. However, some arguments (Fernández, 2002) point out that ETH is not radically different from the structural accounts discussed in the previous chapter. It relies heavily on structural principles for determining the attachment ambiguity except in ñcertainò constructions, such as RC attachment.

The simplified account for the 26c) type of structural ambiguity resolution is that the preference of attachment should be relevant to the statistical value, e.g., the corpus analyses. Mitchell, Cuetos and Corley (Mitchell et al., 1992) have used a small-scale corpus study comparing English and Spanish resolution for ambiguity with two NP-sites, and have found that the results are identical with those obtained in experiments with native speakers, and that in analysis of corpora, NP_1 is preferred in Spanish and NP_2 is preferred in English. Similar testing in French two-site RC attachment has also given results in line with the predictions of the theory (mentioned in Brysbaert & Mitchell, (1996)). In addition, the theory gives a reasonable explanation for differences in individualsô preferences in the same language.

A key question to the ETH is the definition of ñfrequencyò. In other words, how does the processor know what size grouping is ñfrequent enoughò to activate the decision of construction. This has been referred as *the grain-size problem* (Mitchell et al., 1995). Mitchell et al. (1995) propose that the problem is related to higher-order analyses (i.e., what is the appropriate level to store the structural information). Fine-grained models consider detailed individual properties of sentences (cf. MacDonald et al., 1994), but they may face the ñsparse data problemò if too many categories are involved. Thus, although coarse-grained models make inaccurate decisions, Mitchell et al. (1995) suggest that they are used in initial processing. Specific choices of the grain-sizes can be chosen by the processor in a relatively later stage.

The figure below lists some possible grain-sizes:

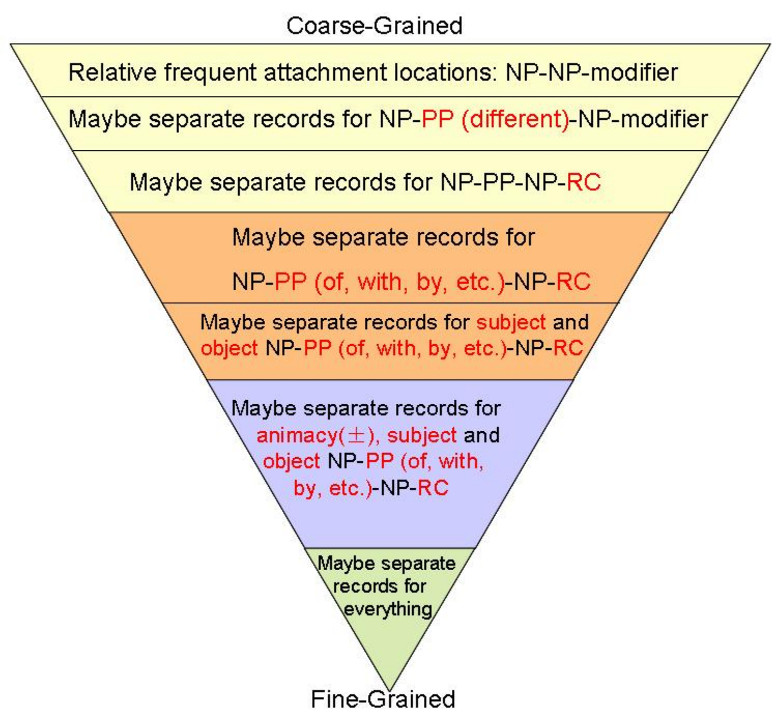


Figure 13 Possible Grain-sizes

3.2.2 The Competition-integration Model

Unlike the ETH model, competition models (e.g., Bates & MacWhinney, 1989; MacDonald et al., 1994; McClelland & Rumelhart, 1981; McRae et al., 1998;

Spivey-Knowlton, 1996) propose that sentence processing evaluates the syntactic alternatives basing on both linguistic and non-linguistic constraints. All available relevant information is then integrated by the comprehension system to compute the best satisfactory interpretation of a sentence to fit those constraints. It is claimed to ñpredict that conceptual information that is correlated with different syntactic alternatives can play a central role in guiding even the earliest moments of language comprehension in general, and ambiguity resolution in particularò. (McRae et al., 1998, p. 284).

Among the variation of the competition models, McRae et al.ôs competition-integration model (McRae et al., 1998) proposes that the input of information is on the word-by-word (or phrase-by-phrase) bases. Each new input (words or segments) brings new competition to the processing system. In McRaeôs model, the initial weights are set differently according to each constraint (cf. Spivey-Knowlton, 1996, initial weights are identical.). Taking another garden-path sentence (similar to 2)) as an example.

33) a. *The crook* arrested by the detective was guilty of taking bribes.
    b. *The cop* arrested by the detective was guilty of taking bribes.

The constraints taking into account for process the above two sentences by McRaeôs model are: initial NP thematic fit, Main clause bias, verb tense/voice, by-bias, the-bias, and agent NP thematic fit. The NP *the crook* is considered a good patient of the verb *arrested*, and consequently a bad agent; and vice versa for the NP *the cop*. The verb *arrested* can be a past-tense, which is biased to a main clause, or a past-participle, which is biased to a reduced relative clause. McRae et al. computed the weight (un-normalised) of different constraints based on either corpus data or empirical results as in the table below:

Table 2 Initial constraint inputs used in the simulation of McRae et al. (1998, p. 293)

Constraint	Good Patients		Good Agents	
	Relative	Main	Relative	Main
Initial NP thematic fit	5.0 (.1)[28]	1.0 (.1)	1.5 (.1)	5.3 (.1)
Main Clause bias	.08	.92	.08	.92
Verb tense / Voice	.67 (.03)	.33 (.03)	.67 (.03)	33 (.03)
By-bias	.8	.2	.8	.2
The-bias	.875	.125	.875	.125
Agent NP thematic fit	4.6 (.2)	1	4.6 (.2)	1

The model suggests that the processing is circular. Weights are computed for the activations on each constraint for the possible interpretation choices. When the activation of one type of interpretation reaches a certain criterion, the competition ends. When all constraints favour one type of interpretation, the competition is fast; when constraints are more balanced, the processing reaches an indecisive stage, thus, processing difficulty occurs.

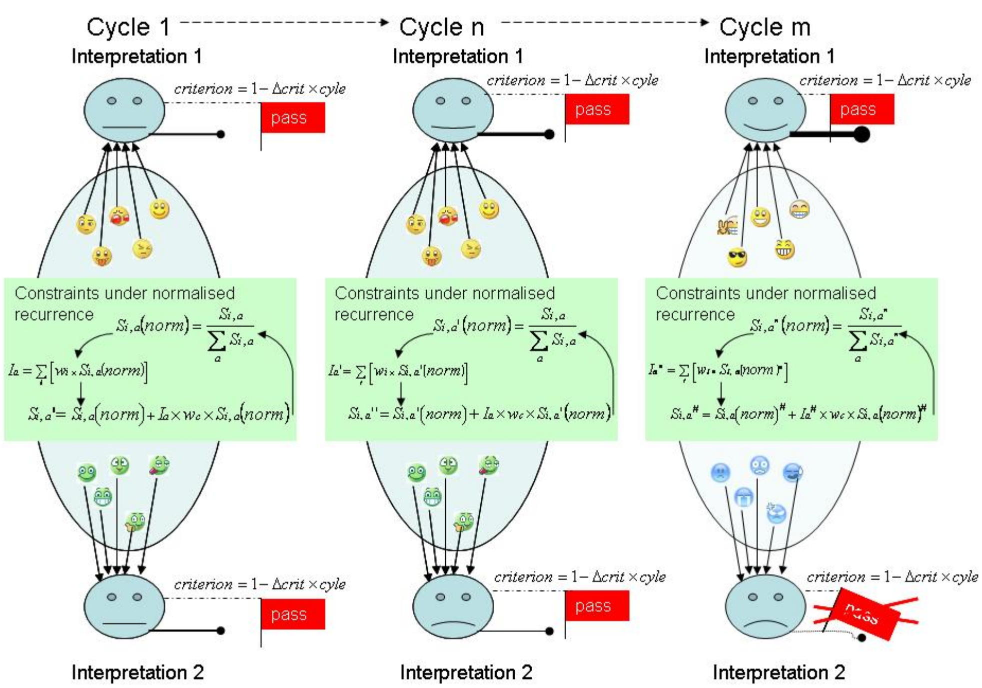


Figure 14 McRae's Model's illustration

[28] In the bracket, there are standard errors.

Figure 14[29] illustrates that for each segment, at each cycle, the competition model employs two systems working simultaneously to choose a successful interpretation: a. an increase in the competition power (normalised recurrence), and b. a reduction in the dynamic criterion level. After each cycle, the likely-to-win interpretation will be strengthened by the integration of the supported constraints, whereas the likely-to-lose interpretation will be weakened. The criterion for selecting one interpretation is dynamic, and is decreased after each cycle, which should ease the decision making of interpretation.

The empirical work of simulation and human data matching will be further discussed in Chapter 4. However, a couple of theoretical deficiencies of this model have to be pointed out. First, it is not clear what constraints should be included (cf. the grain-size problem). Without knowing all the constraints and their weights, there is no way to falsify this model (Pickering & van Gompel, 2006). Second, the constraint-satisfaction is based on the assumption that the activation spreads through a network modulated by weights on the connections, it does not require a resource-limited active search process (MacDonald & Seidenberg, in press). However, language procedure is by all means a type of cognitive and perceptual related activity, and it has been proven by normal and brain-impaired human data that syntactic structure and meaning play important roles in memory for speech (Slobin, 1971). Although the model accounts for constraints on resources (such as memory) by carrying over only a portion of the weights from the previous phase, it cannot account for the variation of memory accessibilities (i.e., direct access and slow search process). Moreover, some research (e.g., McElree et al., 2003) has shown that the memory retrieval in sentence processing follows the direct mechanism. The current constraint-based models do not specify these.

3.2.3 Summary

In accounting for language processing at the sentence level, the parallel models are more *lexicalist*. They assume that syntactic components are actually stored within each

[29] As this thesis does not attempt to focus on modular simulations, the formulae for calculation the activities and the dynamic criterion in each cycle will not be explained in detail.

lexical item (MacDonald et al., 1994). There are also slight variations in this type of models in terms of the time course when the constraints are considered. E.g., Simplicity in Structure Building (SSB) (Gorrell, 1995) assumes that as the parser may work most efficiently by adopting a simplest structure allowed in the construction, information of plausibility can be used. Incremental Interactive Model (IIM) (Altmann & Steedman, 1988) suggests that although all grammatical structures are generated parallel modularly, (cf. Gibson, 1991; Gibson, Pearlmutter et al., 1996; Hemforth et al., 1998), the parsing is interactive with semantic context effects. The Concurrent Model (Boland, 1997a, , 1997b) assumes that syntactic and semantic processing are not strictly ordered, but the syntactic generator is modular. A criticism towards this model is that the generation of the syntax alternatives occurs prior to other information, but the choice of the syntax is simultaneous to other information. However, there are also possibilities when the syntax alternatives are not generated, but rather activated (i.e., the structures do not need to be generated, but are already in existence) at the same stage as other information. Parallel models should specify whether the structure is constructed or merely activated.

These models (i.e., SSB, IIM, and Concurrent Model) are not statistical based. The common assumption of all those models is that parsing enters into the processing in the same time as other information processing.

3.3 Parsing Posterior to Other Processing Types

Syntactic processing being triggered at a later stage of sentence processing is no new statement. Before structuralism became a fashion with the boom of Chomskyan syntactic theories, syntactic processing was treated ñcursorilyò. For instance, Bever (1970) suggests that sentence processing, being within a normal perceptual processing group, is possibly carried out with little regard for actual structure; Schank (1973) also suggests that humans do not necessary compute syntactic analyses to understand a sentence. However, with the common acknowledgement of the existence of the ñlanguage facultyò and possibly a universal linguistic account, this type of proposal has become less acceptable. Thus, even though parallel models base their assumptions on all kinds of factors in the initial parsing stage, syntactic processing is still among them.

Nevertheless, language is a perceptual behaviour (Bloomfield, 1936), so we have to perceptually organise speech or text as made up of discrete sounds or graphs and higher-order units. Taking sentence 34) (Townsend & Bever, 2001, p. 158) as an example:

34) The horse races and wins.

It is clear from its written form that the sentence consists of five word units. However, consider this sentence being said by some one, the input has become:

35) Thehorseracesandwins.

Moreover, crossslinguistically, some languages do not even separate segments in their written form:

36) 那匹馬參加賽跑並且勝利了。

This means that the processor has to segmentise the input into separate units to be able to carry out further processing. This type of segmentation cannot be simply categorised as a lexical level of processing, because information involved in segmentation can be as complicated as prosody, context, semantics and structure, etc.

In this thesis, two disconnected theories are introduced. Their common scheme is that syntactic processing is affected by processing of other information.

3.3.1 The Implicit Prosody Hypothesis (IPH)[30]

A prosodic account is proposed by Fodor (1998) to seek resolution from another source of sentence parsing strategy. It proposes that the crosslinguistic differences in parsing attachment ambiguity lie in the differences of prosodic weight of the attaching constituent relative to the host constituent. It postulates the ñanti-gravity lawò (J. D. Fodor, 2002a, , 2002b), which states that when the attaching constituent is prosodically

[30] Similar to the ETH, IPH was not proposed directly for an LAST account. However, the fact is that this proposal claims that prosodic information has to be processed first to be able to continue to process syntax and meaning and hence being categorised in the LAST group in the current thesis.

heavy, the processor tends to attach it to the higher constituent in the syntactic tree, and vice versa. Under this anti-gravity law, the Late Closure effect can be interfered with.

Later, Fodor (1998; 2002a; 2002b) generalized her prosodic account in to the *Implicit Prosody Hypothesis* (IPH). Under the assumption that even in silent reading, the prosodic analysis would still be carried out (i.e., ñimplicitlyò), she proposes that the default prosodic pattern used by the processor during silent reading influences ambiguity resolution at any time. The processor has a bias to attach to the potential host with the most natural prosodic pattern.

Under this account, the length is specifically important. Compare 37) below:

37) John said Sue will leaveé
    a. yesterday! (Garden-path effect)
    b. when the pompous soprano had finally finished singing the aria. (Ok)
    c. , yesterday. (Ok.)

Sentence 37a) will activate a Garden-path effect because parser attached the adverbial phrase (AP) *yesterday* to the locally available host site, the VP *will leave*, but since it is an adverbial that only associates with the past tense, the local VP does not seem to be appropriate. The parser then reanalyzes it and switches to the VP at a higher position in the structure. However, as the attachment phrase is longer for sentence 37b, according to IPH, the parser attaches the longer AP directly to the higher VP, which does not activate the Garden-path effect.

IPH argues that normally longer-attachment (i.e., non-localð not following the locality principle in 10) above) strategy is not a direct strategy; rather, it is the segmentation of the phrase that affects the processorõs decision. Consider the sentences in 37 (cf. a & c). In sentence 37c), a comma indicates a longer intonation break. In this situation, IPH also predicts that the processor will attach the AP to the higher VP and thus no Garden-path effect is activated. Indeed, in daily conversation, it is very common for speakers to use intonation breaks to disambiguate. Fernández (2002) argues that in fact a longer constituent is more likely to be an independent prosodic unit, and therefore it is freer to be attached non-locally. Another example is RC attachment ambiguity as in sentence 26c above. A structure is suggested as in 38) below:

38) a. NP_1 PP **NP_2** **RC** (NP-low preference, short RC without break)

b. **NP_1** PP NP_2, **RC** (NP-high preference, long RC or with intonation break)

Thus, the sentence in 26c (the English version) can be represented in 39) below:

39) a. The journalist interviewed *the daughter* of *the colonel* who had had the accident. (NP-low preference, short RC without break)

b. The journalist interviewed *the daughter* of *the colonel*, who had had the accident. (NP-high preference, with intonation break (or long RC, but not in this case))

This proposal has been tested crosslinguistically, and the empirical evidence will be reported in Chapter 4.

IPH approaches sentence processing from a whole new perspective. As psycholinguistic research differs from pure linguistic research in the sense that from acquiring sentential information to achieving the meaning embedded in the information, the processing is inevitably involving aspects both linguistic and non-linguistic. Early psycholinguistic research intentionally attempts to avoid the non-linguistic factors (Frazier, 1978), making the processing of sentences appear to be pure-linguistic. IPH's approach to sentence processing has cast heavy doubt on pure-structural models.

3.3.2 The Late Assignment of Syntax Theory (LAST)

The last theory being introduced in this thesis is called LAST (Townsend & Bever, 2001). This theory was originally used to try to explain acoustic sentence processing data, as in 35) above. According to LAST, a simplified version of segmentations of acoustic input should follow a template in 40) (Townsend & Bever, 2001, p.160):

40) Input → Sequence of sounds → Words → Phrases

Because every language has ñits phonetic, phonological, lexical and syntactic constraints that severely limit the possible sequencesò (ibid.), LAST is inspired by an early speech recognition framework (Halle & Stevens, 1964), proposing that in sentence level, processing can be ñanalysis-by-synthesisò.

The procedure of LAST is proposed as 41)

41) LAST

i. Assigning a likely meaning:

a. Lexical recognition;

b. Phrase segregation;

c. Assigning a configurationally syntactic structure;

ii. Forming the syntax;

iii. Syntax generation and matching.

Very similar to the parallel models, LAST is largely based on lexical information. It is not surprising because the initial meaning assignment heavily depends on the lexical itemsô properties.

In stage 41i), further three sub-stages have been defined: i.*a*, the process is passive. Lexical items can be segmented in this phase. Townsend and Bever (T&B) also suggest that in this stage, the isolated lexical items can be related into a likely syntactic structure. 41i*b*) is a stage where separate phrases have been built, based on functional words or morphological cues. Townsend and Bever suggest that phrases can be categorised in this phase 41i*c*) and that this is a stage where phrases are integrated within a higher-order framework in order to be able to reach the ñmeaningò level. Thus, possible arguments and predicates should be specified. Because of the similarity of 41i*c*) and ñrealò syntax form, Townsend and Bever name the output of 41i*c*) *pseudo-syntax*, and claim that the assignment in this stage is statistical-based instead of rule-based; therefore the pseudo-syntax is ñassociativeò, instead of ñcategoricalò (cf. syntactic categories).

In stage 41ii), formal syntax is constructed based on the *pseudo-syntax* formed in stage 41i). Stage 41iii) is a *checking* stage. If the assignments of syntax in stage ii match the ñoriginal inputò, the computation is terminated; if not, then the whole procedure starts again.

There are two major problems with LAST. First, the initial stage is not defined clearly. In fact, heavy linguistic (or rather syntactic) related terminologies are used, so that it is not substantially distinct from syntactic processing. Moreover, as Townsend and Bever claim that the initial stage is statistical-based, it faces the same ñGrain-size Problemò as ETH. To tackle this problem, they propose that the grain-size should be ñsentence-level grammarò. It is not surprising that some may feel that ñit looks like syntax, smells like syntax; it is syntaxò[31]. The second problem with LAST is that, although the authors claim that their theory integrates with the newly developed MP, it still recognizes the assumption that there is an S-Structure, which MP theory no longer differentiates. The best suggestion for replying to these criticisms is that: do not recycle the remaining of syntax in the first stage!

A diagram (borrowed from Townsend & Bever, 2001, p. 163) provides a clearer explanation of the LAST system:

[31] Personal communication with Edson T. Miyamoto (2005), Hong Kong.

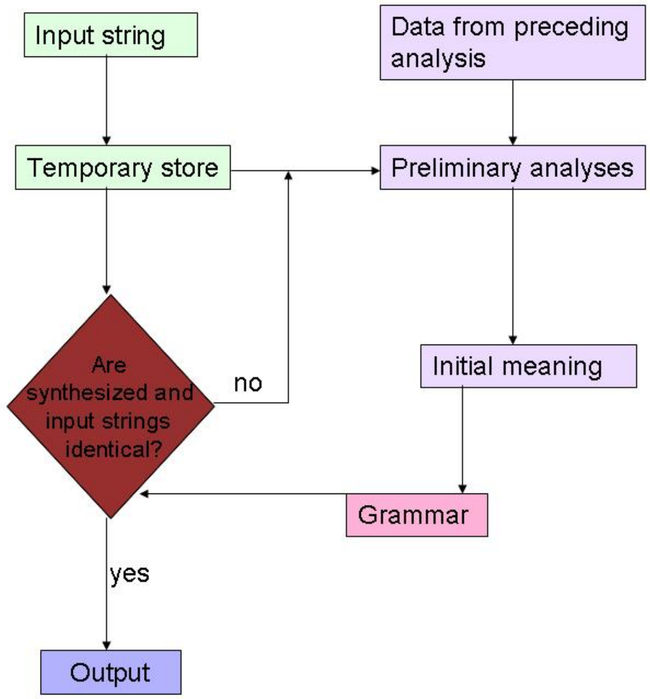


Figure 15 An Illustration of LAST at Sentence Level

Without the messy syntactic terminologies, Figure 15 clearly shows that it is in the later stage that grammar is taken into concern. In fact, as discussed earlier, sentences are a product of behaviour, whereas linguistic rules are descriptive of the products. It is not necessary that behaviour has to follow the regulation of its products. Instead, behaviour should have its own rules, which might be identical as its products, or might not be so. As the aim of the current thesis is to some extent in agreement with LAST, future discussion will be provided in later chapters.

3.3.3 Summary

There is strong suspicion that the initial stage of sentence processing is not syntactic. IPH seems to have grasped some convincing non-syntactical constraint, whereas LAST attempts to persuade other researchers in a more theoretical sense. Taking the development of the linguistic theory into serious concern, it seems that pure syntactic structure cannot transfer signal into meaning initially. The proportion of semantic component in the stage of ñpreparation for syntactic constructionò can be treated as the frontier between the three Chomskyan theory generations. This development in the linguistic theories should be taken into account in more psycholinguistic models. A separation between linguistic rules and psycholinguistic human behaviour should also be clarified.

Moreover, a theory that claims that syntactic processing does not necessary disclaim the innateness and the universality of the UG. In fact, the innateness can be embedded in many aspects involving perceptual processes of language (as a substantial object). This will be further discussed in the third part of this thesis. In the next chapter, common psycholinguistic research methodologies will be introduced, and in Chapter 5, evidence supporting or disproving the aforementioned theories will be evaluated.

Chapter 4

The Methodologies

The issues most psycholinguistic study attempts to solve are *which, what*, *when, how* and *why*. In much of cognitive psychology research, canonical conditions can provide information, but only to a limited extent. It is these problematic sentences that help to specify those questions. For example, in a globally ambiguous sentence, one can ask ñ*which attachment is preferred and why is it like that?*ò In a locally ambiguous sentence, one can ask ñ*when does the ambiguity happen and how does the re-analysis proceed?*ò In a non-canonical sentence, one can ask ñ*what is the problem?*ò etc.

In order to answer those questions, relatively precise measurements need to be conducted. In general, these measurements can be divided into *Off-line* and *On-line*.

4.1 Off-line

A typical task of off-line study is questionnaires. In a questionnaire, some judgment on the meaning of a sentence can be identified. For example, for globally ambiguous sentences, using multiple-choice questions can provide general information about the attachment preference, as in sentence 42i); for non-canonical sentences, using naturalness rating can provide information of which type of structures are naturally allowed in a particular language, as in 42ii). Sometimes questionnaire studies involve sentence production tasks, e.g., complete a sentence task, as in 43). This is based on the assumption that the preferable structure in comprehension should match the corpus data from production.

42) i. I saw an iceberg on the shoulder of a giant.

Question: What is on the shoulder of a giant? ( )

a. A person; b. An iceberg c. No one

ii. Please rate the naturalness of the sentence according to your native language from 1 to 5 (1 as completely unnatural; and 5 as completely natural).

The horse raced past the barn fell. ( )

43) Please finish the sentence below:

The crook arrested _____________. (sentence used by McRae et al., 1998)

The detective arrested by ___________.

Off-line measurements provide information of the final stage of language processing: the decisions. Although in most of the cases, an experimenter would ask the participants to respond ñas soon as possibleò, the tasks involved in off-line measurements are such that they allow the participants to adjust their response. Thus, data from an off-line study cannot fulfil the demand for answering the questions involving initial stages of sentence processing, and therefore on-line measurements are predestined to play a key role in psycholinguistic research.

4.2 On-line

To monitor real-time sentence processing, psycholinguists have applied many techniques from cognitive psychology research. On-line experimental methods include means from simple stimulus-reaction (SR) to measure recognition up to word (or at most phrase) unit, to the measurement of eye-movement and even brain activity. Those methods consist of analysing the cognitive activity as it unfolds the activity outcomes (Rouet & Passerault, 1999). In the current thesis, only closely related methods will be introduced.

4.2.1 Self-Paced Reading

A Self-paced Reading task is a type of reading time measuring task. In the task, one segment of the text (such as a word, a phrase or even a sentence) is exposed to the subjects on a computer screen while the other parts are masked. After the subjects have completed reading this bit, they are required to press a button and a new segment is exposed while the previous one is masked. The duration between the new segment appearing and the next button being pressed is recorded[32]. It is assumed this time course can reflect some of the linguistic and cognitive processes for reading and comprehending language (Aaronson & Ferres, 1984; Mitchell, 1984).

An example of possible segmentations is displayed below:

44) i. Equal segmentation length

The crook / arrested by / the detective / was guilty / of taking / bribes.

Segment 1 Segment 2 Segment 3 Segment 4 Segment 5 Segment 6

(segmentation used by McRae et al., 1998)

ii. Unequal segmentation length

Die Leiterin des Kochkurse / bestimmt, / daβ / Erika / die dosen / öffnet.

Segment 1 Segment 2 Seg. 3 Seg. 4 Seg. 5 Seg. 6

The director of the cookery course/decide/that / Erika / the cans / opens.

(Sentence used by Weyerts et al., 2002)

◆ Typical Tasks

There are mainly three types of self-paced reading task: the moving window task (Just et al., 1982), the stationary window task and the pointing task (Haberlandt, 1994; Mitchell, 1984).

In the moving window task, the computer screen is filled with patterns of dashes that correspond to letters of a word, and spaces that represent the spaces between words in place of text. When a word is processed, the next word will appear in the next dashed segment and the previous one will become dashed again. The dashes can imply the length of the coming word (as well as the passed ones, even though one is not supposed to regress).

In the stationary window task, the segments will appear at the same position in the screen (normally in the centre of the screen) and the course of pressing button is repeated. This is one of the most commonly used types of tasks nowadays.

[32] Some researchers view it as a time course between the two ñsuccessive button pressesò (Mitchell, 1984).

The pointing task allows subjects to regret. In this type of task, subjects use a mouse-like device to point to a location on the screen to reveal the current segment, while at the same time the rest of the segments are masked. The mask can imply the length of the word only. The subjects are supposed to point forward to reveal the entire sentence but they are allowed to regress to the previous word. The report of the regression has been found to be relatively low, similar to those from eye-tracking tasks (Haberlandt, 1994; Mitchell, 1984).

◆ Evaluation

The self-paced reading task can be seen to be relatively sensitive to on-line processes in reading. The processing load can be tested throughout the sentences and the reading time of any particular point in the sentences can be singled out and measured. It is very economical. As for technical support, free software such as *DMDX* (developed by J. Forster & K. Forster, at the University of Arizona) under the Windows operating system or *PSYSCOPE* (Cohen et al., 1993) on any Mac system performs well. A further benefit is convenience. Since there is no extra equipment needed, researchers can run this task on a laptop whenever and wherever is suitable for the participant. This benefit is particularly essential in language studies since cross-language studies sometimes require travel abroad to gather data from monolinguals.

Nevertheless, the biggest problem of self-paced reading measurement is the *naturalness* of the language presentation. First of all, natural reading is not presented in segments. It is also not clear why one type of segmentation should be chosen over another (cf. 44) above). Second, most of the self-paced reading tasks do not allow regression, despite the fact that free regressive reading is quite common in natural reading scenarios. It does not mean this will affect earlier processing but may well cause extra noise in the following process, since the subjects are ñforcedò to move on. It is not difficult to understand that distortions occur when they are forced. Last, self-paced reading generally involves button-pressing to receive new information. This means there are actually two tasks involved ï reading and button pressing. Thus, the ñreading timeò in the measurement might be affected by other non-reading-related factors such as button pressing speed. The downside notwithstanding, no proof has been shown that the self-paced reading task has provided misleading data (Mitchell, 2004). Thus, this technique remains a valuable and maybe even a ñtrailblazingò one in the field of psycholinguistic research.

4.2.2 Eye-tracking

The eye-tracking technique itself can be applied to many different cognitive and perceptual tasks. In psycholinguistic research, the assumption is that when text is difficult, readers fixate for a longer time, move their eyes at shorter distances with each saccade, and make more regressive movements (Frazier & Rayner, 1982; Rayner et al., 1989). This technique has two subcategories in the field of language study (Carreiras & Clifton, 2004): pure reading tasks, and tasks involving scene viewing when listening to speeches (Altmann & Kamide, 1999; Sussman & Sedivy, 2002). This thesis focuses on the former.

Although the presentation of the sentence does not need to be segmentised, it is important for the researcher to clarify the interest-area of the measurement for analyses purpose. An eye-tracking system consists of at least two main components: head-mounted cameras to record the eye movement, and a computational system to convert the recorded movement into numerical values that are mapped into the region of interest, which can be further computed and analysed. However, as the whole reading behaviour involves an enormous amount of information, such as interest-area, fixation time, fixation location, saccade, etc., handling the data needs careful and detailed plans.

◆ Data Extraction

It is necessary to discuss what kind of data will be extracted from the system. As the eye tracker has recorded all potentially relevant information, one has to decide to extract maximally informative data, but as economically as possible. Traditionally in language processing experiments, interest has mostly focused on five kinds[33] of information,

[33] *Regression proportion* (or *number of regression*) is sometimes reported (Pickering, 2004). It is the percentage of

namely, *first fixation time*, *first pass time*, *regression path time*, *reconstruction time* and *total time*.

First fixation time (Inhoff, 1984) is literally the very first fixation in an interest area. It can be seen as a measurement on lexical access, if the interest area is a one word region.

First pass time (Rayner et al., 1989) is the sum of the fixations on an interest area from the first entry until the area is left in a backward or forward direction for the first time. This measurement reflects the behaviour of processing when the information is first completely viewed, if one makes an assumption of immediacy.

Regression path time (Brysbaert & Mitchell, 1996; Konieczny et al., 1995; Mitchell et al., 1999; van Gompel et al., 2005) is the reading time associated with regressive reading behaviour, which is calculated as the "sum of the durations of fixations in a region (for the first time), plus the fixations from that region to earlier parts of the text until the region is left in a forward direction" (Vonk & Cozijn, 2003). This is looked on as one of the most interesting for sentence processing since it is believed to reflect a complete processing of a region for the first time. Connectionists (McRae et al., 1998) believe that it is the behaviour when the processor is balancing the weights of the processing information, while supporters of the Garden path theory (Frazier & Clifton, 1996) would regard it as the time to go back to check and re-analyse the previously parsed information. Whatever it is, it is expected to represent the processing difficulties in the region.

Although some (e.g., Pickering, 2004) consider this measurement to be a "late" one, ñbecause the reader may repeatedly fixate the beginning of the sentence after fixating the target wordò, however, in this paper it is considered to be a relatively earlier stage, since it is the **first time** when the reader has parsed the interest area to a sufficient degree that they are prepared to input new information.

Reconstruction time (Rayner et al., 1989) is also called *second pass time*. It is

all first-pass movement to an earlier part of the sentence material. It is possible to measure regression proportion from a region to its previous region, as well as from a later region into the current region. However, these measurements will not be introduced in detail, as it is not used in the experimental measurement of the current thesis.

calculated by subtracting first-reading time from the total reading time. This measurement is said to reflect a relatively later stage of reading difficulty. It is the re-reading procedure.

The total time (Just & Carpenter, 1980) is the sum of all fixations in an interest area throughout the time of reading the whole sentence. It is said to reflect even later effects of reading difficulty.

These traditional measurements are based on a strict assumption that all comprehension processes start when the word/region is viewed and last as long as the region is fixed (Just & Carpenter, 1980). Nevertheless, it does not include the duration while the eyes are moving around, namely saccade time. It is argued by Vonk and Cozijin (2003) that since reading is not a linear process, where some words are skipped but their meaning is still processed, and there is no evidence to prove that the parser pauses during saccade movement, a consideration of saccade movement is necessary. According to them, the duration of processing might be increased by parsing difficulties at the end of a region, but also would be decreased right after the difficulty because of facilitated integration. Thus a cancellation of the RT change would happen. Only if saccades are taken into account, a more informative comparison could be made. It is suggested that saccade time should be added into the measurement introduced above.

Furthermore, Vonk and Cozijin have suggested a new measurement, which they refer to as ***forward reading time***. It is calculated as the first pass time on a region but until the region is left into a *forward* direction. If regression occurs, it would be counted as missing data. Vonk and Cozijin (2003) propose that the first fixation durations that are immediately followed by a regression are smaller than those prior to a forward saccade. Thus, by introducing forward reading time, the "cancellation" caused by regression should be eliminated. Additionally, Vonk and Cozijin also suggest that saccade time should be considered in each measurement introduced above, as they postulate that although no new information is taken during saccade movement, the mind does not stop thinking. Thus the processing time should include saccade movement time. This suggestion will be tested and evaluated in Chapter 6.5.

◆ Evaluation

Eye-tracking technique has been claimed to be ñone of the *best ways*ò (Rayner et al., 1989, p. 23) to study moment-to-moment language processing in reading (e.g., Carreiras & Clifton, 2004; Hyönä et al., 2003; Mitchell, 2004; Sussman & Sedivy, 2002). Unlike the self-paced reading (introduced in 4.2.1), eye-tracking is more sophisticated in the sense that it reflects a more ñnormalò reading behaviour (Rayner et al., 1989). It does not need to present segmented sentences to a reader and with new developments in the eye-tracking system, an eye-tracker is very comfortable to wear[34]. This means the simulation of reading in a laboratory can be very close to natural reading scenarios. Moreover, it is also relatively cheap compared to EEG and brain imaging. In a word, it has all the benefits that self-paced reading measurements lack.

The disadvantages of this technique mostly lie in the fact that reading behaviour is messy and complicated. Rayner et al. (1989) have suggested two critical issues concerning the validity of the usage of eye movement data of reading. One is the *Perceptual Span* (PS), and the other is the *Eye-mind Span* (EMS).

PS is the region where useful information can be obtained at a particular fixation point. For example, it has been reported that content words longer than six letters will not be skipped normally, whereas short functional words (which are normally short words) have much less fixation rate of 19% -- 38% (i.e., can be skipped most of the time) (Just & Carpenter, 1980; Rayner & Duffy, 1988; Rayner et al., 1989). However, it is not very clear whether the effects are caused by the lexical identity or pure length. Moreover, it has also been reported that a region that readers use to extract information covers from no more than 3-4 character spaces to the left of fixation, but about 15 character spaces to the right of the fixation. However, it is also reported that to identify a word, a fixation has no more than 5-7 character spaces to the right (Rayner et al., 1989). These reports have been agreed as of the basic facts of eye-movement. Nevertheless, these facts are based on English data only. Although crosslinguistic research has begun to supplement

[34] But cf. Mitchell (2004) for a review on earlier techniques, which requires a head-clam or a bite-bar to fix the head during recording.

the field, the lack of evidence is severe.

EMS concerns how tightly the mind and the eyes are linked. Just and Carpenter (1980) made two assumptions: immediacy and eye-mind. They argue against the eye-mind span, and propose that the reader can complete all processing associated with a given word (also see Boland, 2004). However, Rayner et al. (1989) argue that parafoveal preview and *spillover* effects are evidence that at least the mind-eye match is not perfect.

If EMS can be neglected, the eye movement data can be regarded as a perfect reflection of moment-to-moment reading. However, to what extent EMS is effective needs further research and proof. In the cognitive research field, there is a tendency to link eye-tracking and brain activity measures. There is no such literature to date in the field of psycholinguistics. The combination of the two techniques will surely bring interesting and convincing data to the field. However, this is beyond the scope of the current thesis.

To recapitulate, the issues related to eye-tracking do not seem to come from the technique itself, but rather from unrevealed theoretical concerns. A more serious problem associated with this technique is that there is no serious theoretical background for identifying the eye-movement with purely syntactic / semantic processing (Boland, 1997a). Nevertheless, these issues should be dealt with urgently. The current thesis attempts to balance the PS in the experiments, and makes the assumption of the *immediacy hypothesis* in terms of the EMS.

4.2.3 Other Measurements

Brain activity measurements have been given significant attention in the past two decades. Although the current thesis does not include this type of methodology in the experiments, the results from these measurements by other psycholinguists and neurolinguists have provided a large amount of evidence supporting different theories introduced in the last chapter. Moreover, as the phonetic measurement has contributed to the field to date in terms of the Implicit Prosody Hypothesis (IPH), some trivial introduction relating to this technique will also be made here.

◆ EEG / ERPs

ERPs (Event-Related Brain Potentials) are tiny voltage changes that can be measured by electroencephalography (EEG), a procedure that measures electrical activity of the brain through the skull and scalp. Those activities are said to be triggered and time locked by sensory stimuli or by cognitive processes, such as word recognition and sentence processing (See Brown & Hagoort, 1999 for an introduction).

In terms of language study, there is some consensus that a negative peak of electricity flow can be triggered around 400 ms (N400) after an onset of a semantic surprise. Likewise, a positive peak of electricity flow is said to be triggered around 600 ms (P600) after an onset of a syntactic surprise (Kutas & van Petten, 1994).

However, it is debatable how reliable those effects are relating to language processing. For example, N400 is closely related to the ñprocessing costò. Thus, for a sentence 45) below:

45) The zebra ran away, chased by
    a. a lion   b. a rabbit   c. a walkman

A greater N400 effect can be triggered by a < b < c. Moreover, there is also research suggesting that some music (Koelsch et al., 2004) and pictures (West & Holcomb, 2002)can also trigger N400 effects. Furthermore, sometimes morpho-syntactic incongruence does not trigger P600 (e.g., Münte et al., 1998).

◆ Phonetic Measurements

This is mainly used in the prosody-related researches. The fundamental frequency peaks (F_0) and sound stress are normally measured in sentence production, or manipulated as acoustic input (e.g., Quinn, 2001; Quinn et al., 2000). It can also be combined with other tasks such as the eye-tracking task (such as in the Snedeker & Trueswell (2003) experiment).

◆ Evaluation

The advantage of measuring brain activity is that this type of measurement has provided insight and evidence to link the language function and the neuron system directly[35] (Carreiras & Clifton, 2004). However, the ERPs effects are not certain in the current state of this type of research. Even so, there are great expectations that EEG / ERPs may bring more valuable findings to the psycholinguistic field.

The introduction of the measurement of sound systems has widened the sentence processing field. The manipulation and control of the prosodic and intonational influence have suggested that not only syntactic cues are available in language processing.

4.3 Summary

Before recapitulation, one problem should be stressed here: most of the techniques introduced above involve the measurement of reading time. However, it is well-known that psychological performances are generally sensible to the *speed-accuracy tradeoffs* (SAT) (i.e., the faster participants respond, the more errors they might make) (see Harley, 2001, for an introduction). Although SAT can be used in psycholinguistic tasks itself (e.g., McElree et al., 2003), researchers have to be careful about the precise instructions given to participants.

To sum up, these techniques introduced in Chapter 4 can provide dynamic information to the psycholinguistic research. The measurement of the procedure of sentence processing can thus be taken from the very moment when the stimulus arrives in the brain, till the very last moment when decisions are made on the questionnaire. In the next chapter, evidence provided by large number of experiments using the techniques

[35] Although some claims that the Blood oxygenation level dependency *indirectly* reflects neural activity (Osterhout et al., 2004).

introduced above will be discussed, in the context of the sentence processing theories introduced in Chapter 3.

Chapter 5

Some Evidence

During the past three decades, numerous studies have been carried out, using the methodologies introduced in Chapter 4 to test the theories introduced in Chapter 3 in order to find out how the human language processing mechanism performs. This chapter will introduce evidence supporting or ñfalsifyingò the theories in the order of the timecourse when syntactic processing (i.e., parsing) is involved[36], focusing on two key points: the timecourse and universality of parsing. These two points are closely related because many parsing theories take the Universal Grammar (UG) as a motivation of parsing and endeavour hard to find universal processing strategies.

As there is no consensus on the separation of syntactic processing and semantic processing, the current thesis categorises two types of processing into the syntactic processing: general structure integration (derived from the Garden-Path Theory Group), and processing following the UG Theories (as this is supposed to be the key theoretical background for the universality of syntax).

5.1 Parsing Prior to Other Types of Processing

Theories of this type presume that sentence processing follows Chomskyan UG theories. The Garden-Path Theory Group (GPTG) assumes that syntactic structures should be assigned in an initial stage in order to be able to acquire the meaning of a sentence, whereas modular theories propose that sentence parsing should be modular following linguistic theories.

[36] Please note that the current thesis makes no attempt to cover all the previous studies. Only those closely related to the models will be focused in this chapter.

5.1.1 Delayed Assignment of Pragmatic Constraints?

A direct comparison of time course between syntactic processing and semantic (pragmatic) processing begins with Rayner, Carlson and Frazierǿs (1983) eye-tracking experiment with sentences such as 46) below:

46) a. The florist sent the flowers was very pleased.
    b. The performer sent the flowers was very pleased.
    c. The performer who was sent the flowers was very pleased.
    d. The performer sent the flowers and was very pleased with herself.

Both sentences in 46a) and 46b) include a reduced relative clause (RRC) *who was*. However, according to ñworld-knowledgeò, a performer is a more typical recipient of ñsending flowersò, whereas a florist is a more typical agent of ñsending flowersò. Furthermore, the verb phrase (VP) can be interpreted as the main verb of a sentence, or a past-participle indicating a RRC. 46c) is the control sentence where the relative clause (RC) *who was* is intact. 46d) is an active sentence, which researchers believe to be ñless plausible on pragmatic grounds than the reduced relative analysisò (Rayner et al., 1983, p. 361). Rayner et al. argue that if the pragmatic cues are used in the initial sentence processing, Garden-path effects should be expected for 46a) and 46d) at the disambiguation region *was very* (i.e., longer reading time, longer fixation duration and increasingly high number of eye-movements). However, if only syntactic cues (such as *Minimal Attachment* (MA)) are used initially, then 46a) and 46b) will activate the Garden-path effects.

Rayner and colleagues found that sentences with RRC were Garden-pathed initially: the reading times were longer for sentence 46a) and 46b) in the disambiguation region; compared to sentence 46c) and 46d), there were also increased number of regressive fixations initiated after the ambiguous string in those sentences. But there were no Garden-path effects (see 3.1.1 for an introduction to these effects) for sentence 46d), which led the authors to believe that the result contradicts the hypothesis from the theories that pragmatic cues are initially used. Moreover, total reading times showed a pragmatic advantage. Sentences with typical recipients (e.g. 46b)) have shorter reading time than atypical recipients (e.g., 46a)).

Thus, the authors claimed that the pragmatic constraints are not taken into account in the initial processing stage. Only syntactic cues are used initially in sentence processing, but pragmatic information can be used to interact with the parser in a later stage of sentence processing.

This study is a pioneer of research on time course of sentence processing. However, the design of the sentences has been criticised to be ñindirectò (e.g., Ferreira & Henderson, 1991), as the disambiguation region *was very pleased* does not directly follow the ambiguity region *sent*. A Noun Phrase (NP) *the flower* was inserted in between the two critical regions. Thus, the semantic / pragmatic cues within a sentence might not be strong enough to activate an effect to over-write that from the syntactic cues.

However, follow-up studies of this type suggest inconsistent results. In an eye-tracking experiment, Ferreira and Clifton (1986) used animacy of the NP prior to the ambiguous VP, followed by a disambiguation region of a preposition phrase (PP) *by+NP*:

47) a. <u>The witness</u> **examined** *by the lawyer* turned out to be unreliable.
b. <u>The evidence</u> **examined** *by the lawyer* turned out to be unreliable.

They found that the reading time at the disambiguation region (PP) was longer for sentences with RRC than those with full RC; no interaction of animacy was found at this region. However, the reading time at the VP region was longer for 47b) than 47a). Ferreira and Clifton argue that the results showed that the awareness of the thematic unfitness cannot rule over the initial syntactic resolution, resulting in a garden-path effect.

Conversely, a similar study by Trueswell et al (1994) has provided different scope also from an eye-tracking study. With stronger biased material similar to 47), they found clear interactions between animacy and syntactic ambiguity at the disambiguation region (i.e., the PP region). Trueswell (1996) even found that the difficulty in reading on the ambiguous and disambiguation region is strongly correlated with the frequency of the use of the verb form, i.e., reading difficulty is more likely to be activated in the association of an inanimate NP and a verb whose V+*ed* form is more frequently used in the past tense; but when the V+*ed* is more frequently used as a past-participle, the difficulty can be reduced.

To sum up, to date, it is not very certain from the data that processing other than syntactic information is delayed.

5.1.2 Parsing is Modular?

In testing the universality of parsing, there are also studies that directly apply syntactic principles in sentence processing. A good example is provided by Sturt (2003), with two eye-tracking experiments examining the Government & Binding (GB) theory, using anaphoric references.

According to GB (and also Minimal Attachment (MA)), a very important syntactic relation in a sentence is *Constituent-Commend* (C-commend). This means that a bound constituent must be C-commended by an appropriate antecedent. For an anaphoric binding, there are three principles:

48) **Binding Principles**

A. Reflexives (and reciprocals, such as ñeach otherò, ñhimselfò) must always be bound in their domains.

B. A pronoun must never be bound within its domain.

C. R-expressions[37] must never be bound.

Thus, in sentence 49a), the pronoun *himself* is bound to Peter according to Principle A; in sentence 49b), the pronoun *him* is bound to John according to Principle B; whereas 49c) is ungrammatical, unless Peter / John is someone else (i.e., neither John or Peter in the sentence, but someone else who has the same name), according to Principle C.

[37] R-expressions are referential expressions: non-pronoun, uniquely identifiable entities, such as NP "the president", or proper names such as "Peter".

49) a. John thinks that Peter hates *himself*.

b. John thinks that Peter hates *him*.

c. *John thinks that Peter hates *Peter / John*.

Sturt (2003) focused on Principle A, and used a set of sentences shown in 50)

50) a. Jonathan was pretty worried at the City Hospital. He remembered that the surgeon had pricked *himself* with a used syringe needle. There should be an investigation soon.

b. Jennifer was pretty worried at the City Hospital. She remembered that the surgeon had pricked *himself* with a used syringe needle. There should be an investigation soon.

c. Jonathan was pretty worried at the City Hospital. He remembered that the surgeon had pricked *herself* with a used syringe needle. There should be an investigation soon.

d. Jennifer was pretty worried at the City Hospital. She remembered that the surgeon had pricked *herself* with a used syringe needle. There should be an investigation soon.

The names *Jonathan* and *Jennifer* represent a gender character in the sentence. The stereotype of a gender character *the surgeon* is used for an expectation of male character. The reflexive pronouns *himself* and *herself* are used to match the antecedents. Thus, in sentence 50a), the attachment is ambiguous, but the accessible match according to GB Principle A is *the surgeon*. In sentence 50b), only the binding allowed by GB is a good match. In sentence 50c), the gender does not match any of the antecedents. In sentence 50d), the gender only matches the antecedent that is forbidden by GB.

Early measurements including *first fixation reading time* and *first-pass reading time* on the reflexive pronoun and a region right after the pronoun (the squired regions in 50a are the critical regions) showed reliable gender effects: the reading times were faster when the anaphor gender matched the stereotyped gender of the antecedent that is allowed by the GB (such as 50a & b). However, there were no early gender effects for the antecedents that are forbidden by the GB. However, later measurements of *second-parse reading time* on the critical regions showed some gender effects where the antecedents are forbidden by the GB (such as 50c & d).

In a follow-up study, Sturt used sentences such as 51)

51) The surgeon who treated Jonathan had pricked *himself* with a used syringe needle.

Thus, although the reflexive *himself* is still C-commended by the NP *The surgeon* syntactically, the binding has a longer linear distance. Sturt attempts to use this type of sentences to rule out the possibility in the previous experiment that the effects were merely as a result of Recency.

The two experiments replicated each other, leading Sturt to conclude that the GB principles have been applied extremely early on in the parsing. Although other factors have effects on sentence processing, they appear only in a much later stage of processing.

Sturtõs results notwithstanding, the initial stage GB effects may be questioned at a closer look at the data, as he failed to provide data on early interaction between the GB principle and gender attachment. If the GB-prohibited antecedents really have no gender effects on reading time, an interaction between the two factors should not be expected. This type of interaction was found in later processing, which suggests that the GB principle was abandoned, as well as that the effects of the GB-prohibited genders had been activated.

Other principles have also been tested. Based on English data, Pritchett (1992) has provided evidence that each lexical item is attached as it is encountered, and attachment is motivated by principles of grammar (e.g., the theta-criterion and binding principles, cf. the Extensive Projection Principle (EPP)[38] introduced in 13) in Chapter 2.3). This construction is satisfied by a ñhead-drivenò parser. It is not difficult to understand from

[38] EPP is a very important principle in GB and MP because it is the key principle to form a tense constituent; therefore it is also the key principle to form the main verb predicate constituent.

a simple English sentence such as 52) below:

52) Peggy gave a beer to Pat.[39]

As the noun *Peggy* is processed before the theta-assigner (verb, *gave*), and the other arguments *a beer* and *Pat* are available after the verb has been processed, it is very easy for the head-driven parser to construct the sentence structure.

When this theory is applied to languages with SOV order, such as German and Japanese, a ñdelayò effect at an object NP is predicted because the head of the phrase *VP* has not been introduced at the object position. However, Miyamoto (2002) has provided evidence from three self-paced reading experiments that the case markers following the NP are the determiner of the processing, but not the VP. This means that the important principle of EPP is not initially used in sentence processingð at least not crosslinguistically. Although EPP is just a single principle in UG, because it is the key principle of forming the main verb phrase constituent, violation of this principle makes it very difficult to validate other phrase construction procedures.

Thus, it is unsure whether principles of the UG are initially applied in sentence processing.

◆ Discussion

In terms of what sentence processing is involved in the initial stage, 5.1.1 and 5.1.2 have provided positive and negative evidence for the parsing-initial theories that are based on Chomskyan theories. It seems that regardless of whether the claim is based on syntactical rules or syntactical principles (grammar), the evidence of initial parsing is not adequate. Although some studies have shown support of syntactic processing theories (e.g., GPTG, Recency, etc.) at re-analyses stages (e.g., Meng & Bader, 2000; Sturt et al., 2002), it is still a mystery that what is the timecourse of information processing in the initial processing stage.

[39] Sentence used by Kamide (1998)

5.1.3 The Crosslinguistic Disaster

Being a *universal* account, two important points about the UG-based theories are the universality and the syntactic rules or principles. A universal account should have general rules and principles that can be applied to different languages. As for the rule-based models (such as GPTG), the universality should be: Minimal Attachment (MA) and Late Closure (LC) and other rules of this type; as for principle-based models, the universality should be grammar (whereas the parameter-setting accounting for language differences is the individual language grammar).

◆ The Issue

Disastrous evidence against this universality was raised by Cuetos and Mitchell (1988), when the crosslinguistic differences between Spanish and English speakers in the resolution of ambiguous relative clause (RC) attachment was examined. The sentence type introduced in 26c will be re-printed here for the convenience of the argument.

53) Crosslinguistic differences (sentences used by Cuetos & Mitchell, 1988):

i. The journalist interviewed the daughter of the colonel *who had had the accident*. (English ï *the colonel* had the accident)

ii. El periodista entrevisto a la hija del coronel *que tuvo el accidente*. (Spanish ï *la hija* (the daughter) had the accident)

Preferences of the two languages have been studied both by questionnaire (sentence completing task) and self-paced reading tasks. The Spanish data provide strong evidence against LC. This result has been confirmed by many other studies using different methodologies including eye-tracking, self-paced reading and questionnaires (e.g., Carreiras & Clifton, 1999; Cuetos et al., 1996; Frazier & Clifton, 1996; 1997).

However, a later study (Gilboy et al., 1995) found that when the NP_1 is a non-human term (e.g., the book of the girl), the preference in English can shift from NP_2 to NP_1. This is possibly because when the complex NPs are mixed with human / non-human configuration, the relative pronoun *that* in English is biased to the non-human host siteð as another relative pronoun *who* is more commonly used (in English) to refer to human agents. In addition to that, Clifton (Clifton, 1988, reported in Frazier & Clifton, 1996) manipulated 24 sentences used in Cuetos & Mitchellôs (1988) experiment, with variations of the position of the RC (subject vs. object position), the preposition between the two NPs (*of* vs. *near*), the relative pronoun (*who* vs. *that* vs. *which*) and the animacy of the NPs. An example of the manipulation of position is given below:

54) a. The journalist interviewed the daughter of / near the colonel who had had the accident. (Object position)
b. The daughter of / near the colonel who had had the accident was interviewed. (Subject Position)

Results from this study showed that the choice of NP2 were stronger when the two nouns in the complex NPs were matched in animacy, and when the preposition was *near*.

Recent research on RC attachment ambiguity has shown that at least in some languages (e.g., German and Spanish), the attachment preference varies depending on the position of the complex NPs (Hemforth & Schimke, 2003), while in some other languages (e.g., English and French) it does not. Hemforth & Schimke argue that the reason for these language differences lie in the topicalisation function of the preverbal position. In languages like Spanish or German, the preverbal NP has a more marked topical status than in English and French.

◆ Some Defence from the GPTG

The Relativized Relevance (RR) is the defence that the GPTG group proposed to settle the complex data from RC attachment (See section 3.1.1, P.36). RR suggests that all other information (e.g., grammatical, informative and appropriate information in the discourse) being equal, the parser will prefer construing a phrase ñas being relevant to the main assertion of the current sentenceò (Frazier, 1990b, p.321). In other words, parsers process a sentence initially based on syntactic principles such as LC; then consider other information at a post-syntactic phase, using principles such as RR. Therefore, the explanation for the NP_1 attachment preference from this theory is that the Spanish parser might have first attached the RC to the local NP according the LC, but then have broken the first attachment and moved the RC to attach to the NP near the VP according to the RR. Evidence to this principle is that there should be some evidence for the bias shift, for example, an extra time cost.

De Vincenzi & Job (1993; 1995) reported data from Italian monolinguals in favour of this account. They carried out a self-paced reading experiment followed by a question about the preference of the relative clause, using materials as in 55) below:

55) a. Lôavvocato diffida / del padre / della ragazza /che si è tradito / al processo.
The lawyer suspects the father of the girl who betrayed[masc] (himself) at the trila.
b. Lôavvocato diffida / del padre / della ragazza /che si è tradita / al processo.
The lawyer suspects the father of the girl who betrayed[fem] (herself) at the trila.
Question: Che si è traditi? Ragazza o padre?
Who self-betrayed? Girl or father?

The disambiguation is realised by a morphological change of the gender-marked past participle *che si è tradito*(*a*). A self-paced reading task showed a shorter reading time of 55b), which suggests that a NP_2 attachment is easier to process on-line, therefore initially there was a NP_2 attachment bias. However, the answers to the question task showed that NP-high is preferred; indicating that at a later stage, the preference of the attachment has switched to NP_1.

Construal theory (see Chapter 3.1.1, P.40) also attempts to tackle this problem. It claims that RCs as non-primary phrases will be associated to the current processing domain, and the specific interpretation will vary depending on information contained within the association site and within the RC. This is supported by the experiment conducted by Gilboy, Sopena, Clifton & Frazier (Gilboy et al., 1995).

In their experiment, both Spanish and English monolinguals were asked to fill in questionnaires containing sentences with two potential NPs for an ambiguous RC. The relationship between the NPs is listed in detail in Table 3 below:

Table 3 NP Types in Gilboy et al.'s (1995) Experiment

Type	Sub-type	Example	Status of NP-low
A	Substance	A sweater of wool	Non-referential[40], Argument of NP_1
	Quantity	A cup of sugar	
B	Kinship	The daughter of the colonel	Referential, argument of NP_1
	Function	The assistant of the lawyer	
	Possessives	The museum of the city	
	Inherent possessives	The window of the room	
	Representational	The picture of the lake	
Bô	Alienable possessives	The book of the kid	Referential, non-argument of NP_1
C	With / near é prep.	The sauce with the food.	Referential, non-argument of NP_1, prep is a theta-assigner

According to Construal, the prediction would be: when NP_2 is referential, it would be a more attractive site, which means that there should be a NP_2 bias for complex NP type B and C; when NP_2 is not the argument of NP_1, as in type Bô and C, the second NP is also a more preferred site than the first one; when the preposition is a theta-assigner, as in type C, a NP_2 attachment is the most favourite; this leaves Type A to be the least favourite for NP_2 preference.

The results of both languages (English and Spanish) supported this prediction. However,

[40] Referential Nouns are the nouns that introduce discourse entities into a discourse model corresponding to already existing discourse entities (Gilboy et al., 1995, P136)

the language difference that English has a NP_2 preference while Spanish has a NP_1 preference, which has been established empirically by many other researchers (e.g., Carreiras & Clifton, 1999; Cuetos et al., 1996; Frazier & Clifton, 1996; 1997) was not shown. It is not clear why there is a lack of systematic differences. One possible argument is that the manipulation of those complex NPs types might have over ruled some of the language features.

One major difference between the *Construal* and *Garden-Path* theories is that the theta assigner is of primary importance according to the former. If NP_2 is an argument of NP_1, as in 53) (the *OF*-condition), the NP_1 will be the head of NP and the RP should be attached to a region incorporating both of the NPs. Frazier and Clifton suggest that in this situation the preference of the attachment should be chosen either in accordance with structural principles such as RR, or non-structural ones, e.g. the specific language alternative unambiguous form, i.e. the *Gricean Maxim of Clarity* (Grice, 1975). In agreement with the *Gricean Maxim of Clarity*, the different preferences between English and Spanish result in the genitive alternative structure that is unambiguous in English if NP_1 modification is intended, compare 56a) & b):

56) a. the servant of the actress who was on the balcony

b. the actress' servant who was on the balcony

Because the 56b) form exists in English, 56a) in English shows a preference for NP_2. By contrast, there is no such alternative form in Spanish. Thus, NP_1 is favoured by the other principles, such as RR instead of the principle of the *Gricean Maxim of Clarity*.

However, when NP_2 is not an argument of NP_1, such as the *NEAR*-condition in 54), the PP *near the colonel* becomes the current processing domain, so that only NP_2 is the available host of the RC and an NP_2 preference is found for this condition in many languages, even including Spanish, which was earlier regarded as an NP_1 -bias language by many psycholinguists.

The other two approaches from syntax-first theories come from Recency / predicate proximity (RP/PP) and Attachment-binding (AB) (See Section 3.1.2). As the advantage of RP / PP for the 3-sites RC attachment has been explained earlier (See 3.1.2), the current section will introduce the evidence of AB theory briefly.

The evidence for AB was provided by an experiment in German (Hemforth et al., 1998), comparing potential complex NPs with RC attachment, where anaphoric processes are required; and complex NPs with PP attachment, where anaphoric processes are not required, as in 57) below:

57) a. Die Tochter der Lehrerin, die aus Deutschland Kam, traf Klaus.
The daughter of the teacher, who[fem] from Germany came, met Klaus.
b. Die Tochter der Lehrerin aus Deutschland traf Klaus.
The daughter of the teacher from Germany met Klaus.

As the relative pronoun is often omitted in English but rarely in Spanish and German, the English has an NP_2 bias, whereas German and Spanish has an NP_1 bias. Hemforthôs German data has confirmed her theory (Hemforth et al., 1998).

◆ Further Problems

The issue with RR has not been solved just by the Italian data. The problem is that no convincing evidence has yet been found crosslinguistically: in Spanish, parsers attach high at all stages (Cuetos & Mitchell, 1988), while in French (Pynte & Colonna, 2000; Zagar et al., 1997), parsers can be trained to have a low attachment by within-task exposure with sentences containing preposition *avec* (with). This lack of empirical support is the crucial weakness of this theory.

The problem with *Construal* is that as more languages are studied, especially the Afrikaans (Brysbaert & Mitchell, 1996) and Dutch (Brysbaert & Mitchell, 1996; Mitchell & Brysbaert, 1998; Mitchell et al., 1999) data cannot confirm this theory. Afrikaans is similar to English in having a frequently used Saxon genitive form. But an offline study[41] showed an NP_1 bias in this language. Contrariwise, in Dutch, where there are three forms for a modifier, including the Saxon genitive form (although the use

[41] Communication between Swanepoel and Mitchell, 1995.

of the Saxon genitive in Dutch is relatively rare), the crucial finding is that, NP_1 is preferred even in a sentence that is the exact translation of English material. These findings indicate that at least the *Gricean* account does not ñact indiscriminately to shift the attachment from NP_1 to NP_2 in each and every language that happens to have an ambiguous Saxon formò (Mitchell & Brysbaert, 1998); so that there are phenomena involved in sentence processing that the *Construal* theory cannot construe.

A problem with RP / PP is that it cannot account for the shift of preference when the preposition between the two nouns has changed (cf. evidence by Clifton (1988) and Gilboy et al. (1995)).

Similarly, crosslinguistic data have cast doubt on AB. Not only do the existing data for Romanian (Ehrlich et al., 1999) and Croatian (LovriÍ, 2003) not confirm the attachment-binding prediction, but English/Spanish bilingual studies (Fernández, 1999) also show a clear difference between L1-English speakers and L2-English speakers whose L1 is Spanish. The native English speaker shows stronger NP_2 preference than the Spanish-English Speaker (no matter whether they are early or late learners). The flaw of this model, as argued by (Mitchell & Brysbaert, 1998), is that such a difference between L2 speakers and L1 users should not exist if the cross-linguistic difference was only due to this model.

◆ Summary[42]

Although some of the evidence supports the parsing-first theories, there is also evidence that challenges these theories. The evidence introduced above casts doubt on the universality of parsing in two ways: not only has the structure construction in initial stages been challenged, but also the processing based on the UG principle. The crosslinguistic RC attachment crisis has even thrown doubt on universality in general, as well as the universal *locality* law. If the foundation of the syntax-first theories has been challenged, how reliable will the theories built upon it be?

[42] The *Exposure Tuning Hypothesis* and *Implicit Prosody Hypothesis* also attempt to explain this issue. They will be introduced in 5.2.2 and 5.3.1.

As a large amount of debate is centred on the RC attachment preference problem, the second part of this thesis will try to provide some evidence from Chinese data. Moreover, evidence for other theories will be introduced briefly in the next sections.

5.2 Parsing Parallel with Other Processing Types

This thesis reports two main studies associated with the parallel models. One is the Competition-integration Model, as it provides a good fit between modelling and human data. Moreover, it addresses the general issue of timecourse of parsing. The other is the Exposure-based Tuning Hypothesis (ETH), as it closely relates to the aforementioned RC-attachment preference issue.

5.2.1 Evidence for Competition-integration Model (CIM)

CIM proposes that all constraints (semantic cues, syntactic cues, frequencies, etc.) have been entered into a recurrent system, where the strength of supporting one interpretation can be calculated. Simultaneously, a dynamic criterion of accepting one interpretation is reduced at a ratio of φcrit (recall Figure 14, P.52). The φcrit is set within a range between 0.005-0.01 so that the reducibility of the criterion level is mild[43].

◆ The Evidence

Sentences from 33) are recreated below:

58) a. *The crook* / arrested by / the detective / was guilty / of taking bribes.

Region1 Region2 Region3 Region4 Region5

b. *The cop* / arrested by / the detective / was guilty / of taking bribes.

The simulation by the model predicts that at Region 2, the good patient in 58a) causes reading difficulty, but the difficulty is reduced from Region 3 onwards. The sentence with good agent in 58b), on the contrary, has an advantage at Region 2, but the

[43] As mentioned in Chapter 3, this thesis makes no attempt to get into detail of the simulation procedure this introduction of simulation is very brief. For a detailed explanation of the CIM, please check 3.2.2.

advantage disappears from Region 3 onwards.

Another simulation monitoring the GPTG, in which the syntactic cues have been given an advantage in entering the system. The prediction is that at Region 2, both sentences have similar reading behaviour. However, the reading difficulty rises rapidly for 58a) in Region 3, but the difficulty rises slowly for 58b) in the same region. The resolution to the reading difficulty for 58a) is solved at Region 4, but continues to increase for 58b).

McRae et al. (1998) compared these simulations to human self-paced reading results. The human data showed clear reading difficulty for the good patient sentences as in 58a) at Region 2, but it soon turned into an advantage at Region 3; whereas the good agent sentences, in 58b), had a clear advantage in Region 2, but this soon turned into a disadvantage from Region 3. It is a relatively good fit with the simulation created by the CIM.

The authors also simulated the GPTG by delaying input other than structure cue with longer or shorter time (i.e., cycles in the simulation). The predictions of GPTG with either types of delay did not provide any good fit with the human data.

McRae et al.ǒs results suggest that the timecourse of structure processing should not be prior to the other cues.

◆ The Issue?

A more recent study by van Gompel and Pickering (2001) has attempted to provide evidence against GPTG and CIM. In their study, van Gompel and Pickering used the sentences below:

59) i-a. The hunter killed the dangerous poacher *with the rifle* not long after sunset.
i-b. The hunter killed the dangerous leopard *with the rifle* not long after sunset.
i-c. The hunter killed the dangerous leopard *with the scars* not long after sunset.
ii-a. The hunter killed only the poacher *with the rifle* not long after sunset.
ii-b. The hunter killed only the leopard *with the rifle* not long after sunset.

ii-c. The hunter killed only the leopard *with the scars* not long after sunset.

Sentences 59i-a) & ii-a) are ambiguous because the PP can be attached to either the VP or the NP. In sentences 59i-b) & ii-b), the PP is forced to attach to the VP, whereas in sentences 59i-c) & ii-c), the PP is forced to attach to the NP. Sentences in 59i) have been found to have a VP-attachment bias, whereas sentences in 59ii) are indecisive.

Being a fixed-choice model, GPTG predicts that when the disambiguation is inconsistent with an initial choice of analysis, parsing difficulty occurs. Thus, in the example of 59), a) should be as easy as the analysis that provides the more preferred interpretation (i.e., b) or c)). On the other hand, being a competition model, van Gompel and Pickering predicts that the ambiguity should create more ñcompetitionò, so that the reading time should be longer for condition a). However, van Gompel and Pickering found a delayed effect of difficulty in reading 59i-c) than 59i-a) & i-b); and 59ii-a) was the easiest compared to 59ii-b) & ii-c), which was not predicted by either theory.

Nevertheless, Green and Mitchell (2006) have provided a series of computer simulations that are based on CIM with materials used by van Gompel et al. They have shown the actual prediction from CIM has an ambiguity advantage, which fits the human data reported by van Gompel and colleagues. This has suggested the absence of real evidence against CIM.

5.2.2 ETH

Since Cuetos & Mitchell (1988) discovered the crosslinguistic differences for the RC attachment ambiguity resolution, a hypothesis of sentence processing based on input frequency has been postulated. As discussed in 3.2.1, with a small-scaled corpus analyses, the hypothesis has been confirmed.

Regardless, as other theories discussed above, ETH cannot account for all findings either. Earlier, Mitchell et al. (1988, see also in 1995) manipulated an experiment so that two groups of Spanish speaking children were tested. One group was exposed to NP_1 attachment sentences and the other to NP_2 attachment for two weeks. In the end, the

NP$_1$-exposed group showed the expected result in that they preferred attachment to the NP$_1$ site; however, the NP$_2$ group did not show any bias on attaching to NP$_2$ site at all. While this might be explained by saying that the children in the NP$_2$-exposed group may have already been acquainted with the first NP attachment preference in Spanish to the extent that the material used by the authors to ñtuneò them was not enough to alter their choice, the theory still cannot explain the findings reported by Gibson et al. (1996). They found that the corpus analysis for the three-site RC-attachment ambiguity, as in 30), does not match the on-line and off-line experiments in English or Spanish. Besides, Gibson and Schütze (1999) again reported that the disambiguation preferences in NP conjunction do not mirror the corpus frequency.

These un-matching corpus analyses might be due to incorrectly chosen ñgrain sizeò, or the range of record-keeping detail commitment. In principle, there should be a correct ñgrainò corresponding to a record-keeping in its database. This principle notwithstanding, even Mitchell and his colleagues (1995) cannot provide a satisfactory choice for one. In fact, in Gibson and Schützeôs report (1999), where they tried to find the ñfinest grain sizeò, it turned out there were not enough sentence items left in the whole corpus. Indeed, the value of the ñgrain-sizeò itself is a paradox, since highly detailed records would lead to a bin with no data and contrariwise. It is vital for ETH to find out the correct ñgrain-sizeò.

A further challenge to ETH comes from corpus data in Dutch (Mitchell & Brysbaert, 1998). While experimental results suggest that Dutch is a NP-high preferred language, the data from four Dutch newspaper and magazine corpora showed only 31% of sentences with NP-high attachment bias feature.

Later investigations of Dutch corpora have shown that the preference for low attachment prevails only because of the modulated characteristics of the complex NPs (Desmet et al., 2002). Their data showed a pattern of NP preference as in table 4 below:

Table 4 NP Preference in Dutch Corpus Study

NP-high	NP-low	Preference

Human	Human	NP-high
Human	Nonhuman	NP-low
Nonhuman	Human	NP-high
Nonhuman	Nonhuman	NP-low

This means that the preference is adjusted according to the human/non-human features of the item in the NPs. This is a good explanation for the unmatching Dutch data.

Fernández (2002) argues that the unmatching corpus studies in Dutch and English may reflect the tendencies of speakers of those languages to attach low on the initial analysis, because in production, ñthere might be less time for reanalysis (and therefore also less time for post-syntactic factors, including Tuning processes, to apply)ò (ibid.). However, it could also be argued that there are differences between written language and spoken language. Instead of having less time for reanalysis, articles published could have been modified again and again by the writers. This means, in opposition to the suggestion made by Fernandez, it could also be suggested that those corpus data are not ñnaturallyò produced, but show the effects of reanalysis by the author. Moreover, it should also be taken into serious account what kind of corpus and what kinds of participants are chosen. In other words, the unmatching data could be because the participants taking the experiments were not exposed to the corpora studied.

Furthermore, Desmet & Gibson (2003) investigated the corpus data in more detail and found that after eliminating the disambiguation using pronoun ñoneò, the discrepancy between the corpus data and the human online data disappeared. This result shows that Gibson and Schütze's (1999) early finding is actually incorrect. More positive data comes from a more recent study (Desmet et al., 2006), which takes lexical properties of the NP host sites (i.e., animacy and concreteness) into account. The study comprises corpus analysis and an eye-tracking experiment that reveal a relatively good fit of the frequency patterns.

It seems that there is lack of real evidence against ETH. Nevertheless, ETH cannot explain how the processing routines are affected by frequency counts. It does not answer the question why locality is the prevalent preference in constructions other than RC ambiguity resolution, such as PP attachment and other phenomena (i.e., why PP attachment follows locality, but RC ambiguity has variation in the first place).

In a word, all this implies that pure tuning is not adequate either, and that at least some non-statistical factor(s) will influence the decision of initial attachment.

5.2.3 Discussion

Recent studies in sentence processing have provided more and more evidence that different constraints can have very early effects on sentence comprehension. Among them, thematic fit (Aoshima et al., 2004; Liversedge et al., 2003), context effects (Su, 2004), and even NP types (Gordon & Johnson, 2004) can be very efficient in early decisions of the choice from ambiguous alternatives. Corpus analyses also confirmed this point of view (Roland et al., 2006). Moreover, a study with eye-tracking and ERPs in Hebrew (Deutsch & Bentin, 2001) has shown an early N400 wave in the VP NP incongruent situation with animacy effects, and also an early P600 wave in the incongruent situation with case-marking effects. Deutsch & Bentin then argue that their results support constraints-based models.

It seems quite difficult to eliminate all other constraints except the syntactic cues in the initial stage of sentence processing. Nevertheless, studies of this kind do not mention to what degree other constraints are effective in comparison with the syntactic constraint. Even in McRae et al.ôs simulation, later entrance of syntactic cues was not included. It is understandable because the dominant view in the psycholinguistic field is that syntactic structures should be initially assigned in sentence comprehension based on the UG. However, some comparison studies between first and second English users (Chipere, 2006) show that the native speakers of a language may perform worse in grammatical comprehension, compared to second language speakers. This study has not only thrown doubt on the Chomskyan language performance / competence view, but has also raised the question whether it is necessary that comprehension is based on building up structures for a native speaker.

5.3 Parsing Posterior to Other Processing Types

The claim that parsing is posterior to other processing types is rather controversial in the current psycholinguistic field. However, there is proof supporting this view.

Two ERP-studies by Kim and Osterhout (Kim & Osterhout, 2005) have provided compelling evidence that semantic information is assigned earlier than syntactic information. In their studies, anomalous sentences such as ñ*The meal was devouring…*ò were used. The critical verb *devour* is ambiguously anomalous: it can be syntactically ill-formed (i.e., devour-*ed* instead of devour-*ing*), but semantically canonical if the NP *the meal* is the Theme of the verb *devour*; this verb can also be semantically ill-formed (i.e., meals cannot devour things) but syntactically canonical if the NP is interpreted as an Agent. As the NP is preceding the VP, the type of the ill-form (i.e., semantic or syntactic) on the verb is rendered by the opposite type of processing (i.e., syntactic or semantic) on the NP. Kim and Osterhout reported replicated results of N400 effect (which is associated with semantic anomalies) on *devouring*. This suggests that the processing on the NP has a semantic interpretation in an early stage.

Although some other studies have provided data that can be explained by this type of theory, researchers attempt to explain their data in a way that the results favours at most those parallel models (e.g., Boland, 1997a, in general discussion) so that other processing types only ñinteractò with parsing. The current section provides some evidence and augments for this type of theory.

5.3.1 Implicit Prosody Hypothesis (IPH)

IPH states that a default prosodic contour projected in silent reading favours the syntactic analysis associated with the default prosody for the construction (See Chapter 3.3.1, p.51). This thesis interprets the evidence for IPH as a support of late-parsing theory and this will be discussed after introducing the evidence.

Quinn et al (2000) have provided good evidence for IPH in English, French and Arabic.

In the experiment, Quinn et al. recorded the fundamental frequency peaks (F_0) and the first stress-bearing word of the RC, while asking some native speakers to read aloud sentences with RC attachments forced to attach to either NP_1 or NP_2.

Their results show that with a shorter RC in the final position of the sentence, as in 38) and 39), there is a lack of F_0 prominence in the RC, with either condition (forced to first or second). However, data in English show that with a longer RC, the RC is prominent in the case of forced NP_1 attachments, and not prominent in the NP_2. Data in French and Arabic is more complex but also gives evidence supporting the assumption that the more independent an attachment constituent is, the more likely it is attached to non-local host, while the differences between those three languages are mainly from the prosodic feather differences from those languages.

In account of the ambiguous RC attachment resolution, IPH predicts that same strategies are used as with other types of attachment. RC attachment studies in English (Quinn, 2001; Quinn et al., 2000; Walter et al., 1999), Arabic (Abdelghany & Fodor, 1999), Croatian (LovriÍ, 2003), French (Pynte & Colonna, 2000), and German (Hemforth et al., 1999), etc., have collected evidence supporting the IPH. More specifically, these studies are in favour of the postulate that longer attachments are likely to be attached to a higher host site. Arguments from those studies are that this longer-higher strategy is not a direct strategy; rather, it is the segmentation of the phrase that affects the parser’s decision. However, this criticism has indeed proved that intonational breaks are important factor in encoding and decoding meaning in word strings such as sentences.

Beyond RC-attachment studies, Snedeker & Trueswell (2003) conducted a series of referential communication tasks, where a speaker and a listener were involved. They were shown a bag of toys that were arranged by the cooperation of the speaker and the listener. The speakers were shown some sentences such as:

60) a. Tap the frog with the flower. (Ambiguous)

b. Tap the frog by using the flower. (Unambiguous)

c. Tap the frog that has the flower. (Unambiguous)

The experimenter demonstrated the arrangement following the instructions, indicating what the speaker should instruct the listener to do, so that the speaker could understand which correct instructions should be given even in an ambiguous situation as in 60a. Acoustic and phonological analyses were carried out to compare the prosody use by the speaker in ambiguous and unambiguous conditions. Later, the listenerôs eye movements were recorded while viewing the objects (toys) during the instructions.

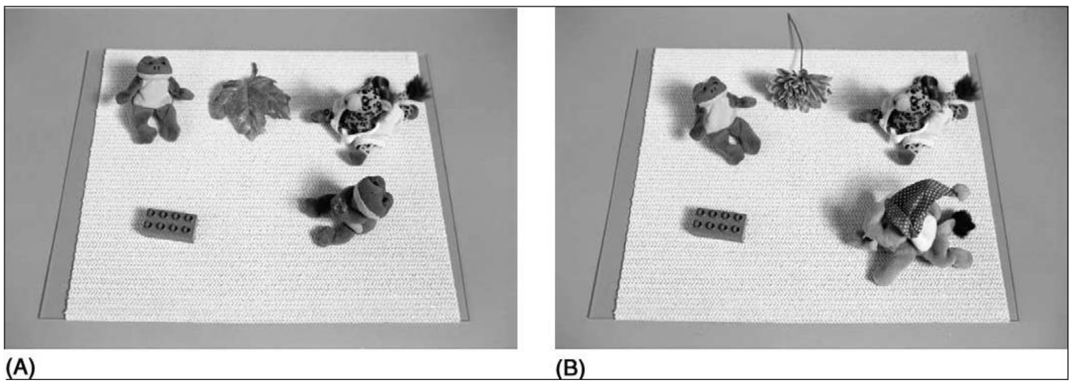


Figure 16 Scenes Used in Snedeker & Trueswell (2003) Experiment

The finding of this study is that both speakers and listeners use prosodic cues to interpret meaning. Moreover, the speakerôs prosody affected the listenerôs interpretation prior to the onset of an ambiguous phrase. The findings also suggest that speakers provide cues when needed, whereas listeners use prosodic cues when available.

Although IPH itself does not claim that prosodic processing is prior to syntactic parsing, evidence has shown that with prosodic cues, both global or local ambiguous structures might not even be activated, that is, prevent the ambiguity (See Jun, 2002, for a review). Moreover, some research on reanalysis (Bader, 1998) has shown that if the reanalysis of local ambiguity requires reconstruction of prosodic breaks, the structure processing will be more difficult than in the case when reconstruction of prosodic structure is not needed. However, the constraint-based models would not predict this difficulty, if all the constraints are entered at the same time. As discussed earlier, the prosodic cue cannot enter the system later than structural cues, and therefore the only possibility is that prosodic processing occurs earlier than the structure processing.

5.3.2 Late Assignment of UG Theory?

One factor that leads researchers to believe in UG Theories is the syntactic dependent processing of empty categories, (see Section 2.2, P.18). The Extended Projection Principle (EPP, see Section 2.3, P.21) requires that argument structure of lexical items must be preserved and all sentences must have a subject. Movement traces and null pronominal (PRO) are said to be bound by their antecedent to form a syntactic chain. After movements, the traces serve as a silent copy of their antecedent whereas PRO shares the referential properties of its antecedent, even though they appear to be an empty category in the structure. This is called the Empty Category Principle (ECP).

Some psycholinguists (Aoshima et al., 2004; Bever & McElree, 1988; Felser et al., 2003; Frazier & Clifton, 1989; Frazier et al., 1983; Gibson, 1998; Gibson & Hickok, 1993) assume that empty categories share the semantic properties of its antecedent and the meaning of the antecedent can be reactivated at the position where a gap exists. The filler of the gap has to be kept in working memory until the subcategoriser is encountered. This indicates that processing difficulty will occur when the distance between the filler and its subcategoriser is increased. However, there is also evidence that the unbound dependencies are processed immediately (when the main verb is encountered), and in an incremental manner (Traxler & Pickering, 1996).

To test the application of the ñgap-fillerò theory from ECP, Boland (1997a) conducted a series of four experiments on the ñgapò associated with wh-questions. The conditions she used are listed below:

61) a. Ambiguous with two arguments

Which *salad* did Jenny *toss* [name: BILL]

b. Ambiguous with three arguments

Which *baseball* did Jenny *toss* [name: BILL]

c. Unambiguous with two arguments

Which necklace did Nancy *inspect* [name: SAM]

d. Unambiguous with three arguments

Which necklace did Jenny *describe* [name: SAM]

According to Government & Binding (GB), the syntactic subcategorisation to the verb used in 61a) & b) *toss* is ambiguous between two possible frames:

62) a. toss <NP>

b. toss <NP PP>

Semantically, *toss* has two meanings: *mix lightly* and *throw*.

Moreover, the verb in 61c) *inspect* takes two arguments only, whereas *describe* in 61d) can take three arguments. However, syntactically the subcategorisation of the two verbs is listed in 63) below:

63) a. inspect <NP>

b. describe <NP PP>

Thus, Boland made predictions according to three models that initially allow syntactic access:

64) i. Structural competition models[44]

61b=61a<61d=61c

ii. Constraints-based models

61b<61a=61d=61c

iii. Garden-path models

61b=61a=61d=61c

However, if the lexical meaning is accessed first, the predictions for the ambiguous conditions do not differ. For the unambiguous condition, the difference between 61c) & d) should not be significant either, as *SAM* can be the possible direct argument of the verbs. The difference between the ambiguous and the unambiguous condition is a bit difficult to predict by lexical-first theories. However, according to other studies (e.g.,

[44] Although Boland has proposed a *concurrent* model, it does not differ much from models that hypothesise that all syntactic structures are activated at the same time. In this current thesis, they are categorised as the structural-competition model.

Green & Mitchell, 2006; van Gompel & Pickering, 2001; van Gompel et al., 2005), an ambiguity advantage can be expected. Thus, a first assignment of meaning theory (such as LAST) would predict the same as the structural competition model, as in 64i).

Boland's finding confirmed the prediction from 64i). She claims that only subcategorisation information constrained the initial syntactic representations, and GB or MP allows subcategorisation information in the lexical entry to be used for syntactic structure construction. However, the finding cannot distinguish between the Structural-competition model and a model such as Late Assignment of Syntax Theory (LAST).

Boland also used multiple tasks such as auditory-naming and the auditory-lexical decision task. The naming task is said to be *sensitive* to syntactic processing, whereas the lexical decision task is said to be *sensitive* to semantic processing. These techniques are questionable, as they are not purely sentence processing tasks. To what extent does the interference between the task and sentence processing occur? Boland's data cannot answer this. As her data analyses are based on naming and lexical decision task performance, it is questionable whether the results can purely represent sentence processing, or just purely lexical processing.

To conclude, Boland's data cannot rule out the predictions given by late assignment of syntax theory. It is still an open question of the stage of syntactic processing. However, Boland's data also raise the possibility of late assignment of syntax. Although few (Townsend & Bever, 2001) claim to support such later parsing theories, the evidence for other claim suggests that late parsing is a sustainable hypothesis.

5.4 Discussion

This chapter has reviewed many studies in the past 20 years. There are two foci: the time-course of syntactic processing and the challenge to universality by the crosslinguistic data on the ambiguous RC attachment resolution. Although many psycholinguists believe and try to interpret their findings from a Chomskyan linguistic point of view, the evidence for syntax-first models is chaotic. As more and more data suggest at least some other cues can be initially accessed in terms of sentence processing. The syntax-first models however face a drastic challenge.

One possibility that early data supported the syntax-first models is that the materials used in laboratories are not *natural* enough. Indeed, sentences like 65) can be solved rapidly by context, intonational breaks, etc.

65) a. While Mary was mending the socks fell off her lap.
    b. Someone shot the servant of the actress who was on the balcony.

Moreover, how often (natural) are sentences like the famous ones in 66) produced in peopleǿs daily life?

66) a. The horse raced past the barn fell.
    b. The rat the cat the dog chased chased tastes rotten.

This is a key issue with the current sentence processing study, as the goal is to find natural sentence processing mechanisms. The laboratory sentences in 66) is good at detecting necessary structural building / re-building, but the missing concept is whether it is necessary to build structure in *natural* circumstances, especially initially.

The next part of the thesis will provide empirical evidence, with a series studies on the RC-attachment in Chinese to provide another language parameter for the investigation of universality, and further experiments on the issue of the time course of syntactic processing.

Part II

Tips of the Iceberg

Chapter 6

Tip One: Relative Clause Attachment Ambiguity

Part I has reviewed several contemporary psycholinguistic models for sentence processing. One of the main arguments between those models is in the timecourse of syntactic processing. As modern psycholinguistic models rely heavily on Chomskyan linguistic theories, the dominant view in the field is that syntactic processing should be assigned in an initial stage of the whole sentence processing procedure (even if it is not the first stage prior to all other information processing). Moreover, the processing should be accounted crosslinguistically because according to Chomskyan School, there should be a universal grammar. One key weakness of these claims is that they are based on a limited amount of language data. A wide range of studies of two-site relative clause (RC) attachment preference (e.g., sentences such as 26c), 38), 39), and 53)-57) above) has cast doubt on those theories and questioned the universality of syntax in Chomskyan theories. For the convenience of the readers, a sample sentence of this type is repeated below:

67) Someone shot the servant of the actress who was on the balcony.

NP_1 Prep NP_2 RC

The potential attachment sites (NP_1 and NP_2) for the modifier RC are both plausible in sentence 67) above. However, some languages, such as Spanish, German, Dutch, French, Russian, Thai, Japanese, Afrikaans, Croatian, Greek, Italian, Polish, and Brazil Portuguese, have an NP_1 attachment tendency, whereas others such as English, Danish, Norwegian, Swedish, Romanian and Arabic are suggested to have an NP_2 preference.

The theories and studies on RC attachment preference introduced in Part I suggest that at least some processing mechanisms are language-specific rather than universal. Although this does not answer the question of the timecourse of syntactic processing, it has questioned the baseline of universal grammar (and hence the theoretical background of contemporary psycholinguistic studies). The current thesis attempts to investigate the timecourse of human sentence processing and the starting point chosen is the RC attachment ambiguity. This is because if the theoretical background of the psycholinguistic study has been questioned, it is very difficult to build up further models based on the old system. As it has been suggested that there is an urgency to provide more language data to test the universality of the linguistic theory, this thesis attempts to provide some Chinese data regarding this issue.

Afterwards, the general issue of the timecourse of syntactic processing will be further tested and discussed in Chapter 7 and Part III of this thesis.

6.1 Chinese and the RC Attachment Ambiguity Issue

This section introduces some basic properties of the Chinese language, especially in relation to the RC attachment ambiguity problem. Some predictions in terms of the attachment preference in Chinese from the aforementioned theories will also be introduced. Finally, a self-paced reading study investigating the RC attachment preference in Chinese will be reported.

It is important to stress that Chinese has an ideographic writing system, and each symbol is equivalent to a morpheme. Modern Chinese (Standard Mandarin) derives from ancient Chinese, which had a different syntactic system in the written form from its spoken form. However, since the written form of this language was used primarily by people in high positions in a hierarchical society, it was respected and handed down in classical literature. It is naturally difficult to find out what the spoken language was like in ancient times; nevertheless, the modern Chinese have retained many phrases with syntactic structures of the ancient written language from literature and used them as components of modern Chinese. Due to this historical factor, the syntax of modern Chinese is a collection of syntactic rules and anomalies from a number of origins. It is difficult to apply some of the basic syntactic categorizations and divisions to modern Chinese since it is in fact a language composed of conflicting structures.

6.1.1 Word Order in Chinese

Word order is one of the basic functions of a language, but the question of whether Chinese is an SVO language (Huang, 1982) or an SOV language (e.g., H. Lin, 1999) has been subject to controversy for decades, and no consensus has been reached. It is beyond the scope of this thesis to summarize a whole literature of debates on this question; instead, in the current article, we will follow the constraints of Chinese word order as described in 68) (Li, 1990, p. 11):

68) The Chinese Word Order Constraint:

a. Chinese is head-final except under the requirement of Case assignment.

b. Case is assigned from left to right in Chinese.

c. A Case assigner assigns at most one Case.

Simply speaking, Chinese appears to be ñhead-initialò within VPs and preposition phrases, but strictly ñhead-finalò in NPs and postposition phrases.

Since Chinese is not an actively inflective language, word order is very important in determining the cases, states, etc. of a sentence. In some cases, changes of word order may even radically change the meaning of the whole sentence. However, because of the feature of the mixed word order constraint, in some other cases, the word order appears to be flexible[45].

6.1.2 The Syntactic Structure of Chinese Relative Clause

In Chinese, a relative clause always precedes the term it modifies (i.e., in Chinese, the relative clause is always attached to the left of a noun.), with no intonational break between the relative clause and the ñheadò noun. It always requires the use of the postposition *de*（的）[46], but lacks an equivalent among English relative pronouns, as in

[45] Although this is topic is out of the scope of the current paper, some further discussion will be provided in Section 6.6.5 concerning *topicalization* in Chinese.

[46] In many Chinese grammar books, *de* is not defined as a postposition, but rather an unclear *particle* (Yip Po-Ching & Don Rimmington, 1997). However, since its usage is equivalent to *no* in Japanese, *de* is best considered a postposition. Since a postposition is a particle, there is nothing contradictory about this. For consistency with other studies, *de* is defined as postposition in this article (further discussion please see Appendix F).

69):

69) a. **ni yao fu de** qian （你要付的錢）

you have to pay de money.

the money (that) you will have to pay

b. **tamen qu zhongguo de** na yi tian （他們去中國的那一天）

they went China de that day

the day (that) they went to China

The problem that arises from this kind of structure is as follows: when the modifier precedes a complex NP, such as structure 70a) below, in which both of the NPs are at the right side of the RC, a structural ambiguity is created similar to example 67). 70b) below is the Chinese ambiguous sentence equivalent to 67).

70) a. NP+VP+RC-de+NP_1-de+NP_2

b. 某人開槍打死了站在陽臺上的女演員的僕人。

mouren (NP) kaiqiang dasi-le (VP) zhanzai yangtaishang-de(RC) nüyanyuan (NP_1) -de puren (NP_2).

Someone (NP) shot (VP) standing-on-balcony-de (RC) actress (NP_1) -de servant (NP_2).

Someone shot the servant of the actress who was on the balcony.

As is shown in 70), the available potential heads for the RC are both the larger NP headed by NP_2 僕人(puren *servant*) and the smaller NP headed by NP_1 女演员 (nüyanyuan *actress*). Figure 17 illustrates the tree diagrams for the syntactic structures of 70b). 17a) is the structure for NP_1 attachment ï the PP that contains the RC is immediately under the domain of the lower NP. Contrariwise, in Figure 17b), the PP that contains the RC is under the domain of the higher NP, which can be understood as an attachment of RC to NP_2. Notice that since Chinese appears to be ñhead-finalò in NPs and PPs (similar to Japanese), the first NP (NP_1) that appears in the sentences is actually NP_{low} because it is the lower NP in the hierarchical constituent structure, as is shown in (21), and contrariwise, the NP_2 is then NP_{high} (cf. Miyamoto et al., 1999). For head-initial languages such as English and Spanish, however, the corresponding sites in the hierarchical structures are: NP1 for NP-high and NP2 for NP-low. As the position in the hierarchical structures for languages with different word orders remain constant, from now on a reference for the hierarchical position (i.e. NP-high and NP-low) will be used.

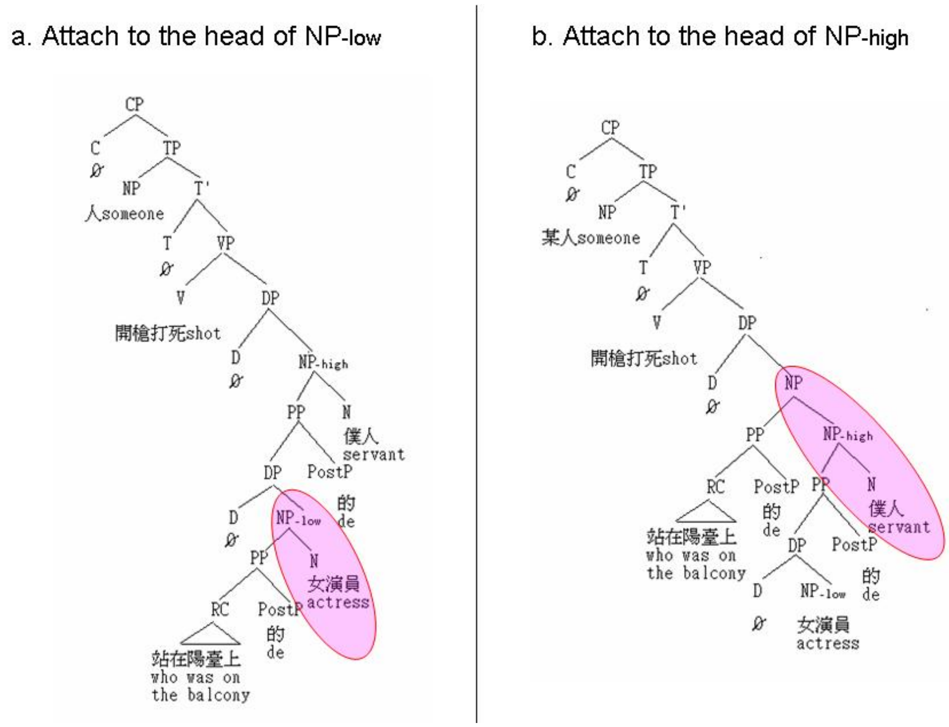


Figure 17 Relative Clause Attachment Preference Structures in Chinese

Moreover, as Chinese is not an actively inflective language, the structural ambiguity as in 70), cannot be disambiguated simply by inflective ways such as tense or number, cf. 71):

71) a. Lôavvocato diffida del padre <u>della ragazza</u> che ci è <u>tradita</u> al processo.

The lawyer suspects the father of the girl who betrayed [fem] (herself) at the trial. (Gender agreement)

b. Someone shot the <u>servants</u> of the actress who <u>were</u> on the balcony. (Number agreement)

The only way to disambiguate this type of sentence in Chinese is to modify the structure by changing the word order. For instance, one can advance the NP-low upon the RC so that the RC is preceding only the NP-high; now, since in Chinese the relative clause always precedes the term it modifies, the RC then only modifies NP-high. Taking 71b) as an example, the word order can be shifted as shown in 72):

72) 某人開槍打死了女演員的站在陽臺上的僕人。
mouren kaiqiang dasi-le nüyanyuan-de zhanzai yangtaishang-de puren.
Someone shot actress-de standing-on-balcony - de servant.
Someone shot the actressô servant who was on the balcony.

The tree diagram of 72) is illustrated in Figure 18 below:

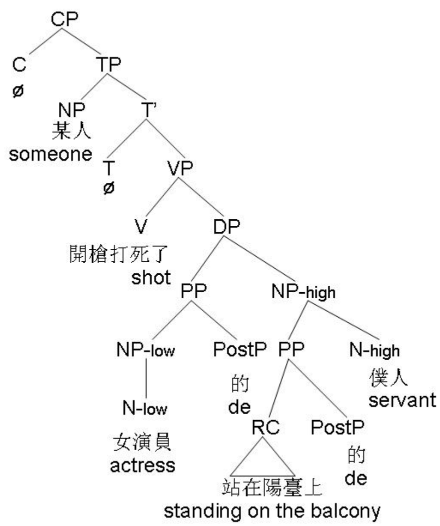


Figure 18 An Alternative NP-high attachment Structure

The structure described above shows evidence that the disambiguation in Chinese can be done by movement, but not by inflection. This special feature of the Chinese language, although it cannot be considered identical to the Saxon genitive form, does not rule out the resemblance of the RC to other languages. The Chinese sentence in 72) is comparable to the English unambiguous form shown in 56b) (in 5.1.3).

Moreover, comparing 70) and Figure 17, one can find that the syntactical structure and the linear structure are not identical in terms of the ñdistanceò between the constituents of the sentences. In other words, it appears that the main assertion of the sentence, 僕人 (puren *servant*) in 70b), does not directly follow the predicate 開槍打死了

(kaiqiangdasi-le *shot*); Yet structurally, this main assertion is still the nearest one to the predicate. These features will be used to make some special predictions in the next section.

6.2 Predictions on Chinese RC Attachment Preference

Not all theories and models introduced in Part I aim to address the RC attachment preference issue. Closely related theories (mainly introduced in Chapter 3) include the Garden Path Theory Group (GPTG; i.e., Garden Path Theory *(GPT*), Relativized Relevance (*RR*) and *Construal*), Recency Preference / Predicate Proximity (*RP*/*PP*), Attachment and Binding (*AB*), Exposure-based Tuning Hypothesis (*ETH*) and the Implicit Prosody Hypothesis (*IPH*). (See Chapter 3 & 5)

6.2.1 The Garden Path Theory Group (GPTG)

According to the GPTP principles introduced in 23) (P.36), a relative clause in Chinese will always cause a Garden-Path effect, since it has the special word order feature described in 68): Chinese is SVO and appears head-initial as far as case assignment is concerned (i.e. Chinese appears to be ñhead-initialò within VPs and preposition phrases, but strictly ñhead-finalò in NPs and postposition phrases). When the processor encounters the VP in a structure like 70a), it should always expect to receive an NP next as an argument. However, when the RC is received, the processor will be surprised and forced to ñre-analyseò the sentence and wait for the head of the NP to appear. In the case of processing relative clauses with complex heads, the situation is more complicated: when the processor first encounters the lower NP, it will assume this NP to be the head of the NP because both the Locality Principle in 10) and the Phase Impenetrability Condition (PIC) in 12) (P.28) have been satisfied. The processor can then terminate the VP processing, in principle. However, it then receives the postposition *de* and is informed that this first-met NP is not the actual head of the complex NP, so the processor has to ñre-analyseò the structure again until there are no more PPs or NPs occurring. Thus, the *GPT* predicts that the RC should attach to NP-low.

Nevertheless, although NP-low is closer to the case assigner VP in a linear structure as shown in 70), NP-high is the actual head of the argument. Therefore, the refined theory of *RR* will make a prediction that there is a shift from attaching to NP-low to the higher NP in the later stage because NP-high is closer to the head of the VP syntactically. Besides, NP-high is in the same processing region as the constituent to which the RC is initially attached, which satisfies the requirement of being re-construed (De Vincenzi & Job, 1993). If this were true, a significant bias of NP-high attachment would be expected in the results of off-line experiments, as in Japanese (see Kamide & Mitchell, 1997, for a discussion); but an NP-low attachment should be favoured in an on-line experiment in the early stage, followed by a diversion indicated by some other evidence (e.g., time delay) at a later stage. Nonetheless, this prediction has already been challenged by Shenȏs (2002) finding from an off-line questionnaire study that there is an NP-low attachment preference in Chinese.

Like in many other languages, the Chinese RC with complex NP heads falls into the principle of the ***Construal*** theory demonstrated in 27b. It predicts that if NP-low is an argument of NP-high, i.e., if NP-low is the substance or quantity of NP-high, a bias towards NP-high should exist; but if NP-low is not an argument of NP-high, such as in the *near* condition in Type 3 in Table 3 (P.83), there should be a bias to NP-low. However, inasmuch as discussed in the last section, Chinese has one and only one alternative unambiguous form for the RC, as in 72), which is relevant to the English genitive form, as in sentence 56b, i.e., a form where only one NP (NP-high, in this Chinese case) is attached. It follows that Chinese is more likely to apply the *Gricean Maxim of Clarity* (P.84) than the *RR*. This also means that when it comes to the ambiguous form, NP-low is more preferred than NP-high. Therefore, *Construal* theory predicts that there is a preference for low attachment, but when NP-low is an argument of NP-high, the low site will be less favoured than otherwise.

6.2.2 Recency Preference / Predicate Proximity (*RP/PP)*

For a two-site relative clause attachment as in 70), what can be suggested according to this theory is that the *PP* theory will postulate an NP-high preference, as the NP is syntactically (i.e., structurally) closer to the predicate. In this sense, the cost for NP-high is lower than NP-low. However, *RP* will postulate an NP-low preference for similar reasons explained in the predictions for *GPT* in the previous section. Unfortunately, we cannot make a convincing further postulation from this theory because there is no convincing parameter available since it is not known whether the parameter setting for the higher preference is strong enough to rule over the low preference.

However, although this hypothesis is weak and yet to be proved, we assume that the sum-up cost for NP-low is lower than for NP-high. The reason comes from the linear structure of Chinese relative clauses like 70a). Unlike many other languages that have been studied (e.g., English, Spanish, German and Japanese), where the position of the head of a phrase is generally unchanging, Chinese has a special word order, in which the head of a phrase appears in a different location depending whether it is in VPs or NPs. Regardless of the fact that NP-high is syntactically the head of the NP, it appears later than NP-low and does not directly follow or precede the verb in contemporary discourse. Thus, the memory decay effect will affect not only the processing of RC, but also of the VP, and hence the cost of processing the farther site, NP-high, will rise. If this were true, then NP-low is likely to be the lower cost one, and therefore more likely to be preferred in the case of RC attachment.

6.2.3 Attachment and Binding (AB)

It has been noted in the beginning of section 6.1.2 that there is no relative pronoun in Chinese. Recall that the *Attachment-binding* theory is also called *Relative Pronoun Drop Model* (see Section 3.1.2, P.44) because the preference of the attachment is determined by the anaphoric binding of the relative pronoun. The anaphoric binding is rather weak in Chinese ï even weaker than in English as the relative pronouns are sometimes used (recall that English is a weak anaphoric-binding language according to Hemforth et al. (1998)). *AB* predicts that there should be a strong bias towards the nearer NP, which is NP-low in this case

6.2.4 The Exposure-based Tuning Hypothesis (ETH)

ETH cannot be used to make predictions before any frequency data for that language have been collected. Unfortunately, no frequency data for RC attachment resolution in Chinese is available to the best of our knowledge, and therefore such predictions could not be made. Instead, a mini test was designed to supplement the corpus.

Subjects: 15 Chinese native speakers who will not participate in the main task.

Material and Procedure: Each of the subjects was asked to create 5-10 sentences with the derivational unity of the structure *RC de NP-low de NP-high*. The sentences should be natural and unambiguous. As introduced above, the only syntactical way to disambiguate the RC attachment sentences in Chinese is to change the word order. However, if the respondents used an unambiguous word order, the preference for the structure *RC de* NP-low *de* NP-high will not be clear. Thus, the subjects were asked to make semantically unambiguous sentences with a structure identical to the sample structures, e.g., for NP-low preference, i) is an example sentence:

i) 我弄脏了那个戴眼镜的同学的衣服。

Wo nong zang-le nage daiyanjing de tongxue de yifu

I made dirty-le the wearing glasses-de student-de clothes.

I stained the clothes of the student who wears glasses.

In i), since only people can wear glasses, the RC will attach to the lower NP site, *the student,* because this is the only valid potential site. Contrariwise, in ii), only NP-high is valid because only the light can twinkle, not the city:

ii) 我晚上在山顶上看见了闪烁的城市的灯光。

Wo wanshang zai shanding shang kanjian-le shanshuo de chengshi de dengguang.

I in the evening on the top of the mountain saw-le the twinkling de city de light.

I saw from the top of the mountain the light of the city that twinkled in the evening.

Analysis: 30.5% of the sentences are NP-high attachment sentences while 61% of the sentences are NP-low attachment sentences. 8.5% of the sentences are ambiguous sentences.

This task suggests that Chinese speakers use NP-low preferred sentences most frequently. The sentences created by the subjects are considered to reflect the characteristics of everyday language output and input to the parser. Thus, albeit tentatively, it is suggested that *ETH* predicts that there should be an NP-low attachment preference in Chinese, although there should be some relationship between other factors (e.g., age, second language, second language level, dialect, etc.) and the preference.

6.2.5 The Implicit Prosody Hypothesis (IPH)

There are many intonational cues such as stress, the fundamental phonetic peak (F_0), intonational breaks and length. As the current study does not focus on phonetic analyses, an easier approach to IPH is the cue of intonational breaks.

As standard RC in Chinese use ñdeò to connect the RC to the modifiees, it is not proper to insert a comma or any type of marker of intonation break between them. However, an intonation break can still exist without a marker after a long modifier, especially in sentence production. Moreover, in Chinese ñdeò can sometimes be dropped in a closer relationship between a modifier and its modifiee. Cf. 73):

73) a.某人開槍打死了站在陽臺上的女演員的僕人。
mouren kaiqiang dasi-le zhanzai yangtaishang-de nüyanyuan-de puren.
Someone shot actress-de standing-on-balcony - de servant.
(Ambiguous).

b. 某人開槍打死了站在陽臺上的 女演員(的)僕人。
mouren kaiqiang dasi-le zhanzai yangtaishang-de nüyanyuan-(de) puren.
(The servant is on the balcony).

c. 某人開槍打死了站在陽臺上的女演員 的僕人。
mouren kaiqiang dasi-le zhanzai yangtaishang-de nüyanyuan-de puren.
(The actress is on the balcony).

As ñsyntactic structure does not determine the size of a pause, although it may affect the probability of a boundaryò (Watson & Gibson, 2004, p. 715), the segmentation break may be ñbiggerò than a normal intonational break, but it should not affect the syntactic structure boundary. Thus, if there is an intonational break between NP-low and ñdeò, NP-low should be more preferred, and vice versa. IPH also predicts that the length of the RC can affect the preference[47].

6.3 Study 1: Self-Paced Reading on Chinese RC Attachment Preference

Most of the theories discussed above, including *GPT*, *Construal*, *RP/PP*, *AB*, *ETH* suggest that there is likely to be a low-attachment preference in Chinese; only *RR* predicts that this low-attachment bias will be switched to a high-attachment bias in a later stage. Although *RR*ô prediction has been questioned by Shenô (2002) off-line data, the current study attempts to investigate whether there is any sign of preference change in on-line tasks. *IPH* predicts that the preference will be affected by the breaks between NP-low and ñdeò. As the segment not yet revealed cannot join the current intonational group, the prosodic cue of intonational breaking can be tested by manipulating the segmentation of ñdeò in a self-paced task. In this section, a self-paced reading study will be reported to provide evidence for or against those predictions in relation to the relative clause attachment preference in Chinese.

6.3.1 Method

◆ **Participants**

40 volunteer native Chinese speakers participated[48], aged between 16 and 55: the people of age group 16-23 were mostly students and those at 24-55 were employed. At the time of the experiment, they were all living in the city of Nanjing, China. 17 of them were from the Jinling High School, 20 of them were from two local newspaper companies and three of them were current university students who graduated from Jinling High

[47] Also, a shorter RC will attract NP-low more than a longer RC. However, as the length of the RC was not controlled at the time when the study was carried out, no further investigation will be focused on this point

[48] In fact, 43 participants participated, but three of them later were excluded. One of them turned out to have Korean as his first language; the other two were eliminated because they were found to have used their mobile phones during the experiment.

School. Their language background was similar[49]: their first language was the Nanjing Dialect, a branch of Standard Mandarin, and their second language was English, which however was not very fluent according to a self-assessment from a questionnaire[50]. They were paid RMB￥5 each for their participation.

◆ **Materials**

24 test sentences were used in the experiment. Each sentence had six versions, varying in terms of compatibility of the RC with the alternative host-sites: RC consists with both NP sites 74a), NP-high 74b), and NP-low 74c) (segmentation points marked ñ/ò). Each type was further divided into two versions according to the point of occurrence of the local structural and prosodic cue, namely, the postposition *de*, the structural signal of modifier in Chinese. As Chinese cannot be inflected, the disambiguation was achieved by semantic congruence.

74) a. RC consistent with both NPs

某人/開槍打死了/和客人一起/站在陽臺上(/的)/女演員(/的)/男僕。

Mouren / Kaiqiangdasi-le / hekerenyiqi / zhanzaiyangtaishang (/-de) / nüyanyuan (/-de) / nanpu.

someone / shot / the-with-guests / on-the-balcony (/-de) / Actress (/-de) / male servant.

Someone shot the male servant of the actress who was on the balcony with the guests.

b. RC consistent with NP-high

某人/開槍打死了/和妻子一起/站在陽臺上(/的)/女演員(/的)/男僕。

Mouren / Kaiqiangdasi-le / heqiziyiqi / zhanzaiyangtaishang (/-de) / nüyanyuan (/-de) / nanpu.

someone / shot / the-with-wife / on-the-balcony (/-de) / Actress (/-de) / male servant.

Someone shot the male servant of the actress who was on the balcony with his

[49] The downside of this is that their similar background made it difficult to test possible factors affecting the preference according to the ETH.

[50] The students were asked to estimate their average score (full score being 100) of weekly English test at school on a four-band scale, ranging *a. 0-50, b. 50.5-70, c. 70.5-85, and d. 85.5-100*.

wife.

c. RC consistent with NP-low

某人/開槍打死了/和丈夫一起/站在陽臺上(/的)/女演員(/的)/男僕。

Mouren / Kaiqiangdasi-le / hezhangfuyiqi / zhanzaiyangtaishang (/-de) / nüyanyuan (/-de) / nanpu.

someone / shot / the-with-husband / on-the-balcony (/-de) / Actress (/-de) / male servant.

Someone shot the male servant of the actress who was on the balcony with her husband.

Each sentence had 5 to 7 (7-9 when *de* was a single segment) segments, with the average length of a segment being 4 characters: each NP being modified was about 2-3 characters in length, and the RC was about 4-10 characters in length (when it was longer than 6 characters, it was broken into two segments so that the segment before the NPs were within the range of 3-5 characters[51]). The word frequency of the two NPs in a sentence was controlled by considering those pairs that had the same band of magnitude (e.g., 0-9, 0-99, 100-999é) of matches in search engine[52] results. Thus, it would minimize the possible noise caused by the difference in length or frequency of the segments. The sentences were created by the experimenter, with the rate of naturalness ranging from 1.5 to 4, on the scale from 1 (very natural) to 5 (very unnatural), and the average rating was 2.26, ranked by voters who did not participate in the on-line experiment.

Moreover, 66 filler sentences were used. The lengths of the filler sentences were not particularly controlled, varying from 3 to 9 segments, because it is not vital to the result of the experiment. 16 of them were derived from the filler sentences used in the off-line study and 50 of them were chosen from standard Chinese literature books. Among those sentences, 29 of them contained one or more relative clauses, to make it more difficult for the participants to guess the strategy of the experiment.

[51] Admittedly, the downside of this control is that it eliminated the possibility of testing the length effects in IPH. However, this makes sure that the reading time in the segment before the NPs will not be affected by the uncontrolled length factor.

[52] The search engine is www.baidu.com

In addition, 28 yes/no comprehension questions were presented one following each of the 28 sentences that were chosen from the test and filler materials by the examiner (8 from test materials and 20 from filler materials).

- **Procedure**

The subjects were asked to give some personal information, specifically their age, gender, occupation, dialect, second language and what they considered their second language level to be, etc. (students were also asked for further language test scores). They were assured that their personal information would be treated confidentially and would be destroyed after the writing up of the experiment is finished.

The participants were first given a summary of the procedure and purpose of the experiment by the examiner. Later, they were introduced to a 14.1ò laptop screen, where a brief introduction was shown, similar in content to the examinerôs initial explanation. Once they had understood the procedure, they started the experiment with eight practice trials. They were allowed to stop to ask questions during this time. Once the practice trials were finished, they were introduced to the actual experiment. Each of the participants encountered a mixture of test sentences: four sentences of each type and the filler sentences in a pseudorandom order. Sentences were shown segment by segment from the beginning to the end, in the middle of a fixed computer window. Participants were asked to read one segment and press button Y to reveal the next segment until they finished reading the whole sentence, where a full stop was shown to signal the end of the sentence. Where there was a comprehension question, they were supposed to respond by pressing button Y for a positive answer and N for a negative answer. The feedback of ñ*Wrong! Please pay more attention!*ò (equivalent in Chinese) would be activated only when the response was wrong. The reading time (RT) of each segment and the error rate of the questions were recorded automatically through the DMDX programme (developed by J. Forster & K. Forster, at the University of Arizona).

There was one break in the middle of each experiment. The total duration of the experiment varied from 20 minutes to 30 minutes.

6.3.2 Results

Firstly, error rate was checked for each participant. The passing rate was set to be less than 14.2% error (which means four out of the 28 questions). This was to ensure that the participants were actually reading the sentence content. All of the 40 participants passed this check. The within-subjects effects were the RC attachment consistency types (as in 74a, b & c) and the appearance of *de*; the between-subjects effects were occupation and gender[53].

Based on most of the predictions and the result of the off-line questionnaire study by Shen (2002), which suggest an NP-low attachment bias for Chinese relative clause, it was assumed that the primary result should be shown on the reading time (RT) of the 5th (or 6th if *de* is a single segment) segment (NP-low). In other words, if the NP-low attachment bias was implemented on-line, increasing reading latencies were expected in condition 74b) (Type B) but not for conditions 74a) (Type A) and 74c) (Type C). Furthermore, the *de* appearance effects were expected that when *de* was attached to the segment of NP-low, cues for RC and warning of more information to come should be activated. The direction of the interaction of the reading time by *de* is hard to predict at the moment; however, the effects should mainly occur on the NP-high condition. This is because without the presentation of *de*, the first NP region can be seen as the end of the sentence; however, it is not the end of the sentence as there is no sentence ending marker (i.e., the full stop.). When the RC is forced to attach to NP-high, there might be a tendency for the reader to try to reveal further sentence components to solve the mismatch, whereas the single region of *de* can give the reader time and a chance to think; thus, when it comes to the final region, NP-high attachment is almost expected. Therefore, the trade-off effects by *de* will make NP-high attachment most favoured when *de* is a single region. When *de* is presented within a segment (e.g., within the first NP region), there is a sign of new coming component, but there is no region between the problematic region and the new region, the trade-off effects should not be expected and therefore the NP-high attachment bias should be the most difficult type to read. The RT

[53] The age factor is not considered as the between-subjects effect because the age and occupation are very strongly correlated.

of segments 1, 2, 3 and 4 were not analysed because they were considered to be of no research interest in this experiment.

6.3.2.1 General Results

The mean RTs for the first NP (i.e., NP-low) site and the second NP (i.e., NP-high) site were submitted to an analysis of variance with repeated measures for both subject and material analyses[54]. The first within subjects factor was the RT at the crucial regions under different RC attachment consistency conditions, namely, Type A, when RC was ambiguous (NP-both); Type B, when RC was forced to attach to NP-high and Type C, when RC was forced to attach to NP-low; the second within-subjects factor was the appearance of *de*, namely, whether it was attached to NP-low or detached as a single segment. Figure 19 below shows general mean RTs of two potential NP sites, with the consideration of the factor of RC type only.

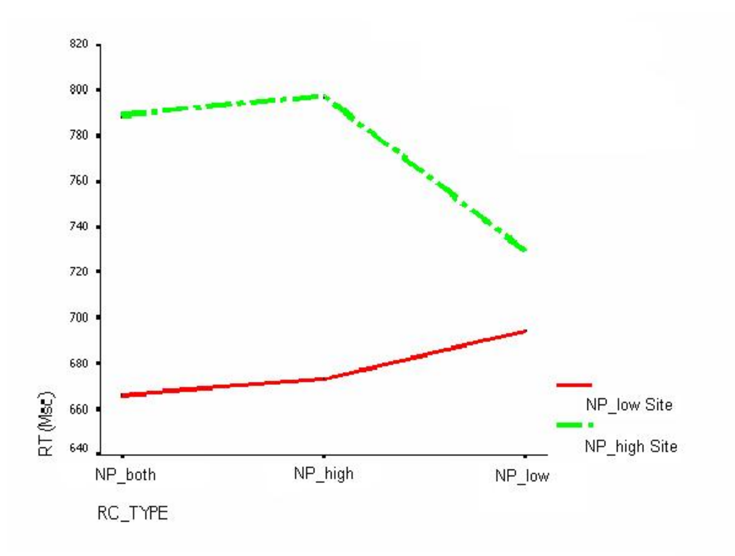


Figure 19 RTs at the NP sites in different Attachment Types

ANOVA shows that the RC consistency Types had no significant effect on the critical

[54] In psycholinguistic studies, it is generally agreed that there should be at least 6 items for each condition. For the current study (as well as Study 2 & 3), there are only 4. Therefore, the insignificance of the item analyses can be caused by lack of power in the design. However, the current thesis will follow the rule of discussing the effects with both F1 and F2 to be significant so that the conclusion can be safely generalised to both bigger participant pool and new materials.

NP-low site (Fs < 1, *p*s > .5). For the NP-high host site, the effects were not significant for the subject analyses (F_1 (2, 78) = 2.047, *p* = .146), although it was significant for material analyses (F_2 (2, 46) = 4.102, *p* = .023). This means that the results cannot be generalised to a wider range of participants.

Between-subject factors (gender and occupation) were also tested and no significant effects or interactions were found (Fs < 2.1, *p*s > .1).

6.3.2.2 Results of *de*

Table 5 shows the mean reading time of the segment of *de*. The average RT on this site was 542ms. A 3*3 ANOVA was carried out comparing the RT of three segments NP-low, *de* and NP-high, and the three attachment biases Type A, Type B and Type C. The results showed that there was no significant difference between the three segments (Fs < 1). This means that the reading time on this single character segment was not different from other sites with more characters. This indicates that this site was not simply skipped. Because there were no clear interactions of segment by attachment type (Fs < 2.3, *p*s > .07), this will not be focused on here.

Table 5 Mean RT of "de" (ms)

	Mean	Std. Deviation	N
NP-both	532.2527	139.76858	40
NP-high	534.1720	183.61786	40
NP-low	560.5858	267.71323	40

Another ANOVA was performed on the single word segment of *de*. There was only one within-subject factor: the RC attachment consistency type. However, no significant differences were found on this site (Fs < 1, *p*s > .5).

6.3.2.3 Results of the Interaction involving *de*

Figure 20 shows the reading times at the two NP sites with ñdeò either as part of the segment or ñdeò as a single segment.

ANOVA suggests that there was no significant interaction of RC consistency type * *de* appearance at NP-low site (Fs < 1, *p*s > .5). Since there was no interesting result showing at the first NP site, the following will focus on the NP-high site.

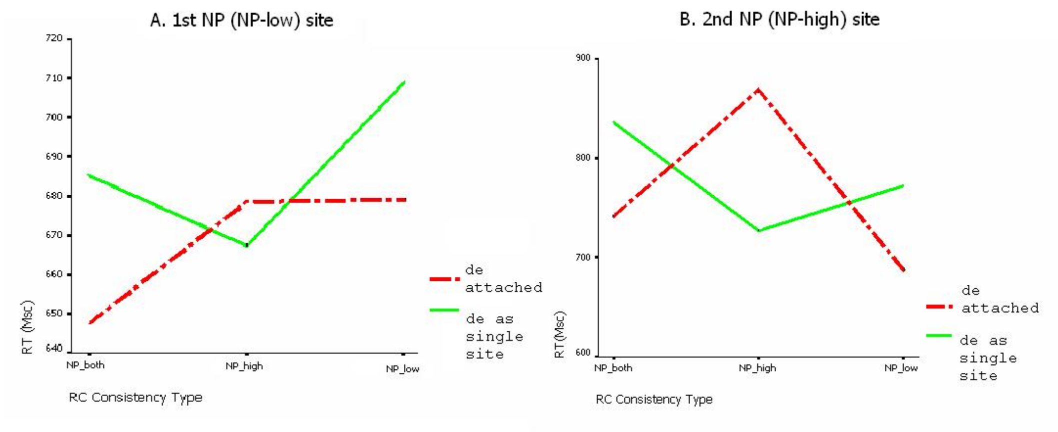


Figure 20 RTs for Different NP Sites with/without "de"

The mean RTs at NP-high are shown on the right of Figure 20. An initial analysis suggested that there was some significant interaction of the RC attachment types * the appearance of *de* (F_1 (2, 78) = 4.693, p = .014; F_2 (2, 46) = 5.128, p = .011). Further contrasts show that Type B was read longer than the average of Type A & Type C (F_1 (1, 39) = 6.618, p = .017; F_2 (1, 23) = 4.789, p = .039), but there was no significant difference between the RC Consistency Type A and Type C (Fs< 3, *p*s> .06). Moreover, the appearance of *de* interacts with the reading time: when *de* was detached, the RT for Type B was much shorter than when *de* was attached to the first NP (F_1 (1, 39) = 6.856, p = .013; F_2 (1, 23) = 11.393, p = .003). However the appearance of *de* does not interact with the RT for Type A & Type C (Fs < 1.3, *p*s > .1).

Furthermore, between-subject factors (gender and occupation) were tested and there were no significant effects (Fs < 2.1, *p*s > .1). However, there was a marginal interaction of RC types * *de* * occupation (F (2, 72) =3.12, *p*= .050).

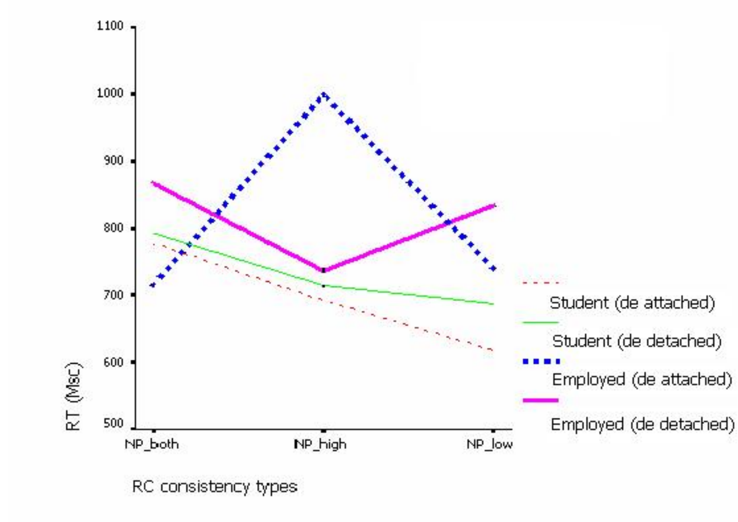


Figure 21 RTs of Different Occupation Groups

Further contrasts showed that when de was attached to the first NP, the employed group had difficulty reading sentences with RC with NP-high attachment bias (F (1, 36) = 5.271, *p* = .028), but not when RC was attached to NP-high (F (1, 36) = .366, *p* = .549). ANOVAs with separate groups suggest that the student group had a significant RC consistency effect (F (2, 32) = 3.640, *p* = .038), but no significant interaction of RC types * *de* (F (2, 32) = .177, *p* = .838); however, the employed group showed a significant interaction of RC types * *de* (F (2, 44) = 4.67, *p* = .014), but no significant RC consistency effects (F (2, 44) = 1.098, *p* = .342).

6.3.3 Discussion

The results from this self-paced reading study need to be dealt with cautiously. First of all, the pattern of reading difficulty in terms of the RT was consistent with the prediction of NP-low bias in some way, i.e., when RC was forced to attach to NP-high as in Type B, the RT was longer than ambiguous (Type A) or NP-low (Type C). There were also interactions of the appearance of *de* and the consistency types. However, this prediction was problematic in the sense that this pattern was not shown on the predicted crucial 5th or 6th segment (the first NP), but rather on the second NP.

To account for this, one suggestion is the contribution of both the gap-filler effects (chapter 5.3.2) and *spillover* effects. Chinese is an SVO language and the second NP (NP-high) is at the sentence ending position (when it is the object of a sentence). Thus, there were two gaps between the verb and the main object 75):

75) 某人開槍打死了站在陽臺上的女演員的男僕。

Mouren Kaiqiangdasi-le [gap1] zhanzaiyangtaishang ï de [gap2] nüyanyuan (pseudo-filler) ï de nanpu (filler).

someone shot [gap1] on-the-balcony –de [gap1] Actress (pseudo-filler)–de male servant (fillerõ).

Someone shot the male servant of the actress who was on the balcony with the guests.

Since Chinese RC has a structure like **[…] + gap + filler**, it requires more time at the filler position (which is NP-high here) in order to process all the information stored in the gap(s) prior to the filler[55]. Among the various filler-gap hypothesis, the *Active Gap Hypothesis* (AGH) (C. Lin & Fong, 2005) is particularly proposed to account for the Chinese RC condition. According to AGH, there should be a slowdown at the gap phrase (here, the RC and the NP-low), and the processor actively seeks a filler to satisfy a gap held in memory. Because of the memory recourse cost, once the processor detects the filler, it will assign this filler to the gap as soon as possible. However, in the case of the object NP, the actual filler (the head of the NP) is marked by a full stop. Thus, although the first NP (NP-low) will be assigned as pseudo-filler whether the RC is consistent with this NP, as long as the full stop did not appear, it is possible to assume more information is coming, i.e., the structure building is open until the next segment is processed. This could suggest some possibility of effects at a later point (in this case, the second NP) that reflects the delayed previous parsing.

Another possible explanation is simply that the (dis)ambiguation in the three conditions is semantic rather than structural. This is not a new phenomenon in terms of self-paced reading studies (McRae et al., 1998). It has captured the ñspirit of the garden-path model in that the non-structural information was delayed by a regionò (ibid., p. 299).

[55] Cf. the normal head-initial languages (such as English) where the structure is **filler + […] + gap**.

However, this does not account for the interaction of *de* and RC consistency.

The results in Figure 20 show that the interaction of *de* and RC consistency makes it more difficult to process NP-high attachment bias at the NP-high site when *de* is attached to the NP-low site. One may think that the existence of this RC cue can predict new coming information more strongly when *de* is attached than when *de* is detached. However, when *de* is attached to NP-low, there is only one segment before NP-high (NP-low-*de* / NP-high), but when *de* is detached, there are two segments before NP-high (NP-low / *de* / NP-high). When RC is inconsistent with NP-low, the spillover effects from the previous segment(s) are stronger for one segment than for two segments. Since there was no difference between the reading time of the segment of *de* and the two NP sites, it can be suggested that the processor is actually processing during this site.

Assuming the above explanations were true, the results of this self-paced reading task can support an NP-low bias for the RC attachment ambiguity in Chinese, which in turn supports the prediction from the models suggested in chapter 6.2, except for the *RR* theory. However, the result from between-subjects and RC attachment consistency interactions is very difficult to explain. It appears to suggest that different groups of people with similar language background show different processing patterns. According to the results, the employed people appear to have an NP-low bias, but the students were really good at seeking the unambiguous solutions: their RT was longer for the ambiguous sentences whether the RC was consistent with NP-low or NP-high. It seems that the students did not process the sentence segment by segment, but tried to store all the previous segments and process them together at the end.

The only account relating to this point is the experience based explanations (ETH). High school students in China face the heavy burden of college entrance examinations and they are “buried” in textbooks and exercise books, unlike the employed people, who have no need to read such material again. The textbooks and exercise books tend to use unambiguous sentences so that the students do not make mistakes in exams caused by misunderstanding of the texts. If the materials the students use do not have a RC bias, students would be better in reading unambiguous sentences than ambiguous sentences.

The employed group read ñnormalò texts; hence their result fits the general pattern of Chinese RC attachment. Moreover, the students are more competitive than working people in terms of seeking the ñright answerò. Because there were comprehension questions for some sentences, the students cared more about the feedback just like their real-world study environment, which have affected the result.

Nevertheless, those conditions were not controlled in this study and therefore no conclusion can be drawn based on the results other than that it is possible that Chinese RC attachment may have an NP-low bias, and the bias is delayed by approximately a region. The next two studies focus on these two points, namely the attachment bias and the delayed effects, combining with one new proposal of topicalisation (Hemforth & Schimke, 2003), with a view to answering some of these unresolved questions.

6.4 Study 2: Website-based Questionnaire for English and Chinese Speakers

Study 1 has tested several of the sentence processing theories and models discussed in 6.2. One further proposal relating to the RC attachment ambiguity resolution called *Topicalisation* was proposed by Hemforth and Schimke (2003). They propose that, at least in some languages (e.g., German and Spanish), the attachment preference varies depending on the position of the complex NPs (Hemforth & Schimke, 2003), while in some other languages (e.g.，English and French) it does not. An example of such sentence constructions can be seen in 76) below:

76) a. Complex NPs in the subject position: *The servant of the actress who was on the balcony looked at the sky.*
b. Complex NPs in the object position: *Someone shot the servant of the actress who was on the balcony.*

Hemforth et al. (2003) argue that the reason for this language difference lies in the topicalisational function of the preverbal position. In languages like Spanish and German, the preverbal NP has a more marked topical status than in English or French.

Accordingly, a topicalisation effect is expected to be found in both Chinese and English when the complex NP is passive subject vs. active subject, since the passive subject has a strong mark of topical status (Radford, 2001).

Moreover, the subject position in Chinese has a topicalisation feature similar to that in Spanish and German:

77) a. Descriptive: 我打人了。

Wo da ren-le.

I beat-up someone. (I have beaten up someone.)

b. Passive: 人被我打了。

Ren *bei wo* da-le.

Someone by me beaten-up. (Someone *was* beaten up *by me*.)

c. Topicalisation: 人我打了。

Ren wo da-le.

Someone I beaten-up. (I have beaten up someone.)

77a) is an active descriptive sentence. The default word order is SVO (see Chapter 6.1.1 for a reference). 77b) is a passive sentence, where the default word order is O 被 SV in Chinese. The PP (被(bei)+NP, equivalent to *by+NP* in English) is a typical marker of passive sentences in Chinese. However, the word order in 77c) is OSV, which is similar to 77b), but it is not a passive sentence because it does not have the passive marker 被 (bei, *by*). This word order is not grammatical in Chinese unless it is used to emphasise or to *topicalise* the object of the sentence. This means that the preverbal position is reserved for the topical status in Chinese. According to the proposal by Hemforth et al., a position effect caused by the topicalisation feature should be expected in Chinese just as in German and Spanish.

Two studies (Study 2 and Study 3) were carried out, attempting to investigate the proposal of *topicalisation*. Study 2 asks participants to choose between the potential heads for the RC attachment in a web-based questionnaire. All of the potential NPs of the attachment of RC are in the preverbal position in this study. The comparison is between passive and active expressions. As the experiments reported by Hemforth et al.

only tested the interaction of attachment preference by position (Hemforth & Schimke, 2003) due to topicalisational feature in one language (French), a further prediction from their proposal would be that an interaction of attachment * position * language should be expected when comparing two different types of languages (according to their topicalisation status in the preverbal position). Moreover, Hemforth et al.ôs experiment design focused on the position effects, but it did not cover the question of how is the general attachment preference affected by topicalisation directly (i.e., attachment * topicalisation) within a language. Study 2 aims to make up for this inadequacy.

Study 3 traced participantsô eye movements to find out whether the potential reading difficulty is caused by topicalisation. The design is similar to Hemforth et al.ôs, involving different positions and topicalisation levels. The details of the study are reported in the next section in more detail.

6.4.1 Method

- **Participants**

144 volunteers participated, of whom 72 were native English speakers and 72 were native Chinese speakers, aged between 16 and 65. They accessed the experiment through the Internet and were not paid for their participation.

- **Materials**

The experiment comprised 24 test sentences with the complex NP being at the subject position of the sentence in English and Chinese. The main clause complex NPs were presented in three different forms as 78) below: A. Active subject complex NP; B. Passive subject complex NP; and C. Complex NP as subject for the verb ñbeò. Each type was further divided into two versions according to the thematic assignment difference by varying the preposition (English) or postposition (Chinese) *of* (the main clause thematic assignee) and *next to* (the RC thematic assigner) in the RC.

78) a. active subject complex NP (Active)

站在陽臺上的女演員（/身邊）的男僕往天上看。

zhanzaiyangtaishang-de nüyanyuan(/shenbian)-de nanpu wangtianshangkan.

The servant *of* / *next to* the actress who was on the balcony looked at the sky.

b. Passive subject complex NP (Passive)

站在陽臺上的女演員（/身邊）的男僕被人開槍打死了。

zhanzaiyangtaishang-de nüyanyuan(/shenbian)-de nanpu beirenkaiqiangdasile.

The servant *of* / *next to* the actress who was on the balcony was shot dead.

c. Complex NP as subject for the verb ñbe (是)ò (Controlð where the VP cannot be passivised)

站在陽臺上的女演員（/身邊）的男僕真是酩酊大醉。

zhanzaiyangtaishang-de nüyanyuan(/shenbian)-de nanpu zhenshimingdingdazui.

The servant *of* / *next to* the actress who was on the balcony was drunk.

20 of the test sentences were derived from an off-line RC attachment study conducted by Shen (2002) and were modified to fit the requirements of the current study. The word frequency of the two NPs in each sentence was controlled as in the previous experiment. The English version is a direct translation of the Chinese one. This was designed to minimize noise effects resulting from differences in the familiarity or frequency of the potential choices. The sentences were rated for naturalness ranging from 1 to 5, on the scale from 1 (very unnatural) to 5 (very natural), and the average rate being 3.65, by raters who did not participate in the experiment.

In addition, 46 filler sentences (including 14 other types of ambiguous sentences and 14 unambiguous sentences containing RCs as foils) were used. All of them were derived from the filler sentences used in Study 1.

Following each sentence, a multiple-choice question was asked according to potential attachments of RC. Participants were expected to choose one from the two potential answers and a filler-answer. For example, for sentence 78) above, the question and choices would be 79)

79) Who was on the balcony? (   )

a. the actress   b. the servant   c. the murderer

The whole experiment consisted of nine pages and the actual test occupied six pages, with four test sentences on each page. Each of the participants encountered a mixture of test sentences containing four sentences of each type (recall the sentence versions introduced above). And the all the sentences were presented in a fixed order[56].

◆ **Procedure**

Visitors to the website address were attracted using emails and some internet forums. Participants were asked to enter the website voluntarily. Once they entered the page, they could choose the language they wanted to use and then read an introduction in their chosen language. They were encouraged to choose their native language. Answers in non native languages were discarded in the final analysis.

In Study 1, there were some occupation effects on the attachment preference. With reasons being unclear, it would be interesting to see whether the effects would be replicated by further studies, and whether there would be other effects caused by other between subjects factors. Therefore, before the test page, participants were asked to fill in some personal information as to their age, gender, occupation, second language and how they considered their second language level (on a 5-point scale from very poor to native speaker fluency level), etc. The participants were promised that their personal information would be treated confidentially and would be destroyed after the experimental report had been compiled.

Each participant started the experiment with three practice trials. By clicking the ñnextò button at the bottom of the page, they could continue the experiment. They were not allowed to return to any of the previous pages to change their answers (but they were permitted to change answers in the current page, even though they were not encouraged to spend too long on one trial.). They were not told whether there was a ñcorrectò answer or not. At the end of the experiment, by clicking ñsubmitò button the result would be emailed to the experimenter via a cgi-mail based system.

[56] The sentence order in each page was inputted pseudo-randomly but once the page was made, it would not change any more.

◆ **Programme**

The website programme was written using the web design programme *Dreamweaver MX*. The site contained three web pages: the first and second pages were simply linked to each other but the third page was further divided into seven layers so as to appear to the participants as seven separate pages for the purpose of data collection.

6.4.2 Results

According to previous studies in English and Chinese, NP-low preference was expected for both languages; moreover, an interaction of attachment preference by different preposition / postposition (i.e., of (的) and next to (旁邊的)) were expected (according to Gilboy et al., 1995; Shen, 2002), with *of*-condition to have less NP-low effects. Furthermore, according to the topicalisation proposal of Hemforth et al. (2003), stronger NP-low preference for 78b), compared with the average of 78a) & 78c) was expected to be caused by topicalisation. The predictions involving the factor of languages are rather complicated. There are two possibilities: i). since the NPs are all at the preverbal position, no interaction of attachment preference by topicalisation by language is expected because in condition 78b, both languages have been topicalised by the passive expression; and ii). Chinese speakers have a stronger NP-low preference for 78b) than English speakers because the default topicalisational preverbal position reinforces the topicalisation effects by the passive expression in Chinese, whereas no such reinforcement occurs in English. This means that English should have no stronger NP-low preference than Chinese by all means.

6.4.2.1 Results for Chinese

An average of more than 63% preferred the lower NP, as shown in Table 6 below:

Table 6 Percentage of NP-low Choices in Chinese

	Mean	Std. Deviation	N
Active, of	71.8750	30.70771	72
Active, next to	84.3750	24.62304	72
Passive, of	63.1944	33.03120	72

Passive, next to	82.2917	26.68198	72
Be, of	69.0972	32.36376	72
Be, next to	84.0278	27.29132	72

The mean percentage of choices of the preferred NP site was submitted to an analysis of variance with repeated measures for both subject and the material analyses. The first within-subject factor was the attachment bias (i.e., NP-high, NP-low[57]); the second within-subjects factor was the topicalisation status, namely, active, passive and prior to *be*; the third within-subject factor was the difference of PP, *of* and *next to*. The between-subjects factors were age, gender, occupation, second language and their second language level.

ANOVA shows that the preference for NP-low was significant (Fs >74.624, *p*s < .001). There was also significant interaction between attachment preferences by PP variations (Fs >35.512, *p*s < .001), with *of*-condition being voted lower than the *next-to*-condition. However, it was not significant for the main effect differences of the topicalisation status (Fs <1, *p*s > .4) or for the PP variations (Fs <2.19, *p*s > .1). Table 6 showed that when the postposition was the RC thematic assigner, more NP-low choices were made. Furthermore, no other interactions (besides the aforementioned) were shown from the analyses (Fs <2.3, *p*s > .1).

As for the differences recorded between the participants, there were no significant results found. (Fs < 1.3, *p*s > .2). Consequently, most of the between-subjects factors (except their first language) will not be discussed further.

6.4.2.2 Results for English

The percentage of the NP-low choices for English speakers was shown in Table 6 below:

Table 7 Percentage of NP-low Choices in English

	Mean	Std. Deviation	N
Active, of	77.7778	28.9352	72
Active, next to	84.3750	24.6230	72
Passive, of	76.0417	28.8992	72

[57] Since only 7 out of 144 answers (4.8%) made the choice of the filler NP, they were coded as missing data in the analyses.

Passive, next to	90.6250	19.8752	72
Be, of	74.3056	29.9566	72
Be, next to	84.0278	26.6384	72

ANOVA shows that the NP-low preference was significant (Fs >189.091, *p*s < .001). There was also a significant interaction between attachment preferences by PP variations (Fs >22.013, *p*s < .001), with *of*-condition being voted lower than the *next-to*-condition. Similar to the Chinese data, Table 7 shows that when the postposition was the RC thematic assigner, more NP-low choices were made. But as in Chinese, it was not significant for the differences of the topicalisation status (Fs <1.051, *p*s > .3) or for the PP variations (Fs <2.7, *p*s > .1). Furthermore, no other interactions (besides the aforementioned) were shown from the analyses (Fs <3.4, *p*s > .07).

6.4.2.3 Comparing Chinese and English

Table 8 below showed the average choices on NP-low for both Chinese and English speakers:

Table 8 Percentage of NP-low Choices in General

	Mean	Std. Deviation	N
Active, of	71.8750	30.70771	144
Active, next to	84.2014	24.5139	144
Passive, of	69.4444	31.6289	144
Passive, next to	86.1111	24.3405	144
Be, of	71.7014	31.1838	144
Be, next to	83.8542	26.8492	144

First of all, an overall NP-low bias were shown by ANOVA (Fs >293.429, *p*s < .001). There were also significant interactions of attachment by participantsôfirst language (F_1 (1, 142) = 4.038, p = .046; F_2 (1, 46) = 4.590, p = .037), and attachment by PP differences (Fs > 62.578, *p*s < .001). The interaction of attachment preference by PP by first language was significant for subject analyses (F_1 (1, 142) = 5.275, p = .023), but this result cannot be generalised to the wider participant pool (F_2 (1, 46) = 2.790, p = .102) (see footnote 54 (p.117)). Contrarily, the interaction of attachment by topicalisation variation by first language was not significant by subject analyses (F_1 (1, 142) = 1.832, p = .162), although it was significant by material analyses (F_2 (1, 46) =

3.627, p = .031). As there were 72 participants in the study, this is strong enough to reject this interaction as it cannot be generalised to a wider participant pool. These two interactions will not be focused on later.

Nevertheless, contrasts show that there was some marginal interaction of attachment by topicalisation by PP differences by language (F_1 (1, 142) = 3.067, p = .082 (marginal); F_2 (1, 46) = 5.380, p = .025), when comparing passive condition and the average of active and *be* conditions. There was no such interaction when comparing active and *be* conditions (Fs < 1.09, *p*s > .3). With a closer look at Table 6 & 7, it becomes clear that there were less NP-low choices in Chinese when passive expression were used; on the contrary, more NP-low choices were made in English in this condition. This suggests that although topicalisation alone had no particular effects on the RC attachment preference, it might affect the preference considering the thematic assigners of the NPs (i.e., the verb or the PP).

As the current analyses mixed English and Chinese data, the first languages were submitted for a between-subjects analysis. However, there were no significant results shown (Fs < 1.5, *p*s > .2).

6.4.3 Discussion

The overall results replicated the results of previous studies in both Chinese and English: there was a strong bias towards NP-low for the RC. Moreover, changing the PP in the RC affected the preference of the attachment. Both Table 6 & 7 above showed clearly that both English and Chinese participantsô choice of NP-low rose when PP was the thematic assigner of the relative clause, in this case, the 旁邊的 / *next to* condition. This corroborates the design of the current study.

As the purpose of this study was to test the proposal of *topicalisation*, the interactions of attachment by topicalisation and attachment by topicalisation by language should be most interesting. Hemforth et al.ôs proposal predicts a stronger NP-low bias for the topicalised condition; furthermore, the topicalised condition in Chinese should have no less NP-low bias than English.

The former hypothesis was not confirmed by the data. The latter interaction was marginal, but the actual stronger NP-low bias was from the English data, which is the opposite of the prediction. Although it is still too early to claim Hemforth et al.ôs proposal is false, the real cause of the position*language interaction in their study is very suspicious. The current web-based questionnaire could not replicate their study because of the design issue. Moreover, the questionnaire study allows rethinking and reconsideration. It is possible that the loose control of timing covered up some earlier effects.

The next study continues to focus on *topicalisation*, making use of a more similar design to Hemforth et al.ôs study, i.e., comparing the attachment preference at different positions (i.e., subject position and object position) of the sentence. Moreover, the next study uses the methodology of eye tracking to explore the online behaviour of the sentence processing of RC attachment issue.

6.5 Study 3: An Eye Tracking Experiment on *Topicalisation*

Since Study 2 only demonstrated participantsô final decision of the NP attachment, other factors might cause people to change their initial preference when choosing the answer. Even though there was no reliable topicalisation effect found in Study 2, the possibility of initial effects left unrevealed by the testing methodology cannot be ruled out. Moreover, Hemforth et al. only found an interaction of attachment (i.e., NP-low or NP-high) by position (i.e., subject or object), but because their proposal states that the topicalisation feature is embedded in a special language property, the interactions such as attachment by language or attachment by position by language may be expected. To clarify, Chinese should have a stronger NP-low bias than English when the complex NP is at the preverbal position because of the topicalisation feature in the language. To investigating these possibilities, the current study made use of the eye tracking technique and changed the factor of PP variation to a factor of position difference (i.e., subject and object). Reading time and eye movements were recorded, analysed and interpreted.

6.5.1 Method

◆ **Participants**

72 volunteers participated, of whom 36 were native English speakers and 36 were native Chinese speakers[58], aged between 18 and 43. All of them were students at the University of Exeter. 36 of them were from the School of Psychology, 17 from the School of Finance and 19 from other departments. 27 were male and 45 were female.

◆ **Materials**

The experiment comprised English and Chinese versions of 24 test sentences. Following the design of Study 1, each sentence had six conditions: with RCs semantically forced to attach to i) both NPs (ambiguous); ii) NP-high, and iii) NP-low, where the complex NPs, being the patient of the action, were at either a) subject or b) object position. Examples are shown in 80) below:

80) a. **Subject position**

i) NP-both

The male servant of the actress *who was on the balcony with the guests* was shot to death.

和客人一起站在陽臺上的女演員的男僕被人開槍打死了。

He kerenyiqi zhanzai yangtaishang-de nüyanyuan de nanpu beiren kaiqiang dasile.

ii) NP-high

The male servant of the actress *who was on the balcony with his wife* was shot to death.

和妻子一起站在陽臺上的女演員的男僕被人開槍打死了。

He qiziyiqi zhanzai yangtaishang-de nüyanyuan de nanpu beiren kaiqiang dasile.

iii) NP-low

58 Initially there were 48 Chinese speakers participated, but the data for 12 participants had to be discarded due to incorrect screen resolution settings on the testing system.

The male servant of the actress *who was on the balcony with her husband* was shot to death.

和丈夫一起站在陽臺上的女演員的男僕被人開槍打死了。

He zhangfuyiqi zhanzai yangtaishang-de nüyanyuan de nanpu beiren kaiqiang dasile.

b. **Object Position**

i) NP-both

Someone shot the male servant of the actress *who was on the balcony with the guests*.

某人開槍打死了和客人一起站在陽臺上的女演員的男僕。

Mouren kaiqiang dasile *he kerenyiqi zhanzai yangtaishang-de* nüyanyuan de nanpu.

ii) NP-high

Someone shot the male servant of the actress *who was on the balcony with his wife*.

某人開槍打死了和妻子一起站在陽臺上的女演員的男僕。

Mouren kaiqiang dasile *he qiziyiqi zhanzai yangtaishang-de* nüyanyuan de nanpu.

iii) NP-low

Someone shot the male servant of the actress *who was on the balcony with her husband*.

某人開槍打死了和丈夫一起站在陽臺上的女演員的男僕。

Mouren kaiqiang dasile *he zhangfuyiqi zhanzai yangtaishang-de* nüyanyuan de nanpu.

The sentences were derived from Study 1 and were modified to fit the requirements of the current study. The word frequency of the two NPs in a sentence was controlled similarly to the previous two studies, as described in section 6.3.1. The English version is a direct translation of the Chinese. The sentences were checked by native speakers, and rated for naturalness on a scale of 1 to 5, on the scale from 1 (very unnatural) to 5 (very natural). The average rates were 3.96 for the Chinese version and 3.05 for the English version. In the preverbal position 80a), all the sentences used passive expression to reinforce the topicalisation status.

Moreover, 46 filler sentences were taken from Studies 1 & 2. In addition, 28 yes/no questions were presented according to the content of some of the sentences. Among those questions, 9 were from experiment sentences and 19 were from filler sentences. Participants were expected to choose one correct answer after they finished reading each of the sentences, e.g., for sentences 80i) above, the question and choices would be 81):

81) Was the servant dead?

The whole experiment was counterbalanced with each condition appearing in four experimental sentences for each participant only once. All material was presented in a pseudo-random order for each participant.

◆ **Procedure**

The randomising programme was implemented as an Excel macro and was run before each experiment. The pseudo-random order was transferred into the Eyelink programme. Each participant was taken into a test cubicle equipped with two computers and an eye tracking machine. They were asked to sit down in front of one main computer connected to the eye-tracking helmet and a game controller. Then they were asked to put on a helmet with one camera adjusted to their right pupil. After training for ten minutes, they should have learnt to calibrate and validate the camera by concentrating on an adjusting dot in the middle of the screen. Once the camera was set, they were introduced to the actual experimental material. They were introduced to four practical trials including button pressing: a big round button was pressed when they finished reading a sentence and decided to go on; a left control button was pressed indicating the answer ñnoò while a right control button indicating ñyesò. A calibration dot appeared on the screen to readjust the camera following each sentence; the big round button on the handset was pressed by the participants themselves to finish the recalibration. When the dot disappeared a new sentence would start. They were allowed to ask questions during the practical trials. Once the real test started, the examiner left the test cubicle and let the participants read the sentences on the screen completely without disturbance. When the whole task was completed, the participants alerted the examiner to re-enter the room to take off the helmet for the participants and save the data in an .EDF file.

The whole experiment normally took up to 20-30 minutes depending on the reading speed of the participants. They were paid ʐ3 for their participation. In five cases, the experiment took about 40-60 minutes because of difficulties setting up the camera. Those five participants were paid ʐ6. The reward was paid after the participants finished the experiment.

- **Apparatus**

The apparatus applied in this experiment was the Eyelink® II eye tracker developed by SR Research Ltd., connected to two Dell computers. The eye tracker has a head mounted system with two miniature cameras mounted on a comfortable padded helmet and an extra camera in the middle of the helmet to determine the central position of the head. The two eye cameras allow binocular eye tracking with built-in illuminators in each of them. The screen was set to the resolution of 800*600 pixels and each sentence was placed at the screen coordinate (67, 175) with the size of 676 * 385 pixels. The examinerȏs computer was equipped with the Eyelink® II set-up and control programme so that all the calibration and validation could be controlled through this screen.

6.5.2 Results

According to the *topicalisation* proposal of Hemforth et al. (2003), it was predicted that there should be a stronger NP-low preference in both languages when the head of the complex NPs at subject position is topicalised by being a patient of the action (i.e., passive expression). Moreover, an attachment by position interaction was expected to be found in Chinese also because it is a language similar to German and Spanish, whose subject position has a topicalisation feature. There should also be interactions of attachment * position * language, when comparing English and Chinese, with Chinese preverbal position being most biased to NP-low.

Furthermore, since the two languages both prefer NP-low attachment, a reliable difference in reading time between condition was expected 80ii) comparing the mean of condition 80i) and 80iii) in both positions, whereas no difference was expected between condition 80i) and 80iii). This can be seen as a simple confirmation of the validation of the design.

6.5.2.1 Data Extraction

The current study attempts to explore relatively early stages of sentence processing. Therefore the main interest focused on three kinds of information, namely, *first pass time*, *regression path time* and *total time*. The predicted direction of the effects is that the longest reading time occurs at the NP-high attachment bias condition for both English and Chinese throughout the time measurements. Moreover, Chinese should have a stronger NP-low bias at the subject position due to the topicalisation feature of this language. Moreover, as Vonk and Cozijin (2003) argue, since reading is a nonlinear process where some words are skipped but their meaning is still processed, and there is no evidence to prove that the parser pauses during saccade movement, a consideration of saccade movement is necessary. Consequently, each measurement including saccade time was extracted, together with the measurement of *forward reading time* (Vonk & Cozijn, 2003) to investigate whether there was a significant amount of processing happening during saccade time. The definition of these measurements has been introduced in Section 4.2.2 (P.65).

6.5.2.2 Results for Chinese

According to Chinese structure, specifically, RC-de + NP-low-de + NP-high, the disambiguation region would be the first NP (NP-low). This is because when the RC has been processed, the processor will look for a potential NP to modify. When it encounters NP-low, if it is attachable, the attachment will be natural and smooth; but if it is not attachable, the processor will be confused by the difficulty and either go back to check the previous information, or carry on to look for another possible NP (see Study 1, section 6.3). Thus, an increased decision time on NP-low was expected. However, the effects were not shown on NP-low but NP-high in Study 1. The possibility of a spillover effect was considered and in this experiment data on both NPs were analysed.

Repeated measures of ANOVAs were carried out for both subjects (F_1) and material analyses (F_2) for each of the measurements. Both types of analyses contained three attachment conditions (80i (ambiguous), 80ii (NP-high bias), and 80iii (NP-low bias)) as first within subjectsô factors and two positions (80a (subject position), and 80b (object position) as second within subjectsô factors. The between subjects factors were not of much interest because no effects were found in Study 2. Moreover, as ANOVA shows that in the current study those differences were again unreliable ($Fs < 1$, $ps > .05$), it will not be concentrated on in this study. There are two main contrasts carried out. Explicitly, contrast 1 compares ii) with the average of i) and iii), whereas contrast 2 compares i) and iii).

◆ **Data Analyses on NP-low**

First Pass Time

The mean first pass fixation with and without saccade is listed in Table 9 below:

Table 9 First Pass Time for NP-low in Chinese

First Pass Time for Chinese low NP

	Mean	Std. Deviation	N
first pass fixation subject NP_both	298.3148	113.47954	36
first pass fixation object NP_both	297.1296	115.87460	36
first pass fixation subject NP_high	297.7593	93.17690	36
first pass fixation object NP_high	322.9815	125.61230	36
first pass fixation subject NP_low	292.3241	74.13777	36
first pass fixation object NP_low	324.3333	102.41633	36

First Pass Time with Saccade for Chinese NP_low

	Mean	Std. Deviation	N
first pass fixation with saccade subject NP_both	317.6759	118.38231	36
first pass fixation with saccade object NP_both	329.4352	122.86454	36
first pass fixation with saccade subject NP_high	333.2685	97.14204	36
first pass fixation with saccade object NP_high	344.3889	126.34761	36
first pass fixation with saccade subject NP_low	317.2130	81.98363	36
first pass fixation with saccade object NP_low	351.6852	103.41784	36

Statistically, there was nothing reliable ($Fs<2.6$, $ps> .10$) from the results whether considering saccades or not. There was no significant difference in any contrasts ($Fs <2.6$, $ps> .10$). It seems that at the early processing stage, the processor did not spend a significant amount of time on any particular condition.

Forward Reading Time

Table 10 below shows forward reading time with and without consideration of saccade:

Table 10 Forward Reading Time for NP-low in Chinese

Forward Reading time for Chinese low NP

	Mean	Std. Deviation	N
forward reading fixation subject NP_both	312.5806	133.24036	31
forward reading fixation object NP_both	294.2258	93.12415	31
forward reading fixation subject NP_high	334.1828	126.36056	31
forward reading fixation object NP_high	301.1828	121.25995	31
forward reading fixation subject NP_low	323.0753	104.53012	31
forward reading fixation object NP_low	324.6129	107.48461	31

Forward Reading Time with Saccades for Chinese low NP

	Mean	Std. Deviation	N
forward reading fixation with saccade subject NP_both	316.0707	146.67357	33
forward reading fixation with saccade object NP_both	325.1616	103.14937	33
forward reading fixation with saccade subject NP_high	358.0101	137.65336	33
forward reading fixation with saccade object NP_high	328.7778	144.22818	33
forward reading fixation with saccade subject NP_low	346.9091	131.92125	33
forward reading fixation with saccade object NP_low	346.7374	107.79365	33

Similarly, there was nothing reliable (Fs<1.2, *p*s> .20) from the results whether considering saccades or not. There were no significant difference in any contrasts (Fs <1.2, *p*s> .20). It seems the measurement of forward reading time did not reflect much during processing at an earlier stage at all.

- Regression Path Time

Table 11 below lists the mean regression path time with or without adding saccade times:

Table 11 Regression Path Time for NP-low in Chinese

Regression Path Time in Chinese low NP

	Mean	Std. Deviation	N
regression path fixation subject NP_both	947.6944	742.83138	36
regression path fixation object NP_both	994.6296	722.70517	36
regression path fixation subject NP_high	813.7870	561.37810	36
regression path fixation object NP_high	1074.500	629.27199	36
regression path fixation subject NP_low	1192.630	760.42340	36
regression path fixation object NP_low	781.1889	539.54304	36

Regression Path Time with saccades in Chinese lo w NP

	Mean	Std. Deviation	N
regression path fixation with saccade subject NP_both	995.3426	758.57527	36
regression path fixation with saccade object NP_both	1073.454	789.13067	36
regression path fixation with saccade subject NP_high	901.3241	594.55561	36
regression path fixation with saccade object NP_high	1140.194	670.73582	36
regression path fixation with saccade subject NP_low	1291.204	840.42662	36
regression path fixation with saccade object NP_low	831.7833	530.87182	36

Very similar to results from above, there was no significant attachment condition difference from regression path time (Fs<1, *p*s > .60). However, there were significant attachment and position interactions from subject analysis (no saccades: F_1 (2, 70) =6.21, p= .003; with saccades: F_1 (2, 70) =6.34, p= .003). Unfortunately, it seems those interactions could not be generalized to other material, since the material analysis did not confirm this result (without saccades: F_2 (2, 44) =1.75, p= .187; with saccades: F_2 (2, 44) =1.73, p=1.90) (see footnote 54 (p. 117)). Analogously, subject contrasts showed a significant interaction of attachment by position when comparing condition ii) (NP-high attachment) to the average of Conditions i) (NP-both) and iii) (NP-low) (no saccades: F_1 (1, 35) =7.06, p= .012, with saccades: F_1 (1, 35) =6.20, p= .018). There was also a significant interaction of attachments by position when comparing condition i) and iii) (no saccades: F_1 (1, 35) =5.40, p= .026, with saccades: F_1 (1, 35) =6.47, p= .016). However, those interactions could not be generalized to other materials (Contrast 1: no saccade F_2 (1, 22) =1.54, p= .220, with saccade F_2 (1, 22) =1.50, p= .233; Contrast 2: no saccade F_2 (1, 22) =1.98, p= .173, with saccade F_2 (1, 22) =2.00, p= .171) (see footnote 54 (p. 117)).

Total Time

The mean total time with and without adding in saccade times is displayed in table 12 below:

Table 12 Total Time for NP-low in Chinese

Total Time for Chinese low NP

	Mean	Std. Deviation	N
total fixation subject NP_both	748.9352	342.86986	36
total fixation object NP_both	786.2037	348.47694	36
total fixation subject NP_high	881.1204	378.63171	36
total fixation object NP_high	976.0556	415.62865	36
total fixation subject NP_low	753.7963	297.02398	36
total fixation object NP_low	741.9074	282.04241	36

Total Time with Saccades in Chinese low NP

	Mean	Std. Deviation	N
total fixation with saccade subject NP_both	801.0463	362.16857	36
total fixation with saccade object NP_both	891.0556	401.64287	36
total fixation with saccade subject NP_high	990.1111	447.56757	36
total fixation with saccade object NP_high	1068.074	454.95621	36
total fixation with saccade subject NP_low	824.7870	329.98066	36
total fixation with saccade object NP_low	819.9167	307.13688	36

There were significant differences for different attachment conditions (no saccades: F_1 (2, 70) =8.45, p= .001; F_2 (2, 44) =6.37, p= .004; with saccades: F_1 (2, 70) =8.24, p= .001, F_2 (2, 44) =7.20, p= .002). However, there were no position effects (Fs< 1.6, ps> .2). Moreover, there were no interactions of attachment by position either (Fs<1, *p*s> .5). Contrasts showed significant difference between condition ii) and the average of Conditions i) and iii) (no saccades: F_1 (1, 35)=12.43, p= .001, F_2 (1, 22) =9.74, p= .005; with saccades: F_1 (1, 35) =11.66, p= .002, F_2 (1, 22) =11.56, p= .003); but no significant difference between i) and iii) (no saccades: F_1 (1, 35)= .26, p= .613, F_2 (1, 22) = .28, p= .603; with saccades: F_1 (1, 35) = .300, p= .587, F_2 (1, 22) = .23, p= .635). The pattern of the contrasts results suggested an NP-low bias, which will be further discussed in the Section 6.5.3.

◆ **Data Analyses on NP-high**

As the reason for analysing this region only comes from Study 1 (that the effects were shown in the region of NP-high), it is in principle not the critical disambiguation region. Generally speaking, in the current study, the effects on this region were of minor significance. From subject analyses, there were significant position effects (no saccades: F_1 (1, 35) =6.51, p= .015, with saccades: F_1 (1, 35) =5.98, p= .020) for regression path time; similar position effects were also found for total reading time (no saccades: F_1 (1, 35) =5.21, p= .029, with saccades: F_1 (1, 35) =5.03, p= .031), but no other effects or interactions were found (F_1s<2.5, *p*s> .08). Nevertheless, they were not confirmed by material analyses at all (F_2s<2.0, *p*s> .06) (see footnote 54 (p. 117)).

This suggests that the eye tracking experiment may have eliminated the spillover effects and picked up the correct disambiguation region. Therefore, analyses on higher NP sites in Chinese will not be included in further analyses in the current study but will be discussed later.

6.5.2.3 Results for English

According to English structure, specifically, NP-high of NP-low + RC, the disambiguation region would be the RC. This is because when the RC appears after the potential modifiees and has been processed, the processor will actively look for a suitable NP to modify. Repeated measures of ANOVAs were carried out for both subjects (F_1) and material analyses (F_2) for each of the measurements. The within subjects factors were the same as in the Chinese analyses, containing three attachment conditions (i); ii) and iii)) as first within subject factors and two positions (a and b) as second within subject factors. The contrasts carried out were the same as in the case of the Chinese data. Moreover, the English participantsô difference was not reliable (Fs<1, *p*s> .1), so it will not be discussed in further detail.

First Pass Time

The mean first pass time is shown in Table 13 below including consideration with and without saccades:

Table 13 First Pass Time for NP-low in English

First Pass Time for English

	Mean	Std. Deviation	N
first pass fixation subject NP_both	460.8148	320.61224	36
first pass fixation object NP_both	337.3241	152.43891	36
first pass fixation subject NP_high	351.5370	138.82316	36
first pass fixation object NP_high	373.0833	124.44545	36
first pass fixation subject NP_low	396.2685	160.69418	36
first pass fixation object NP_low	357.1852	127.32437	36

First Pass Time with Saccades in English

	Mean	Std. Deviation	N
first pass fixation with saccade subject NP_both	532.5741	352.09291	36
first pass fixation with saccade object NP_both	410.6944	166.20776	36
first pass fixation with saccade subject NP_high	420.0278	148.51499	36
first pass fixation with saccade object NP_high	428.5833	141.46980	36
first pass fixation with saccade subject NP_low	467.8889	180.00972	36
first pass fixation with saccade object NP_low	421.4259	137.32926	36

Statistically, there were some significant position effects from subject analyses (no saccades: F_1 (1, 35) =4.53, *p*= .04; with saccades: F_1 (1, 35) =4.34, *p*= .045). There seemed to be some reliable interaction of attachment by position as well (no saccades: F_1 (2, 70) =4.39, *p*= .018; with saccades: F_1 (2, 70) =3.06, *p*= .058 (marginal)). But those effects appeared to be only restricted to the current material because the results were not confirmed from material analyses (F_2s<2.3, *p*s> .14) (see footnote 54 (p. 117)). There was no significant difference in either subject or material contrasts (Fs <2.6, *p*s> .10). Similarly, subject contrasts showed significant interaction between position and attachment when comparing condition ii) and the average of Conditions i) and iii) (no saccades: F_1 (1, 35) =5.94, *p*= .02, with saccades: F_1 (1, 35) =4.18, *p*= .048). But there was no significant interaction between position and Condition when comparing i) and iii) (no saccades: F_1 (1, 35) =2.89, *p*= .098, with saccades: F_1 (1, 35) =1.98, *p*= .168). However, those interactions could not be generalized to other materials, as material analyses did not confirm the significances (Contrast 1: no saccade F_2 (1, 23) =3.11, p= .091, with saccade F_2 (1, 23) =1.90, p= .182; Contrast 2: no saccade F_2 (1, 23) =1.13, p= .298, with saccade F_2 (1, 23) = .96, p= .337) (see footnote 54 at p. 117).

Forward Reading Time

Table 14 below shows forward reading time with and without consideration of saccades:

Table 14 Forward Reading Time for NP-low in English

Forward Reading Time for English

	Mean	Std. Deviation	N
forward reading fixation subject NP_both	643.6970	440.03766	22
forward reading fixation object NP_both	472.9697	289.20037	22
forward reading fixation subject NP_high	416.3788	181.70688	22
forward reading fixation object NP_high	393.6515	156.13728	22
forward reading fixation subject NP_low	456.6818	187.14232	22
forward reading fixation object NP_low	446.3788	188.92742	22

Forward Reading Time with Saccades in English

	Mean	Std. Deviation	N
forward reading fixation with saccade subject NP_both	735.7879	499.09833	22
forward reading fixation with saccade object NP_both	567.5758	315.56655	22
forward reading fixation with saccade subject NP_high	502.9848	190.34548	22
forward reading fixation with saccade object NP_high	445.9697	167.12302	22
forward reading fixation with saccade subject NP_low	541.3939	181.88570	22
forward reading fixation with saccade object NP_low	530.3939	210.16289	22

There were neither significant attachment nor position effects from subject analyses (Fs <3.5, ps> .06) when saccades were not taken into consideration. But there seemed to be some marginal attachment effects when saccade times were added (F_1 (2, 42) =3.47, p= .055), and some marginal position effects as well (F_1 (1, 21) =4.22, p= .053). But those effects appeared to be only restricted to the current material because the results were not confirmed from material analyses (F_2s <1.5, ps> .30) (see footnote 54 (p. 117)). There were no reliable interactions (Fs <2.5, ps> .08). Contrasts from subject analyses showed significant differences when comparing condition ii) and the average of Condition i) and ii) (no saccades: F_1 (1, 21) =4.92, p= .038, with saccades: F_1 (1, 21) =4.99, p= .037), but no significant difference between condition i) and ii) (no saccades: F_1 (1, 21) =1.88, p= .185, with saccades: F_1 (1, 21) =2.42, p= .135). Contrasts also showed that when saccade times were not considered, there was some interaction of attachment by position when comparing condition i) and iii) (F_1 (1, 21) =4.36, p= .049).

But the interaction disappeared when saccade times were taken into account (F_1 (1, 21) =3.03, *p*= .097). Again, those effects and interactions were not confirmed by material analyses (F_2s<1, *p*s> .60) (see footnote 54 (p. 117)).

In general, there were some significant results in the subject analyses, but nothing was significant in the material analyses, which suggests that those differences and interactions were only limited to the experimental material.

- Regression Path Time

Table 15 below lists the mean regression path time with and without adding in saccade times:

Table 15 Regression Path Time for NP-low in English

Regression Path Time in English

	Mean	Std. Deviation	N
regression path fixation subject NP_both	1119.917	485.86821	36
regression path fixation object NP_both	1825.037	1081.04150	36
regression path fixation subject NP_high	979.2778	427.34660	36
regression path fixation object NP_high	1709.213	810.00426	36
regression path fixation subject NP_low	1006.074	795.15746	36
regression path fixation object NP_low	2124.583	1177.84279	36

Regression Path Time with Saccades in English

	Mean	Std. Deviation	N
regression path fixation with saccade subject NP_both	1291.120	539.07385	36
regression path fixation with saccade object NP_both	2023.417	1151.64041	36
regression path fixation with saccade subject NP_high	1142.917	480.51047	36
regression path fixation with saccade object NP_high	1835.898	804.99088	36
regression path fixation with saccade subject NP_low	1170.963	875.19299	36
regression path fixation with saccade object NP_low	2306.880	1311.39662	36

The results showed strong position effects with less regression time at subject position (no saccades: F_1 (1, 35) =52.21, *p*< .001, F_2 (1, 23) =12.05, *p*= .002; with saccades: F_1 (1, 35) =42.86, *p*< .001, F_2 (1, 23) =11.96, *p*= .002). Neither attachment effects nor position*attachment interactions were found (Fs<2.5, *p*s> .08) (see footnote 54 (p. 117)). Subjects contrasts showed significant differences when comparing condition ii) and the average of i) and iii) (no saccades: F_1 (1, 35) =4.11, *p*= .050, with saccades: F_1 (1, 21) =4.29, *p*= .046). But there was no significant difference between condition i) and iii) (no saccades: F_1s<1, *p*s> .4). Those contrasts could not be generalized to other materials though (F_2s<1.4, *p*s> .2) (see footnote 54 (p. 117)).

- Total Time

The mean total time with and without adding in saccade times is displayed in table 16 below:

Table 16 Total Time for NP-low in English

Total Time for English

	Mean	Std. Deviation	N
total fixation subject NP_both	1032.194	594.01688	36
total fixation object NP_both	995.6389	501.28328	36
total fixation subject NP_high	890.7407	452.50648	36
total fixation object NP_high	1027.519	396.42275	36
total fixation subject NP_low	885.6852	459.01220	36
total fixation object NP_low	1039.481	395.97923	36

Total Time with Saccades for English

	Mean	Std. Deviation	N
total fixation with saccade subject NP_both	1201.398	680.82775	36
total fixation with saccade object NP_both	1151.009	554.16922	36
total fixation with saccade subject NP_high	1065.528	512.73115	36
total fixation with saccade object NP_high	1181.241	495.53899	36
total fixation with saccade subject NP_low	1045.935	513.86930	36
total fixation with saccade object NP_low	1175.389	478.66553	36

Statistically, there were no reliable results ($Fs<3$, $ps> .08$) whether considering saccades or not. There was no significant difference in any contrasts ($Fs <2.6$, $ps> .08$). It seems in the end the processor did not spend more time than average in any particular condition.

6.5.2.4 Comparing Chinese and English

Same repeated measures of ANOVAs were carried out for all participants on the within subject factors. In addition, the participantsô first language was submitted as a between subjects factor.

- First Pass Time

The mean first pass fixation time is shown in Table 17 below including consideration with and without saccades:

Table 17 First Pass Time for NP-low

First Pass Time

	first language	Mean	Std. Deviation	N
first pass fixation subject NP_both	English	460.8148	320.61224	36
	Chinese	298.3148	113.47954	36
	Total	379.5648	252.41805	72
first pass fixation object NP_both	English	337.3241	152.43891	36
	Chinese	297.1296	115.87460	36
	Total	317.2269	135.95462	72
first pass fixation subject NP_high	English	351.5370	138.82316	36
	Chinese	297.7593	93.17690	36
	Total	324.6481	120.47091	72
first pass fixation object NP_high	English	373.0833	124.44545	36
	Chinese	322.9815	125.61230	36
	Total	348.0324	126.68374	72
first pass fixation subject NP_low	English	396.2685	160.69418	36
	Chinese	292.3241	74.13777	36
	Total	344.2963	134.82624	72
first pass fixation object NP_low	English	357.1852	127.32437	36
	Chinese	324.3333	102.41633	36
	Total	340.7593	115.91324	72

First Pass Time with Saccades

	first language	Mean	Std. Deviation	N
first pass fixation with saccade subject NP_both	English	532.5741	352.09291	36
	Chinese	317.6759	118.38231	36
	Total	425.1250	282.36163	72
first pass fixation with saccade object NP_both	English	410.6944	166.20776	36
	Chinese	329.4352	122.86454	36
	Total	370.0648	150.77640	72
first pass fixation with saccade subject NP_high	English	420.0278	148.51499	36
	Chinese	333.2685	97.14204	36
	Total	376.6481	132.03466	72
first pass fixation with saccade object NP_high	English	428.5833	141.46980	36
	Chinese	344.3889	126.34761	36
	Total	386.4861	139.75867	72
first pass fixation with saccade subject NP_low	English	467.8889	180.00972	36
	Chinese	317.2130	81.98363	36
	Total	392.5509	158.24863	72
first pass fixation with saccade object NP_low	English	421.4259	137.32926	36
	Chinese	351.6852	103.41784	36
	Total	386.5556	125.70690	72

For within-subjects factors, there were neither significant attachment nor position effects from this analysis (F_1s<1.2, *p*s> .30), no matter whether saccades were considered. There seemed to be some reliable position by first language interaction when no saccades were counted (F_1 (1, 70) =6.79, *p*= .010; F_2 (1, 45) =3.78, *p*= .058 (marginal)). The data show that the subject position reads faster for Chinese (with an average of 295.67ms) than the object position (average 314.33ms), whereas the subject position reads faster for English (355.67ms) than the object position (402.23ms). This interaction could not be generalized to other materials when saccades were counted (F_1 (1, 70) =6.38, *p*= .014; F_2 (1, 45) =3.07, *p*= .087) (see footnote 54 (p. 117)). Subject contrasts also showed significant interaction between attachments by position when comparing condition ii) and the mean of the average of Condition i) and iii) with no saccades involved (F_1 (1, 70) =4.95, *p*= .029), but the interaction disappeared when saccades were involved (F_1 (1, 70) =2.31, *p*= .133). Moreover, those interactions could not be generalised to other materials (*p*s> .06) (see footnote 54 (p. 117)).

For between-subjects effects, there were significant differences between Chinese and English native speakers (no saccades: F1 (1, 70) =13.06, *p*= .001; F2 (1, 45) =8.87, *p*= .005; with saccades: F1 (1, 70) =25.73, *p*< .001; F2 (1, 45) =19.61, *p*< .001).

- Forward Reading Time

Table 18 below shows forward reading time with and without consideration of saccade:

Table 18 Forward Reading Time for NP-low

Forward Reading Time

	first language	Mean	Std. Deviation	N
forward reading fixation subject NP_both	English	643.6970	440.03766	22
	Chinese	312.5806	133.24036	31
	Total	450.0252	339.95756	53
forward reading fixation object NP_both	English	472.9697	289.20037	22
	Chinese	294.2258	93.12415	31
	Total	368.4214	216.06887	53
forward reading fixation subject NP_high	English	416.3788	181.70688	22
	Chinese	334.1828	126.36056	31
	Total	368.3019	155.61991	53
forward reading fixation object NP_high	English	393.6515	156.13728	22
	Chinese	301.1828	121.25995	31
	Total	339.5660	142.98345	53
forward reading fixation subject NP_low	English	456.6818	187.14232	22
	Chinese	323.0753	104.53012	31
	Total	378.5346	157.68549	53
forward reading fixation object NP_low	English	446.3788	188.92742	22
	Chinese	324.6129	107.48461	31
	Total	375.1572	157.31797	53

Forward Reading Time with Saccades

	first language	Mean	Std. Deviation	N
forward reading fixation with saccade subject NP_both	English	735.7879	499.09833	22
	Chinese	316.0707	146.67357	33
	Total	483.9576	390.74590	55
forward reading fixation with saccade object NP_both	English	567.5758	315.56655	22
	Chinese	325.1616	103.14937	33
	Total	422.1273	243.71330	55
forward reading fixation with saccade subject NP_high	English	502.9848	190.34548	22
	Chinese	358.0101	137.65336	33
	Total	416.0000	174.51749	55
forward reading fixation with saccade object NP_high	English	445.9697	167.12302	22
	Chinese	328.7778	144.22818	33
	Total	375.6545	162.92903	55
forward reading fixation with saccade subject NP_low	English	541.3939	181.88570	22
	Chinese	346.9091	131.92125	33
	Total	424.7030	180.06763	55
forward reading fixation with saccade object NP_low	English	530.3939	210.16289	22
	Chinese	346.7374	107.79365	33
	Total	420.2000	179.74224	55

For within-subjects factors, there were some significant position effects in subject analyses (no saccades: F_1 (1, 69) =6.16, p= .016, with saccades: F_1 (1, 51) =4.82, p= .033), but not in material analyses (F_2s<1, ps> .5). There seemed to be some reliable attachment by first language interactions in subject analyses (no saccades: F_1 (2, 102) =4.12, p= .027; with saccades: F_2 (2, 106) =5.51, p= .008), but those effects appeared to be only restricted to the current material because the results were not confirmed by material analyses (F_2s<1.5, ps> .2) (see footnote 54 (p. 117)). Subject contrasts showed significant differences in attachment when comparing condition ii) and the average of i) and iii) (no saccades: F_1 (1, 51) =4.69, p= .035); with saccades (F_1 (1, 53) =4.71, p= .035). Moreover, there were some interactions between attachments by position when comparing condition ii) and the average of Conditions i) and iii) with saccades being considered (F_1 (1, 53) =6.52, p= .014), but the interaction disappeared when saccades were excluded (F_1 (1, 51) = .41, p= .523). There were also some interactions between attachments by position when comparing Condition i) and iii) (no saccades: F_1 (1, 51) =5.34, p= .025, with saccades: F_1 (1, 53) =4.70, p= .035). Unfortunately, those interactions could not be generalized to other materials (F_2s<3, ps> .06) (see footnote 54 (p. 117)).

The between-subjects effects (language differences) were still significant (no saccades: F_1 (1, 51) = 33.29, p< .001; F_2 (1, 41) =13.50, p= .005; with Saccades: F_1 (1, 53) =50.27, *p*< .001; F_2 (1, 41) = 27.01, *p*< .001).

Regression Path Time

Table 19 below lists the mean regression path time with and without adding in saccade times:

Table 19 Regression Path Time for NP-low

Regression Path Time

	first language	Mean	Std. Deviation	N
regression path fixation subject NP_both	English	1119.917	485.86821	36
	Chinese	947.6944	742.83138	36
	Total	1033.806	629.20934	72
regression path fixation object NP_both	English	1825.037	1081.04150	36
	Chinese	994.6296	722.70517	36
	Total	1409.833	1004.18665	72
regression path fixation subject NP_high	English	979.2778	427.34660	36
	Chinese	813.7870	561.37810	36
	Total	896.5324	502.31768	72
regression path fixation object NP_high	English	1709.213	810.00426	36
	Chinese	1074.500	629.27199	36
	Total	1391.856	787.88936	72
regression path fixation subject NP_low	English	1006.074	795.15746	36
	Chinese	1192.630	760.42340	36
	Total	1099.352	778.17612	72
regression path fixation object NP_low	English	2124.583	1177.84279	36
	Chinese	781.1889	539.54304	36
	Total	1452.886	1133.54403	72

Regression Path Time with Saccades

	first language	Mean	Std. Deviation	N
regression path fixation with saccade subject NP_both	English	1291.120	539.07385	36
	Chinese	995.3426	758.57527	36
	Total	1143.231	670.14837	72
regression path fixation with saccade object NP_both	English	2023.417	1151.64041	36
	Chinese	1073.454	789.13067	36
	Total	1548.435	1090.66999	72
regression path fixation with saccade subject NP_high	English	1142.917	480.51047	36
	Chinese	901.3241	594.55561	36
	Total	1022.120	550.34099	72
regression path fixation with saccade object NP_high	English	1835.898	804.99088	36
	Chinese	1140.194	670.73582	36
	Total	1488.046	814.81404	72
regression path fixation with saccade subject NP_low	English	1170.963	875.19299	36
	Chinese	1291.204	840.42662	36
	Total	1231.083	854.07086	72
regression path fixation with saccade object NP_low	English	2306.880	1311.39662	36
	Chinese	831.7833	530.87182	36
	Total	1569.331	1240.29684	72

The within-subject effects showed significant position effects (no saccades: F_1 (1, 70) =30.10, *p*< .001; F_2 (1,45)=8.03, *p*= .007; with Saccades: F_1 (1, 70) =25.73, *p*< .001; F2 (1, 45) =7.64, *p*= .008) and a significant interaction between position * first language (no saccades: F_1 (1, 70) =35.41, *p*< .001; F_2 (1, 45) =11.34, *p*= .002; with Saccades: F_1 (1, 70) =30.90, *p*< .001; F_2 (1, 45) =10.45, *p*= .002). Moreover, when saccades were included, there were significant interactions of attachments by position by language (F_1 (2, 140) =5.54, *p*= .006; F_2 (2, 90) =2.99, *p*= .057(marginal)). The data show that the subject position reads faster for Chinese than the object position, whereas the subject position reads faster for English than the object position. Subject contrasts showed an interaction between attachment by position and by language when comparing condition ii) and the average of i) and iii) (no saccades: F_1 (1, 70) =5.31, *p*= .024, with saccades: no saccades: F_1 (1, 70) =5.75, *p*= .019). There was also some interaction between attachment by position when comparing i) and iii) (no saccades: F_1 (1, 70) =5.41, *p*= .023, with saccades: F_1 (1, 70) =5.40, *p*= .023). Unfortunately, those contrasts could not be generalized to other materials (*p*s> .05) (see footnote 54 (p. 117)).

The between-subjects effects (language differences) only appeared to be significant with saccades being taken into consideration (F_1 (1, 70) =30.52, *p*< .001; F_2 (1, 45) =10.45, *p*= .002).

- Total Time

The mean total time with and without adding in saccade times is tabulated in table 20 below:

Table 20 Total Time for NP-low

Total Time

	first language	Mean	Std. Deviation	N
total fixation subject NP_both	English	1032.194	594.01688	36
	Chinese	748.9352	342.86986	36
	Total	890.5648	502.23160	72
total fixation object NP_both	English	995.6389	501.28328	36
	Chinese	786.2037	348.47694	36
	Total	890.9213	441.42494	72
total fixation subject NP_high	English	890.7407	452.50648	36
	Chinese	881.1204	378.63171	36
	Total	885.9306	414.28726	72
total fixation object NP_high	English	1027.519	396.42275	36
	Chinese	976.0556	415.62865	36
	Total	1001.787	404.10060	72
total fixation subject NP_low	English	885.6852	459.01220	36
	Chinese	753.7963	297.02398	36
	Total	819.7407	389.56720	72
total fixation object NP_low	English	1039.481	395.97923	36
	Chinese	741.9074	282.04241	36
	Total	890.6944	372.77164	72

Total Time with Saccades

	first language	Mean	Std. Deviation	N
total fixation with saccade subject NP_both	English	1201.398	680.82775	36
	Chinese	801.0463	362.16857	36
	Total	1001.222	577.74826	72
total fixation with saccade object NP_both	English	1151.009	554.16922	36
	Chinese	891.0556	401.64287	36
	Total	1021.032	498.03960	72
total fixation with saccade subject NP_high	English	1065.528	512.73115	36
	Chinese	990.1111	447.56757	36
	Total	1027.819	479.35883	72
total fixation with saccade object NP_high	English	1181.241	495.53899	36
	Chinese	1068.074	454.95621	36
	Total	1124.657	475.74348	72
total fixation with saccade subject NP_low	English	1045.935	513.86930	36
	Chinese	824.7870	329.98066	36
	Total	935.3611	442.99761	72
total fixation with saccade object NP_low	English	1175.389	478.66553	36
	Chinese	819.9167	307.13688	36
	Total	997.6528	437.58895	72

For within-subjects factors, there were some significant attachment effects in subject analyses (no saccades: F_1 (2, 140) =3.13, *p*= .047, with saccades: F_1 (2, 140) =3.48, *p*= .034). There were also some position effects when saccades were not included in the analyses (F_1 (1, 70) =4.21, *p*= .044). Those effects were not found in material analyses (F_2s<1.5, *p*s> .2) (see footnote 54 (p. 117)). Subjects analyses also showed some interactions between attachment * first language (no saccades: F_1 (2, 140) =5.34, *p*= .006, with saccades: F_1 (2, 140) =4.52, *p*= .013), but they were not confirmed by material analyses either (F_2s< .1, *p*s> .1) (see footnote 54 (p. 117)). Subject contrasts showed significant differences in attachment when comparing condition ii) and the average of i) and iii) (no saccades: F_1 (1, 70) =5.35, *p*= .024); with saccades (F_1 (1, 70) =5.17, *p*= .026), but not between i) and iii) (Fs <1, *p*s> .3). Moreover, there were some interactions between attachment by first language when comparing condition ii) and the average of i) and iii) (no saccades: F_1 (1, 70) =10.65, *p*= .002); with saccades (F_1 (1, 70) =7.82, *p*= .007), but not between i) and iii) (Fs<1, *p*s> .6). Unfortunately, those interactions could not be generalized to other materials (F_2s<3, *p*s> .06) (see footnote 54 (p. 117)).

Similar to regression path time analyses, the between-subject effects (language differences) only appeared to be significant with saccades being taken into consideration (F_1 (1, 70) =8.43, p= .005; F_2 (1, 45) =5.59, p= .022).

6.5.3 Discussion

The results reported in the last section seem rather complex. To sum up the results that were confirmed by both of F_1 and F_2, a simplified report is presented below:

a. For Chinese general results, there were significant attachment effects in total time on the region of lower NP. Contrasts showed patterns of significantly longer reading time for condition ii), where RC is forced to attach to NP_{high}, but no significant differences between condition i) NP_{both} and iii) NP_{low}. These results were significant for data including and excluding saccades;
b. For English results, only position effects in regression path time were significant in both analyses, with/without consideration of saccades;
c. For overall results, position effects were found for data with or without saccade times in regression path time; interactions between position*first language were found in first path time only without saccades time and regression path time with/without saccades; between subjects language effects were found in first pass time and forward reading time when saccades were not included, and in all measurements (first pass time, forward reading time, regression path time and total time) when saccades were included.

6.5.3.1 Chinese Results and the Methodology

First of all, there was no evidence of interactions of attachments by position in any of the measurements in Chinese. This means the topicalisation feature in Chinese did not lead to any change in preference. However, the results showed an attachment effect in total time. As introduced above, total time is regarded as reflecting a later stage of processing. Failing to produce any significant differences or interactions in other earlier measurements implies that Chinese speakers might store the information and only process it when enough information has been collected.

This is indeed in line with Study 1, where no effects on the critical region (NP-low) could be found, and instead all the expected effects were shown on a later region (NP-high). In Study 1, experimental sentences were segmented and participants could only read one region at a time. A button had to be pressed to discover the next region. No regression was allowed, as the segments move only in a forward direction. When the ambiguity at NP-low was encountered, besides waiting to check back to the previously processed materials, the only other choice the participants had was to press the button and move onto the next region to get more information that would be helpful for processing. The next region however, being NP-high, carries a lot of information, as it is the head noun for the complex NP and the main argument of the verb, etc. Most importantly, it is closer to the disambiguation region the processor has just passed and the quicker the processor arrives at this region, the more use could be made of it, since the shorter time taken to get to the region implies less memory cost. But how quick could it be? It would take no less time than normal processing in a region because that would be the minimal time people could spend to get the information. The processor then starts its ñrealò processing time on the NP-high region, integrating the information taken till this stage. That is possibly the reason why the time spent on this region could reflect the difficulty of processing an earlier region.

In the eye-tracking experiment, on the other hand, sentences were shown in their entirety. Participants did not need to wait for the next region to be shown and were allowed to do as many regressions as they needed. This is believed to be more natural in terms of reading behaviour in the real world. The lack of difference in earlier processing stages such as first pass, or in the main processing stages such as regression, suggests that the awkwardness of an attachment was not concentrated on at these stages. However, it does not mean that the awkwardness was not discovered, as total time spent on the critical region does reflect an NP-low bias (see Figure 22 below), confirming the results of Study 2. The reason that reading time differences were found only on the actual critical region (NP-low) instead of a later region (NP-high) would be the freedom of regressions. But it seems the Chinese speakers prefer not to process the difficulty before gathering more information. Thus, instead of regressive processing from the critical region right when the difficulty occurs, they move on to new regions and return to the critical ones in a later stage. This suggests that the processing in Chinese follows a ñstore-up then processò model, instead of an on-site immediate one.

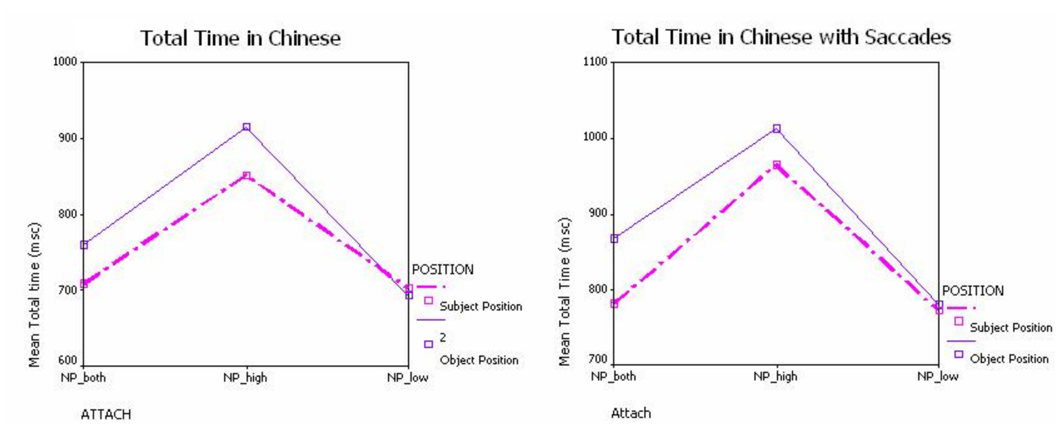


Figure 22 Total Time in Chinese

Noticeably, the materials of Study 1, Study 2 and the current study all used semantic disambiguation instead of a syntactic one. Study 2 showed the attachment effects in a region later than the critical one, whereas the current study showed the attachment effects in a very late time stage of processing. One suggestion from the Chinese data would be that thematic or semantic differences only serve at a later stage of processing. The lack of differences in earlier parsing stages might be because the syntactic structures in all those attachment conditions are indeed the same. This will be further discussed with the English data later.

6.5.3.2 English Results and the Methodology

At first glance, the finding of position effects might be a good ñsupportò of *topicalisation*: With a topiticalised subject, position effects were found, even though the position effects were not reported in other studies (Hemforth et al. 2003), where the subject was not topicalised. However, this only means that the reading time in the subject position is shorter than in the object position. What should really be expected by *topicalisation* are reliable interactions between attachments by position, which were not confirmed by the results. Moreover, the lack of attachment effects throws further doubt on the credibility of the situation, since many earlier studies (Cuetos & Mitchell, 1988, Gilboy et al., 1995, etc.) have recorded NP-low bias in English. Figure 23 below is an example of such a result measured by regression path time. It is clear that the subject position required less time than the object position, but NP-low bias did not show any processing advantage (for detailed results see Section 6.5.2.3). It could just be because at the subject position there was less information to regress to than at the object position, but not anything as complicated as *topicalisation*.

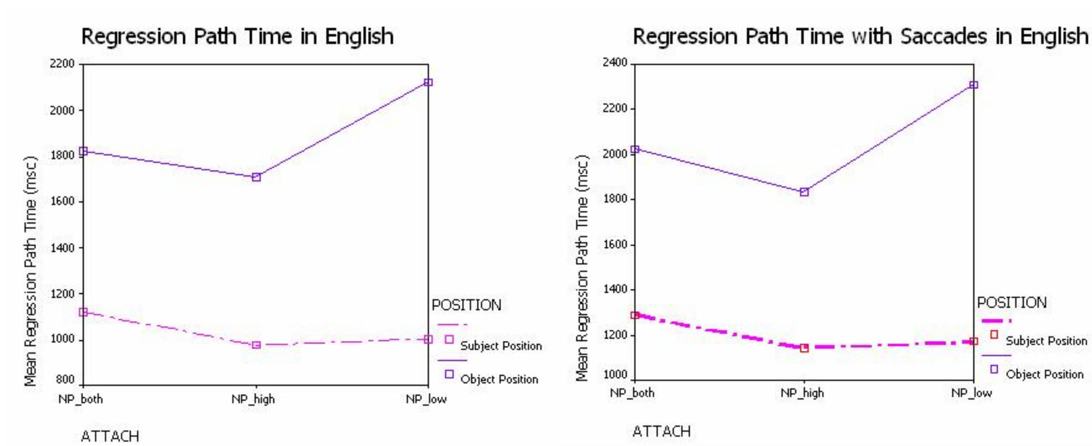


Figure 23 Regression Path Time in English

One explanation for this would be similar to the aforementioned proposal in section 6.3.3: The thematic or semantic differences were not paid enough attention to in an earlier stage of the process. (Most of early studies in English used syntactic disambiguation.) A possible explanation is that the processing difficulty from semantic and thematic mismatching may be not as clear as syntactic differences.

However, it should be noted that since there were some effects from subject analyses, although not confirmed by material analyses, there might be some inaccuracy in the degree of ambiguity in the translation.

6.5.3.3 Overall Results

The main findings from the overall results are the between-subjects effects and the interaction they produce. The general language effects mean that Chinese reading times are generally shorter than English reading times (See Table 17-20). This might be caused by the differences in language form and word length. Chinese is an ideographic language while English is an alphabetic language. Due to this feature difference, the Chinese version tends to fit in to a shorter space than the English counterpart. However, these features were not of particular interest nor informative in terms of the RC attachment preference issue. The main effect should show up as part of the interaction of semantic manipulations of the attachment by position within and between the languages. Unfortunately, these interactions were not confirmed: Although the subject analyses did provide some attachment effects and interactions of attachment by position and by language, these interactions cannot be generalised to new materials. Although the ambiguity degree of manipulation may vary by translation, failing to find interactions in the material analyses makes the claim for *topicalisation* less credible.

The above did not confirm the prediction from *topicalisation*, nor could it be explained on the basis of *topicalisation*.

6.5.3.4 The Measurements

In an earlier part of this thesis (sections 4.2.2 and 6.5.2.1), saccades and forward reading time were introduced. It is in contrast to the previous definition (Just & Carpenter, 1980) that comprehension only happens when the information is viewed and fixed. Under the consideration that processing might be taking place even during saccade movements, saccade times were added to the measurements. Forward reading times were introduced as a measurement to be more ñinformativeò than the traditional first pass time because it can help to reduce the ñcancellationò of regression effects (Vonk and Cozijn, 2003).

The current study has considered both sides of the argument. A comparison of measurements including and excluding saccades is in tables 21 and 22 below:

Table 21 Measurements in English and Chinese (subjects analyses)[59]

		English NP-low	Chinese NP-low
First Pass	position	F(1,35)=4.53, p= .04	n.s
	position * attach	F(2,70)=4.39, p= .018	n.s
Forward reading	attachment	n.s	n.s
	position	n.s	n.s
	position * attach	n.s	n.s
Regression Path	attachment	n.s	n.s
	position	F(1,35)=52.21, p< .001	n.s
	position * attach	n.s	F(2,70)=6.21, p= .003
Total	attachment	n.s	F(2,70)=8.45, p= .001
	position	n.s	n.s
	position * attach	n.s	n.s
First Pass with saccade	attachment	n.s	n.s
	position	F(1,35)=4.34, p= .045	n.s
	position * attach	F(2,70)=3.06, p= .058	n.s
Forward reading with saccade	attachment	F(2,42)=3.47, p= .055	n.s
	position	F(1,21)=4.22, p= .053	n.s
	position * attach	n.s	n.s
Regression Path with saccade	attachment	n.s	n.s
	position	F(1,35)=42.86, p< .001	n.s
	position * attach	n.s	F(2,70)=6.34, p= .003
Total with saccade	attachment	n.s	F(2,70)=8.24, p= .001
	position	n.s	n.s
	position * attach	n.s	n.s

Table 22 Measurements of Overall Results[60]

		Subject NP-low	Material NP-low
first pass	position*fi_lang	F(1,70)=6.97, p= .010	F(1,45)=3.78, p= .058
	attach*position	F(2,140)=4.47, p= .015	n.s
	fi_lang	F(1,70)=13.06, p= .001	F(1,45)=8.87, p= .005
forward reading	attach*fi_lang	F(2,102)=4.12, p= .027	n.s
	position	F(1,69)=6.16, p= .016	n.s
	fi_lang	F(1,51)=33.29, p< .001	F(1,41)=13.50, p= .005
regression path	position	F(1,70)=30.10, p<. 001	F(1,45)=8.03, p= .007
	position*fi_lang	F(1,70)=35.41, p< .001	F(1,45)=11.34, p= .002
	fi_lang	F(1,70)=24.16, p< .001	n.s
	attach	F(2,140)=3.13, p= .047	n.s
	attach*fi_lang	F(2,140)=5.34, p= .006	n.s
Total	position	F(1,70)=4.21, p= .044	n.s

59 Due to the size of the table, only subject analyses are shown here.
60 Due to the size of the table, only results with some significance are shown here.

	fi_lang	F(1,70)=4.95, p= .029	n.s
first pass with saccade	position*fi_lang	F(1,70)=6.38, p= .014	n.s
	fi_lang	F(1,70)=26.70, p< .001	F(1,45)=19.61, p< .001
	attach	F(2,106)=3.13, p= .057	n.s
	attach*fi_lang	F(2,106)=5.51, p= .008	n.s
forward reading with saccade	position	F(1,51)=4.82, p= .033	n.s
	fi_lang	F(1,53)=50.27, p< .001	F(1,41)=27.07, p< .001
regression path with saccade	position	F(1,70)=25.73, p<. 001	F(1,45)=7.64, p= .008
	position*fi_lang	F(1,70)=30.90, p< .001	F(1,45)=11.53, p= .001
	attach*position*fi_lang	F(2,140)=5.54, p= .006	F(2,90)=2.99, p= .057
	fi_lang	F(1,70)=30.52, p< .001	F(1,45)=10.45, p= .002
total with saccade	attach	F(2,140)=3.48, p= .034	
	attach*fi_lang	F(2,140)=4.52, p= .013	
	fi_lang	F(1,70)=8.34, p= .005	F(1,45)=5.59, p= .022

Those two tables illustrate the fact that including saccades and introducing the new measure of forward reading time does not provide steady effects. Sometimes it can increase the power of the results, but sometimes it decreases the power. It seems that the current experiment tends to support the traditional definition of reading process. However, further study is needed to investigate processing during saccade time, but it is beyond the scope of the current thesis.

6.6 Conclusions of Tip 1

At the beginning of this chapter, it was mentioned that the current thesis attempts to investigate the timecourse of sentence processing and the belief in the universality of syntax processing, starting by providing Chinese data on RC attachment preference. Meanwhile, English data involving *topicalisation* was added in comparison to the Chinese data. In this section, some conclusions will be drawn and further studies will be suggested.

6.6.1 RC Attachment in Chinese

In Chapter 6.2, a group of theories accounting for the RC attachment ambiguity resolution were introduced and predictions from them accounting for Chinese data have been provided. Essentially, most of the theories, except Relativised Relevance (*RR*), predict that according to Chinese language structure, there should be an NP-low preference in this language. The Implicit Prosody Hypothesis (IPH) further proposes that if there is an intonational break, there should be a stronger NP-low preference. Moreover, *topicalisation* proposes that there should be a more significant NP-low preference in Chinese when the RC is at the subject position compared to the object position.

Results from Studies 1, 2 and 3 have confirmed that NP-low preference exists in Chinese. In Studies 1 and 3, the sentences with RC forced to attach to NP-high took much longer reading time than NP-low attachment. In Study 2, the majority of the choices to NP-low were made in the questionnaire.

During the whole sentence processing, no preference transfer between high and low NPs were made. This has falsified the prediction of *RR*. Moreover, although the intonational break *de* in Study 1 has interacted with the preference, unlike the prediction from IPH, the intonational break did not strengthen the preference (See Figure 20 in section 6.3.2.3, P.118). One possible explanation is that the break of segmentation is not the same as a natural intonational break. Nevertheless, this cannot be tested by the current study and IPH will not be further discussed in this thesis. In terms of *topicalisation*, the results were consistent with Chinese data (i.e., consistent NP-low preference from the questionnaire study and the eye tracking study) that topicalisation does not alter the preference of the attachment. For the English data, however, the questionnaire results were in line with the Chinese data. Nonetheless, *topicalisation* predicts that for the crosslinguistic comparison, the topicalised condition in Chinese should have no less NP-low bias than English, but stronger NP-low bias was found in English than Chinese. This has thrown strong doubt on the *topicalisation* proposal. Unfortunately, as the generally agreed NP-low preference in English has not been replicated in the eye tracking study, there is still scope for further refinement.

6.6.2 The Delayed Effects in Chinese

For the Chinese data, in Study 1, the effects of the attachment bias were found a region delayed, whereas in Study 3, the effects were found in later stages of processing. As discussed in 6.3.3, the delay in Study 1 cannot simply be explained by the defect of the methodology of self-paced reading. This delay has not been widely reported in the previous RC attachment studies. Notably, previous studies of this kind used a syntactic mismatch to create a temporary ambiguity, whereas the current studies used a semantic mismatch. Now the question is whether the apparent delay effect is caused by a feature specific to Chinese or by semantic parsing occurring later than syntactic parsing.

This brings us back to the main topic of the current thesis, explicitly, the timecourse of sentence processing. It cannot simply be concluded from the RC attachment data, as other possible explanations have not been excluded. Moreover, as the RC attachment preference has become conclusive after the three studies reported above, the next chapter will focus only on the issue of delay and the timecourse of syntactic processing and semantic processing.

Chapter 7

Tip Two: Timecourse of Syntax Processing

The discussion in Tip 1 started from the relative clause attachment ambiguity in Chinese. The predictions from a number of theories were investigated. These included the Garden Path Theory Group (GPTG), the Structural-competing groups (namely, Recency Preference / Predicate Proximity (RP / *PP*), Attachment Binding (AB) and Topicalisation Proposal), Exposure Tuning Hypothesis (ETH) and Implicit Prosody Hypothesis (IPH). The NP-low bias in Chinese has confirmed most of those theories, including the GPTG (except the Relativized Relevance (RR)), the RP / *PP*, the AB and the ETH. Conspicuously, most of those confirmed theories belong to the first processing of syntax theories. A simple explanation of the delayed effects might be that semantic processing occurs at a later stage than syntactic processing. However, none of the designs in the three studies in Chapter 6 has tested the prediction of the late assignment of syntax theories. The mismatching result for the IPH cannot falsify this theory because the artificial segmentation break may not carry features of a natural intonational break, and this discrepancy may have affected investigation of the hypothesis. Moreover, as there were no English data in Study 1, and no effects found for English data in Study 3, it is not sure whether the apparent delay effect is caused by a feature specific to Chinese only. Therefore, no conclusions can be drawn with regard to the late assignment of semantic processing.

It has been reviewed in Chapter 2 that the Universal Grammar believers consider sentence processing a procedure of assessing crude syntax → assigning semantics → final syntax refinement (i.e., syntactic processing is then the beginning as well as the end of sentence processing). However, it has also been reviewed in Chapter 3 that theories of this type have difficulty in explaining acoustic processing such as 34) in Chapter 3.3 presented orally (and the acoustic sentence is re-presented in 82) below):

82) Thehorseracesandwins.

Moreover, crossslinguistically, some languages do not even separate segments in their written form, and Chinese is one of the languages of this type:

83) 那匹馬參加賽跑並且勝利了。

It has been argued in 3.3 that the processor has to segmentise the input into separate units to be able to carry out further processing. This type of segmentation does not simply belong to a lexical level of processing, because information for segmentation includes prosody, context, semantics and structure.

The cause of the delay effects in Study 1 and 3 has become the motivation for this chapter: the results of two further eye-tracking experiments directly comparing semantic processing and syntactic processing crosslinguistically (i.e., English and Chinese) will be reported. If the delay effects were caused by late assignment of semantic processing, a clear delay in the semantic manipulated sentences should be established for both languages; but if semantic processing occurs no later (or even earlier) than syntactic processing, this trend should also be observed in both languages. Moreover, if the latter were true, interaction of processing by language might also be observed, because some of the language specific features should provide an explanation for the delayed effects in Chinese that have not been widely reported in other languages. Thus, one would expect no significant anomaly effects in Chinese in early stages because of the potential extra information processing (e.g., word segmentation), but relatively significant anomaly effects (at least semantic anomaly effects, if syntactic processing occurs no earlier than semantic processing) should be expected in English in early stages already.

7.1 Study 4: An Eye-tracking Experiment on Timecourse of Syntactic and Semantic Processing

The current study does not attempt to scrutinise the theories introduced and tested in Chapters 5 and 6 (especially those specifically accounting for RC attachment ambiguity resolution). The focus is on the timecourse of syntactic and semantic processing, by comparing reading times of passive sentences with semantic and syntactic anomaly. The methodology of eye-tracking makes it possible to investigate the issue in a more natural way, e.g., the lack of artificial segmentation makes it possible for segmentation in silent reading to occur if necessary. Although the methodology cannot directly test the Implicit Prosody Hypothesis (IPH), if segmentation in silent reading did occur, extra processing time should be detected, consequently supporting the late assignment of syntax theories introduced in Section 3.3.

7.1.1 Method

◆ **Participants**

36 volunteers participated, of whom 18 were native English speakers and 18 were native Chinese speakers, aged between 18 and 37. All of them were either students or staff at the University of Exeter. 13 of them were from the School of Psychology, 11 from the School of Finance, six from the School of Modern Languages, and the other six were from other departments. Nine were male and 27 were female.

◆ **Materials**

The current study makes use of 24 passive sentences with a) no anomaly, b) semantic anomaly, and c) syntactic anomaly, in both Chinese and English. It comprises two parts: an eye-tracking study to investigate on-line reading behaviour and a questionnaire study to ask how acceptable a sentence is in the participantsô native language. The example sentences are given below.

84) a. **Normal**

John broke his ankle when he fell off the ladder. He was carried *by his friend* to a hospital.

約翰從梯子上摔了下來，摔壞了腳踝。他被他的朋友一路擡到醫院。

Yuehan cong tizishang shuaile xialai, shuaihuaile jiaohuai. Ta *bei tade pengyou* yilu taidao yiyuan.

b. **Semantic Anomaly**

John broke his ankle when he fell off the ladder. He was carried *by his cake* to a hospital.

約翰從梯子上摔了下來，摔壞了腳踝。他被他的蛋糕一路擡到醫院。.

Yuehan cong tizishang shuaile xialai, shuaihuaile jiaohuai. Ta *bei tade dangao* yilu taidao yiyuan.

c. **Syntactic Anomaly**

John broke his ankle when he fell off the ladder. He was carried *his friend* to a hospital.

約翰從梯子上摔了下來，摔壞了腳踝。他他的朋友一路擡到醫院。

Yuehan cong tizishang shuaile xialai, shuaihuaile jiaohuai. Ta *tade pengyou* yilu taidao yiyuan.

The two languages are comparable, because the anomaly appears to be at the same level of syntactic structure in both languages, as shown in Figure 24 below.

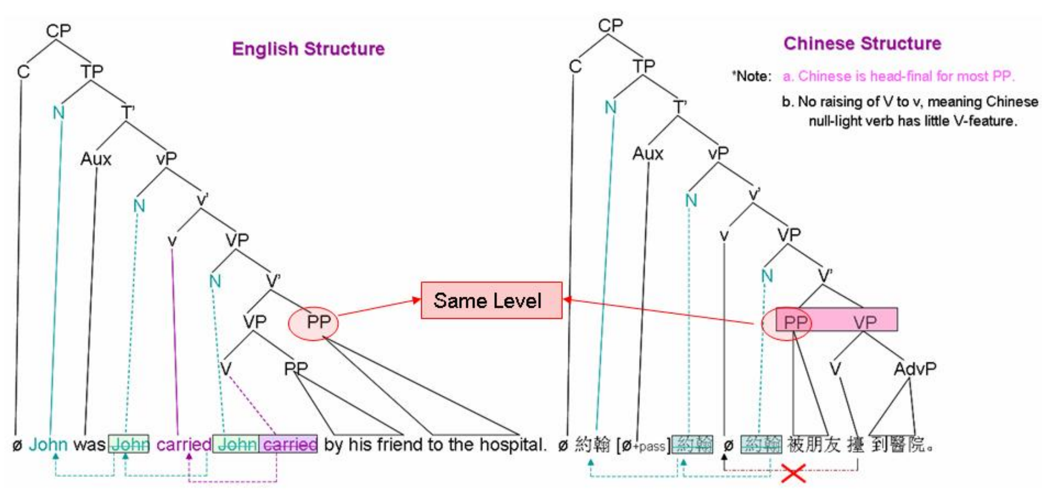


Figure 24 Correct Passive Structures of English and Chinese

To put it simply, a sentence is centred by the predicate (and it is normally a verb phrase (VP)). This predicate will assign the subject and object (direct object) of the sentence. In a SVO language such as English and Chinese, the normal form in a simple sentence would be subject (normally a NP) + VP + object (normally another NP). However, when this sentence is expressed in a passive form, the object of the VP is moved to the subject position (pre-verbal) to be emphasised (or to be topicalised). In both languages, the marker of the real subject (i.e., the agent) is no longer the NP itself: it has to be combined with a preposition (in English, it is *by*, whereas in Chinese, it is 被). Meanwhile, a number of movements occur in the syntactic structure to form the passive expression. As details of the movement are not the focus of the current paper, they will not be discussed further here. After all the movements, the direct object NP has been moved to be the head of the first Tense Phrase, which is the preverbal position and acts as the subject of the sentence, while the PP has become the head of the second TP to mark the real subject of the predicate.

In the case of semantic anomaly as in 84b), the structure has not been changed, but since *the cake* is inanimate and thus semantically cannot be the agent of an action, it is wrong for both languages to use this NP as the subject.

In the case of syntactic anomaly as in 84c), the structure has some changes as shown in Figure 25 below. If the preposition *by* or 被 has been taken out, the PP will be changed into an NP. Since the position of the subject of the sentence has already been filled by another NP (the direct object), but as the subject of the predicate, it has to be a PP that fills the head of the second TP position. In line with the rules of passive sentences, the NP cannot fit into the structure. Therefore, a syntactic anomaly occurs. It is clear that the anomaly occurs at the same structure level in both languages as shown in Figure 25.

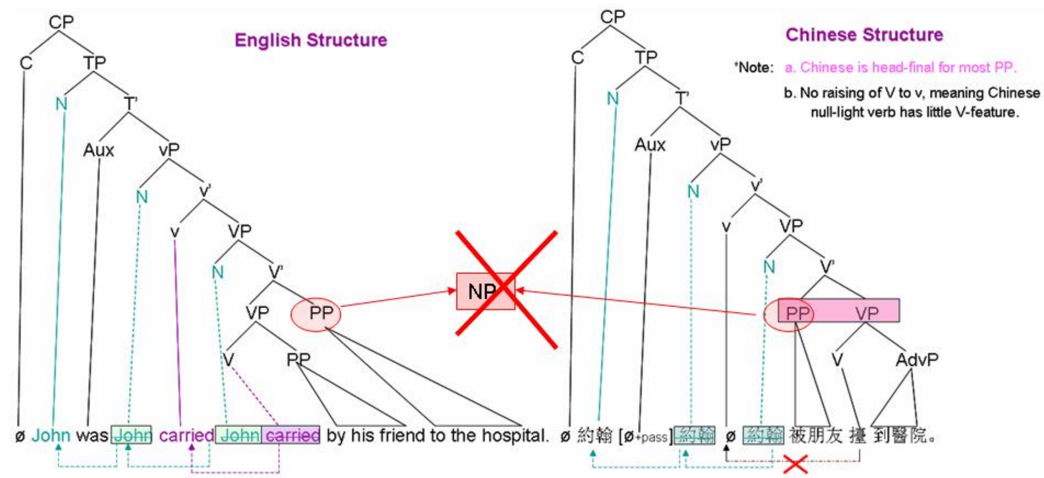


Figure 25 Incorrect Structures in English and Chinese

The font used in the English material was Arial, size 18, bold typeface. The font used in Chinese was 楷體_GB2312, size 20 and bold. The average interest area (IA) size for the PP was about 65*111 pixels for English, and 72*118 pixels for Chinese. The average IA sizes for the NP without ñ*by*ò or ñ被ò was about 50*110 pixels for English and 59*118 pixels for Chinese.

The word frequency of the head of the agent NPs in a sentence was controlled in the same way as in Study 1. The English version was a direct translation of the Chinese as before. The sentences were checked with native speakers, with the rate of naturalness ranging from 1 to 5, on a scale from 1 (very unnatural) to 5 (very natural). The results showed that for the normal condition, both Chinese and English had an average of 4.23 and above, while for the anomalous conditions, both languages had an average of below 2.03.

Moreover, 66 filler sentences (including 14 ambiguous sentences as foils) were used to make the participants unaware of the test materials. All of them were derived from the filler sentences used in Study 1. Similarly, 28 yes/no questions were asked based on the content of some of the sentences. Among those questions, 9 were from the current experimental sentences and 19 were from filler sentences. Participants were expected to press the correct button to answer the question after they finished reading each of the sentences, e.g., for sentences 84) above, the question and choices would be 85):

85) Has John broken his arm?

The whole experiment was counterbalanced with each condition appearing in eight experimental sentences for each participant only once. Al the material was presented in a pseudo-random order created by a Perl programme.

◆ Procedure

The randomising programme was implemented as a Perl script and was run before each experiment. The pseudo-random order was saved as a text file, and then transferred into the Eyelink programme. Each participant was taken into a test cubicle equipped with two computers and an eye tracking machine. They were asked to sit down in front of one main computer connected to the eye-tracking helmet and a game controller. Then they were asked to put on the helmet with one camera adjusted to their right pupil. Calibration of the programme normally took 10 minutes (longer in some cases), including camera adjustment, calibration and validation. During this time, the participants learnt to concentrate on an adjusting dot in the middle of the screen and their calibration and validation should pass a certain value (which would be indicated by a feedback of ñgoodò, ñfairò or ñpoorò). Once the calibration and validation were both good, they were introduced to the experimental material. The general instructions for the experiment and button pressing were displayed as part of the practical trial so that the participants could learn and practise the button pressing and self-calibration at the same time. This procedure included: a big round button was pressed when they finished reading texts of a sentence and decided to go on; a left control button was pressed indicating the answer ñnoò while a right control button indicating ñyesò. After finishing each trial, a dot in the centre of a screen would appear to adjust the calibration and validation following each sentence; the big round button was pressed while the participants were focusing on the dot in exactly the same way as in their calibration and validation training. When the dot disappeared a new screen would appear (with a new sentence). They were introduced to eight practical trials, including reading and question answering. They were allowed to ask questions during this time. Once they finished the practical trials, the examiner left the test cubicle and allowed them to read the sentences on the screen without disturbance. When the participants had finished the whole task, the experimenter was informed and re-entered the room to take off the helmet and save the data in an .EDF file.

After the eye-tracking experiment, the participants were asked to fill in a questionnaire to rate the naturalness of the sentences. The sentences presented on the form were exactly the same sentences they just read on the screen, in the exactly same order. Since a questionnaire study can reflect a later stage of parsing, the ñpriming effectò from the eye-tracking task is not considered to be huge, especially as the participants were told to fill in the form as quickly as possible.

The whole experiment normally took up to 25-35 minutes depending on the reading speed of the various participants. They were paid ʐ4.50 for their participation. In four cases, the experiment took about 40-60 minutes because of a longer setting up time. Those five participants were paid ʐ5-6. The reward was paid after the participants finished the experiment.

◆ **Apparatus**

The same machine and system were used as described in Section 6.5.1.

7.1.2 Results

According to syntax-first theories, in both languages, in earlier stages of sentence processing, 84c) condition should be more difficult to process, than either a) or b). That is because c) is syntactically wrong, and if syntax is processed first, the processor would notice the problem right away. Condition a) & b) have no syntactic anomaly, so in the syntactic processing period, there should not be any difference between these two conditions.

However, LAST[61] predicts that in both languages, in earlier stages b) is more difficult to parse, compared to a) and c), because b) is semantically wrong. Since the meaning is first to be parsed, the violation of the semantics would cause a problem in processing. Conditions a) and c) should not show differences in the earlier stages, since there is no violation in these two conditions in this period of processing.

If the delay found in Chapter 6 is a pure language difference issue, parsing difficulty would be shown in English in the earlier stages, but not in Chinese.

All the assumptions predict that for the eye-tracking experiment, in the later stages, the data will show more difficulty in parsing condition b) and c) than a), since a) is the only condition without any kind of violations.

As for the rating task, all assumptions predict that the rating is lower for b) and c) than a), while it is not possible to predict the difference between b) and c).

The measurements of the eye-tracking data have been introduced in Chapter 4. In chronological order, they are *first fixation time*, *first pass time*, *regression path time*, the *reconstruction time*, and *total time*. The first fixation and first pass times are regarded as early parsing stages, whereas the reconstruction time and the total time are regarded as later ones in the on-line measurements. The regression path time is the first time that the parser finished processing a region and moved on to new material, so it can be regarded as a relatively early stage. Because Study 3 has not provided steady results from measurements involving saccade, these measurements (e.g., the *forward reading time*, and other measurements plus saccade time) will no longer be included in this experiment.

The measurement for the rating task is the score given to the sentences by the participants according to what they felt about the naturalness of the sentences in their languages. The lower rating can be related to bigger anomalous effects. This reflects an

[61] LAST in Chapter 3 & 5 only stands for one theory called *Late Assignment of Syntax Theory* by Townsend and Bever (2001). However, as described early in this chapter, the focus is only on the timecourse of the syntactic and semantic processing but not to scrutinise each theory in detail. From now on, LAST is used to refer to all theories that support parsing posterior to other types of sentence processing.

even later stage of processing.

7.1.2.1 Chinese Results

The interest areas focused on in the eye-tracking task were the PP (by + NP) and its anomalous counterparts. Repeated measures of ANOVAs were carried out for both subjects (F_1) and material analyses (F_2) for each of the measurements, as well as the rating scores. Both types of analyses contained three PP (or its counterparts) conditions as the within-subjects factors. Two main contrasts were carried out as: condition b) compared with the average of a) & c); then a) compared with c), and a further planned one-tail contrast t-test between b) & c) was carried out later.

◆ **First Fixation Time**

The mean first fixation time is listed in Table 23 below:

Table 23 Mean First Fixation Time for Chinese

	Mean	Std. Deviation	N
FFix Normal	224.5556	46.01435	18
FFix Semantic	238.9444	43.45176	18
FFix Syntactic	227.5000	39.25295	18

In this very early stage, there were no reliable main effects ($Fs<1.7$, $ps> .12$), nor were there any significant differences in any contrasts ($Fs<3.4$, $ps> .08$). It seems that at this processing stage, the processor did not spend a significant amount of time on any particular condition.

◆ **First Pass Time**

The mean first pass time is listed in Table 24 below:

Table 24 Mean First Pass Time for Chinese

	Mean	Std. Deviation	N
FPass Normal	381.5278	150.69296	18
FPass Semantic	472.4167	179.94274	18
FPass Syntactic	405.2778	169.57067	18

Statistically, there were no reliable main effects of the PP manipulation. (F_1 (2, 34) = 2.81, p= .074; F_2 (2, 46) = 1.75, p= .186). However, with a closer look at the contrasts, there was a significant difference between condition b) and the average of a) & c) from the subject analyses (F_1 (1, 17) = 8.622, p= .009); but this effect was only marginal in terms of material analyses (F_2 (1, 23) = 3.326, p= .081), further contrast between b) and c) show no significant differences (ts < 1.3, ps > .1 (one-tailed). No significant difference between a) & c) was found (Fs <1, ps > .6).

This implies that in the second early stage of the processing, the stimuli do not seem to be strong enough to trigger effects that can be generalised to either new participants or new materials.

◆ **Regression Path Time**

The mean regression path time is listed in Table 25 below:

Table 25 Mean Regression Path Time for Chinese

	Mean	Std. Deviation	N
RPath Normal	1992.666	1212.0236	18
RPath Semantic	2341.805	897.8041	18
RPath Syntactic	2046.166	1085.5529	18

In this stage, there were no reliable main effects (Fs<1.5, ps> .23), nor were there any significant difference in any contrasts (Fs<2.3, ps> .14). The tendency for the difference found in the earlier stage seemed to have disappeared. It seems at this processing stage, the processor did not spend a significant amount of time on any particular condition.

◆ **Reconstruction Time**

Table 26 Mean Reconstruction Time for Chinese

	Mean	Std. Deviation	N
Reconstruction Normal	1992.666	1212.0236	18
Reconstruction Semantic	2341.805	897.8041	18
Reconstruction Syntactic	2046.166	1085.5529	18

Table 26 above shows the mean reconstruction time for the Chinese speakers. It was apparent and also statistically confirmed that there were reliable main PP effects (F_1 (2,

34) = 29.662, $p < .001$; F_2 (2, 46) = 23.408, $p < .001$)). Further contrasts showed that there were significant differences between condition b) and the average of a) & c) (F_1 (1, 17) = 42.277, $p < .001$; F_2 (1, 23) = 31.673, $p < .001$), as well as reliable differences between a) and c) (F_1 (1, 17) = 14.788, $p = .001$; F_2 (1, 23) = 12.453, $p = .002$). The one-tailed comparison between b) and c) also showed significant difference (t_1 (17) = 4.4, $p < .001$; t_2 (23) = 3.3, $p = .001$).

From Table 26 and the statistical results, it is safe to conclude that condition b), which is the semantic anomaly, required a much longer reading time for Chinese speakers, compared to the other two conditions; whereas condition c), which is the syntactic anomaly, needed a much longer reading time than condition a) (the normal condition), but much less time than condition b). The normal condition did not cause any reading difficulty.

This stage seems to be the main parsing stage in Chinese in light of the fact that strong effects started to appear.

◆ **Total Time**

The mean total time is displayed in Table 27 below:

Table 27 Mean Total Time for Chinese

	Mean	Std. Deviation	N
Total Normal	909.972	553.85139	18
Total Semantic	1877.777	522.80854	18
Total Syntactic	1353.2778	638.7337	18

The effects in this stage were very similar to the previous one. There were reliable main PP effects (F_1 (2, 34) = 45.525, $p < .001$; F_2 (2, 46) = 20.264, $p < .001$)). Further contrasts showed that there were significant differences between condition b) and the average of a) & c) (F_1 (1, 17) = 67.111, $p < .001$; F_2 (1, 23) = 27.490, $p < .001$), with b) read the slowest; and reliable differences between a) and c) (F_1 (1, 17) = 20.554, $p < .001$; F_2 (1,23) = 10.168, $p = .004$), with a) read the quickest. The one-tailed comparison between b) and c) also showed significant difference (t_1 (17) = 5.4, $p < .001$; t_2 (23) = 3.2, $p = .002$), with b) costing longer reading time than c).

This suggests that the PP anomalous effects were very clear at the later stages of the parsing progress for Chinese speakers.

◆ **Rating for Chinese**

The mean total time is displayed in Table 28 below:

Table 28 Mean Rating for Chinese

	Mean	Std. Deviation	N
Rating Normal	4.1528	.3919	18
Rating Semantic	2.0764	.8414	18
Rating Syntactic	3.1736	.8544	18

As an off-line rating task can only reflect an even later stage of processing, the PP effects on the naturalness score are unsurprisingly consistent with the other two later stage measurements introduced above. There were reliable main PP effects (F_1 (2, 34) = 54.039, $p < .001$; F_2 (2, 46) = 100.477, $p < .001$)). Further contrasts showed there were significant differences between condition b) and the average of a) & c) (F_1 (1, 17) = 70.539, $p < .001$; F_2 (1, 23) = 134.783, $p < .001$), with b) rated as the least natural expression. Reliable differences between a) and c) were also shown (F_1 (1, 17) = 29.780, $p < .001$; F_2 (1, 23) = 53.127, $p < .001$), as unsurprisingly a) was the most acceptable form.

7.1.2.2 English Results

The English analyses repeat the Chinese ones. Whether they are consistent with the Chinese ones is a key question to the universality of the sentence processing strategy, and will be discussed in a later section.

◆ **First Fixation Time**

The mean first fixation time is listed in Table 29 below:

Table 29 Mean First Fixation Time for English

	Mean	Std. Deviation	N
FFix Normal	197.5278	29.99729	18
FFix Semantic	194.0000	33.37399	18

FFix Syntactic	190.6111	34.65738	18

As in Chinese, in this very early stage, there were no reliable main effects (Fs<1, *p*s> .5), nor were there any significant differences in any contrasts (Fs<1, *p*s> .4). The processor did not spend a significant amount of time on any particular condition.

◆ **First Pass Time**

The mean first pass time is listed in Table 30 below:

Table 30 Mean First Pass Time for English

	Mean	Std. Deviation	N
FPass Normal	338.9444	118.39663	18
FPass Semantic	402.4722	135.10350	18
FPass Syntactic	334.0833	118.50766	18

Somehow, surprisingly, there were already reliable main effects of the PP manipulation. (F_1 (2, 34) = 3.808, *p*= .034; F_2 (2, 46) = 6.704, *p*= .006). Moreover, contrasts showed that there was a significant difference between condition b) and the average of a) & c) (F_1 (1, 17) = 5.625, *p*= .03; F_2 (1, 23) = 8.571, *p*= .008), showing that b) required a longer reading time. However, there was no difference between a) and c) (Fs < 1, *p*s> .5). The one-tailed comparison between b) and c) also showed significant difference ((one-tailed) t_1 (17) = 2.5, *p*= .013; t_2 (23) = 2.6, *p*= .008), with b) reading longer than c).Taking this and the fact that b) is longer than the average of a) and c), it is safe to conclude that b) cost the longest reading time among the three conditions.

This indicates that in the second early stage of the processing, the stimuli in English were strong enough to trigger some effects that can be generalised to both new participants and new materials. This also means the effects appear to occur earlier than those in Chinese.

◆ **Regression Path Time**

The mean regression path time is listed in Table 25 below:

Table 31 Mean Regression Path Time for English

	Mean	Std. Deviation	N

RPath Normal	1578.8056	850.93352	18
RPath Semantic	1732.0833	1664.07640	18
RPath Syntactic	1603.3889	800.66609	18

The main effects that appeared in the previous measurements seemed to have disappeared in the subject analyses in this stage (F_1 (2, 34) = .084, p= .91), but it remained in the material analyses (F_2 (2, 46) = 5.482, p= .017). Subject analyses showed no difference in all the contrasts (Fs<1, ps> .7), but some of the differences remained in the material analyses, such as the difference between b) and the average of a) & c) (F_2 (1, 23) = 7.187, p= .013). Since these effects can only be generalised to new materials, but not new participants, they are not considered to be safe enough to be reported as reliable effects.

◆ **Reconstruction Time**

Table 32 Mean Reconstruction Time for English

	Mean	Std. Deviation	N
Reconstruction Normal	375.3611	275.21215	18
Reconstruction Semantic	865.8611	574.57907	18
Reconstruction Syntactic	643.8056	366.00880	18

Table 32 above is the mean reconstruction time for English speakers. Similarly to the Chinese data, in this stage the main PP anomaly effects were significant (F_1 (2, 34) = 9.489, p= .001; F_2 (2, 46) = 10.414, p< .001)). Further contrasts show that there were significant differences between condition b) and the average of a) & c) was also reliable (F_1 (1, 17) = 11.017, p= .004; F_2 (1, 23) = 16.848, p= .001), meaning that b) cost longer reading time. Moreover, there were also significant differences between a) and c) (F_1 (1, 17) = 7.157, p= .016; F_2 (1, 23) = 3.734, p= .070 (marginal)). The one-tailed comparison between b) and c) also showed significant difference (t_1 (17) = 2.05, p= .03; t_2 (20) = 2.0, p= .03), with b) reading longer than c). It is safe to conclude that b) is the most difficult condition to read.

◆ **Total Time**

The mean total time is displayed in Table 33 below:

Table 33 Mean Total Time for English

	Mean	Std. Deviation	N
Total Normal	714.3056	282.14135	18
Total Semantic	1268.3333	652.58132	18
Total Syntactic	977.8889	382.81725	18

The effects in this stage were very similar to the Chinese one. There were reliable main PP anomaly effects (F_1 (2, 34) = 12.756, $p < .001$; F_2 (2, 46) = 14.370, $p < .001$)). Further contrasts showed that there were significant differences between condition b) and the average of a) & c) (F_1 (1, 17) = 15.333, $p = .001$; F_2 (1,23) = 19.781, $p < .001$), and reliable differences between a) and c) (F_1 (1, 17) = 8.100, $p = .011$; F_2 (1,23) = 7.423, $p = .012$). The one-tailed comparison between b) and c) also showed significant difference (t_1 (17) = 2.6, $p = .008$; t_2 (23) = 2.7, $p = .006$), with b) reading longer than c). These results suggest that a) cost shortest reading time, while b) cost longest reading time.

◆ **Rating for English**

The mean total time is displayed in Table 34 below:

Table 34 Mean Rating for English

	Mean	Std. Deviation	N
Rating Normal	4.2014	.3765	18
Rating Semantic	1.8194	.5787	18
Rating Syntactic	2.8750	.8291	18

In English, the PP effects on the naturalness scores were also consistent with the other two later stage measurements introduced above. There were reliable main PP effects (F_1 (2, 34) = 132.883, $p < .001$; F_2 (2, 46) = 258.424, $p < .001$)). Further contrasts showed that there were significant differences between condition b) and the average of a) and c) (F_1 (1, 17) = 294.981, $p < .001$; F_2 (1, 23) = 304.313, $p < .001$), and reliable differences between a) and b) ((F_1 (1, 17) = 323.997, $p < .001$; F_2 (1, 23) = 501.190, $p < .001$). Combining the data from Table 34 above, it is safe to conclude that condition a) was the most acceptable form, while condition b) was the least acceptable form.

7.1.2.3 Comparing Chinese and English

Repeated measures of ANOVAs were carried out for all participants on both subjects (F_1) and material analyses (F_2) for each of the measurement. The within subject factors contain three PP (or its counterpart) conditions. The between subjects factor is the first language, where the number of each category is well balanced. In addition to the contrasts carried out in the previous analyses, further contrasts on condition a) and b), as well as b) and c) were performed. The most interesting results from these contrasts lie in the interaction of language * conditions. By looking at these interactions, the different behaviours between the two languages can be observed.

◆ **First Fixation Time**

The mean first fixation time is listed in Table 25 below:

Table 35 Mean First Fixation Time

	LANGUAGE	Mean	SD	N
FFix Normal	English	197.5278	29.99729	18
	Chinese	224.5556	46.01435	18
	Total	211.0417	40.66103	36
FFix Semantic	English	194.0000	33.37399	18
	Chinese	238.9444	43.45176	18
	Total	216.4722	44.46892	36
FFix Syntactic	English	190.6111	34.65738	18
	Chinese	227.5000	39.25295	18
	Total	209.0556	41.00867	36

Despite no within-subjects main effects or PP by language interactions (Fs < 1.2, *p*s> .3), and no differences by contrasts (Fs < 3.1, *p*s> .1), there were significant differences between Chinese and English speakers (F_1 (1, 34) = 12.034, *p*= .001, F_2 (1, 46) = 8.521, *p*= .005). It is clear from Table 13 that Chinese speakers read slower than English speakers.

◆ **First Pass Time**

The mean first pass time is listed in Table 36 below:

Table 36 Mean First Pass Time

	LANGUAGE	Mean	SD	N
FPass Normal	English	338.9444	118.39663	18
	Chinese	381.5278	150.69296	18
	Total	360.2361	135.29484	36
FPass Semantic	English	402.4722	135.10350	18
	Chinese	472.4167	179.94274	18
	Total	437.4444	160.78201	36
FPass Syntactic	English	334.0833	118.50766	18
	Chinese	405.2778	169.57067	18
	Total	369.6806	148.63076	36

There were significant PP anomaly main effects in this measurement (F_1 (2, 68) = 6.048, p= .004; F_2 (1, 23) = 7.047, p= .001). Further contrasts showed significant differences between b) and the average of a) & c) (F1 (1, 34) = 14.034, p= .001; F2 (1, 46) = 11.381, p= .002). However, there was no difference between a) and c) (Fs <1, ps> .5). This means that participants, in general, took longer time to process b), while a) took the shortest time. It is also important to remember that the difference in processing time for a) and c) were not very significant, as reported earlier. This will be further discussed in the next section. The one-tailed comparison between b) and c) also showed significant difference (t_1 (35) = 2.9, p= .003; t_2 (47) = 2.8, p= .004), with b) reading longer than c).

The significant between-subjects language effects had disappeared in this stage (Fs<2.3, ps > .1). Furthermore, there was no PP by language interaction either in the main effects or in the contrasts (Fs < 1, ps > .1). This means the two language groups behave very similarly.

◆ **Regression Path Time**

The mean regression path time is listed in Table 37 below:

Table 37 Mean Regression Path Time

	LANGUAGE	Mean	SD	N
RPath Normal	English	1578.8056	850.93352	18
	Chinese	1992.6667	1212.02361	18
	Total	1785.7361	1053.21337	36
RPath Semantic	English	1732.0833	1664.07640	18
	Chinese	2341.8056	897.80414	18
	Total	2036.9444	1353.55926	36

RPath Syntactic	English	1603.3889	800.66609	18
	Chinese	2046.1667	1085.55298	18
	Total	1824.7778	966.52284	36

In this stage, the main effects that appeared in the previous measurement were reduced to a nonsignificant level. For the within-subjects factor, the subject analyses showed nothing significant, either in the main effects, contrasts or interactions (F_1s < 1, *p*s > .4), even though material analyses showed some main effects (F_2 (2, 92) = 6.898, *p* = .002). There were also differences in some of the contrasts in the material analyses only: significant differences between b) and the average of a) & c) (F_2 (1, 46) =9.440, *p*= .004). However, the result cannot be generalised to a wider range of participants.

As for the between-subjects effect, it was not reliable either. Even though the subject analyses showed significant language differences (F_1 (1, 34) = 4.303, *p*= .045), it was not confirmed by material analyses (F_2 (1, 46) = 1.069, *p*= .307).

To sum up, the effects in this stage cannot be generalised to either a wider participant sample pool, or other materials.

◆ **Reconstruction Time**

Table 38 Mean Reconstruction Time

	LANGUAGE	Mean	SD	N
Reconstruction Normal	English	1578.8056	850.93352	18
	Chinese	1992.6667	1212.02361	18
	Total	1785.7361	1053.21337	36
Reconstruction Semantic	English	1732.0833	1664.07640	18
	Chinese	2341.8056	897.80414	18
	Total	2036.9444	1353.55926	36
Reconstruction Syntactic	English	1603.3889	800.66609	18
	Chinese	2046.1667	1085.55298	18
	Total	1824.7778	966.52284	36

Table 38 above is the mean reconstruction time. Not violating results patterns in English and Chinese in the reconstruction period, there were significant PP anomaly main effects (F_1 (2, 68) = 36.398, *p*< .001, F_2 (2, 80) = 30.715, *p*< .001). Further contrasts showed that the significant differences were between condition b) and the average of a) and c) (F_1 (1, 34) = 47.501, *p*< .001, F_2 (1, 40) = 44.056, *p*< .001), between a) and c) (F_1

(1, 34) = 21.543, $p< .001$, F_2 (1, 40) = 14.295, $p= .001$), between condition b) and c) (F_1 (1, 34) = 19.983, $p< .001$, F_2 (1, 40) = 15.636, $p< .001$), as well as between condition a) and b) (F_1 (1, 34) = 58.459, $p< .001$, F_2 (1, 40) = 64.563, $p< .001$). The one comparison between b) and c) has also shown significant difference (t_1 (35) = 4.4, $p< .001$, t_2 (44) = 3.7, $p< .001$), with b) reading longer than c). From Table 38 above, it is clear that b) required the longest reading time while a) required the shortest.

Between the participants, language differences were also significant (F_1 (1, 34) = 6.829, $p= .013$, F_2 (1, 40) = 6.206, $p= .017$), but this only means that Chinese speakers read slower than English speakers. Moreover, there was an interaction of difference between condition a) and b) by language (F_1 (1, 34) = 4.668, $p= .038$, F_2 (1, 40) = 4.417, $p= .042$). This suggests that in this stage, the reading time difference between the semantic anomalous condition and the normal condition is larger for Chinese speakers than for English speakers.

◆ **Total Time**

The mean total time is displayed in Table 39 below:

Table 39 Mean Total Time

	LANGUAGE	Mean	SD	N
Total Normal	English	1578.8056	850.93352	18
	Chinese	1992.6667	1212.02361	18
	Total	1785.7361	1053.21337	36
Total Semantic	English	1732.0833	1664.07640	18
	Chinese	2341.8056	897.80414	18
	Total	2036.9444	1353.55926	36
Total Syntactic	English	1603.3889	800.66609	18
	Chinese	2046.1667	1085.55298	18
	Total	1824.7778	966.52284	36

The effects in this stage are very similar to the Chinese one. There were reliable main PP anomaly effects (F_1 (2, 34) = 12.756, $p< .001$; F_2 (2, 46) = 14.370, $p< .001$)). Further contrasts show that there were significant differences between condition b) and the average of a) & c) (F_1 (1, 17) = 15.333, $p= .001$; F_2 (1,23) = 19.781, $p< .001$), reliable differences between a) and c) (F_1 (1, 17) = 8.100, $p= .011$; F_2 (1,23) = 7.423, $p= .012$), between condition b) and c) (F_1 (1, 34) = 30.644, $p< .001$, F_2 (1, 46) = 17.691, $p< .001$), and between condition a) and b) (F_1 (1, 34) = 85.008, $p < .001$, F_2 (1, 46) = 65,546, $p < .001$). These results suggested that a) cost the shortest reading time, while b) cost the longest reading time.

The results from the total time were very similar to that of the reconstruction time. There were significant PP main effects (F_1 (2, 68) = 51.895, $p < .001$, F_2 (2, 92) = 34.263, $p < .001$). Further contrasts showed that the significant differences were between condition b) and the average of a) and c) (F_1 (1, 34) = 68.519, $p < .001$, F_2 (1, 46) = 46.646, $p < .001$), and between a) and c) (F_1 (1, 34) = 27.548, $p < .001$, F_2 (1, 46) = 17.418, $p < .001$). The one-tailed comparison between b) and c) has also shown significant difference (t_1 (35) = 5.4, $p < .001$, t_2 (47) = 4.2, $p < .001$), with b) reading longer than c).The reading patterns suggest a consistency with the previous analyses, with b) to be the most difficult to read, therefore costing more reading time, whereas a) was the easiest, and taking the least amount of time to read.

Between the participants, language differences were also significant (F_1 (1, 34) = 6.755, $p = .014$, F_2 (1, 46) = 19.369, $p < .001$). They were differences between condition b) and c) by language (F_1 (1, 34) = 5.098, $p = .030$, F_2 (1, 46) = 3.892, $p = .055$ (marginal)), and the interaction of difference between condition a) and b) by language (F_1 (1, 34) = 4.668, $p = .038$, F_2 (1, 40) = 4.417, $p = .042$). The subject analyses showed some PP anomalous effects by language interaction (F_1 (1, 68) = 3.852, $p = .026$), but these could not be generalised to other materials (F_2 (1, 92) = 2.543, $p = .084$ (marginal)).

- **Rating**

The mean total time is displayed in Table 40 below:

Table 40 Mean Rating

	LANGUAGE	Mean	SD	N
Rating Normal	English	4.1528	.39191	18
	Chinese	4.2014	.37656	18
	Total	4.1771	.37959	36
Rating Semantic	English	2.0764	.84141	18
	Chinese	1.8194	.57876	18
	Total	1.9479	.72357	36
Rating Syntactic	English	3.1736	.85442	18

Chinese	2.8750	.82916	18
Total	3.0243	.84347	36

There were significant main PP anomaly effects (F_1 (2, 68) = 162.098, p< .001; F_2 (2, 92) = 305.892, p< .001)). Further contrasts showed that there were also significant differences between condition b) and the average of a) and c) (F_1 (1, 34) = 239.041, p< .001; F_2 (1, 46) = 384.892, p< .001), and reliable differences between a) and c) (F_1 (1, 34) = 86.120, p< .001; F_2 (1, 46) = 195.765, p< .001). Moreover, there was also a significant difference between a) and b) ((F_1 (1, 34) = 353.485, p< .001; F_2 (1, 46) = 716.970, p< .001). The reading patterns replicated the previous analyses.

There were no between subjects language differences (Fs<1, ps> .5). The subject analyses showed no reliable interaction of PP effects by language (F1s<2.5, ps> .17). This means the two participant groups behaved the same in terms of the naturalness of the material.

7.1.3 Discussion

The results reported in last section seem rather complex. To sum up the results that were confirmed by subject-analyses and material analyses, a simplified report in chronological order and by language is presented below in table 41:

Table 41 Result Outline

Time Blocks	Language	Results
A. First Fixation Time	a. Chinese	• Nothing significant
	b. English	• Nothing significant
	c. Mix	• Language difference
B. First Pass Time	a. Chinese	• Nothing significant
	b. English	• PP anomaly effects; • Significant longer reading time for condition b)>a) and b)>c); • No difference between a) and c)

	c. Mix	• PP anomaly effects; • Significant longer reading time for condition b); • No difference between a) and c)
C. Regression Path Time	a. Chinese	• Nothing significant
	b. English	• Nothing significant
	c. Mix	• Nothing significant
D. Reconstruction Time	a. Chinese	• PP anomaly effects; • Significant longer reading time for condition b)>c)> a)
	b. English	• PP anomaly effects; • Significant longer reading time for condition b)>c)> a)
	c. Mix	• PP anomaly effects; • Significant longer reading time for condition b)>c)> a); • Language difference; • Anomaly * language interaction between condition a) and condition b)
E. Total Time	a. Chinese	• PP anomaly effects; • Significant longer reading time for condition b)>c)> a)
	b. English	• PP anomaly effects; • Significant longer reading time for condition b)>c)> a)
	c. Mix	• PP anomaly effects; • Significant longer reading time for condition b)>c)> a); • Language difference; • Anomaly by language interaction between condition a) and condition b), and between condition b) and c)
F. Rating[62]	a. Chinese	• PP anomaly effects; • Significant lower naturalness scores for condition b)>c)> a)
	b. English	• PP anomaly effects; • Significant lower naturalness scores for condition b)>c)> a)
	c. Mix	• PP anomaly effects; • Significant naturalness scores for condition b)>c)> a); • No language difference; • No interaction

◆ Stage A

In the very early stage A, the only effect was language difference. Figure 26 below shows that Chinese readers were much slower than English readers:

[62] Although rating is not technically part of the online time blocks, it is considered to be an even later stage here.

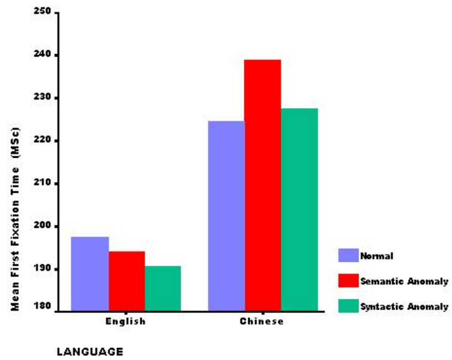


Figure 26 First Fixation Time

Considering the size of the Chinese text regions—not significantly larger than its English counterpart—why did it take so much longer for the Chinese to read? One simple explanation is that the Chinese readers are generally slow readers. However, what caused the slow speed in reading is questionable. Another explanation is that at this stage some processing occurs in Chinese, but not in English. One possible process of this kind is word segmentation (as mentioned in earlier sections). This suggestion can be reinforced in later discussion.

- **Stage B**

In the relatively early stage B, the Chinese reading time remained nonsignificantly affected by the PP types. However, it was no longer the case for the reading time of the English group and the mixed group. The mixed group showed a similar effect to that of the English one, mainly because the pattern of the Chinese data is the same as the English one. Figure 27 below sketches out the pattern of reading time:

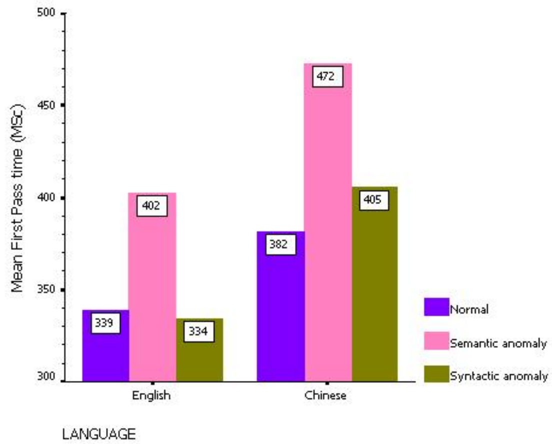
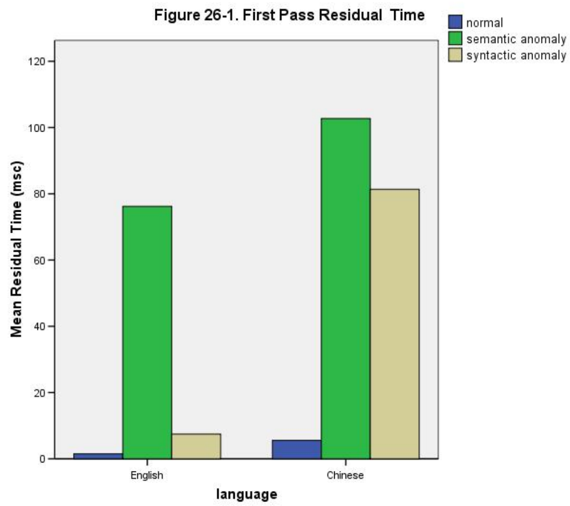


Figure 27 First Pass Time

The figure shows that in both languages, semantic anomaly cost the longest reading time. Since the difference between conditions a) and c) was not significant, one can conclude that in this stage the syntactic anomaly was not recognized. The processing strategy in the syntactically anomalous condition was not much different from the one in a normal condition. This is clear evidence that the English speakers picked up the effects earlier than the Chinese speakers. However, their merely picking up the effects caused by semantic anomaly suggests that the semantic information is processed to a higher level than the syntactic information by this stage.

One concern of this time measurement is that it might be sensitive to the critical region size. In the current experiment, the length between the three anomaly conditions has not been strictly controlled. This might reduce the reliability of the effects in very early stages, as low-level eye movement procedures might be affected by the physical layout of the material. Therefore, a length correction is carried out and same analyses performed on the original data are repeated on the residual time.

The residual time is calculated as this: a sample of first pass time was taken from all normal sentences with different length of the interest areas. A linear regression was performed for each participant, predicting reading time from length (measured in characters). For each data point in the analyses, the residual value is the difference between the predicted first-pass time and the actual time measured by the experiment. A positive value means that the actual time is longer than the predicted value, and vice versa. The residual values (as in the graph below) are then entered into ANOVA as described in section 7.1.2.



For Chinese data, there were main effects of the PP manipulation from the subject analyses (F_1 (2, 34) = 3.49, p= .042), but it cannot be generalised to other materials (F_2 (2, 46) = 1.957, p= .153). 1-tailed t test showed that there was significant difference between a) normal and b) semantic anomaly (t_1 (17) = 2.706, p = .007; t_2 (23) = 2.706, p= .007), but no other differences were confirmed (ts < 1.7, ps > .06).

For English data, there were significant main effects of the PP manipulation from the subject analyses (F_1 (2, 34) = 3.344, p= .047; F_2 (2, 46) = 3.404, p= .042). 1-tailed t test showed that there was significant difference between a) normal and b) semantic anomaly (t_1 (17) = 1.99, p = .032; t_2 (23) = 2.052, p= .026). Differences between b) semantic and c) syntactic anomalies have also been confirmed (t_1 (17) = 2.15, p = .023; t_2 (23) = 2.098, p= .024). However, there was no difference between a) and c) (ts < 1, ps > .8)

For the mixed data, there were significant PP anomaly main effects in this measurement (F_1 (2, 68) = 5.858, p= .004; F_2 (1, 23) = 4.226, p= .018). However, there is no language difference (Fs <1.3, ps > .2), nor PP anomaly * language interaction (Fs <1.07, ps > .3). 1-tailed t test showed that there was significant difference between a) normal and b) semantic anomaly (t_1 (35) = 3.343, p = .001; t_2 (47) = 2.798, p= .003). Differences between b) semantic and c) syntactic anomalies have also been confirmed (t_1 (35) = 1.8, p = .04; t_2 (47) = 1.719, p= .046). However, there was no difference between a) and c) (ts < 1.6, ps > .05).

The above means, with the length correction, there is also a tendency of early main effects in Chinese. Moreover, the difference between normal and semantic condition has become significant. All the other effects replicated the analyses using the untrimmed raw data. This also means that the uncontrolled length is not a major problem in the design and therefore no further residual analyses need to be carried out.

Therefore, it is safe to conclude that at the early stage such as first pass time, the semantic anomaly has been spotted by the processor for both English and Chinese speakers. However, the semantic anomaly seems to remain unnoticed.

- **Stage C**

This is a middle stage. During the regression time, the effects from the main stimuli and the contrasts were all reduced for both languages. If one does not look at a later stage, it is not clear what exactly was happening here. But since all the effects reappeared in the later stages, this middle stage can be looked on as a period when the syntactic anomaly started to be noticed and processed. Figure 28 below shows the reading time for this stage:

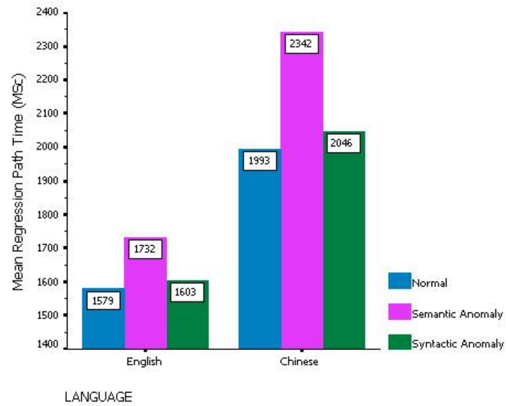


Figure 28 Regression Path Time

Noticeably, the pattern of reading time is as before, that condition b) cost the longest reading time. But the picking up of the anomaly in condition c) has cancelled out some of the effects caused by condition b).

◆ **Stage D**

This is considered as the first later processing stage. It is also the first time when significant effects were shown in Chinese. Figure 29 below is an illustration of the two languages at this stage. The two languages behave very similarly according to this figure:

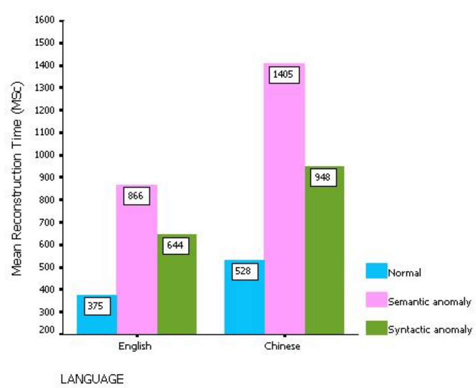


Figure 29 Reconstruction Time

The reading time pattern (from longest to shortest reading time: condition b)>c)>a)) suggests that condition b) is the most difficult to process, while condition a) is the easiest. There were some language differences though, as Chinese readers spent a longer time reading.

The interaction of the difference between a) and b) by language means a sudden rouse of semantic anomaly recognition in Chinese. Recall that Chinese did not show much effect in earlier stages. This sudden growth in recognition in Chinese can be seen as a result of cumulative processing of the stimuli of semantic anomaly. This does not mean that the processing of syntactic anomaly was not taking place. In fact, it is statistically reliable that condition c) cost shorter reading time than b), but longer than a). The lack of interaction of contrasts involving c) by language implies that the processing of syntactic anomaly of the two languages was at a similar pace, while the semantic anomaly process that was found earlier in English was now starting to grow in Chinese. This is in line with the assumption given above, that Chinese processing is delayed compared to its English counterparts. Moreover, one can also see the fact that semantic processing takes place earlier than syntactic processing in both languages

- **Stage E**

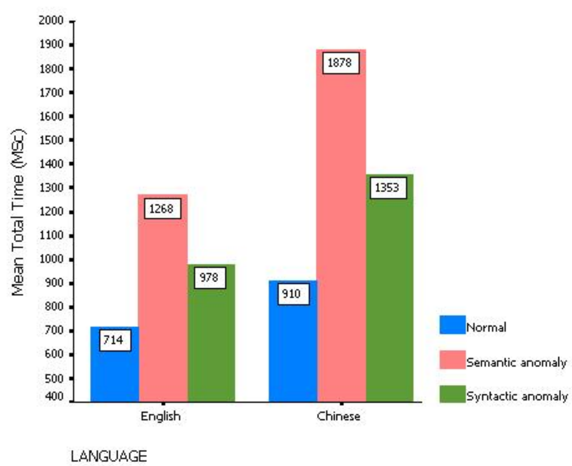


Figure 30 Total Time

The pattern and effects of either English or Chinese, and the language differences will not be introduced here, since they are identical to Stage D discussed above. The results of interest are the interactions of condition a) and b) difference by language, and b) and c) difference by language. Figure 30 above shows that in this stage, the reading time on both the semantic and syntactic anomaly increased much faster in Chinese than in English. While the processor kept picking up the semantic anomaly, the syntactic anomaly effects were amplified. This again suggests that the Chinese readers could process the anomaly, but at a much later stage. And once the anomaly was found, it kept affecting the reading time stoutly for a period of time. Even though the effects in English at this stage were still very robust, it was stronger in Chinese.

◆ **Stage F**

As for the rating task, the indication from the results is consistent with the later stages from the eye-tracking task. Figure 31 below demonstrates the pattern:

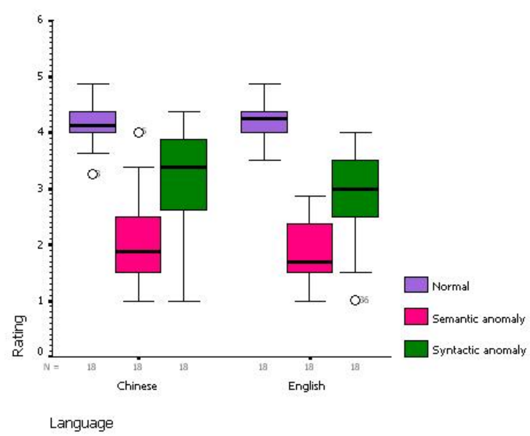


Figure 31 Rating

The lowest scores in condition b) suggest that in both languages semantically anomalous sentences were the least acceptable. Even though the significant difference between b) and c) means that the syntactic anomaly was far more acceptable than syntactic anomaly, the significant difference between a) and c) means that c) was still judged as an incorrect sentence. This shows not only evidence of the judgment of the naturalness of the stimuli, but also that the materials used in the experiments were credible – the anomalous stimuli were not rated the same as the normal control condition.

There were no language differences or interactions in this task. This suggests that at a very late stage, the processing and judgment towards the anomaly were universal.

◆ **General Discussion**

To sum up, there was no sign of syntactic anomaly effects in early stages in either language. These anomaly effects were picked up in much later stages. The results do not support syntax-first assumptions, but are consistent with *LAST*.

Moreover, the Chinese reading time is consistently longer than the English reading time; the effects of semantic processing also appeared later in Chinese than English – recall that the English effects were already shown in the first pass time stage, unlike in Chinese. This suggests that there is a delay in Chinese processing compared to English. This interaction is in line with previous RC attachment studies. One possible explanation of this delay is word segmentation. As Chinese is a language without clear lexical marks, it is important to identify words any sentence processing level can occur. As discussed earlier, syntax-first assumptions are not good at explaining phenomena of this type (another similar processing would be acoustic sentence processing).

One criticism against *LAST* from syntax-first believers is the definition of the ñpseudosyntaxò. Since this pseudosyntax is ñprocessedò (or as it were ñassignedò) at an early stage of the processing, it is in contrast to what the syntax-first assumptions claim. According to Townsend and Bever, the ñ*pseudosyntax* consists of the immediate initial processes that isolate major phrases, differentiate lexical categories, and assign initial thematic relationsò (2001, P.187). It ñinvolves recognition of function morphemes and lexical categoriesò. Depending on the frequency of the sentence patterns and the subcategorisation properties of verbs, it assigns words to phrases and then to thematic roles. This definition is rather wide; and some (e.g. Miyamoto, personal communication, 2005) would claim that the pseudosyntax is already a ñrealò syntax.

Townsend and Bever's definition of pseudosyntax is problematic also in terms of explaining the current data. Without assigning a proper syntactic structure, one would expect the syntactic anomaly to read faster even than the normal condition in the very early stage, because according to Townsend and Bever's definition, in the pseudosyntax assignment stage, the frequency of the argument to a verb phrase (VP) should be an NP, rather than a PP. In that case, the syntactically anomalous stimulus is more proper than the normal condition. However, the current results suggest that the processor did not find PP more difficult to process than NP. Rather, it is the mismatch between the argument and the verb that causes the problem.

One possible suggestion is to decrease the range of the definition in terms of the pre-syntactic stage. Instead of calling it "pseudosyntax", it can be simply called ñpre-syntax processingò (PSP). This stage is involved in all kinds of processing EXCEPT syntactic cues: The processor will identify words and their functions based on lexical meanings and frequency. Then the processor will try to combine the words next to it and assign possible meanings based on frequency of word order and combination. This is not a simple stage of lexical processing because it involves processing of words and word groups; however, it is not a stage of syntactic processing either, since there is no structure being constructed at this time.

In the case of this study, in English, when be+v-ed is processed already as a passive form (because the passive form is the most frequent form for this combination), the incoming unanimated NP has become unacceptable, even though there is a *by* to mark the passive subject. However, if the coming input is a possible NP, skipping a function word such as "by" is allowed, since the function word itself does not bear much of the meaning while the meaning of the centre of the sentence (verb) is clear. This again suggests that in earlier stages, the meaning is very important to the processor. Word identification for English in written form is relatively easy, since there are spaces between words.

The Chinese case is more complicated, since there are no spaces between words. One strategy is to read a couple of characters further down the sentence, and then identify words by the pre-syntax processing. This of course involves more processing strategies than English, such as the use of memory, etc. This may be why the Chinese processor needs a longer reading time than the English one in the very early stages. Once the word segment is identified, the parser can carry on with further processing.

In Chinese, the passive subject PP appears in front of the verb. When 被 is identified after an NP, the passive form is identified at the same time, according to the frequency of the combination. So an unanimated NP will be improper right after 被. However, when 被 is omitted, since the verb has not been processed, and the combination of NP+NP is not common in Chinese, one would expect the effects to appear right away. The results from the current study seem to be inconsistent with this proposal. Therefore, inevitably, the grain-size problem in Chinese processing has to be discussed – the time spent on the word segmentation operation.

Does the processor store the whole sentence, or process it bit by bit? If it is the latter, how much is "a bit"? This question cannot be solved by the current study. However, common sense suggests that this segmentation is done bit by bit, because of the cost on memory. It is impossible to store up a complete long sentence and still remember the segmentation at the beginning of the sentence. As for the amount of information that is needed for storing, one suggestion is one to three characters after a word. Because one is the minimum character to form a word, while two to three characters are the common size of a Chinese word, one can already tell whether the next character is connected to the previous characters from this information. Yet these suggestions are already out of the scope of the current study. More experiments need to be carried out to address the above questions.

Coming back to the topic of the inconsistency of the results, our proposal agrees that the verb is the centre of the sentence. When it is NP+NP in Chinese, the processing of the second NP stays in the stage that the second word is unidentified. Following the second NP, the new incoming information is the verb. Because the verb is the key to the meaning of a sentence and the second NP is just prior to the verb, the meaning connection between the second NP and the verb becomes very close. It does not mean this kind of connection is necessarily correct (although it is incorrect in this case). In the early stage, once the meaning is connected, the processor can move on to further information. Indeed, the Chinese results did not violate this proposal.

7.1.4 Conclusion and Criticism

The current study shows contrary results to the predictions from syntax-first assignment theories. With some modifications to the *LAST*, the results can be well explained. Moreover, the delayed effects in Chinese in the current study and the previous RC attachment studies can be explained as a language specific feature. This feature may be the unclear segmentation of words in Chinese.

However, some (personal conversation with Miyamoto, 2005) have presented the criticism that the stimuli sizes are not balanced in that experiment. It was said that in the syntactic anomaly, dropping the two-letter preposition *by*, or its equivalent one-character Chinese word 被, is too small to be noticed (or very easy to be skipped), compared to the manipulation of the Noun Phrase in the semantic anomaly. This leads to a serious question: is the lack of evidence of early syntactic anomaly actually caused by an early syntax processing so that people have self-mended the anomalous part by adding a *by* there when they "skip" the preposition?

The next study tries to answer this question. The methodology of the new experiment is very similar to the previous one, with some modifications to the material to "balance" the size of stimuli across the two languages.

7.2 Study 5: Timecourse of Syntactic and Semantic Processing Continued

As stated in Chapter 7.1.3, the motivation for this experiment was to refine the design of Study 4. If the results can be replicated by a more careful balanced design, the support of *LAST* would be more robust.

7.2.1 Method

◆ **Participants**

A total of 36 volunteers participated, of whom 18 were native English speakers and 18 were native Chinese speakers, aged between 18 and 45. Most of them were either students or staff at the University of Exeter. 19 of them were from the School of Psychology, 16 from the School of Finance, and one is a visiting family member of a student from China. 10 were male and 26 were female.

◆ **Materials**

The materials were very similar to the previous study. There were 24 passive sentences with a) no anomaly, b) semantic anomaly, and c) syntactic anomaly, in both Chinese and English. Sentences for condition a) and c) remained the same as in the previous study, but condition b) changed. Instead of using unanimated NPs to achieve the semantic anomaly, the prepositions *by* and 被 were changed by other prepositions, *at* and 在.. The example sentences are given below.

86) a. **Normal**

John broke his ankle when he fell off the ladder. He was carried *by his friend* to a hospital.

約翰從梯子上摔了下來，摔壞了腳踝。他被他的朋友一路擡到醫院。

Yuehan cong tizishang shuaile xialai, shuaihuaile jiaohuai. Ta *bei tade pengyou* yilu taidao yiyuan.

b. **Semantic Anomaly**

John broke his ankle when he fell off the ladder. He was carried ***at*** *his friend* to a hospital.

約翰從梯子上摔了下來，摔壞了腳踝。他在他的朋友一路擡到醫院。.

Yuehan cong tizishang shuaile xialai, shuaihuaile jiaohuai. Ta *zai tade pengyou* yilu taidao yiyuan.

c. **Syntactic Anomaly**

John broke his ankle when he fell off the ladder. He was carried *his friend* to a hospital.

約翰從梯子上摔了下來，摔壞了腳踝。他他的朋友一路擡到醫院。

Yuehan cong tizishang shuaile xialai, shuaihuaile jiaohuai. Ta *tade pengyou* yilu taidao yiyuan.

The comparability of the two languages has been discussed in the previous chapter, so it will not be repeated here. However, one point worth explaining is the anomaly for the semantics in condition 86b). There has been some criticism[63] that:

i) Not using *by*, but *at*, actually changes the deep (D)-structure of the sentence, so the sentence is no longer syntactically identical.

ii) As a preposition is a function word (functor), it carries very little lexical meaning, but rather grammatical functions, condition b) is actually syntactically wrong, but semantically intact.

The debate of the relationship between semantics and syntax is out of the scope of the current paper. Nevertheless, the two criticisms are answerable. It has been made clear in the beginning of this thesis (see Chapter 2) that the development of the universal grammar has followed an incremental course of eliminating the so-called ñD-structureò. However, Criticism i) does not stand at any stage of the development of the UG (namely, the Transformational Grammar (TG), the Government and Binding (GB) and the Minimalist Programme (MP)).

First, TG claims that the D-structure of a passive sentence is the active form. A transformation from the active form to the surface (S) ï structure includes transferring the subject to the object position, transferring the object to the subject position and adding ñbeò and ñbyò. (See Figure 5 in Chapter 2, P.24). Although TG claims that the transformation is not bi-directional, we have to trace back the D-structure to be able to answer i).

[63] An anonymous comment presented at the Camling conference, 2006.

As sentence 86a) is a normal passive sentence, its D-structure should be 87):

87) … *His friend* carried him to a hospital.

If 86b) does not share the same D-structure, then its own D-structure should be something like 88):

88) … (Someone) carried him to a hospital at *his friend*.

According to the Government and Binding (GB) theories introduced in Chapter 2.2, a passive sentenceôs D-structure is itself.

In the ñtransformedò passive sentences 86a) above when *his friend* is interpreted as the acting agent, 87) is a perfect sentence. However, if *his friend* in 86b) is not interpreted as the agent, sentence 88) does not make sense (so it is anomalous before and after the transformation). Therefore, the NP *his friend* has to be interpreted as the agent and 88) becomes 89):

89) … *His friend* carried him to a hospital.

Apparently, 89) is identical to 88), which means the two sentences should share the same deep structure. *At* in the S-structure makes it problematic in tracing back the original D-structure because it is anomalous, and it is exactly the design of the experiment.

Secondly, according to GB, the D-structure of a passive sentence is itself (See Chapter 2, Figure 6, P.25). The transformation from D-structure to S-structure is to change the accusative form to its nominative form and to move *be* to the head of the IP, both leaving a trace of movement. In this sense, there is no transformation on the preposition phrase *by* + NP. Similarly, if the PP is *at* + NP, there is no transformation on this structure either. The two sentences 86a) & b) have identical structure, and there is no change of PP in their structure, therefore the D-structure has not been changed, even the head of the PP is *at* instead of *by*.

Lastly, MP does not emphasise on D-structure any more, and therefore the issue raised in Criticism i) is moot.

As to Criticism ii), even though function words do not carry rich lexical meanings (Radford, 2001), it does not mean that they are meaningless. Indeed, they are rich in semantic terms. For example, *at* can be used, according to the *Merriam-Webster Online*:

90) a) -- used as a function word to indicate presence or occurrence in, on, or near;

b) -- used as a function word to indicate the goal of an indicated or implied action or motion;

c) -- used as a function word to indicate that with which one is occupied or employed;

d) -- used as a function word to indicate situation in an active or passive state or condition;

e) -- used as a function word to indicate the means, cause, or manner;

f) -- used as a function word to indicate the rate, degree, or position in a scale or series;

g) -- used as a function word to indicate age or position in time.

However, among all the meanings this dictionary gives, there is no such item that *at* can be used *as a function word to indicate through the agency or instrumentality of*. Yet, this item can be found for *by*.

The key reason that b) was first mis-transformed into 88) lies in the mismatching of the semantics of *at* in the sentence. Because it lacks the meaning for indicating an agent, the NP following this preposition (or to say, the NP within the PP) cannot be interpreted correctly. Thus, the mistake lies in the semantic term, but not the syntactic term.

Notably, there are some linguists (J. D. Fodor, 1977) who consider that the deep structures of sentences themselves contain enough representations of meaning, so that the relation between deep structure and meaning is simply that of identity. If this is the case, our statement still stands: A correction of changing *at* into *by* is needed because *at* is anomalous in its meaning in the sentence, or the deep structure is also anomalous. The identical deep structures of 86a) and b) could not have been shown without this correction. Hence, this anomaly is chosen in our testing manipulation.

The fonts, sizes, frequencies and the filler sentences and questions were controlled in exactly the same as in Study 4 so it will not be repeated here.

◆ **Procedure and Apparatus**

The procedure and apparatus were exactly the same as in Study 4 and their descriptions will not be repeated here.

7.2.2 Results

As the purpose of this experiment is to answer the ñstimulus balanceò problem, the prediction is that the results of the current experiment should replicate the previous study, despite of the change of manipulation.

The results of the previous study show that at earlier stages, in English, it is more difficult to parse the semantic anomalous condition b), but no particular parsing difficulty for the syntactic anomalous condition c); however, there are no anomalous effects in Chinese in the early stages. Interactions of anomaly by language in the later stages suggest that the parsing of the two languages in the early stages is different; besides, in Chinese the semantic anomaly is picked up earlier than the syntactic anomaly, although the significant effects occur within the same stage. In later stages, both in English and in Chinese, it is more difficult to parse condition b) and c) than the normal condition a), with the parsing of condition b) to be the most difficult. The delayed effects suggest that an extra parsing procedure of word identification is needed in Chinese.

Furthermore, it is expected that in the very early stages, there should be parsing difficulty for condition b) than a) already, if the semantic information was assigned first. It is possible to make this kind of comparison in the current study, and this will be discussed later in the report.

The data extraction was the same as the previous experiment, namely, the five chronological stage of processing (i.e., first fixation time, first pass time, regression path time, reconstruction time and total time). Moreover, rating task results were also recorded as a very late stage of processing.

7.2.2.1 Chinese Results

Firstly, the interest areas being focused on in the eye-tracking task were the PP (by + NP) and its anomalous counterparts. Three-leveled repeated measures of ANOVAs were carried out for both subjects (F_1) and material analyses (F_2) for each of the measurements, as well as the rating scores. The three levels of the within-subject factors were the normal sentences with *by* + NP, the semantic anomalous sentences with *at* + NP and the syntactic anomalous sentences without either preposition. Two post hoc contrasts were carried out as: condition b) to compare with the average of a) & c); then a) to compare with c).

- **First Fixation Time**

The mean first fixation time is listed in Table 42 below:

Table 42 Mean First Fixation Time for Chinese

	Mean	Std. Deviation	N
FFix Normal	210.2500	34.79446	18
FFix Semantic	208.6111	33.10703	18
FFix Syntactic	213.2222	28.34567	18

In this very early stage, there were no reliable main effects ($Fs<1$, $ps> .6$), nor were there any significant differences in any contrasts ($Fs<1$, $ps> .5$). At this processing stage, the parser did not spend a significant amount of time on any particular condition in Chinese.

- **First Pass Time**

The mean first pass time is listed in Table 43 below:

Table 43 Mean First Pass Time for Chinese

	Mean	Std. Deviation	N

FPass Normal	374.0833	151.32680	18
FPass Semantic	386.6667	152.37020	18
FPass Syntactic	346.8611	130.57011	18

Statistically, there were no reliable main effects of the PP manipulation (Fs<1, *p*s> .6). There were no differences from the contrasts either (Fs<1.2, *p*s> .2). This means that there were no anomaly effects in the second early stage of processing in this language.

◆ **Regression Path Time**

The mean regression path time is listed in Table 44 below:

Table 44 Mean Regression Path Time for Chinese

	Mean	Std. Deviation	N
RPath Normal	1713.8333	887.56751	18
RPath Semantic	1940.2778	1238.26073	18
RPath Syntactic	1776.2500	878.06132	18

In this stage, there were no reliable main effects (Fs<1.6, *p*s> .28), nor were there any significant differences in any contrasts (Fs<2.7, *p*s> .11). It seems at this processing stage, the parser still has not picked up the anomalies in the sentences.

◆ **Reconstruction Time**

Table 45 Mean Reconstruction Time for Chinese

	Mean	Std. Deviation	N
Reconstruction Normal	515.0556	393.37857	18
Reconstruction Semantic	1196.3889	997.96650	18
Reconstruction Syntactic	814.4444	466.58197	18

Table 45 above is the mean reconstruction time for the Chinese speakers. Statistically, there were reliable main PP effects (F_1 (2, 34) = 6.966, *p*= .013; F_2 (2, 46) = 23.408, *p*< .001)). Further contrasts showed that there were significant differences between condition b) and the average of a) & c) (F_1 (1, 17) = 6.102, *p*= .024; F_2 (1, 23) = 31.673, *p*< .001), and reliable differences between a) and c) (F_1 (1, 17) = 17.232, *p*= .001; F_2 (1, 23) = 12.453, *p*= .002). The one-tailed comparison between b) and c) has also shown significant difference (t_1 (17) = 1.9, *p*= .065 (marginal), t_2 (23) = 3.3, *p*= .003), with b) reading longer than c).

Table 45 and the statistical results replicated the results from the previous experiment. It has proven that condition b), which is the semantic anomaly, cost much longer reading time for Chinese speakers, compared to the other two conditions; whereas condition c), which is the syntactic anomaly, cost much longer reading time than condition a) (the normal condition), but much less time than condition b). The normal condition was the easiest to read so the reading time was the shortest among the three conditions.

The results also confirmed that this stage is the main parsing stage in Chinese, as the effects start to appear.

◆ **Total Time**

The mean total time is displayed in Table 46 below:

Table 46 Mean Total Time for Chinese

	Mean	Std. Deviation	N
Total Normal	889.1389	449.76644	18
Total Semantic	1583.0556	1100.25367	18
Total Syntactic	1161.3056	521.48583	18

There were reliable main PP effects (F_1 (2, 34) = 7.246, p= .01; F_2 (2, 46) = 18.609, p< .001)). Further contrasts showed that condition b) read significantly longer than the average of a) & c) (F_1 (1, 17) = 6.870, p= .018; F_2 (1, 23) = 20.356, p< .001), and reliable differences between a) and c) (F_1 (1, 17) = 10.442, p= .005; F_2 (1, 23) = 12.568, p= .002), with a) read the fastest. The one-tailed comparison between b) and c) has also shown significant difference (t_1 (17) = 2.1, p= .025, t_2 (23) = 3.1, p= .003), with b) costing longer than c).

This suggests that the PP anomalous effects were very clear at the later stages of the parsing progress for Chinese speakers.

◆ **Rating for Chinese**

The mean total time is displayed in Table 47 below:

Table 47 Mean Rating for Chinese

	Mean	Std. Deviation	N

Rating Normal	3.8681	.5448	18
Rating Semantic	2.4931	.7218	18
Rating Syntactic	2.4167	.7376	18

There were reliable main PP effects (F_1 (2, 34) = 47.611, *p*< .001; F_2 (2, 46) = 154.336, *p*< .001)). Further contrasts showed there were also significant differences between condition b) and the average of a) & c) (F_1 (1, 17) = 22.221, *p*< .001; F_2 (1, 23) = 27.874, *p*< .001). Reliable differences between a) and c) were also shown (F_1 (1, 17) = 68.497, *p*< .001; F_2 (1, 23) = 192.614, *p*< .001), with a) unsurprisingly being the most acceptable form.

Moreover, as it is the latest stage of the processing, two further contrasts were performed: there was also a reliable difference between condition a) and the average of b) & c) (F_1 (1, 17) = 66.059, *p*< .001; F_2 (1, 23) = 301.251, *p*< .001). However, there was no reliable difference between b) and c) (Fs<1, *p*s> .7). The results suggest that in the very late stages, the participants regarded the two anomalous conditions less acceptable than the normal condition in there native language.

7.2.2.2 English Results

As in Study 4, the English analyses replicated the ones in Chinese. For some special cases further contrasts were conducted and are reported below.

◆ **First Fixation Time**

The mean first fixation time is listed in Table 48 below:

Table 48 Mean First Fixation Time for English

	Mean	Std. Deviation	N
FFix Normal	178.4722	29.75215	18
FFix Semantic	203.5833	47.52963	18
FFix Syntactic	203.0000	25.12440	18

Unlike in Chinese, there were already some main effects in English in this very early stage (F_1 (2, 34) = 4.308, *p*= .042, F_2 (2, 46) = 4.596, *p*= .015). As this general effect came surprisingly early, more planned contrasts were carried out. Further contrasts show that the significant differences lay between the fixation time of condition a) and the average condition of b) and c) (F_1 (1, 17) = 8.091, p= .011, F_2 (1, 23) = 8.453, p= .008), and between condition a) and c) (F_1 (1, 17) = 18.127, p= .001, F_2 (1, 23) = 6.963, p= .015). There were also marginal differences between condition a) and b) (F_1 (1, 17) = 3.852, p= .066, F_2 (1, 23) = 6.292, p= .020).

These very early effects were surprising because normally it is assumed that in this stage the information is not fully processed, therefore it is normally unexpected to show clear effects. Furthermore, the very early differences between a) and c) seem to contradict *LAST* and the previous study. This will be further discussed and more deeply analysed in the next section.

In fact, there were differences between condition c) and the average of a) and b) by the subject analyses (F_1 (1, 17) = 7.069, p= .017), but this is not confirmed by the material analyses (F_2 (1, 23) = 2.465, p= .130). Since this result cannot be generalised to more materials, they will not be further discussed later.

◆ **First Pass Time**

The mean first pass time is listed in Table 49 below:

Table 49 Mean First Pass Time for English

	Mean	Std. Deviation	N
FPass Normal	311.9722	95.57834	18
FPass Semantic	389.9206	127.01958	18
FPass Syntactic	346.0556	139.97120	18

There were marginal main effects of the PP manipulation. (F_1 (2, 34) = 2.912, p= .068; F_2 (2, 46) = 2.568, p= .088). Further contrasts showed that there were significant differences between b) and the average of a) and c), with b) to cost longer reading time (F_1 (1, 17) = 5.008, p= .039; F_2 (1, 23) = 4.410, p= .047). The one-tailed comparison between b) and c) has also shown significant difference (t_1 (17) = 2.2, p= .02, t_2 (23) = 1.8, p= .04), with b) reading longer than c).Moreover, because of the effects in earlier stages, a contrast comparing a) and b) was also performed. There were significant differences between condition a) and b) (F_1 (1, 17) = 7.892, p= .012; F_2 (1, 23) = 6.044, *p*= .022). However, there was no difference between a) and c) (Fs < 1.05, *p*s> .3). Taking this and the fact that b) is longer than the average of a) and c), it is safe to conclude that a) cost longest reading time among the three conditions. In general, it means that the semantic effects started to appear, but the syntactic effects appearing in the earlier stage have disappeared.

◆ **Regression Path Time**

The mean regression path time is listed in Table 50 below:

Table 50 Mean Regression Path Time for English

	Mean	Std. Deviation	N
RPath Normal	1277.1944	871.65978	18
RPath Semantic	1500.3056	645.88462	18
RPath Syntactic	1368.1944	444.57717	18

The main effects that appeared in the previous measurements seem to have all disappeared in this stage (Fs < 1, *p*s > .4). There were no differences from any of the contrasts either (Fs < 2.4, *p*s > .1). This means that the participants spent relatively the same amount of time in the PP regions among the three conditions.

◆ **Reconstruction Time**

Table 51 Mean Reconstruction Time for English

	Mean	Std. Deviation	N
Reconstruction Normal	276.7222	213.60500	18
Reconstruction Semantic	701.1508	347.01269	18
Reconstruction Syntactic	569.0278	205.58156	18

Table 51 above is the mean reconstruction time for English speakers. In this stage, the main PP anomaly effects were significant again (F_1 (2, 34) = 17.428, $p< .001$; F_2 (2, 46) = 20.907, $p< .001$)). The difference between condition b) and the average of a) & c) was also reliable (F_1 (1, 17) = 14.477, $p= .001$; F_2 (1, 23) = 16.382, $p= .001$), meaning that b) cost significantly longer reading time. Furthermore, there was also a significant difference between a) and c) (F_1 (1, 17) = 23.125, $p< .001$; F_2 (1, 23) = 30.770, $p< .001$), showing a) cost shortest reading time. The comparison between b) and c) has also shown marginal difference (t_1 (17) = 1.65, $p= .058$ (marginal), t_2 (23) = 1.73, $p= .048$), with b) reading longer than c).

This means that in this stage, the syntactic anomaly effects have reappeared, and both the anomaly types have strong effects on the participantsô reading time. People spent more time on the regions in the anomalous conditions, but less in the normal conditions.

◆ **Total Time**

The mean total time is displayed in Table 52 below:

Table 52 Mean Total Time for English

	Mean	Std. Deviation	N
Total Normal	588.6944	259.61573	18
Total Semantic	1091.0714	402.90379	18
Total Syntactic	915.0833	210.16235	18

The effects in this stage are very similar to the reconstruction stage. There were reliable main PP anomaly effects (F_1 (2, 34) = 24.668, p< .001; F_2 (2, 46) = 21.929, p< .001).. There were significant differences between condition b) and the average of a) & c) (F_1 (1, 17) = 24.570, p< .001; F_2 (1, 23) = 18.994, p< .001), and between a) and c) (F_1 (1, 17) = 24.811, p< .001; F_2 (1, 23) = 18.994, p< .001). The one-tail comparison between b) and c) has also shown some significant difference (t_1 (17) = 2.4, p= .007, t_2 (23) = 1.9, p= .02), with b) reading longer than c).These results suggested that a) cost least reading time, while b) cost most reading time.

◆ **Rating for English**

The mean total time is displayed in Table 53 below:

Table 53 Mean Rating for English

	Mean	Std. Deviation	N
Rating Normal	4.4097	.3424	18
Rating Semantic	2.9365	.6262	18
Rating Syntactic	2.8889	.8345	18

The English results for rating are very similar to the Chinese ones. There were reliable main PP effects (F_1 (2, 34) = 57.746, p< .001; F_2 (2, 46) = 82.313, p< .001)). Further contrasts showed that there was a significant difference between condition a) and the average of b) and c) (F_1 (1, 17) = 80.983, p< .001; F_2 (1, 23) = 413.029, p< .001). There was also a significant difference between condition b) and the average of a) and c) (F_1 (1, 17) = 37.942, p< .001; F_2 (1, 23) = 27.874, p< .001), and a reliable difference between a) and c) (F_1 (1, 17) = 68.166, p< .001; F_2 (1, 23) = 192.614, p< .001). As in Chinese, there was no reliable difference between condition 2) and 3) (Fs< 1, ps> .7). This suggested that condition a) was the most acceptable form, while condition b) and c) were equally unacceptable.

7.2.2.3 Comparing Chinese and English

Similar to the previous analyses, repeated measures of ANOVAs were carried out for all participants on both subjects (F_1) and material analyses (F_2) for each of the measurements. The within subjects factors contained three PP (or its counterparts) conditions. The between subjects factor was the first language, where the number of each category was well balanced. Contrasts were the same as the previous analyses.

- **First Fixation Time**

The mean first fixation time is listed in Table 54 below:

Table 54 Mean First Fixation Time

	LANGUAGE	Mean	SD	N
FFix Normal	English	178.4722	29.75215	18
	Chinese	210.2500	34.79446	18
	Total	194.3611	35.74425	36
FFix Semantic	English	203.5833	47.52963	18
	Chinese	208.6111	33.10703	18
	Total	206.0972	40.44923	36
FFix Syntactic	English	203.0000	25.12440	18
	Chinese	213.2222	28.34567	18
	Total	208.1111	26.90223	36

The PP anomalous effects were not significant (Fs < 2.5, ps > .08), but there was a marginal difference between condition a) and the average of b) and c) (F_1 (1, 34) = 5.366, p= .027, F_2 (1, 46) = 3.585, p= .065 (marginal)). There was also a significant difference between condition a) and c) (F_1 (1, 34) = 6.970, p= .012, F_2 (1, 46) = 4.025, *p*= .05). This result is similar to the English first fixation time results, in that it seemed that the processor picked up the syntactic anomaly first. Nevertheless, further analyses and discussion about it will be performed in Section 7.2.3. The reading time differences between Chinese and English speakers were approaching the critical margin (F_1 (1, 34) = 3.475, *p*= .071, F_2 (1, 46) = 3.629, *p*= .063).

Moreover, the difference between condition a) and c) varied between Chinese and English from the subject analyse (F_1 (1, 34) = 4.282, *p*= .046), but it was not confirmed by the material analyses (F_2 (1, 46) = 1.786, *p* = .188). This means that this interaction cannot be generalised to new materials. Furthermore, there were no general anomalous effects, or other differences from the contrasts, or anomaly * language interactions (Fs < 3.2, *p*s > .08). These will not be discussed further.

◆ **First Pass Time**

The mean first pass time is listed in Table 55 below:

Table 55 Mean First Pass Time

	LANGUAGE	Mean	SD	N
FPass Normal	English	311.9722	95.57834	18
	Chinese	374.0833	151.32680	18
	Total	343.0278	128.65407	36
FPass Semantic	English	389.9206	127.01958	18
	Chinese	386.6667	152.37020	18
	Total	388.2937	138.26025	36
FPass Syntactic	English	346.0556	139.97120	18
	Chinese	346.8611	130.57011	18
	Total	346.4583	133.40531	36

There was marginal difference between condition b) and the average of a) and c) (F_1 (1, 34) = 4.387, *p*= .044, F_2 (1, 46) = 3.776, *p*= .058 (marginal)), with condition b) cost longest reading time. The one-tailed comparison between b) and c) has also shown marginal difference (t_1 (35) = 1.76, *p*= .04 (marginal), t_2 (47) = 1.93, *p*= .06 (marginal), with b) reading longer than c). However, the main effects of anomaly or language have all declined in this stage (Fs < 2.19, *p*s> .1). There were no other differences from contrasts, or any language * anomaly interactions (Fs < 2.8, *p*s > .09). This means the two language groups behaved very similarly.

◆ Regression Path Time

The mean regression path time is listed in Table 56 below:

Table 56 Mean Regression Path Time

	LANGUAGE	Mean	SD	N
RPath Normal	English	1277.1944	871.65978	18
	Chinese	1713.8333	887.56751	18
	Total	1495.5139	894.81871	36
RPath Semantic	English	1500.3056	645.88462	18
	Chinese	1940.2778	1238.26073	18
	Total	1720.2917	998.56957	36
RPath Syntactic	English	1368.1944	444.57717	18
	Chinese	1776.2500	878.06132	18
	Total	1572.2222	716.44905	36

In this stage, the main effects that appeared in the previous measurement were reduced to a nonsignificant level. The subject analyses show significant language differences (F_1 (1, 34) = 7.089, p= .012), but it could not be generalised to new materials (F_2 (1, 46) = 1.283, p= .263). Moreover, although there were no main anomaly effects (Fs < 1.8, ps> .1), the material analyses showed some significant differences between b) and the average of a) & c) (F2 (1, 46) = 4.433, p= .041). To sum up, there were no special effects in this stage.

◆ Reconstruction Time

Table 57 Mean Reconstruction Time

	LANGUAGE	Mean	SD	N
Reconstruction Normal	English	276.7222	213.60500	18
	Chinese	515.0556	393.37857	18
	Total	395.8889	334.56067	36
Reconstruction Semantic	English	701.1508	347.01269	18
	Chinese	1196.3889	997.96650	18
	Total	948.7698	778.00810	36
Reconstruction Syntactic	English	569.0278	205.58156	18
	Chinese	814.4444	466.58197	18
	Total	691.7361	376.50372	36

There were significant PP anomaly main effects (F_1 (2, 68) = 15.744, p< .001, F_2 (2, 80)

= 40.394, $p < .001$). There were also differences between condition b) and the average of a) and c) (F_1 (1, 34) = 12.695, $p = .001$, F_2 (1, 40) = 47.510, $p < .001$), and between a) and c) (F_1 (1, 34) = 39.535, $p < .001$, F_2 (1, 40) = 29.964, $p < .001$). The one-tailed comparison between b) and c) has also shown significant difference (t_1 (35) = 2.4, $p = .024$, t_2 (47) = 3.6, $p = .001$), with b) reading longer than c).This can be referred as a proof of longest reading time for condition b) and shortest for a).

Between the participants, language differences were also significant (F_1 (1, 34) = 6.560, $p = .015$, F_2 (1, 40) = 32.882, $p < .017$). However, it only means that the Chinese speakers read slower than the English speakers did, since there was no interaction of anomaly by language (Fs< 1.5 *p*s > .2). There reading time patterns did not vary much.

◆ **Total Time**

The mean total time is displayed in Table 58 below:

Table 58 Mean Total Time

	LANGUAGE	Mean	SD	N
Total Normal	English	588.6944	259.61573	18
	Chinese	889.1389	449.76644	18
	Total	738.9167	392.68798	36
Total Semantic	English	1091.0714	402.90379	18
	Chinese	1583.0556	1100.25367	18
	Total	1337.0635	853.85778	36
Total Syntactic	English	915.0833	210.16235	18
	Chinese	1161.3056	521.48583	18
	Total	1038.1944	411.25575	36

Similar to the reconstruction time, there were significant PP main effects (F_1 (2, 68) = 18.343, $p < .001$, F_2 (2, 92) = 37.377, $p < .001$). There were also reliable differences between b) and the average of a) and c) (F_1 (1, 34) = 16.101, $p < .001$, F_2 (1, 46) = 37.549, $p < .001$), and between condition a) and c) (F_1 (1, 34) = 31.461, $p < .001$, F_2 (1, 46) = 36.878, $p < .001$). The one-tailed comparison between b) and c) has also shown significant difference (t_1 (35) = 2.7, $p = .010$, t_2 (47) = 3.6, $p = .001$), with b) reading longer than c).The reading patterns were consistent with the previous analyses, with b) being the most difficult to read, therefore cost more reading time, whereas a) was the easiest, and cost least reading time.

Between the participants, language differences were also significant (F_1 (1, 34) = 5.151, p= .03, F_2 (1, 46) = 14.141, p< .001), but the language did not vary the PP anomaly reading patterns (no interaction) (Fs < 2.5, ps> .1).

◆ Rating

The mean total time is displayed in Table 59 below:

Table 59 Mean Rating

	LANGUAGE	Mean	SD	N
Rating Normal	English	3.8681	.54482	18
	Chinese	4.4097	.34248	18
	Total	4.1389	.52592	36
Rating Semantic	English	2.4931	.72186	18
	Chinese	2.9365	.62622	18
	Total	2.7148	.70295	36
Rating Syntactic	English	2.4167	.73764	18
	Chinese	2.8889	.83456	18
	Total	2.6528	.81235	36

The results of the mixed data are consistent with the separate data reported above. There were significant main PP anomaly effects (F_1 (2, 68) = 104.861, p< .001; F_2 (2, 92) = 210.878, p< .001)). Further contrasts show that there were significant differences between condition a) and the average of b) and c) (F_1 (1, 34) = 146.261, p< .001; F_2 (1, 46) = 702.551, p< .001). There were also significant differences between condition b) and the average of a) and c) (F_1 (1, 34) = 57.327, p< .001; F_2 (1, 46) = 77.321, p< .001), and reliable differences between a) and c) (F_1 (1, 34) = 136.573, p< .001; F_2 (1, 46) = 408.665, p< .001). There was no reliable difference between condition b) and c) (Fs< 1, ps> .4).

Furthermore, there was also between-subjects language differences (F_1 (1, 34) = 7.977, p< .001; F_2 (1, 46) = 36.880, p< .001), but there were no reliable interaction of PP effects by language (Fs<1, ps> .7). By looking at the figures in Table 18, it is safe to conclude that the Chinese group rated more strictly than the English group, however, the judgement pattern of the anomaly in their native languages did not vary.

7.2.3 Discussion

This section will not only discuss the current experiment, but also compare some of the results to the previous experiment. Firstly the complex results that were confirmed by both subject analyses and material analyses will be summed up; and then be discussed in a chronological order. A simplified results report has been presented in Table 60 below:

Table 60 Results Outline

Time Block	Language	Results
A. First Fixation Time	a. Chinese	• Nothing significant
	b. English	• Marginal PP anomaly main effects; • Shorter reading time a) and the average of b) and c); • Reliable differences between a) and c); • Marginal differences between a) and b)
	c. Mix	• Shorter reading time for a) than the average of b) and c); • Reliable differences between a) and c) • Marginal language differences, with English speakers reading faster than Chinese speakers; • No clear language * anomaly interactions
B. First Pass Time	a. Chinese	• Nothing significant
	b. English	• Marginal PP anomaly effects; • Marginal shorter reading time for a) than the average of b) and c); • Significant longer reading time for condition b) than the average of a) and c); • Significant longer reading time for b) > a) and b)>c) • No reliable difference between a) and c)
	c. Mix	• Significant longer reading time for condition b); • No difference between a) and c); • Marginal longer for b) than c); • No language effects; • No language * anomaly interactions;
C. Regression Path Time	a. Chinese	• Nothing significant
	b. English	• Nothing significant

	c. Mix	• Nothing significant
D. Reconstruction Time	a. Chinese	• PP anomaly effects; • Significant longer reading time for condition b)>c)> a) (the differences between b) and c) are marginally significant)
	b. English	• PP anomaly effects; • Significant longer reading time for condition b) and c)> a), but the differences between b) and c) are not significant
	c. Mix	• PP anomaly effects; • Significant longer reading time for condition b)>c)> a); • Language difference; • No anomaly * language interaction
E. Total Time	a. Chinese	• PP anomaly effects; • Significant longer reading time for condition b)>c)>a) (The differences between b) and c) are marginally significant)
	b. English	• PP anomaly effects; • Significant longer reading time for condition b)>c)> a) (The differences between b) and c) are marginally significant)
	c. Mix	• PP anomaly effects; • Significant longer reading time for condition b)>c)> a); • Language difference; • No anomaly by language interaction
F. Rating[64]	a. Chinese	• PP anomaly effects; • Significant lower naturalness scores for condition b) and c)> a), but no differences between b) and c)
	b. English	• PP anomaly effects; • Significant lower naturalness scores for condition b) and c)> a) , but no differences between b) and c)
	c. Mix	• PP anomaly effects; • Significant naturalness scores for condition b) and c)< a) , but no differences between b) and c); • Language difference; • No interaction

◆ **Stage A**

In Study 4, in the very early stage A, the only effect was language difference. However, in the current experiment, the statistics have shown significant anomaly effects in English and the mixed data[65]. More worrying from the *LAST* point of view, the

[64] Although rating is not technically part of the online time blocks, it is considered an even later stage here.
[65] Since the Chinese results do not show much effect, and it do not differ from the previous experiment, only English

significant differences lie in the difference between condition a) and c), which are the normal condition and the syntactic anomalous condition. Can this invalidate *LAST*? The answer is ñnoò.

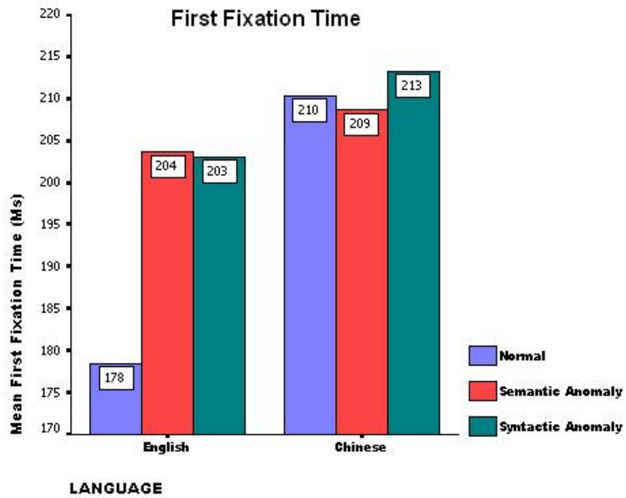


Figure 32 First Fixation Time

Take a close look at the data: it is very clear that the reading time for condition b) in English is 1ms longer from condition c). How can the difference between a) and c) be statistically more reliable than the difference between a) and b)? The only possible explanation for this is that there is more variance in b) than c), as shown in Figure 33 below:

and the mixed data will be focused here.

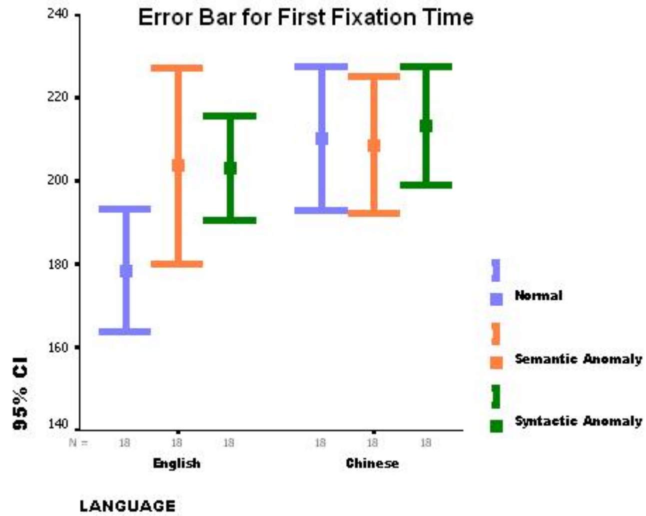


Figure 33 Error Bar for First Fixation Time

Statistically, for the subject analyses, the mean squared error is 597.396 for the contrast of a) and c), but it is 2946.458 for the contrast of a) and b); for the material analyses, the mean squared error is 1912.528 for the contrast of a) and c), but it is 2186.895 for the contrast of a) and b).

Although it is surprising to find effects in such an early stage, the variation did not occur accidentally. The explanation for this would be the stimuli change in the material design.

In the previous experiment, the critical region is PP in the normal and the semantic anomalous conditions, and NP without the preposition in the syntactic anomalous condition. The form of the critical region in condition a) and b) is *by* + NP[66], whereas the form in condition c) is NP. Yet, this NP carries two missions: being a part of an impaired PP, without a head of the phrase; and a comprehensive NP. Thus, all the anomalies were undertaken by the form of an NP in the design.

In the current experiment, the semantic anomaly occurs in the preposition *at*, so the NP following *at* is actually in a post-anomalous location. However, the syntactic condition is NP without a preposition, so the NP in this condition is both an anomalous and a

[66] The preposition *by* was kept in the region because the NPꞌs counterpart in condition 3) is actually a PP, rather than just an NP. The NP is condition 3) is not merely an NP, but an impaired PP.

ñpost-anomalousò region. It is anomalous because this region should have been a PP region; it is ñpost-anomalousò because the NP itself has no violation to the syntactic rule, just like an NP in other conditions. Since previous studies have shown evidence that people may skip function words, it is possible some of the first fixations were not landing on the word *at,* but the NP in the region. As the NP in condition 3) carries more information than the other two conditions, the anomaly form is no longer a simple NP across the conditions. Hence, the comparison involving NP is actually unfair in the measurement of the first fixation time. This problem occurs mostly in this stage because for the later measurements, movements within the full region are allowed and counted, whereas in the first fixation stage, only one fixation is counted, so the very landing point is essential.

Thus, further analyses were performed to compare the first fixation time on the preposition between conditions a) and b) only. Syntax-first theories would predict no differences between a) and b), because the syntactic (not semantic) features of the preposition of a) and b) are identical. If there were differences between a) and b), it means the semantic parsing is occurring. As for the parsing for c), firstly, it is admittedly very difficult to measure a fixation in a ñnonexistentò region (because the anomaly is embedded within the NP). Secondly, it will not be focused on here because *LAST* can be used to explain the parsing of syntactic anomaly as well, which is the frequency of the meaning boundary violation. (Please refer to the general discussion section in the previous experimental report.)

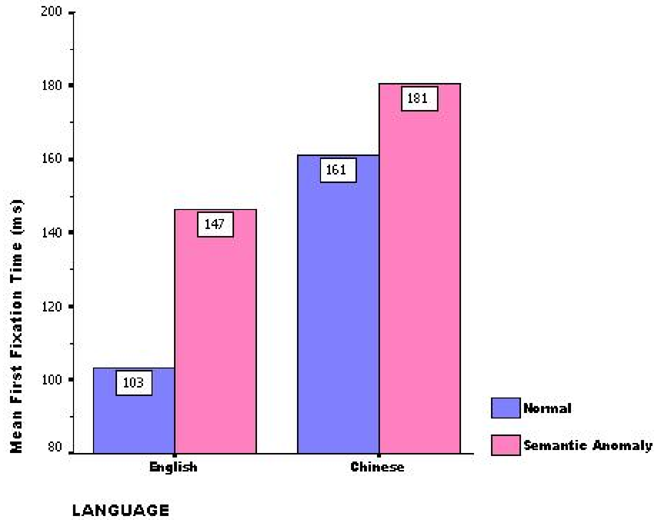


Figure 34 First Fixation Time on the Preposition

The analyses were done in English, Chinese and with mixed data. For the English data, a paired Sample T-test showed reliable differences between a) and b) on the two-letter preposition (t_1 (17) =2.32, p = .033, t_2 (23) =2.95 p = .007). The effects were also reliable for the mixed data (F_1 (1, 34) = 6.351, p = .017, F_2 (1, 46) = 9.890, p = .003). However, this effects was not reliable for Chinese (ts<1.4). There was no language * anomaly interaction either.

Although it was not necessary, the comparisons on this region were further performed for all the other chronological stages. The effects were very similar to the analyses performed for the full PP region: Significantly longer reading time was found for *at* for English and the mixed data in the first pass time; no significant results for the regression path time for English, Chinese or mixed data were found; but significantly longer reading time for the semantic anomalous *at* region than the normal *by* region for reconstruction time and total time for both languages and the mixed data (Fs > 15, ts > 2.15, *p*s < .046). It is a solid fact that the semantic anomaly is picked up by just one fixation at the material in this very early stage.

There were also language differences in most of the stages with Chinese speakers reading significantly slower than English speakers (Fs >7, *p*s< .004).

The Chinese speakers read constantly slower than the English speakers did in both experiments, regardless of the region or the stages. A possible explanation suggested in the previous experiment was an extra stage of word identification in Chinese. Nevertheless, since there were no language*anomaly interaction in the current experiment, it is hard to conclude that the Chinese speakers use a completely new strategy in processing that differs from the English speakersô behaviour. On the other hand, since *LAST* can be used to explain the word segmentation procedure, one can argue that the English and Chinese speakers actually share the same kind of processing strategy that can be derived from *LAST*. At this moment, it is safe to conclude that in the first fixation stage, the differences between conditions a) and b) have falsified the predictions from the syntax-first theories. It is quite clear that in the very early stage, little semantic violations can be picked up by the processor. The semantic processing did not occur later than the syntactic processing.

◆ **Stage B**

In the relatively early Stage B, the Chinese reading time remained insignificantly affected by the PP types. However, it is not the case for the reading time of the English and mixed groups. The mixed group showed a similar effect to that of the English one, because the pattern of the Chinese data is the similar to the English one (with no language differences or interactions). Figure 35 below sketches out the pattern of the reading time:

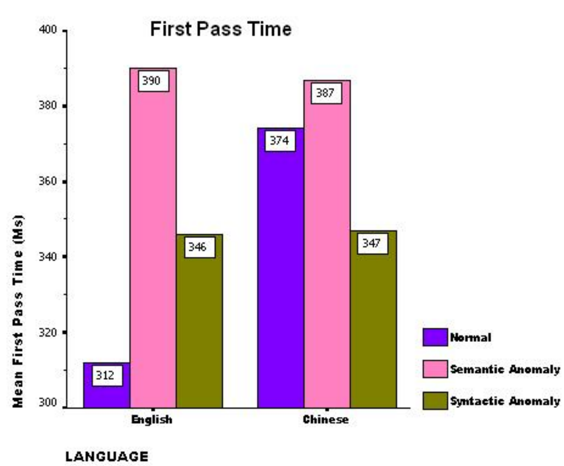


Figure 35 First Pass Time

The figure above shows that in both languages, especially English semantic anomaly cost the longest reading time. Since the difference between condition a) and c) was not significant, one can conclude that in this stage the mysterious syntactic anomaly effects in the first fixation stage have disappeared. Noticeably, the eye fixation time in a region in this stage is no longer just one glance. This also means small forwards or backwards reading time within the region was allowed. This measurement reflects the behaviour of processing when the information from the full region is first viewed, if one makes an assumption of immediacy. Even though in condition c) there is one two-letter word less than the other two conditions, as was argued earlier, the information that the NP carries is equal to the PP.

This, and the results from the preposition analyses in this stage, further suggest that the syntactic effects found in the first fixation time occurs may just be the result of a messy treatment of the region. When the region was first fully exposed to the parser, the parsing strategy in the syntactically anomalous condition was not much different from the one in a normal condition. This is also consistent with the results from Study 4.

◆ **Stage C**

It is consistent with the previous experiment that during the regression path time, the effects from the main stimuli and the contrasts all reduced for both languages. Figure 36 below shows the reading time for this stage:

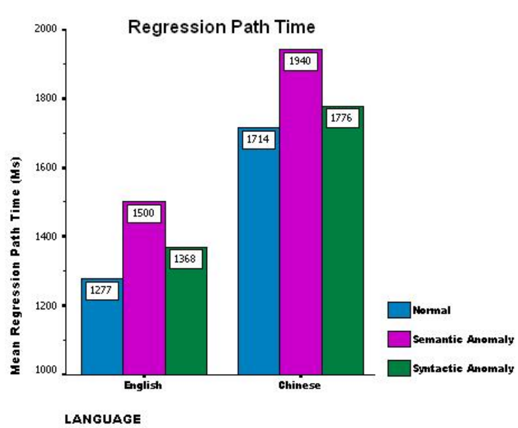


Figure 36 Regression Path Time

Noticeably, the pattern of reading time is the same as the one in the first pass time, that condition 2) cost the longest reading time. Despite that, more regressive movements in all the conditions cancelled out some of the effects caused by condition 2). Even though it was not significant, the differences between 3) and 1) grew, especially in Chinese. This stage, in both experiments, has shown a tendency for a realisation of the syntactic anomaly.

◆ **Stage D**

This is a relatively later processing stage. It is also the first time when significant effects were shown in Chinese. Figure 6 below is an illustration of the two languages at this stage:

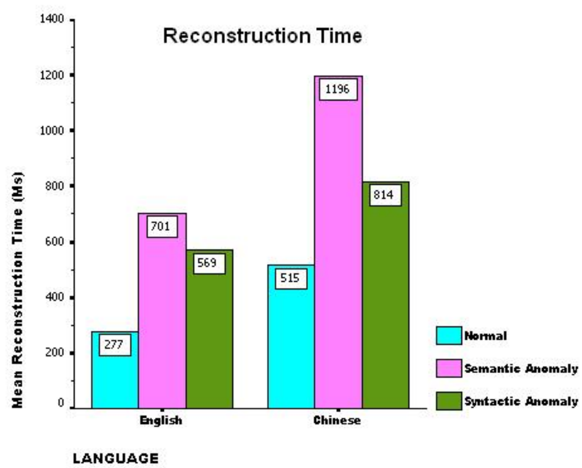


Figure 37 Reconstruction Time

The two languages behave very similarly, according to this figure: The reading time pattern (from longest to shortest reading time: condition b) > c) > a)) suggests that condition b) is the most difficult to process, while condition 1) is the easiest. Notable are the differences between b) and c), although they were not statistically significant. This means that the parsing difficulty of the syntactic anomaly is increasing similarly in English, but the semantic anomaly still attracts more attention than the syntactic anomaly in Chinese.

There were also some language differences, as Chinese readers spent a longer time reading. However, no interaction means that the parsing pattern is not different from the two languages. Recall that in the previous experiment, there was a significant interaction of the difference between a) and b) by language, which suggests a sudden rouse of semantic anomaly recognition in Chinese, and the sudden rouse in Chinese was interpreted as the result of cumulative processing of the stimuli of semantic anomaly. The disappearing of the interaction in the current experiment can be explained by the reduced stimuli power of the semantic anomaly. Because the semantic anomalous region was an NP of at least two words in the previous experiment, but a preposition of only two letters in the current experiment, it is reasonable to require less processing time on a much smaller region.

To sum up, as the semantic anomaly effects appeared in English in earlier stages, the syntactic anomaly was noticed and resulted in longer reading time than the normal one in this reconstruction time. On the other hand, as no anomaly was effective in Chinese in any of the early stages, both semantic and syntactic anomaly started to show effects in the second parsing stage. Longer reading time in the semantic anomalous condition suggests that semantic anomaly is more difficult to parse than the syntactic anomaly.

It is arguable that although the Chinese data has shown that there was more difficulty in processing the semantic anomaly, there was no clear evidence from the current experiment that the semantic anomaly was processed earlier than the syntactic anomaly.

To answer this, a close look at the data is needed. Table 61 below is a summary of the mean reading time for the two languages:

Table 61 Mean Reading Time Summary

	LANGUAGE	Mean	Std. Deviation	N
FFIX_NO	English	178.4722	29.75215	18
	Chinese	210.2500	34.79446	18
	Total	194.3611	35.74425	36
FPASS_NO	English	311.9722	95.57834	18
	Chinese	374.0833	151.32680	18
	Total	343.0278	128.65407	36
RPATH_NO	English	1277.194	871.65978	18
	Chinese	1713.833	887.56751	18
	Total	1495.514	894.81871	36
RECON_NO	English	276.7222	213.60500	18
	Chinese	515.0556	393.37857	18
	Total	395.8889	334.56067	36
TOTAL_NO	English	588.6944	259.61573	18
	Chinese	889.1389	449.76644	18
	Total	738.9167	392.68798	36
FFIX_SE	English	203.5833	47.52963	18
	Chinese	208.6111	33.10703	18
	Total	206.0972	40.44923	36
FPASS_SE	English	389.9206	127.01958	18
	Chinese	386.6667	152.37020	18
	Total	388.2937	138.26025	36
RPATH_SE	English	1500.306	645.88462	18
	Chinese	1940.278	1238.26073	18
	Total	1720.292	998.56957	36
RECON_SE	English	701.1508	347.01269	18
	Chinese	1196.389	997.96650	18
	Total	948.7698	778.00810	36
TOTAL_SE	English	1091.071	402.90379	18
	Chinese	1583.056	1100.25367	18
	Total	1337.063	853.85778	36
FFIX_SY	English	203.0000	25.12440	18
	Chinese	213.2222	28.34567	18
	Total	208.1111	26.90223	36
FPASS_SY	English	346.0556	139.97120	18
	Chinese	346.8611	130.57011	18
	Total	346.4583	133.40531	36
RPATH_SY	English	1368.194	444.57717	18
	Chinese	1776.250	878.06132	18
	Total	1572.222	716.44905	36
RECON_SY	English	569.0278	205.58156	18
	Chinese	814.4444	466.58197	18
	Total	691.7361	376.50372	36
TOTAL_SY	English	915.0833	210.16235	18
	Chinese	1161.306	521.48583	18
	Total	1038.194	411.25575	36

The table shows that the difference between semantic anomaly and the normal condition in the English data at stage B is about 78 ms, whereas the difference between semantic anomaly and the normal condition in Chinese at stage C is around 127 ms. However, the standard deviation for the Chinese mean reading time, in the semantic anomalous condition is much higher than English from the regression path time to the total time stages. The SDs in these three stages for condition b) is also higher than the other stages and conditions within Chinese.

This suggests that Chinese processors could not agree in the semantic anomalous condition. When regression happens, re-processing of the material starts. In Chinese, this means that a possible re-segmentation of words is needed again. In the current experiment, the situation is more complicated than the previous one, because the PP is a thematic assigner and the meaning of the preposition relates to the thematic roles of the NP in it. If the preposition is semantically problematic and needs revising, the meaning and the thematic role of the following NP should satisfy the needs to be re-assigned as well. This problem did not occur in the previous experiment, because the preposition *by* was used properly; the thematic role reassignment therefore was not needed. Thus, two possible strategies can be applied in the current study: to move the eyes to the next region to the right side of the PP, and try if it is possible to connect the next character to the just processed NP, or to move the eyes back to a previous region, to re-identify the word(s) the parser just parsed. The former means the regression path time is equal to the first pass time, whereas the latter can last as long as the parser is satisfied with its comprehension.

Because there is no need for identifying a word, the processing for English material is much simpler, comparing to Chinese. Less variation means that it is easier for the data to reach significance. Nonetheless, the semantic anomaly effects did not appear later than the syntactic anomalous ones, which means that semantics were not assigned later in the Chinese data either. The word identification is the hypothesized reason for no earlier signs of semantic effects in Chinese. Moreover, since early assignment of syntax theories cannot explain the word segmentation stage, and *LAST* can perfectly explain this phenomenon, we can conclude that the current experiment did not violate the *LAST*.

- **Stage E**

This is the final stage in the eye-tracking task. Figure 38 illustrates the parsing pattern:

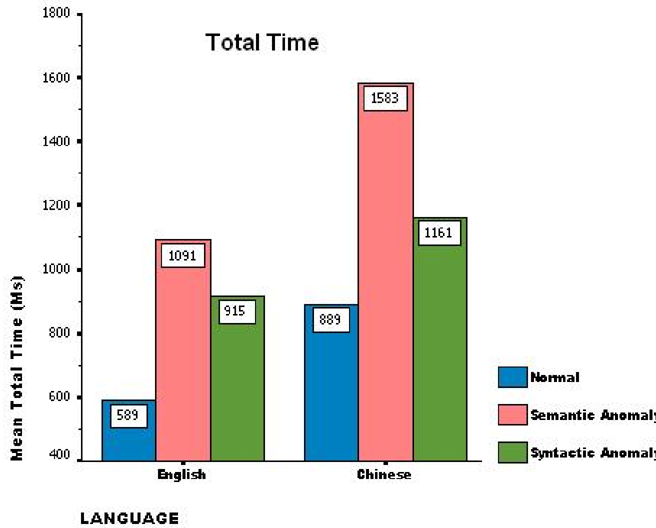


Figure 38 Total Time

The pattern and effects are identical to Stage D discussed above, but it is not exactly the same as in Study 4. Unlike in Study 4, there were no interactions of anomaly by language in the current study.

From Table 61 above, it is clear that the variation of processing time for the semantic condition is very high in Chinese. As discussed earlier, this high standard deviation starts from stage C to the current Stage E. For a similar reason as was given in the previous discussion, any stages that allow a regression would create a parsing stage of re-identification of words in Chinese. If the processor decided to re-identification of a case assigner, it could consequently lead to a reassignment from a single word to pretty much the whole item.

In a word, the current study did not show a very strong early semantic parsing pattern in Chinese, because the design of the material requires the word-identification to be reused to a more complex level, such as reassigning thematic roles. On the other hand, the semantic anomaly remains the most difficult processing condition. *LAST* can successfully explain these phenomena, whereas the early assignments of syntax theories, which predict no early semantic anomaly effects, have been challenged again.

- **Stage F**

The rating results are also slightly different from the previous experiment. Figure 39 below demonstrates the pattern**:**

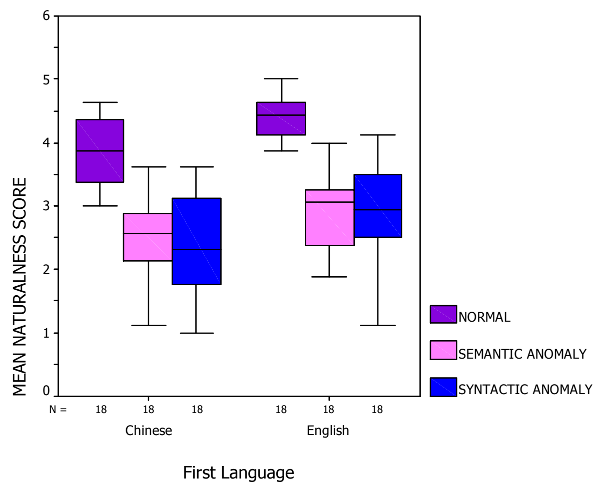


Figure 39 Mean Naturalness Rating Score

Naturally, the normal condition gets the highest scores in both languages, but unlike the previous study, the semantic anomalous condition does not receive the lowest scores. No differences between condition b) and c) meaning in this very late stage, when the participants have read all the materials, and can freely rethink about the sentences, the two anomalous conditions were judged as equally unnatural.

Further analyses were performed to compare the two studies on the rating scores. Table 62 below is a summery of the rating scores in both studies:

Table 62 Mean Rating Scores

	FIRST_LA	STUDY	Mean	Std. Deviation	N
NORMAL	Chinese	Study 1	4.1528	.39191	18
		Study 2	3.8681	.54482	18
		Total	4.0104	.48951	36
	English	Study 1	4.2014	.37656	18
		Study 2	4.4097	.34248	18
		Total	4.3056	.37014	36
	Total	Study 1	4.1771	.37959	36
		Study 2	4.1389	.52592	36
		Total	4.1580	.45579	72
SEMANTIC	Chinese	Study 1	2.0764	.84141	18
		Study 2	2.4931	.72186	18
		Total	2.2847	.80101	36
	English	Study 1	1.8194	.57876	18
		Study 2	2.9365	.62622	18
		Total	2.3780	.82100	36
	Total	Study 1	1.9479	.72357	36
		Study 2	2.7148	.70295	36
		Total	2.3313	.80670	72
SYNTACTI	Chinese	Study 1	3.1736	.85442	18
		Study 2	2.4167	.73764	18
		Total	2.7951	.87533	36
	English	Study 1	2.8750	.82916	18
		Study 2	2.8889	.83456	18
		Total	2.8819	.81992	36
	Total	Study 1	3.0243	.84347	36
		Study 2	2.6528	.81235	36
		Total	2.8385	.84322	72

The general patterns of results are consistent with Study 4, and will not be repeated here. Summing up, the scores for condition b) are significantly lower than c), with a) being the highest (Fs > 30, *p*s < .001). However, it is conspicuous that there was significant anomaly * study interaction between condition b) & c), and between a) & b) (Fs > 50, *p*s < .001). The interaction of anomaly * study between condition a) and c) was marginal (F_1 (1, 68) = 3.516, *p*= .065, F_2 (1, 92) = 9.113, *p*= .003). Moreover, there was a language * study interaction (F_1 (1, 68) = 7.125, *p*= .009, F_2 (1, 92) =28.186, *p*< .001).

Similar comparison in other stages was performed. The study * anomaly, or study * language, or study * anomaly * language interactions were either not able to be generalised to new participants, to new materials or not significant. The results will not be reported in detail here.

This suggests that the change of the material has created a clear augment on the participantsô judgement on the acceptance of the second condition. This can be seen as proof that the stimuli were balanced. It has also reduced the participantsôacceptances on the third condition in Chinese. The reason is not certain, but it seems that the Chinese participantsô scores are more affected by other materials in the same questionnaire, i.e., they tend to consider the acceptance in comparison to the other sentences. It is not clear if this is because when the semantic anomalous NP attracted more attention, they tended to pay more attention to the NPs in other conditions, but when the semantic anomaly focused on the preposition, they tended to notice the lack of preposition more. This is beyond the scope of the current study. However, one conclusion can be made: the change of the semantic anomaly affected in this much later stages, but not in early stages.

7.2.4 Conclusion for Tip 2

To sum up, the current study mostly reproduced effects seen in the previous studies: there are clear early semantic effects in English, and there is clearly more difficulty in processing the semantic anomalous condition in both languages. Although the semantic effects in Chinese do not seem to be as strong as in the previous experiment, this discrepancy can be explained by a proposal that here is a word identification procedure in this language, which is a procedure that can be seen as derived from and explained by *LAST*. Moreover, the existence of the semantic effects in the very early processing stages (with only one fixation) in English, and the lack of evidence of earlier syntactic effects in either language, have challenged the first assignment of syntax theories again. Furthermore, the lack of an early interaction of language * anomaly, suggest that the parsing strategies derived from *LAST* could be universal.

Despite the stimulus size criticism, the *LAST* has not been falsified by the two experiments reported in Chapter 7. Despite the strong explanation to the current data, it is perceptible that one strong assumption needs to be proven solid, which is the processing stage of the proposed word-segmentation in Chinese. This type of stage also exists in acoustic data in languages like English. Nevertheless, it has exceeded the capacity of the current experiments. This topic can be further discussed with the support or disproof of the suggested experiments.

Part III

At LAST

Chapter 8
Recapitulation and Discussion

Contemporary psycholinguistic theories are based on the development of linguistic theories, mainly the Universal Grammar (UG) developed by the Chomskyan School, and the development of cognitive science, where ñmentalò movements (processing of input information) can be measured and analysed. However, they differ from those two fields in the fact that the description of a language grammar cannot be used directly to explain mental processing behaviour, and the outcomes of the measurement using either offline or online tools have to be explained based on the linguistic theories. Without integrating the two fields, linguistic theories cannot explain why the language form is as it describes, whereas the cognitive science of language can be misleading in terms of what mental processing it actually involves. If the psycholinguistic study of the human sentence processing mechanism were to be a hidden iceberg in the sea, reviews of previous studies provided in the current thesis, together with some empirical data has shed light on some tips of this iceberg.

8.1 Theoretical Discussion

The second chapter in this thesis reviewed the most accepted and applied linguistic theory group: the UG. Based on its name, it is not difficult to conjecture that the key points of the theory are universality and grammaticality. UG claims that the universality characteristic is based on the innateness of language of the human race, as the biological structure of a human body provides competent organs for the purpose of acquiring and using language; moreover, a child is born with a mental language faculty, equipped with the competence of the ñdefault grammarò of human language. This ñdefaultò human language further develops into variations of languages in the world by habits and rules of each certain language, and it is the innate competence of language that allows a child to acquire different forms of human language. Thus, language is closely linked to the human mind in an objective way. Although it is not certain what the ñdefault grammarò is, Chomskyan scholars believe that syntactic rules transform the default form into different languages. In searching for these ñdefaultò universal rules, the Chomskyan School does not focus on the rules (grammar) in individual languages, but rather the general rules (i.e., the syntax) that can explain phenomena across all languages. According to this school, sentences are the object of language study and syntax is the skeleton of the sentences. Only by analysing the syntax of a sentence, can the meaning of a sentence be understood.

Based on Chomskyan Schoolôs theories, a group of psycholinguists started searching for the mental movements of sentence processing. Currently, it is generally agreed that sentence processing involves processing of all information available, but what is really subject to dispute is the timecourse of processing the information. The dominant view that syntax processing occurs prior to semantic processing has been established for nearly 30 years. However, the evidence has never been consistent. Some evidence (e.g., Frazier & Clifton, 1996; Frazier & Fodor, 1978; Frazier & Rayner, 1982; Gibson & Hickok, 1993; Hemforth et al., 1998; Konieczny et al., 1995; Pickering & van Gompel, 2006; Rayner et al., 1983; Rayner & Duffy, 1988) supports the syntax-first theory, while other observations show that the processing of other cues (frequency, context and prosody etc.) can be competitive with the syntactic cue (e.g., Green & Mitchell, 2006; MacDonald et al., 1994; McRae et al., 2005; McRae et al., 1998) or even prior to syntax (e.g., Snedeker & Trueswell, 2003; Su, 2004; Townsend & Bever, 2001).

Linguistically, although Chomsky (1958, 1965) has tried to focus on the syntax movement and make a relatively clear distinction between syntax and semantics, it has been admitted that the border between syntax and semantics is hazy. During the development of the UG theory, there has been trend towards accentuating the role of semantics. This improvement considers a key purpose of language: to communicate, i.e., the motivation of producing a sentence is to convey meaning, and the motivation of processing one is to transfer the meaning. Pure syntactic processing without a semantic component does not logically exist on its own. In the most current version of this theory, the minimalist programme (MP), as there is no longer a distinction between D- and S-structures, the relationship between syntax and semantics has then integrated into the same phase. According to MP, the movement of a sentence structure is based on necessity because of the limited capacity of human memory; consequently, the analysis of structure cannot be taken up until finishing processing the input of all sentential components, otherwise many unnecessary movements could be assigned along with the input. These re-analyses and recovery of the structure assignment will overload the working memory with repeated, old information, as well as new incoming information ï possibly exceeding the capacity of the human mindôs resources. If this is the case, syntax analyses in a real sense should only start at a much later phase of processing. So far, the only psycholinguistic theory clearly addressing this issue is the Late Assignment of Syntax Theory (LAST) (Townsend & Bever, 2001), where sentence processing takes place before the real syntax analyses is considered ñpseudosyntax processingò, and likely meanings can be assigned differently prior to the real syntax analyses.

Other psycholinguistic theories, on the other hand, are either based on an older version of the UG principles (e.g., the Garden Path Theory Groups), claiming syntax should be analysed at a very early stage of sentence processing, or not directly based on the UG (e.g., statistical models), claiming that there is no good reason to assume syntax is processed in a different time course from other information.

Taking the Garden Path Theory Group (Frazier, 1978; Frazier & Clifton, 1996) as an example for the syntax-first model, the basic UG principle it follows is locality. However, consider sentence 26c), re-presented here as 91)

91) Crosslinguistic difference (sentences used by Cuetos & Mitchell, 1988):

i. The journalist interviewed *the daughter* of *the colonel* who had had the accident. (English – *the colonel* had the accident)

ii. El periodista entrevisto a *la hija* del *coronel* que tuvo el accidente. (Spanish – *la hija* (the daughter) had the accident)

The difference between the Spanish and English preference, as well as the evidence that prosodic manipulation can directly change the attachment preference, not only challenged the locality law, but it also challenged the universality character of this principle.

Taking the integration competition model (McRae et al., 1998) as an example, McRae et al. tested a syntax-first model and a syntax-simultaneous model. The syntax first model predicted a bad fit with the human data, whereas the simultaneous model predicted a much better fit. Unfortunately, McRae et al. did not find a ñgood reasonò to set a parameter for delaying the syntax processing. If the minimalist programme was considered, a good reason should be available.

As is emphasised in the current thesis, a good psycholinguistic model should not neglect a theoretical or conceptual background of linguistics. Moreover, any debate in the psycholinguistic field should be based on agreed linguistic and psychological concepts. Because the most accepted and developed linguistic theory is the UG, and because many psycholinguistic theories distinguish syntactic processing and semantic processing as defined in certain versions of UG, the concepts of the terminology in those psycholinguistic theories also need to follow the UG rationale and definition. Unfortunately, these concepts were not seriously followed in the field of psycholinguistics. More focus has been placed upon the psychological (i.e., cognitive, perceptual, and even neurological) explanation of the behavioural data (i.e., sentence producing and processing). The current paper attempts to investigate psycholinguistic phenomena, strictly following the definition of the MP syntax theory, providing Chinese data to test the universality character of the UG, as well as some psycholinguistic theories reviewed in Chapters 3, 5 and 6. Moreover, direct comparisons of the timecourse of processing syntax and other information have also been carried out.

8.2 Empirical Discussion

◆ **Tip 1: Crosslinguistic Data of RC Attachment Preference**

Study 1 provided Chinese data and Studies 2-3 provided English and Chinese data on the relative clause (RC) attachment ambiguity resolution (as exemplified in 91) above). A clear NP-low attachment preference was found in the three experiments for both languages.

Study 1 used a self-paced reading task with sentences in Chinese displayed segment by segment, and found that sentences with NP-high attachment took longer to read on the second NP site. The longer reading time shows that the NP-high attachment was not as strongly preferred as NP-low attachment, so a longer processing time was required. However, the reading difficulty was expected to be found on the first NP site, according to the Chinese sentence structure. The fact that the effects were shown on the second NP site in this study suggests there is a delay.

Study 2 used a web-based questionnaire in both English and Chinese. Results from multiple choices among NP-high, NP-low and fillers also confirmed an NP-low bias in both languages. However, the English PP effects found in other studies (e.g, Gilboy et al. (1995)) were not confirmed.

Study 3[67] used an eye-tracking methodology, and found that in Chinese, NP-high attachment resulted in the longest total reading-time (which was at a later stage of processing). This is consistent with the findings from Study 1 that the NP-low attachment was preferred but the effects were delayed (i.e., the processing difficulty was not shown immediately when the crucial site was first shown). However, the attachment bias was not clear for the English data at any stages of processing.

No language and preference interaction was detected from studies 2 & 3. Thus, the discussion here will focus on the Chinese results only.

The NP-low attachment preference in Chinese does not violate the prediction from most of the psycholinguistic theories (see Chapter 6). The special hybrid Chinese RC sentence feature (i.e., SVO in general, but SOV for RCs) makes it more favourable for

[67] This study concurrently tested Hemforth et al.ô*s* *topicalisation* theory. However, as the English data did not replicate previous studies, the topic remains open and will no longer be focused on in this part.

NP-low attachment. NP-low is the closest possible head of the complex NP linearly to both the main verb and the relative clause, and it is accessed earlier than NP-high. In this sense, there should be no conflict in assigning NP-low as the head of the complex NP between the requirement from the main verb assertion and the syntactic ñlocality lawò. However, this has cast doubt on the ñlocality lawò because, according to the syntactic structure, NP-high is still the closest possible head to the main verb assertion of a sentence (see Figures 17 & 18, pp.105-106). In this sense, the conflict between verb and locality should still exist. If a mental parser were to be the sole mechanism for sentence processing, the semantic mismatch should not cause processing difficulty on either NP site initially, as it is not syntactically wrong at this time. Study 1 presented sentences segment by segment with no regression allowed. Assuming semantic processing takes place slightly later than syntactic processing and considering spillover effects, one would expect reading difficulty in the second NP site (NP-high) whenever the attachment was unambiguous (regardless of whether NP-low is the preferred site), simply because NP-high would be the only site to solve any mismatch between the previous and current segments. This certainly does not match the current results from Study 1. In the case of Study 3, where free eye movement was allowed, if semantic processing was slightly later and NP-low was the ñdefaultò attachment, then when the RC was forced to attach to NP-high, there would be reading difficulty in the later stage of reading on the NP-low site. This matches the results from the current thesis. As a result, the results from Tip 1 cannot be satisfactorily explained by applying linguistic rules strictly to the concept of ñsyntax structureò, but applying the linear structure of the Chinese language instead of the syntactic structure, the results invalidate syntax rules. The clear NP-low attachment seems to highlight a paradoxical black hole in the syntax-based theories.

Because the debate above is not conclusive, the delayed effects remain mysterious: they were not simply caused by methodology because delayed effects were constant in Studies 1 and 3, where different tasks were used; nor could they be caused by the proposal that semantics are processed later than syntax because that suggestion does not match the results from Study 1. One proposal is that semantic processing is not delayed in comparison to syntactic processing, and the delay is caused by other factors. The subsequent studies focused on the timecourse of semantic and syntactic processing, because if this issue cannot be clarified, the current data remain unexplainable.

◆ **Tip 2: Timecourse of Semantic and Syntactic Processing**

Studies 4 and 5 used an eye tracking technique, and directly manipulated semantically anomalous and syntactically anomalous sentences to compare with each other as well as normal sentences in English and Chinese. The results of the two experiments were consistent: the measured eye movements showed a very early processing difficulty with semantically anomalous sentences, but slightly later effects (or, at least no earlier effects from Study 5) with the syntactic anomaly in English, whereas the Chinese data showed effects of anomaly in later stages, with semantic anomaly effects appearing slightly earlier than syntactic anomaly. Moreover, the mean reading time was longer in Chinese than in English in both studies.

These two studies show explicit late syntactic anomaly effects in both languages. This suggests that the delayed effects in Studies 1 and 3 could not have originated from the semantic manipulation. As both Study 4 and Study 5 have shown postponed effects in Chinese sentences compared to their English counterparts, it seems that Chinese does not have immediate effects in general (whether in connection with RC attachment ambiguity or with syntactic/semantic anomaly). Thus, it is proposed that there may be extra information in Chinese to process (but not in English). One possible piece of information is the word boundary: as the Chinese written form does not have spaces to separate words or phrases, it is impossible to continue further processing without first clarifying the boundary. The boundary has to be determined by the readers based on frequencies of combination and meaning. Moreover, the decision on a boundary always requires a couple of further characters to be entered, or the right border of the boundary cannot be determined as there is always a possibility to combine the new incoming character within the current constituent. Although testing this proposal is not within the scope of the current thesis, it suggests that syntactic structure cannot be assigned before certain information is available.

In summary, Tip 2 comprises results supporting the Late Assignment of Syntax Theory (LAST) (Townsend & Bever, 2001) but against syntax first theories. However, LAST has rarely been accepted in the field. The next section will focus on this theory.

8.3 Discussion

The data provided in the current thesis supports the Late Assignment of Syntax Theory, or LAST (Townsend & Bever, 2001), as it provides an explanation and rationale (i.e., the MP theory) for a possible later analysis of syntax in sentence processing. Inevitably, a question will be asked about why there is so much evidence supporting the early syntax processing theories. To answer this question, two points should be considered: the methodologies used to provide the ñevidenceò and the explicability of *LAST*.

8.3.1 Methodology

Psycholinguistic experiments are normally conducted in well controlled environments. The advantages and disadvantages of specific testing methods have been reviewed in Chapter 4 and will not be repeated here. In this section, some problems in the material design and the sensitivity of the task measurements will be discussed.

In order to eradicate noise from the data, the materials are normally ñwell controlledò so that the most explicit stimuli can be presented to the participants. This is in some way necessary in terms of serving the purpose of research. However, some experimental designs exaggerate the stimuli so that the effects from the data are actually guided by the design, instead of being a true representation of the process. Sometimes information is lost so that there is no possibility of measuring processing of non-syntax information; in other cases, the stimuli are mixed so that it may not be clear whether the effects really originate from one type of information processing or another.

◆ **Guided design**

A good example of the unnaturally guided design would be sentence 65a) (and reprinted in 92a) below.

92) a. While Mary was mending the sock fell off her lap. (Frasier, 1978)

b. While Mary was mending, the sock fell off her lap.

92a) is a well known Garden-path sentence. When the readers read the NP *the sock*, it will normally be interpreted as the object of the verb *mending*. However, in real literature, a comma can be inserted between the verb and the NP as in 92b) so that the misinterpretation can be minimised and sometimes eliminated. Thus, it is questionable whether the early misinterpretation is directly caused by the structure or the appearance of the sentence.

◆ **Information-poor**

Taking sentence 93) as an example, the NP ñ*her goals*ò is the direct object (DO) of the verb *realised* in a, the head of a reduced sentential complement (SC-0) in b, and the head of a complete sentential complement (SC-that) in c.

93) a. The athlete realised *her goals* through hard training. (Roland et al., 2006)

b. The athlete realised *her goals* would be difficult to achieve.

c. The athlete realised that *her goals* would be difficult to achieve.

Similar to 92a), 93b) would raise a Garden-path problem at the NP *her goals* because it can be interpreted either as in a) or c). According to syntax first processing theories (such as the Garden Path Theory Group), this ambiguity is inevitable, from the principle of Minimal Attachment. However, Roland et al. (2006) use regression modelling applied to the data from the British National Corpus to ñmeasure the amount of specificity of the information available for disambiguation in natural language useò (ibid., p. 245) and found that there is enough information for predicting whether *that* will be used in a given instance, and therefore resolving the DO/SC ambiguity in 92b) (i.e., when the comprehender can predict whether there should be a complementiser *that*, it should be possible to predict whether an NP is a DO or the head of an SC).

The information includes the semantic nature of the post-verbal NP (i.e., the thematic fit mentioned in McRae et al. (1998)) and the structural property information such as (full

NP vs. pronoun, length, frequency, etc.). However, those factors are easily omitted from the design of experiments in laboratories, as often experiments of this sort have a bias of expecting a sentential complement (ñguided designò), and therefore the effects of information other than syntax have rarely been focused on.

◆ **Mixed Stimuli**

Sentences in 94) were used in an ERPs study by Neville et al. (1991)

94) The scientist criticized Maxôs proof of the theorem.

a. Semantic Anomaly:

*The scientist criticized Maxôs event of the theorem.

b. Phrase Structure Violation:

*The scientist criticized Maxôs of proof the theorem.

c. Specificity Constraint Violation:

* What did the scientist criticize Maxôs proof of?

d. Subjacency Constraint Violation:

* What was a proof of criticized by the scientist?

The violation types of b), c), and d) were regarded as the variations of the syntactic anomaly. Neville et al.ôs study compared the four types of anomaly with the normal sentence in 94), and found positive potential at around 300-500msec (N400) after the onset of the stimuli (the underlined word) for the semantic anomaly, whereas negative potentials at around 450-600msec (P600) as well as early 125msec (N125, ELAN) for the syntactic anomaly.

There is no doubt that the anomaly type a) is a good example of the semantic anomaly and it is also used in Study 4 of the current thesis. However, a closer look at the stimuli in b), c), and d) throws doubt on the fairness of the manipulation.

Sentence b) swaps the preposition *of* with the noun *proof*. As a consequence, the structure of the sentence has been broken down into more than one layer, because the specifier *Max's* is required to have a head by the *Phrase-marker Rule*, (introduced in chapter 2, P.20) the *Headedness Principle* (Chapter 2, P.29) and the *Binarity Principle* (Chapter 2, P.29); whereas the prepositional phrase *of*- is required to have a head by the same rules. Moreover, the same rule requires the noun *proof* to have a complimentiser, and the NP *the theorem* to have a head. Thus, one change of word order has rendered the consequent constituents all structurally impaired. Although the measurement in Neville et al.ôs study was on the word *of*, unlike the manipulation in Study 4 and 5 in the current study, the size of the effect is much larger. It is not just one simple structural anomaly, but a conspiracy of many.

Similar arguments apply to type c) and d). Even the authors admit that ñthe Specificity constraint involves both syntax and semanticsò (Neville et al., 1991, p.154). Although one can argue that if the potential wave type was shown as those normally shown for the syntactic anomaly (such as ELAN), prior to the N400 semantics wave, it is questionable whether the outcome is purely caused by only one type of anomaly.

◆ **The Materials in the Current Study**

The issues discussed above suggest that some of the results of psycholinguistic experiments carried out in laboratories may be artefacts of the design. The criticisms above do not attempt to invalidate all of those studies, but the point to be noted is that ñevidenceò can sometimes be misleading.

The materials used in Studies 1-3 are typical laboratory sentences. Although attempts were made to remove any bias towards syntax-first or syntax-simultaneous theories, as the issue of the timecourse of processing was not focused on in those three studies, the predictive information suggested by Roland et al. (2006) was not controlled. However, since the research question is to investigate the ñdefaultò relative clause attachment bias in Chinese, material without rich predictive information can well serve the purpose.

Studies 4-5 explicitly separate the two anomalies and attempt not to stimulate more than one anomaly in the manipulation. In this sense, the current studies have a better control than Neville et al.ôs. Moreover, Studies 4 and 5 do not have a significant amount of problemss with the ñinformation-poorò issue, because the nature of the current studies is not to investigate ambiguity, but rather to spot the anomaly. There is no ambiguity in the materials in the last two experiments; therefore there should be no bias in the data once the anomaly sizes are controlled.

The size of effects has been controlled so that the semantic anomaly and syntactic anomaly both violate only one node in the structure.

95) John was carried **his friend* to the hospital. (Syntactic anomaly)
John was carried **by his cake* to the hospital. (Semantic anomaly in Study 4)
John was carried **at his friend* to the hospital. (Semantic anomaly in Study 5)

The semantic anomaly is very similar to Neville et al.ǿs study, and the reason for changing the NP to the preposition anomaly has been explained in Chapter 7. Notably, there is no anomaly preceding the anomalous regions in the current studies (i.e., the sentence is intact until the verb *carried*). Therefore, there is no extra noncanonical information carried over when the anomalous region starts. Furthermore, this is also confirmed by the naturalness rating task in Study 5, where no difference between the two anomalous conditions was considered. This is the crucial difference between the current studies and Neville et al.ǿs design.

◆ **The Sensitivity of the Task Measurements**

In the case of Neville et alǿs (1991) study, it is also questionable whether the EEG/ERPs measure (the potential transition) is more sensitive to syntactic anomaly than to semantic anomaly. This is somewhat similar to the issue of stimulus size discussed above, but instead of unequal stimulus size, the measurements used by a particular task are more sensitive to detect one type of stimuli than the other. This sensitivity issue can apply to other tasks, including the eye-tracking tasks used in the Study 4 & 5: is the eye movement more sensitive to semantic anomaly? Although there is strong evidence that eye movement can be sensitive to syntactic anomaly (e.g., Rayner et al., 1983; Rayner et al., 1989), this issue remains open. There is no evidence to prove or disprove the imbalance in sensitivity of the measurements at the moment.

8.3.2 Explicability

When Townsend and Bever (2001) raised the topic of *LAST*, they did not provide empirical data from new studies of their own. Instead, they reviewed many previous studies and tried to explain the ñunexplainableò by applying *LAST*. Moreover, even data that was construed by syntax-first or syntax-simultaneous theories can be explained by *LAST*, as discussed in Chapter 5.3.

Chapters 3 and 7 both explained and evaluated the original *LAST*. In Chapter 7, some criticisms and suggestions for modifying the original *LAST* have also been proposed. Before discussing the explicability of this theory, it is important to clarify its claims.

◆ ***LAST***

The current thesis proposes that sentence processing is a procedure that involves processing all information available under the rule of economy, as it is a process consuming mental resources which are limited. Thus, information is processed in order of importance.

As the purpose of human language is to transfer meanings, semantic related information should arguably be the most important. Moreover, as the smallest unit of sentence processing is lexical, identifying word boundaries and assigning meaning to lexical units should be the first stage of sentence processing. As the input of lexicon increases, the processor combines the new incoming lexical information with the earlier ones and identifies simple word phrases based on frequencies of combination and word order. Furthermore, as there is no consensus on the border of syntactic and semantic feature of word categorisation, in the current version of *LAST* (as well as in Townsend and Beverôs version), word categorisation is regarded as semantic information. This is because with the meaning of a word being assigned, it is evident whether it is a description of an object (noun), action (verb) or others (adverbial, adjective, etc.). One may argue that sometimes a word can be a noun or a verb in the same form (at least in English). That is why *LAST* consider frequency of combination and word order as the second most important factor. This is not a simple stage of lexical processing because it involves processing of words and word groups; however, it is not a stage of syntactic processing since there is no structure being constructed at this time. The current thesis names it as the PSP (pre-syntax processing) stage.

As the main thematic assigner of a sentence is the verb, in simple sentences with only one verb (i.e., without either reduced or full clauses), sentence processing may end in the PSP stage. This means that syntax processing may not be necessary for simple sentences such as those in 96) below

96) I have eaten an apple.
    Lisa is a good student.
    Bart is bad.
    We wonœt have an exam tomorrow.
    …

However, sometimes other information is also needed for simple sentences such as 97) below:

97) A: Have you eaten an apple?
    B: No, I havenœt.

It will not be clear what B hadnœt done without knowing what A asked. Although it is a simple conversation, context information is more important than assigning the omitted ñempty categoryò in the syntactic structure in B. Thus, PSP is a stage when sentence can be comprehended without building up the syntactic tree.

Because processing in the PSP stage relies heavily on frequency of combination and word order, it is predicted that active sentences (SVO) are easier to process than passive sentences (O is V-ed by S) in English.

98) I have eaten an apple.
    An apple was eaten by me.

In other cases, nevertheless, complicated sentences can be produced, so they have to be processed. A simple PSP stage will not be sufficient, as the complicated structure costs more mental resources (e.g., memory) so with the help of rules the processing can be simplified and thus more efficient. One symbol of this type of sentences is that there is more than one verb. In this case, the sentences used in Study 1-3 fall into this category. Since sentence structure should be assigned bottom-up according to the UG theory (specifically, the MP theory), syntax processing only occurs after the processor finishes scanning the whole sentence at least once.

This may be against some ñevidenceò, especially the evidence from regression-path time measured by eye-tracking techniques. The regression-path time measures instantaneous reading difficulty, assuming that the eyes move immediately backwards (instead of forward to get more information) when problematic information enters the processor. Supporters of syntax-first theories, (i.e., Garden path theory (Frazier & Clifton, 1996)) would regard it as the time to go back to check and re-analyse the previously parsed information. However, there is no evidence to confirm that the regressive movements are actually caused by syntactic anomaly (or ambiguity). Indeed, Studies 4 and 5 in the current thesis used simple sentences with syntactic anomaly but did not find any reliable regressive movement on the problematic region in English and Chinese. It is noticeable that in Garden-path sentences, many temporary ambiguities are shown and resolved at the second verb, or *by* at the reduced passive RC. In 99) previously mentioned Garden-path sentences are relisted here for the readerôs convience:

99) a. The horse raced past the barn *fell*.
b. While Mary was mending the sock *fell off* her lap.
c. The athlete realised her goal *is* impossible to reach.
d. The defendant examined *by the lawyer* turned out to be unreliable.

In sentences 98a-c), before reaching the ambiguous verb, the sentence can be well processed using information in PSP. Encountering the new upcoming verb violates the frequency of word combination and the processor can be ñsurprisedò. Going back to check previous information and assigning possible syntax structure can be urgent because the harmony of the previously comprehended constituent has been broken, and extra information is needed to solve this problem. It is possible that regressive movements occur without requiring more new information. In sentence like 98d), as the frequent word order is *SVO* for active sentences and *O be V-ed by S* for passive sentences, after processing the verb without *be* preceding it, *by* is not expected. Hence regressive movements can occur.

In Studies 4 & 5, the syntactic anomaly occurs in a passive sentence:

100) * John was carried his friend to the hospital.
　　* 約翰他的朋友擡到醫院。

In English, it is not required to have the PP (by + NP) in a passive sentence. In this sense, *John was carried to the hospital* is a full sentence. Moreover, *carry* is a verb that can take a patient, so *carry* + *his friend* can be expected. Because there is only one verb being processed, the processor is still open-minded to new lexicon information. Therefore no immediate regressive movements occurred in the current studies. In the case of Chinese, as the verb has not appeared yet, the processor is in the ñwait-and-seeò status, so no regressive movement is activated.

Because assigning syntax structure means complex information processing and the initial PSP stage is not sufficient, after syntax processing, the processor needs to be re-assured that the processing was correct. Thus, further syntax generation and meaning matching is required. This is the last stage in sentence processing proposed by *LAST*.

◆ Some Further Issues

The above points reinforce the concept of *LAST* developed in the current thesis and also apply *LAST* to explain some previously demonstrated evidence proposed by syntax-first theories with comparison of the current data. However, there are still some issues that need to be clarified:

- The borderline between syntax and semantics

So far there is no agreement on demarcation between syntax and semantics linguistically. The controversial borderline starts from the word categorisation. In the Government and Binding theory, word category is a syntactic concept; however, the subcategory is heavily based on the semantic feature of that word ï a paradoxical approach.

Although Neville et al.ôs (1991) design has been criticised in the earlier section, their neurological approach to the problem is novel and encouraging. They have found that there is a distinct timing and distribution of the semantic and syntactic effects across the brain: syntactic and semantic processing result in nonidentical patterns of activation, including greater frontal engagement during syntactic processing and larger increases in temporal and temporo-parietal regions during semantic analyses. Although the timing is debatable (cf. Pulvermüller et al., 2001), the distinct localisation of the effects have been replicated by many other studies using fMRI techniques (Newman et al., 2004 etc.). Maybe the borderline debate can be resolved by further neurological approaches with more carefully controlled materials.

- Re-analyses and Parsing

The current thesis proposes a PSP stage of sentence processing. As there is no syntax analysis involved in this stage. Parsing (specifically meaning the syntax processing) is no longer a necessary stage. This means the concept of re-analysis should start with parsing.

Nevertheless, the parser has been accepted as an innate mechanism in the human language faculty. Does the PSP stage disclaim this innate mechanism? *LAST* does not propose so. In fact, the parser can still be innate for language processing. However, when processing extrinsic information, it is not necessary to activate all the parsing mechanisms inside the brain. The parser can still be the sensitive syntax processor, but because it is not always necessary to process syntax, the parser can be at rest when it is not activated. Although the purpose of language is to convey meaning, it is not an isolated object, as it relies heavily on the media of visual, acoustic and even direct tactile sensation. Moreover, language is a rather novel facility in terms of evolutionary history. Even though it is generally agreed that non-human species do not have language, it is undoubted that animal sounds can have simple meanings (as summarised in Chapter 1). This means the parser is a rather new mechanism. It helps with decoding more complex structures, but it may not be necessary in terms of decoding simple ones. Therefore, the current thesis proposes that parsing only starts when re-analyses is necessary.

- Initial Processing Factors

LAST proposes that sentence processing depends on lexical information in the initial stage (PSP), including meaning, function (including thematic assignments), and frequency of combination and word order. So far, it also proposes that possible factors can include context, discourse, and other information such as imaginability and environment. Without specifying what the factors that can be involved in the PSP stage are, the theory is unfalsifiable and hence not a valid scientific theory.

The current thesis proposes that the initial processing factors are lexical and contextual (including discourse and prosody) information. Lexical information has been discussed in earlier sections when introducing *LAST*. The reason of including context information (including discourse and prosody) is that more and more research results suggest that they have an influence on language processing (e.g., Altmann & Kamide, 1999; J. D. Fodor, 2002a; 2002b; Roland et al., 2006; Watson & Gibson, 2004). Noncanonical structures can be resolved with little or no difficulty with enough context information.

It is admitted that these initial processing factors should be further studied and, if possible, stabilised, but that is a topic for further study.

Chapter 9

Implications and Concluding Remarks

The reason that *LAST* can be very explanatory is because *LAST* does not deny the processing of syntax. As the purpose of psycholinguistic study is to investigate the human language processing mechanism and the object of the study should be natural human language, a good theory should be able to account for not only complex sentences as used in Study 1-3, but also simple sentences as in Study 4-5. In Chapter 2, simple sentences as in conversation 17) are repeated here in 101):

101) A: Nice to meet you!
B: You too!

A linguist is certainly able to draw a syntax tree of those sentences with all the movements and empty categories. But the real question a psycholinguist should ask is: do we really need to ñbotherò our brain to work out the structure of the sentences to be able to understand these simple greetings?

Therefore, *LAST* proposes that in order to save mental resources in sentence processing, for simple sentences with only one predicate, the human sentence processing mechanism ï parser will not be activated. The stage before the activation of the parser is called the PSP stage. This proposal is based on universal grammar and the current thesis has provided evidence to support this proposal. This proposal can also explain evidence provided for syntax-first and syntax-simultaneous theories.

9.1 Implications

As *LAST* emphasises the importance of lexical information in the PSP stage, word identification has become a very essential procedure in the PSP stage. The current studies have chosen Chinese as a sample language that bears the feature of no word boundary in its written form. Therefore an extra processing stage of word segmentation may have been detected when comparing to the English counter parts. However, this explanation should be focused and further tested, as the current thesis focuses mainly on the timecourse of the syntactic and semantic processing, but not the word segmentation feature.

If the extra stage of word segmentation is true, it can be a good example of ñeconomical processingò, i.e., if word segmentation is not needed in English, this stage can be skipped during sentence processing, then certainly when assigning structure is not needed, it can be skipped as well.

Moreover, if the word segmentation is needed in Chinese, then maybe within the same language (i.e., English), the processing stages for acoustic data and written form can be different in terms of identification of the word boundary? There is no doubt that different biological organs are involved between the listening and reading tasks. The point here is that there should also be a word segmentation stage for acoustic data.

These implications should be further tested by different methods with more language data, and should be good topics for future psycholinguistic research.

9.2 Future Direction of the Study

The current thesis has provided compelling evidence for *LAST*; however, at the same time it opens new avenues for exploration.

◆ Theoretical Direction

As ñfrequencyò is one of the constraints that *LAST* regards as important information in the PSP stage, the theory faces a similar ñgrain-sizeò issue as the Exposure-based Tuning Hypothesis (ETH). The current thesis hypothesises that the frequency information considered in the PSP stage includes frequency of the lexical meaning relating to the word form, and the frequency of word combination. Future studies could address the question of the extent of the ñfine grainedò information taken in. Especially, as the frequency of word combination is considered, it should be interesting to know what the minimal number of words combined on either the left or the right side of the current word are. The current study makes the assumption that one word to the left and one word to the right of the currently processed word are used as information of frequency of combination. This assumption needs to be corroborated in future studies.

Moreover, McRae et al. (1998) applied the competition model to test syntax-first and parallel theories, but they did not have a persuasive reason to test a syntax-later theory. The current work has provided good motivation to re-set the parameters and see how the model with late assignment of syntax information will fit in with human data.

◆ Empirical Direction

The current thesis has left some proposals, such as the *Filler-Gap* theory and the relative clause in Chinese (see Section 6.3), processing during saccade time (see Section 6.5), and the word segmentation stage, requiring further investigation. Among them, word-segmentation is the closest related topic.

The current thesis reported consistent delay effects in Chinese. It hypothesises that there is an extra word segmentation stage. Further studies should make attempts to substantiate this assumption in Chinese. Moreover, more language data should be collected to test the existence of this stage. This can be done by examining the spoken form (for example in English), or finding other languages with similar features (for example Japanese[68]) in their written form. It would be very interesting to explore the cues for segmenting words. For instance, if space is the cue for English written words, and *hiragana* is the cue for Japanese written words, is there any cue for Chinese written words? Moreover, space and *hiragana* are both visual presentation but prosodically this cue does not exist. What would be the cue for word boundary identification? The current LAST proposes that the cues for word segmentation would be frequencies of the word itself and the combination with the words that are juxtaposed to it. Then the

[68] It is arguable that the task of word identification in Japanese is much easier than in Chinese. Although there is no word boundary in its written form, the Japanese inflection is always in its *hiragana* form.

consequences are that the prosodic combination frequency can be used to determine the acoustic sentence processing. The downside of this proposal is that it cannot give a proper explanation to the ñgrain-sizedò issue of this frequency.

Furthermore, direct measurements of the timecourse of neuron network activation under syntactic and semantic stimuli can always provide important information for this research question. As the last two experiments have provided simple and well controlled materials, applying the methodologies such as EEG/ERPs measurements using the current stimuli could be valuable in further studies. If it is true that N400 is sensitive to semantic surprise, and P600 is sensitive to the syntactic surprise, the negative electricity flow would be expected to be triggered earlier. Moreover, when using the ñbalancedò materials as in Study 5, in the semantic anomalous *at* condition, it is predicted that ELAN (which associates with word-category errors, and normally strongly correlates with P600) should not be activated, as the word-categorisation is correct in this condition. The downside of this is that the stimuli of semantic anomaly may be too weak to activate anything.

Nevertheless, there are good reasons to believe that the aforementioned suggestions can bring interesting evidence for the psycholinguistic field.

9.3 Conclusing Remarks

The current thesis provides linguistic theoretical background to explain the urge to develope a late assignment of syntax theory. It also provides empirical data crosslinguistically (English and Chinese) supporting the *LAST* (Townsend & Bever, 2001) (Studies 4 & 5), as well as data questioning syntax-first theories (Study 1 & 3). Furthermore, it proposes for the PSP stage in sentence processing (in exchange of the pseudo-syntax stage proposed by Townsend and Bever), where lexicon information is focused and processed. The theory follows the newly developed Chomskyan linguistic theory of syntax in theoretical proposal and empirical controls. However, it claims that syntax is not necessary for processing, especially in simple sentences with only one verb. *LAST* does not deny the universality of the UG and the existence of language faculty, instead, it considers the limitation of human mental resources and proposes that sentence processing should also be economical. Thus, only necessary information will be processed in order of importance. This is again consistent with the spirit of the Minimalist Program.

However, the data provided in the current study are only the tips of a big iceberg of the human language processing mechanism. Human language processing is a rather intricate and complicated procedure. To verify the universal grammar and *LAST*, more studies from different dimensions should and will be conducted.

Aaronson, D., & Ferres, S. (1984) The Word-by-word Reading Paradim: An Experimental and Theoretical Approach. In D. E. Kieras & M. A. Just (Eds.), *New Methods in Reading Comprehension Research* (pp. 31-68): Lawrence Erlbaum Associates.

Abdelghany, H., & Fodor, J. D. (1999, September 23-25) *Low Attachment of Relative Clause in Arabic (Poster).* Paper presented at the AMLaP, Edinburgh.

Altmann, G. T. M. (1997) *The Ascent of Babel*. Oxford: Oxford University Press.

Altmann, G. T. M., & Kamide, Y. (1999) Incremental Interpretation at Verbs: Restricting the Domain of Subsequent Reference. *Cognition 73*, 247-264.

Altmann, G. T. M., & Steedman, M. (1988) Interaction with Context During Human Sentence Processing. *Cognition 35*, 1-39.

Aoshima, S., Phillips, C., & Weinberg, A. (2004) Processing Filler-Gap Dependencies in a Head-final Language. *Journal of Memory and Language 51*(1), 23-54.

Atkinson, M. (1992) *Children's Syntax: an Introduction to Principles and Parameters Theory*: Blackwell.

Bader, M. (1998) Prosodic Influences on Reading Syntactically Ambiguious Sentences. In J. D. Fodor & F. Ferreira (Eds.), *Reanalysis in Sentence Processing* (pp. 1-46): Kluwer Academic Publishers.

Bates, E., & MacWhinney, B. (1989) Functionalism and the Competition Model. In B. MacWhinney & E.Bates (Eds.), *The Crosslinguistic Study of Sentence Processing*. Cambridge, UK: Cambridge University Press.

Bever, T. G. (1970) The Cogntive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the Development of Language* (pp. 279-362). New York: Wiley.

Bever, T. G., & McElree, B. (1988) Empty Categories Access their Antecedents During Comprehension. *Linguistic Inquiry 19*.

Black, C. A. 1999. http://www.sil.org/mexico/ling/E002-IntroGB.pdf

Bloomfield, L. (1936) Language or Ideas? In J. Katz (Ed.), *The Philosophy of Linguistics* (pp. 19-25). Oxford: Oxford University Press.

Boland, J. E. (1997a) The Relationship between Syntactic and Semantic Processes in Sentence Comprehension. *Language and Cognitive Progresses 12*(4), 423-484.

Boland, J. E. (1997b) Resolving Syntactic Category Ambiguities in Discourse Context: Probablistic and Discourse Constraints. *Journal of Memory and Language 36*, 588-615.

Boland, J. E. (2004) Linking Eye Movements to Sentence Conprehension in Reading and Listening. In M. Carreiras & J. Charles Clifton (Eds.), *The On-line Study of Sentence Comprehension: Eyetracking, ERPs and Beyond*. New York / Hove: Psychology Press.

Bornkessel, I., McElree, B., Schlesewsky, M., & Friederici, A. D. (2004) Multi-dimensional Contributions to Garden Path Strength: Dissociating Phrase Structure from Case Marking. *Journal of Memory and Language 51*(4), 495-522.

Brown, C. M., & Hagoort, P. (1999) Introductions In C. M. Brown & P. Hagoort (Eds.), *The Neurocognition of Language*. Oxford / New York: Oxford University Press.

Brysbaert, M., & Mitchell, D. C. (1996) Modifier Attachment in Sentence Processing: Evidence from Dutch. *The Quarterly Journal of Experimental Psychology 49*(A), 664-695.

Burling, R. (1993) Primate calls, human language, adn nonverbal communication. *Current Anthropology 34*, 25-53.

Cairns, H. S. (1999) *Psycholinguistics: An Introduction.* Austin, Texas: Pro.ed.

Carreiras, M., & Clifton, C. (1999) Another Word on Parsing Relative Clauses: Eyetracking Evidence from Spanish and English. *Memory & Cognition 27*(5), 826-833.

Carreiras, M., & Clifton, C. (Eds.). (2004) *The On-line Study of Sentence Comprehension: Eyetracking, ERPs and Beyond*. New York / Hove: Psychology Press.

Carroll, D. W. (1994) *Psychology of Language* (2 ed.). Pacific Grove, California: Brooks / Cole Publishing Company.

Chipere, N. 2006. http://cogprints.org/712/00/real.PDF

Chomsky, N. (1957) *Syntactic Structure*. Mouton: The Hague.

Chomsky, N. (1959) Review of Skinner's "Verbal Behavior". *Langauge 35*, 26-58.

Chomsky, N. (1965) *Asects of the Theory of Syntax*. Cambridge, Massachusetts: The M.I.T. Press.

Chomsky, N. (1982) *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge, Mass: The MIT Press.

Chomsky, N. (1986) *Knowledge of Language: Its Nature, Origin, and Use* (1 ed. Vol. 2). New York, Westport, Connecticut, London: Praeger.

Chomsky, N. (1995) *The Minimalist Program* (Vol. 28). Cambridge, Massachusetts / London, England: The MIT Press.

Chomsky, N. (1999) *Derivation by Phase*: MIT Working Papers in Linguistics.

Chomsky, N. (2001) Beyond Explanatory Adequacy: MIT.

Chomsky, N., & Lasnik, H. (1995) The Theory of Principles and Parameters. In N. Chomsky (Ed.), *The Minimalist Program* (Vol. Chapter 1, pp. 13- 128). Cambridge, Massachusetts, London, England: The MIT Press.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993) PsyScope: A new graphic interactive environment for designing pychology experiments. *Behavioral Research Methods, Instruments, and Computers 25*(2), 257-271.

Connell, L., & Keane, M. T. (2006) A Model of Plausibility. *Cognitive Science 30*(1), 95-120.

Cook, V. J. (1993) *Chomsky's Universal Grammar*. New York: Blackwell.

Cooper, D. E. (1973) *Philosophy and the Nature of Language* (Vol. 14). London: Longman.

Cooper, D. E. (1975) *Knowledge of Language*. London / New York: London: Prism / New York: Humanities.

Corley, M. (1995) *The Rôle of Statistics in Human Sentence Processing.* Unpublished PhD, the University of Exeter, Exeter.

Crain, S., & Steedman, M. (1985) On not being led up the garden path: The use of context by the psychological parser. In D. Dowty, L. Karttunnen & A. Zwicky (Eds.), *Natural Language Parsing: Psychological, Computational and Theoretical Perspectives*. Cambridge: Cambridge Unviersity Press.

Cuetos, F., & Mitchell, D. C. (1988) Cross-linguistic Differences in Parsing: Restrictions on the Use of the Late Closure Strategy in Spanish. *Cognition 30*, 73-105.

Cuetos, F., Mitchell, D. C., & Corley, M. (1996) Parsing in Different Languages. In M. Carreiras, J. Garcia-Albea & N. Sabastian-Galles (Eds.), *Language Processing in Spanish*. Hove: Erlbaum.

De Vincenzi, M., & Job, R. (1993) Some Observations on the Universality of the Late-Closure Strategy. *Journal of Psycholinguistic Research 22*(2), 189-206.

De Vincenzi, M., & Job, R. (1995) An Investigation of Late Closure: The Role of Syntax, Thematic Structure, and Pragmatics in Initial and Final Interpretation. *Journal of Experimental Psychology: Learning, Memory and Cognition 21*(5), 1303-1321.

Desmet, T., Brysbaert, M., & De Baecke, C. (2002) The Correspondence between Sentence Production and Corpus Frequencies in Modifier Attachment in Dutch. *Quarterly Journal of Experimental Psychology 55A*(3), 879-896.

Desmet, T., De Baecke, C., Drieghe, D., Brysbaert, M., & Vonk, W. (2006) Relative Clause Attachment in Dutch: On-line Comprehension Corresponds to Corpus Frequencies when Lexical Variables are Taken into Account. *Language and Cognitive Progresses 21*, 453-481.

Desmet, T., & Gibson, E. (2003) Disambiguation Preferences and Corpus Frequencies in Noun Phrase Conjunction. *Journal of Memory and Language 49*, 353-374.

Deutsch, A., & Bentin, S. (2001) Syntactic and Semantic Factors in Processing Gender Agreement in Hebrew: Evidence from Erps and Eye Movements. *Journal of Memory and Language 45*, 200-224.

Diesing, M. (1990) Verb Movement and the Subject Position in Yiddish. *Natural Language and Linguisitc Theory 8*, 41-79.

Ehrlich, K., Fernández, E. M., Fodor, J. D., Stenshoel, E., & Vinereanu, M. (1999, March 18-20) *Low Attachment of Relative Clause: New Data from Swedish, Norwegian and Romanian. (A Poster).* Paper presented at the 12th Annual CUNY Conference on Human Sentence Processing, New York.

Elman, J. L. (1993) Learning and Development in Neural Networks: the Importance of Starting Small. *Cognition 48*, 71-99.

Felser, C., Clahsen, H., & Münte, T. F. (2003) Storage and Integration in the Processing of Filler-Gap Dependencies: An ERP Study of Topicalization and *Wh*-Movement in German. *Brain and Language 87*, 345-354.

Fernández, E. M. (1999) Processing Strategies in Second Language Acquisition: Some Preliminary Results. In E. Klein & G. Martohardjono (Eds.), *The Development of Second Language Grammars: A Generative Approach* (pp. 217-239). Amsterdam and Philadelphia: John Benjamins.

Fernández, E. M. (2002) *Bilingual Sentence Processing: Relative Clause Attachment in English and Spanish* (1 ed. Vol. 29). Amsterdam/Philadelphia: John Benjamins B.V.

Ferreira, F., & Clifton, C. (1986) The Independence of Syntactic Processing. *Journal of Memory and Language 25*, 348-368.

Ferreira, F., & Henderson, J. M. (1991) How is Verb Information Used During Syntactic Parsing? In G. B. Simpson (Ed.), *Understanding Word and Sentence* (pp. 305-330). Amsterdam / New York / Oxford / Tokyo: North-Holland.

Fodor, J. A., & Garrett, M. F. (1967) Some Determinants of Sentential Complexity. *Perception & Psychophysics 2*, 289-296.

Fodor, J. A., Garrett, M. F., & Bever, T. G. (1968) Some Determinants of Sentential Complexity, II: Verb Structure. *Perception & Psychophysics 3*, 453-463.

Fodor, J. D. (1977) *Semantics: Theories of Meaning in Generative Grammar* (Vol. 2): The Harvester Press.

Fodor, J. D. (1998) Learning to Parse. *Journal of Psycholinguistic Research 27*(2), 285-319.

Fodor, J. D. (2002a) *Prosodic Disambiguation in Silent Reading.* Paper presented at the North East Linguistic Society (NELS), GSLA, University of Massachusetts, Amherst.

Fodor, J. D. (2002b, 11-13 April 2002) *Psycholinguistics Cannot Escape Prosody.* Paper presented at the Speech Prosody 2002, Aix-en Provence, France.

Frank, R. (2004) Restricting Grammatical Complexity. *Cognitive Science 28*(5), 669-697.

Frazier, L. (1978) *On Comprehending Sentences: Syntactic Parsing Strategies.* Unpublished PhD., University of Connecticut.

Frazier, L. (1987) Sentence Processing. In M. Coltheart (Ed.), *Attention and Performance* (Vol. 12, pp. 559-586). Hove, England: Erlbaum.

Frazier, L. (1990a) Exploring the Architecture of the Language System. In G. T. M. Altmann (Ed.), *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives*. Cambridge, MA: MIT Press.

Frazier, L. (1990b) Parsing Modifiers: Special Purpose Strategies in the Human Sentence Processing Mechanism. In D. Balota, G. B. F. d'arcais & K. Rayner (Eds.), *Comprehension Processing in Reading*. Hillsdale: LEA.

Frazier, L., & Clifton, C. (1989) Successive Syclicity in the Grammar and the Parser. *Language and Cognitive Progresses 4*, 93-126.

Frazier, L., & Clifton, C. (1996) *Construal*. Cambridge, Mass: The MIT Press.

Frazier, L., & Clifton, C. (1997) Construal: Overview, Motivation, and Some New Evidence. *Journal of Psycholinguistic Research 26*, 277-295.

Frazier, L., Clifton, C., & Randall, J. (1983) Filling Gaps: Decision Principles and Structure in Sentence Comprehension. *Cognition 13*, 187-222.

Frazier, L., & Fodor, J. D. (1978) The Sausage Machine: A New Two-stage Parsing Model. *Cognition 6*, 291-325.

Frazier, L., & Rayner, K. (1982) Making and Correcting Errors During Sentence Comprehension: Eye-movements in the Analysis of Structurally Ambiguous Sentences. *Cognitive Psychology 14*, 178-210.

Gibson, E. (1991) *A Computational Theory of Human Language Processing: Memory Limitations and Processing Breakdown.* Unpublished Ph.D, Carnegei Mellon University.

Gibson, E. (1998) Linguistic Complexity: Locality of Syntactic Dependencies. *Cognition 68*, 1-76.

Gibson, E., & Hickok, G. (1993) Sentence Processing with Empty Categories. *Language and Cognitive Progresses 8*, 147-161.

Gibson, E., Pearlmutter, N., Canseco-Gonzalez, E., & Hickok, G. (1996) Recency Preference in the Human Sentence Processing Mechanism. *Cognition 59*, 23-59.

Gibson, E., & Schütze, C. (1999) Disambiguation Preferences in Noun Phrase Conjunction Do Not Mirror Corpus Frequency. *Journal of Memory and Language 40*, 263-279.

Gibson, E., Schütze, C., & Salomon, A. (1996) The Relationship between the Frequency and the Processing of Linguistic Structure. *Journal of Psycholinguistic Research 23*, 381-405.

Gilboy, E., Sopena, J., Clifton, C., & Frazier, L. (1995) Argument Structure and Association Preferences in Spanish and English Complex NPs. *Cognition 54f*, 131-167.

Gordon, P. C., & Johnson, R. H. a. M. (2004) Effects of Noun Phrase Type on Sentence Complexity. *Journal of Memory and Language 51*(1), 97-114.

Gorrell, P. (1995) *Syntax and Parsing*. New York: Cambridge University Press.

Green, M. J., & Mitchell, D. C. (2006) Absence of Real Evidence Against Competition During Syntactic Ambiguity Resolution. *Journal of Memory and Language 55*, 1-17.

Grice, H. P. (1975) Logic and Conversation. In P. Cole & G. Morgan (Eds.), *Syntax and Semantics 3*. New York: Academic Press.

Haberlandt, K. (1994) Methods in Reading Research. In M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 1-31). San Diego: Academic Press.

Halle, M., & Stevens, K. N. (1964) Speech recognition: A model and a program for research. In J. A. Fodor & J. J. Katz (Eds.), *The Structure of Language: Readings in the philosophy of language*. Englewood Cliffs, NJ: Prentice-Hall.

Harley, T. A. (2001) *The Psychology of Language: From Data to Theory* (2 ed.). Hove / New York: Psychology Press Ltd / Taylor & Francis Inc.

Hemforth, B., Konieczny, L., Scheepers, C., & Strube, G. (1998) Syntactic ambiguity resolution in German. In D. Hillert (Ed.), *Sentence Processing: A Crosslinguistic Perspective. Syntax and Semantix* (Vol. 31, pp. 293-312). San Diego, CA: Academic Press.

Hemforth, B., Konieczny, L., Seelig, H., & Walter, M. (1999, March 18-20) *Case Matching and Relative*

Clause Attachment. Paper presented at the 12th Annual CUNY Conference on Human Sentence Processing, New York, NY.

Hemforth, B., & Schimke, S. (2003) *Length and Position in French (poster presentation).* Paper presented at the AMLaP, Glasgow.

Hockett, C. F. (1959) Animal "languages" and human languages. In J. N. Spuhler (Ed.), *The evolution of man's capacity for culture* (pp. 32-39). Detroit, MI: Wayne State University Press.

Hockett, C. F. (1960) The origin of speech. *Scientific American 203*(October), 89-96.

Hockett, C. F. (1961) The problem of universals in language. In J. H. Greenberg (Ed.), *Universals of language* (pp. 1-29). Cambridge, MA: MIT Press.

Holmes, V. M. (1979) Some Hypotheses About Syntactic Processing in Sentence Comprehension. In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence Processing: Psycholinguistic studies presented to Merrill Garrett.* Hillsdale, N. J.: Erlbaum.

Huang, C.-T. J. (1982) *Logical Relations in Chinese and the Theory of Grammar.* Unpublished Ph.D., MIT, Cambridge, Massachusetts.

Hyönä, J., Radach, R., & Deubel, H. (Eds.). (2003) *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*. Amsterdam-Boston-London-New York-Oxford-Paris-San Diego-San Francisco-Singapore-Sydney-Tokyo: Elsevier.

Inhoff, A. W. (1984) Two Stages of Word Processing During Eye Fixations in the Reading of Prose. *Journal of Verbal Learning and Verbal Behavior 23*, 612-624.

Jun, S.-A. (2002) *Prosdic Phrasing and Attachment Preference.* Paper presented at the 15th Annual CUNY Conference on Human Sentence Processing, UC Santa Barbara.

Just, M. A., & Carpenter, P. A. (1980) A Theory of Reading: From Eye Fixations to Comprehension. *Psychology Review 87*(329-354).

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982) Paradigms and Processes in Reading Comprehension. *Journal of Experimental Psychology: General I*(11), 228-238.

Kamide, Y. (1998) *The Role of Argument Structure Requirements and Recency Constraints in Human Sentence Processing.* Unpublished PhD, the University of Exeter, Exeter.

Kamide, Y., & Mitchell, D. C. (1997) Relative Clause Attachment: Non-determinism in Japanese Parsing. *Journal of Psycholinguistic Research 26*(247-254).

Kim, A., & Osterhout, L. (2005) The Independence of Combinatory Semantic Processing: Evidence from Event-related Potentials. *Journal of Memory and Language 52*, 205-225.

Kimball, J. (1973) Seven Principles of Surface Structure Parsing in Natural Language. *Cognition 2*, 15-47.

Kitakawa, Y. (1986) *Subjects in Japanese and English.* Unpublished Ph.D, University of Massachusetts, Amherst.

Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., & Friederici, A. D. (2004) Music, Language and Meaning: Brain Signatures of Semantic Processing. *Nature Neuroscience 7*, 302-307.

Konieczny, L., Hemforth, B., Scheepers, C., & Strube, G. (1995) PP-attachment in German: Results from Eye Movement studies. In J. M. Findlay, R. Walker & R.W.Kentridge (Eds.), *Eye Movement Research: Mechanisms, Processes and Applications*. New York: Elsevier Science Publishing.

Konieczny, L., Scheepers, C., Hemforth, B., & Strube, G. (1997) The Role of Lexical Heads in Parsing: Evidence from German. *Language and Cognitive Progresses 12*, 307-328.

Koopman, H., & Sportiche, D. (1991) On the Position of Subject. In J. McCloskey (Ed.), *The Syntax of Verb-initial Languages* (Special ed., pp. 211-258): Lingua.

Kuroda, S. Y. (1988) Whether We Agree or Not. *Lingvisticae Investigations 12*, 1-47.

Kutas, M., & van Petten, C. (1994) Psycholinguistics Electrified. In M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics*. New York: Academic Press.

Langford, J., & Holmes, V. M. (1978) *Comprehending Presuppositions in and out of Context*.Unpublished manuscript.

Li, Y. A. (1990) *Order and Constituency in Mandarin Chinese*. Dordrecht / Boston / London: Kluwer Academic Publisher.

Lin, C., & Fong, S. (2005, July 22-23, 2005) *Explaining Filler-gap Factors in Chinese Processor Relativization.* Paper presented at the Workshop on Theoretical East Asian Linguistics, Harvard.

Lin, H. (1999) Word Order Issue in Modern Chinese.

Liversedge, S. P., Pickering, M. J., Clayes, E. L., & Branigan, H. P. (2003) Thematic Processing of Adjuncts: Evidence from an Eye-tracking Experiment. *Psychonomic Bulletin & Review 10*(3), 667-675.

Lovriĺ, N. (2003) *Implicit Prosody in Silent Reading: Relative Clause Attachment in Croatian.* Unpublished PhD, the City University of New York, New York.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994) The Lexical Nature of Syntactic Ambiguity Resolution. *Psychology Review 101*, 676-703.

MacDonald, M. C., & Seidenberg, M. S. (in press) Constraint Satisfaction Accounts of Lexical and Sentence Comprehension. In M. J. Traxler & M. A. Gernsbacher (Eds.), *The Handbook of Psycholinguistics* (2 ed.). London, UK: Elsevier.

Marcus, M. P. (1980) *A Theory of Syntactic Recognition for Natural Langauge*. Cambridge, MA: MIT Press.

Marslen-Wilson, W. D. (1976) Linguistic description and psychological assumptions in the study of sentence perception. In R. J. Wales & E. Walker (Eds.), *New approaches to languages mechanisms*. Amsterdam: North Holland.

McClelland, J. L., & Rumelhart, D. E. (1981) An Interactive Activation Model of Context Effects in Letter Perception: Part 1. An Account of Basic Findings. *Psychology Review 88*(5), 375-407.

McDonald, S. A., Carpenter, R. H. S., & Shillcock, R. C. (2005) An Anatomically Constrained, Stochastic Model of Eye Movement Control in Reading. *Psychology Review 112*(4), 814-840.

McElree, B., Foraker, S., & Dyer, L. (2003) Memory Structures that Subserve Sentence Somphrehension. *Journal of Memory and Language 48*, 67-91.

McRae, K., Hare, M., & Tanehous, M. K. (2005) Meaning Through Syntax Is Insufficient to Explain Comprehension of Sentences With Reduced Relative Clauses: Comment on McKoon and Ratcliff (2003). *Psychology Review 112*(4), 1022-1031.

McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998) Modeling the Influence of Thematic Fit (and Other Constraints) in On-line Sentence Comprehension. *Journal of Memory and Language 38*, 283-312.

Meng, M., & Bader, M. (2000) Ungrammaticality Detection and Garden Path Strength: Evidence for Serial Parsing. *Language and Cognitive Progresses 15*(6), 615-666.

Mitchell, D. C. (1984) An evaluation of subject-paced reading tasks and other methods for investigating immediate processes in reading. In D. E. Kieras & M. A. Just (Eds.), *New Methods in Reading Comprehension Research* (pp. 91-118). Hillsdale, NJ: Lawrence Erlbaum Associates.

Mitchell, D. C. (1994) Sentence Parsing. In M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistic Research* (pp. 375-409). New York: Acadamic Press.

Mitchell, D. C. (2004) On-line Methods in Language Processing: Introduction and Historical Review. In M. Carreiras & J. Charles Clifton (Eds.), *The on-line Study of Sentence Comprehension: Eyetracking, ERPs and Beyond* (pp. 15-33). New York, Hove: Psychology Press.

Mitchell, D. C., & Brysbaert, M. (1998) Challenges to recent theories of language differences in parsing: Evidence from Dutch. *Syntax and Semantics 31*, 313-335.

Mitchell, D. C., Brysbaet, M., Grondelaers, S., & Swanepoel, P. (1999) Modifier attachment in Dutch: Testing aspects of Construal Theory. In A. Kenedy, R. Radach, D. Heller & J. Pynte (Eds.), *Readings as a Perceptual Process*.

Mitchell, D. C., Cuetos, F., & Corley, M. (1992) *Statisitcal vs. Linguistic Determinants of Parsing Bias: Cross-linguistic Evidence.* Paper presented at the CUNY Human Sentence Processing Conference, New York.

Mitchell, D. C., Cuetos, F., Corley, M., & Brysbaert, M. (1995) Exposure-based models of human parsing: Evidence for the use of Coarse-Grained (Nonlexical) statistical records. *Journal of Psycholinguistic Research 24*, 469-486.

Miyamoto, E. (2002) Case Markers as Clause Boundary Inducers in Japanese. *Journal of Psycholinguistic Research 31*(4), 307-347.

Miyamoto, E., Gibson, E., Pearlmutter, N., Aikawa, T., & Miyagawa, S. (1999) A U-shape Relative Clause Attachment Preference in Japanese. *Language and Cognitive Progresses 14*(5/6), 663-686.

Münte, T. F., Say, T., Clahsen, H., & Kutas, M. (1998) Decomposition of morphologically complex words in English: evidence from Event-Related Brain Potentials. *Cognitive Brain Research 7*.

Neville, H., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. F. (1991) Syntactically Based Sentence Processing Classes: Evidence from Event-Related Brain Potentials. *Journal of Cognitive Neuroscience 3*(2), 151-165.

Newman, A. J., Pancheva, R., Ozawa, K., Neville, H., & Ullman, M. T. (2004) An Event-Related fMRI Study of Syntactic and Semantic Violations. *Journal of Psycholinguistic Research 30*(3), 339-364.

Osterhout, L., McLaughlin, J., Kim, A., Greenwald, R., & Inoue, K. (2004) Sentences in the Brain: Event-Related Potentials as Real-Time Reflections of Sentence Comprehension and Language Learning. In M. Carreiras & J. Charles Clifton (Eds.), *The On-line Study of Sentence Comprehension: Eye-tracking, ERPs and Beyond* (pp. 271-307). New York / Hove: Psychology Press.

Petersson, K. M., Forkstam, C., & Ingvar, M. (2004) Artificial Syntactic Violations Activate Broca's Region. *Cognitive Science 28*(3), 383-407.

Pickering, M. J. (2004) Eyemovement and Semantic Composition. In M. Carreiras & J. Charles Clifton (Eds.), *The On-line Study of Sentence Comprehension: Eyetracking, ERPs and Beyond*. New York / Hove: Psychology Press.

Pickering, M. J., & van Gompel, R. P. G. (2006) Syntactic Parsing. In M. J. Traxler & M. A. Gernsbacher (Eds.), *The Handbook of Psycholinguistics* (2 ed.). London, UK: Elsevier.

Pinker, S. (1984) *Language Learnability and Language Development*. Cambridge, MA: Harvard University Press.

Pinker, S. (1999) *Words and Rules: The Ingredients of Language*. London: Weidenfeld & Nicolson.

Pollard, C., & Sag, I. A. (1994) *Head-driven phrase structure grammar*. Chicago, London: University of Chicago Press and CSLI Publications.

Prichett, B. L. (1992) *Grammatical Competence and Parsing Performance*. Chicago: University of Chicago Press.

Pritchett, B. L. (1992) *Grammatical Competence and Parsing Performance*. Chicago: University of Chicago Press.

Pulvermüller, F., Assadollahi, R., & Elbert, T. (2001) Neuromagnetic evidence for early semantic access in word recognition. *European Journal of Neuroscience 13*(1), 201-205.

Pynte, J., & Colonna, S. (2000) Decoupling Syntactic Parsing from Visual Inspection: The case of Relative Clause Attachment in French. In A. Kennedy, R. Radach, D. Heller & J. Pynte (Eds.), *Reading as a Perceptural Process*. Oxford, UK: Elsevier.

Quinn, D. (2001) The Influence of Implicit Prosody on Syntactic Parsing Preferences: Relative Clause Attachment in French and English.

Quinn, D., Abdelghany, H., & Fodor, J. D. (2000, March 30-April 1) *More Evidence of Implicit Prosody in Reading: French and Arabic Relative Clauses. (poster).* Paper presented at the CUNY Conference on Human Sentence Processing, La Jolla, CA.

Radford, A. (1997) *Syntactic: A Minimalist Introduction*. Cambridge: CUP.

Radford, A. (2001) *Foundations of Linguistics* (pp. 195): University of Essex.

Rayner, K., Carlson, M., & Frazier, L. (1983) The Interaction of Syntax and Semantics During Sentence Processing. *Journal of Verbal Learning and Verbal Behavior 22*, 358-374.

Rayner, K., & Duffy, S. A. (1988) On-line comprehension processes and eye movements during reading. In M. Daneman, G. E. MacKinnon & T. G. Waller (Eds.), *Reading research: Advances in theory and practice* (Vol. 6). London: Academic Press.

Rayner, K., Serno, S. C., Morris, R. K., Schmauder, A. R., & Clifton, C. (1989) Eye Movements and On-line Language Comprehension Processes. *Langauge and Cognitive Processes 4*(SI), 21-49.

Roland, D., Elman, J. L., & Ferreira, V. S. (2006) Why is *that*? Structural Prediction and Ambiguity Resolution in a Very Large Corpus of English Sentences. *Cognition 98*, 245-272.

Rouet, J.-F., & Passerault, J.-M. (1999) Analyzing Learner-hypermedia Interaction: An Overview of Online Methods. *Instructional Science 27*, 201-219.

Ruhlen, M. (1987) *A Guide to the World's Languages* (Vol. 1: Classification): Stanford University Press.

Runner, J. T., Sussman, R. S., & Tanenhaus, M. K. (2006) Processing Reflexives and Pronouns in Picture Noun Phrase. *Cognitive Science 30*(2), 193-241.

Schank, R. C. (1973) Identification of conceptualizations underlying natural language. In R. C. Schank & K. M. Colby (Eds.), *Computer Models of Thought and Language*. San Francisco: Freeman.

Seidenberg, M. S., & McClelland, J. L. (1989) A Distributed, Developmental Model of Word Recognition and Naming. *Psychology Review 96*(4), 523-568.

Shen, X. (2002) *Sentence Parsing in Different Languages: A Study of Relative Clause Attachment Ambiguity in Chinese.* Unpublished Master of Arts, the University of Essex, Colchester.

Slobin, D. I. (1971) *Psycholinguistics* (1 ed.). Glenview, London: Scott, Foresman and Company.

Snedeker, J., & Trueswell, J. (2003) Using Prosody to Avoid Ambiguity: Effects of Speaker Awareness and Referential Context. *Journal of Memory and Language 48*(1), 103-130.

Spivey-Knowlton, M. J. (1996) *Integration of Visual and Linguistic Information: Human Data and Model Simulations.* Unpublished Ph.D, Universit of Rochester, Rochester, N.Y.

Steinberg, D. D., Nagata, H., & Aline, D. P. (2001) *Psycholinguistics: Langauge, Mind and World* (2 ed.). Harlow / London / NYC / Reading, etc.: Longman.

Sturt, P. (2003) The Time-course of the application of Binding Constraints in the Reference Resolution. *Journal of Memory and Language 48*(3), 542-562.

Sturt, P., Scheepers, C., & Pickering, M. (2002) Syntactic Ambiguity Resolution after Initial Misanlaysis: The Role of Recency. *Journal of Memory and Language 46*(2), 371-390.

Su, I.-R. (2004) The Effects of Discourse Processing with Regard to Syntactic and Semantic Cues: A Competition Model Study. *Applied Psycholinguistics 25*, 587-601.

Sussman, R. S., & Sedivy, J. C. (2002) The Time-course of Processing Syntactic Dependencies: Evidence from Eye Movements. *Language and Cognitive Progresses 18*(2), 143-163.

Taraban, R., & McClelland, J. L. (1988) Constituent Attachment and Thematic Role Assignment in Sentence Processing: Influences of Content-based Expectations. *Journal of Memory and Language 27*, 597-632.

Townsend, D. J., & Bever, T. G. (2001) *Sentence Comprehension: The Integration of Habits and Rules*. Cambridge, Massachusetts / London, England: The MIT Press.

Traxler, M. J., & Pickering, M. (1996) Plausibility and the Processing of Unbounded Denpendencies: An Eye-Tracking Study. *Journal of Memory and Language 35*, 454-475.

Trueswell, J. (1996) The Role of Lexical Frequency in Syntactic Ambiguity Resolution. *Journal of Memory and Language 35*, 566-585.

Trueswell, J., Tanenhaus, M. K., & Garnesy, S. M. (1994) Semantic Influences on Parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language 33*(285-318).

van Gompel, R. P. G., & Pickering, M. J. (2001) Reanalysis in Sentence Processing: Evidence against Current Constraint-Based and Two-Stage Models. *Journal of Memory and Language 45*, 225-258.

van Gompel, R. P. G., Pickering, M. J., Pearson, J., & Liversedge, S. P. (2005) Evidence against Competition during Syntactic Ambiguity Resolution. *Journal of Memory and Language 52*(2), 284-307.

Vonk, W., & Cozijn, R. (2003) On the Treatment of Saccades and Regressions in Eye Movement Measures of Reading Time. In J. Hyönä, R. Radach & H. Deubel (Eds.), *The Mind's Eye: Cognitive and Aplied Aspects of Eye Movement Research* (pp. 293-311). Amstertam - Boston - London - New York - Oxford - Paris - San Diego - San Francisco - Singapore - Sydney - Tokyo: NH Elsevier.

Walter, M., Clifton, C., Frazier, L., Hemforth, B., Konieczny, L., & Seelig, H. (1999) *Prosodic and Syntactic Effects on Relative Clause Attachment in German and English (Poster).* Paper presented at the AMLaP, Edinburgh.

Watson, D., & Gibson, E. (2004) The Relationship between Intonational Phrasing and Syntactic Structure in Language Production. *Language and Cognitive Progresses 19*(6), 713-755.

West, W. C., & Holcomb, P. J. (2002) Event-related potentials during discourse-level semantic integration of complex pictures. *Cognitive Brain Research 13*, 363-375.

Weyerts, H., Penke, M., Münte, T. F., Heinze, H.-J., & Clahsen, H. (2002) Word Order in Sentence Processing: An Experimental Study of Verb Placement in German. *Journal of Psycholinguistic Research 31*(3), 211-268.

Whitney, P. (1998) *The Psychology of Language*. Boston / New York: Houghton Mifflin Company.

Zagar, D., Pynte, J., & Rativeau, S. (1997) Evidence for Early-closure Attachment on First-pass Reading Times in French. *Quarterly Journal of Experimental Psychology 50A*(2), 421-438.

Appendices

Appendix A: Instruction and Material for Study 1[69]

试验简介与合同

试验题目：语言处理——汉语的定语从句修饰问题
试验目的：本试验的目的在于通过研究自由阅读的速度来调查汉语言处理的过程，特别是定语从句的修饰问题。目前为止，世界上的许多语言，包括罗马语系，日耳曼语系，中东语系，阿尔泰语系等，都已经被研究过。将这项研究扩展到中文应用区域，有助于帮助发展现有的语言处理模式，并且加深理解语言的处理与应用过程。
试验过程：首先，请您根据自己的语言背景回答一些问题。然后阅读一些句子，并且根据句子的内容回答一个问题。请从所提供的选项中以最快的速度选出一个您认为最合适的答案。试验的结果只用于本项语言处理的研究。您所提供的一切资料都会被保密。您的背景问答将在 2004 年 8 月之后被销毁。而您的答案将会以数据形式运算与计算机中，保证不会保留与您个人有关的任何信息。
试验可能的风险：除了可能会觉得无趣之外，暂时没有观察到别的可能风险或者不快。
益处：首先，直接的益处是我们可以了解到汉语——这一是用人口最多的语言的处理方法。而且通过探索汉语言的独特的心理语言学特征，发展综合心理语言学的语言处理模式，有利于于此相关的别的学科，包括语言教学，计算机综合语言处理等等，最终有益于所有的人。
保密：您的参与内容将严格对试验者，与此项研究相关的心理学家，心理语言学家以外的人保密。您的试验结果将以数据的形式储存。任何带有您姓名的文件（包括这份合同）都将被分头储存和处理。您的姓名以及任何关于您的个人信息将不会出现在和这个试验相关的任何报告中。
试验中止：您的参与完全是自愿的，您可以在任何时候退出试验。
问题：您应该在签署合同前就试验进一步提出任何问题。
签名：我已经将试验在这份合同签署前对我的试验参与者解释清楚。我看懂了这个合同的所有内容，在签署一下姓名前，我已获得我的参与者的同意。

Instruction and Consent

TITLE OF PROJECT: Sentence Processing: RC attachment preference in Chinese

PURPOSE OF STUDY: The purpose of this study is to investigate sentence parsing, especially relative clause attachment in Chinese. Many languages have been studied recently, including Roman, Germanic, Middle East and Altaic languages. To expand the sentence parsing data to Chinese could help test and develop the current sentence

[69] The experiment was conducted in China for Chinese monolinguals only. Therefore, all the instructions and materials were in Chinese. However, for the purpose of this thesis, an English translation of the instructions is listed here as well.

processing models, as well as to learn more about how language is processed and produced.

PROCEDURES: This study will start with a short questionnaire about your language background. Then you will be introduced to a computer screen, where there will be a brief introduction about the experiment and instructions for how to continue. Then you can first acquaint yourself with the experiment through practice sentences to give you an idea of how the experiment is conducted. You will be introduced to the experiment part as you go on.

The sentences in the experiment have been cut into bits. You will be asked to press a button to carry on reading as soon as you understand the currently displayed bit. After some of the sentences, you will be asked some yes/no questions. You are expected to answer these by pressing a button according to the content of the text. Only your response time and the correctness of your answers will be recorded.

The experiment will be confidential and will only be used for the purpose of this sentence processing experiment. The questionnaire will be destroyed after August 2003. Your record will be saved as an Excel file without your name or other identifying characteristics.

REMUNERATION: Participants can receive 5-10 yuan after the experiments.

POTENTIAL RISKS AND ETHICAL CONSIDERATIONS: The main risk associated with the experiment is the possibility that some participants will feel slightly bored. No other risks are known by now.

BENEFITS: The direct benefit of this experiment would be to know how this language is processed, which is of particular interest since Chinese has the largest native speaker base of all world languages. Moreover, it is another chance to explore those special and unique Chinese psycholinguistic features, and develop the human language processing models so that they could benefit related fields and applications as language learning, language processing in a computer environment, etc., which may benefit everyone.

CONFIDENTIALITY: The information you give which is recorded will be kept strictly confidential, except as may be required by the law or professional guidelines for psychologists. All information will be identified by an identification code, not your name. Any form that requires your name (e.g., this consent form) will be stored separately from the other material. Your name or other identifying information will never be associated with any research reports or publications that use the results of your questionnaires or interviews.

WITHDRAWAL/PREMATURE COMPLETION: Your participation in the evaluation is voluntary, and you may discontinue at any time, without prejudice.

INVITATION TO ASK FURTHER QUESTIONS: You should ask any questions you have concerning this study before you sign this consent form.

CONSENT: I give my informed consent to participate in this study of **Sentence Processing: RC attachment preference in Chinese**. I have read and understand the consent form. Upon signing below, I will receive a copy of the consent form from the study investigator.

Materials

1. 某人开枪/打死了/和宾客们/一起站在/阳台上的/女演员的/男仆。（both）
   某人开枪/打死了/和妻子/一起站在/阳台上的/女演员的/男仆。（high）
   某人开枪/打死了/和丈夫/一起站在/阳台上的/女演员的/男仆。(low)
   Someone shot the male servant of the actress who was on the balcony <with the guests (both)> <with his wife (high)> <with her husband (low)>.
2. 他爱上了/正在吃/午餐的/歌星的/化妆师。（both）
   他爱上了/正在为/演员们化装的/歌星的/化妆师。（high）
   他爱上了/正在后台/等待表演的/歌星的/化妆师。（low）
   He has fallen in love with the designer of the singer who was <eating lunch (both)> <dressing up the actors (high)> <waiting to perform (low)>.
3. 十分钟后/公布的结果/鼓舞了/在操场上/等候的/学生们的/爸爸妈妈。（both）
   十分钟后/公布的结果/鼓舞了/一直陪在/孩子身边的/学生们的/爸爸妈妈。（high）
   十分钟后/公布的结果/鼓舞了/刚从考场/出来的/学生们的/爸爸妈妈。(low)
   Ten minutes later, the results encouraged the parents of the students who <was waiting in the playground (both)> <was waiting next to their children (high)> <just finished their examinations>.
4. 大家都/不喜欢/那个/模棱两可的/富翁的/演讲。（both）
   大家都/不喜欢/那个/内容空洞的/富翁的演讲。（high）
   大家都/不喜欢/那个/富得流油的/富翁的/演讲。（low)
   Nobody likes the speech by[70] the rich man who (that) was <vague (both)> <abstract (high)> < filthy rich (low)>.
5. 昨天我/看到了/等在/新酒楼前的/老朋友的/轿车。（both）
   昨天我/看到了/停在楼下的/老朋友的/轿车。（high）
   昨天我/看到了/在超市工作的/老朋友的/轿车。(low)
   Yesterday I saw the car of an old friend who was <waiting in front of the new restaurant (both)> <parked downstairs (high)> <working in a supermarket (low)>.
6. 人们希望/亲眼目睹/传说中/所向无敌的/魔法师的/宝剑。（both）
   人们希望/亲眼目睹/传说中/不能出鞘的/魔法师的/宝剑。（high）
   人们希望/亲眼目睹/传说中/能腾云驾雾的/魔法师的/宝剑。(low)
   People hope to see the sword of the magician who (that) <had no enemy (both)> <should not be taken out of the sheath (high)> <could fly[71] (low)>.
7. 村民们/仍然非常/服从/号召大家/抗洪抢险的/村长的/口令。（both）
   村民们/仍然非常/服从/被村委会/修正过了的/村长的/口令。（high）

[70] In Chinese it is ok to use *de* in this occasion.
[71] The modification in Chinese can only be used to describe humanôs supernatural power.

村民们/仍然非常/服从/依法处置/亲生儿子的/村长的口令。（low）

The villagers still obeyed the order of the head of the village who (that) <called on people to fight against the flood (both)> <was modified by the village committee (high)> <sentenced his own son based on the law (low)>.

8. 大家在/背地里/偷偷议论那/嫌贫爱富的/丫头的/父母。（both）
   大家在/背地里/偷偷议论那/拆散女儿恋爱的/丫头的/父母。（high）
   大家在/背地里/偷偷议论那/不肯穿裙子的/丫头的/父母。(low)

   Everyone is gossiping about the parents of the girl who <was/were[72] ill-minded towards money (both)> <forbade their daughter to date (high)> <refused to wear a skirt (low)>.

9. 教务处长/又看了一眼/这个/桀骜不驯的/女学生的/男朋友。（both）
   教务处长/又看了一眼/替女友/顶罪的/女学生的/男朋友。（high）
   教务处长/又看了一眼/已经怀孕/的女学生的男朋友。(low)

   The education minister cast another look at the boyfriend of the female student who <was obstinate and unruly (both)> <sacrificed himself for his girlfriend (high)> <was pregnant (low)>

10. 其实我们/并不喜欢/这些/来自宫廷的/御厨的/菜肴。（both）
    其实我们/并不喜欢/这些/色泽鲜艳/香味扑鼻的/御厨的/菜肴。（high）
    其实我们/并不喜欢/这些/手艺高明的/御厨的/菜肴。(low)

    In fact, we do not like the dish of the chef who (that) <was from the palace (both)> <was colourful and aromatic (high)> <cooked very well (low)>.

11. 孩子们都/十分害怕/那说话粗俗的/学长的/家里人。（both）
    孩子们都/十分害怕/那反对孩子/出门玩耍的/学长的/家里人。（high）
    孩子们都/十分害怕/那勤奋学习的/学长的/家里人。（low）

    The children were afraid of the family members of the senior student who <spoke in an offensive manner (both)> <forbade their child to go out to play (high)> <studied hard low)>.

12. 我读过/那本获得/国家金奖的/摄影师的/游记。（both）
    我读过/那本关于/西藏风情的/摄影师的/游记。（high）
    我读过/那本关心/生态问题的/摄影师的/游记。（low）

    I have read the travel book by that photographer who (that) <won the golden prize (both)> <was about the views in Tibet (high)> <thought of the ecology problems (low)>.

13. 选民们/纷纷不满/那个非常/种族歧视的/政治家的/演说。（both）
    选民们/纷纷不满/那个又长/又枯燥的/政治家的/演说。（high）
    选民们/纷纷不满/那个/长着小胡子的/政治家的/演说。（low）

    Electors were not happy with the speech by the politician who (that) <was very racist (both)> <was long and boring (high)> <grew moustache (low)>.

14. 参观者/正在讨论/那极富/讽刺意味/的/画家/的/新作。（both）
    参观者/正在讨论/即将出版/的/画家/的/新作。（high）
    参观者/正在讨论/擅长国画/的/画家/的/新作。(low)

    Visitors were talking about the new piece of work by the painter who (that) <was very sarcastic (both)> <was about to be published (high)> <was very good at traditional Chinese painting (low)>.

15. 记者采访了/又一次/获得/产品专利的/学生们的/辅导员。（both）
    记者采访了/培养过/在三届全国/大赛中/得奖的/学生们的/辅导员。（high）

[72] There is no number mark in Chinese.

记者采访了/被保送到/重点中学/学习的/学生们的/辅导员。(low)

The journalist interviewed the councillor of the students who <had received another patent for their invention (both)> <educated top prize winners for three years continuously (high)> <was recommended for admission to a key school for higher education (low)>.

16. 大家/终于认出了/那天/在商场/买东西的/女学生的/老爹。（both）
大家/终于认出了/和年轻妻子/在一起的/女学生的/老爹。（high）
大家/终于认出了/参加/选美比赛的/女学生的/老爹。(low)

People finally recognised the father of the female student who <was shopping in the mall (both)> <was with a young wife (high)> <won the beauty contest (low)>.

17. 这乞丐/一不小心/得罪了/坐在轿子里的/长官的/内人。（both）
这乞丐/一不小心/得罪了/正在化妆的/长官的/内人。（high）
这乞丐/一不小心/得罪了/刚刚上任的/长官的/内人。(low)

The beggar incautiously offended the wife of the officer who <was sitting in the sedan chair (both)> <was putting on her make-ups (high)> <just took up his duty (low)>

18. 瞎爷爷/已经听出了/细腻柔和的/姑姑的/嗓音。（both）
瞎爷爷/已经听出了/因为吸烟/而沙哑的/姑姑的/嗓音。（high）
瞎爷爷/已经听出了/在国外呆了/几十年的/姑姑的/嗓音。(low)

The old blind man recognised the voice of the aunt who (that) <was very gentle and feminine (both)> <became hoarse because she smoked heavily (high)> <was abroad for tens of years (low)>.

19. 他折断了/靠在/墙角的/侦探的/拐杖。（both）
他折断了/被扔在/地上的/侦探的/拐杖。（high）
他折断了/正准备/出门的/侦探的/拐杖。(low)

He broke the stick of the detective who (that) <was leaning against the wall (both)> <was thrown on the floor (high)> <was about to go out (low)>.

20. 这个人/一下就/闻到了/令他/难以拒绝的/牛肉的/香味。（both）
这个人/一下就/闻到了/空气中/传来的/牛肉的/香味。（high）
这个人/一下就/闻到了/被炖了/三个小时的/牛肉的/香味。(low)

This person soon smelt the aroma of the beef that <was so irresistible (both)> <filled the room (high)> <had been stewed for three hours (low)>.

21. 同学们/仍不明白/看似/高人一筹的/助教的/解析。（both）
同学们/仍不明白/内容/晦涩难懂的/助教的/解析。（high）
同学们/仍不明白/正在研究/这个课题的/助教的/解析。(low)

The students still do not understand the solution by the tutor who (that) <was seemingly superior (both)> <was incredibly incomprehensible (high)> <was studying this topic (low)>.

22. 参加/婚礼的人/极力赞美/高贵大气的/新娘的/礼服。（both）
参加/婚礼的人/极力赞美/款式新颖的/新娘的/礼服。（high）
参加/婚礼的人/极力赞美/手拿鲜花的/新娘的/礼服。(low)

The wedding guests sang high praise for the dress of the bride who (that) <was elegant and noble (both)> <was designed in a novel way (high)> <was holding a bouquet (low)>.

23. 徒弟们/还是无法/记住/那些/啰里啰唆的/师傅的/教诲。（both）
徒弟们还是无法/记住/曾经反复/在耳边的/师傅的/教诲。（high）
徒弟们/还是无法/记住/年老体迈的/师傅的/教诲。(low)

The pupils still cannot remember the edification by the mentor who (that) <was very repetitive (both)> <was repeated to them every day (high)> <was very old and weak (low)>.

24. 今天我/遇到了/在学校/教书的/舅舅的/老同学。（both）

今天我/遇到了/跟我姨妈/结婚的/舅舅的/老同学。（high）

今天我/遇到了/家里年纪/最小的/舅舅的/老同学。(low)

Today I met the old classmate of my uncle who <was teaching in a school (both)> <married to my auntie[73] (high)> <was the youngest in my family (low)>.

Language-background Questionnaire

语言背景调查

姓名 (Name)：_______________

年级 (Grade, if applicable)：_______________

出生年月(D.O.B)：___________

性别 (Gender)：_______________

1. 您的第二语言是 (Second language)：__________________________
2. 您认为您的第二语言的水平为 (Second language level)：______________
3. 您在课/业余使用什么语言（请选择）(What languages do you normally use)__________
   a. 普通话 (Mandarin) b.南京话 (Nanjing Dialect) c.上海话 (Shanghai Dialect) d.广东话 (Cantonese) 5.其它（请注明）(others)______________
4. 您的英语考试成绩多半在（请选择）(average English score)___________
   a. 0-50 b. 50.5-70 c.70.5-85 d.85.5-100
5. 您平时爱玩电动游戏么？(Are you a game player?) _____________________
6. 您平时爱读外文原版读物，爱看外文原版影视节目么？(Do you like reading foreign books and watching foreign movies?) ____
7. 您爱好阅读么？(Do you like reading?) _____

[73] The words *auntie* / *uncle* in this sentence refer to the siblings of someone's mother.

Appendix B: Instruction and Material for Study 2[74]

Instruction and Consent

Dear everyone,

I appreciate how busy you are but would be very grateful if you would take a moment to read this email. My name is Rachel Shen and I am a Psychology Postgraduate student at Exeter University, currently working on an MPhil/PhD in Psychology.

As part of my research, I am carrying out some an experiment on sentence processing. I would therefore appreciate it if you could spare about 15 minutes to take part in this online academic research survey. This research study is, of course, entirely voluntary, but I am hoping to get a good response, as it is important to have a substantial body of data to analyze.

If you agree to take part, the process is very straightforward:
You simply click the web link below and complete the questionnaire by clicking on the choice you want to make.

You are guaranteed absolute anonymity; no attempt will be made to identify your identity and you will not need to fill in your name on the questionnaire. Your institution is also guaranteed anonymity in the research study. All the data received will be stored on a protected and secured database, which can only be accessed by my supervisor and me.

This survey has a life span of 8 weeks starting from the moment you received this information package in your email box. The survey closes on the 6th of March 2004.

Thank you very much for taking the time to read this message

Any correspondence should be directed to Xingjia Shen at x.shen@ex.ac.uk

Materials

1. The sister of/next to the singer who danced just now is drinking some orange juice in the corner/ is flattered by this handsome young man /is a pretty girl.
   Who danced just now? ( )
   a. The singer b. the sister c. not mentioned
2. The friend of/next to the girl who smokes is now driving a car without a legal

[74] The experiment was conducted in both Chinese and English. However, only the English instructions are presented here. The materials will be presented in both English and Chinese.

licence / was caught by the police for dealing drugs / was an orphan.
Who smokes? ( )
a. The friend b. the girl c. the one who is not driving without a legal licence

3. The wife of/next to the professor who likes Chinese food has applied for a cooking course / has been invited to the ball / is in her 50s.
Who likes Chinese food? ( )
a. The wife b. the professor c. someone in the cooking course
4. The classmate of/next to the girl who always lies has decided not to say a word this time / was found hidden in a small chamber in the hall / is a good swimmer.
Who always lies? ()
a. The girl b. the classmate c. nobody
5. The man of/next to the woman who likes to talk behind others just slaughtered a pig / was caught stealing in the shop / is a slaughterer.
Who likes to talk behind others? ( )
a. The woman b. the man c. others
6. The father of/next to the artist who devoted loads of money to the poor area enjoys planting / is chosen to be the representive of the village / is the representive of the village.
Who devoted loads of money to the poor area? ( )
a. The father b. the artist c. both
7. The wife of/next to the boss who is always hard on others was just shopping in the street / was hated by the employees / was a gambler.
Who is always hard on the others? ( )
a. The wife b. the boss c. both
8. The kid of/next to the woman who did the stupid thing is now crying by the road/ was saved from the water / is the champion of chess in local area.
Who did the stupid thing? ( )
a. The kid b. the woman c. both
9. The housekeeper of/next to the rich woman who was very rude to others had a very bad cold / was robbed yesterday / was actually an illegal child of this family.
Who was very rude? ( )
a. The housekeeper b. the rich woman c. nobody
10. The secretary of/next to the politician who is a racist is preparing for the election / will be fired / is from Germany.
Who is a racist? ( )
a. The secretary b. the politician c. the elected person
11. The boyfriend of/next to the girl who has very long hair painted her portrait / was sent to France to study art / is an artist.
Who has very long hair? ( )
a. The boyfriend b. the girl c. none
12. The female manager of/next to the actress who just dumped her boyfriend decided to quit her job / is asked out by a millionaire / is a busy woman.
Who was just dumped by her boyfriend? ( )
a. The female manager b. the actress c. nobody
13. The husband of/next to the daughter who is about to leave is now packing his bag /

was moved to tears / is a soldier.
Who is about to leave? ( )
a. The husband b. the daughter c. the touching guy

14. The grandfather of/next to the boy who cooks really good food once fought in the battle of Iwo jima / was once forced to join the army in the WWII / was a military chef.
Who cooks really good food? ( )
a. The grandfather b. the boy c. the soldier

15. The parents of/next to the students who was just very excited all went home / is now very shocked / were the parent-committee members.
Who was very excited just now? ( )
a. The parents b. the students c. both

16. The assistant of/next to the journalist who had been to Tibet will publish his own book / is promoted to be the editor of a local newspaper / is the author of the book.
Who had been to Tibet? ( )
a. The publisher b. the journalist c. the assistant

17. The daughter of/next to my friend who is working in the supermarket just graduated from high school / was offered a place in a university / is a high school graduate.
Who is working in the supermarket? ( )
a. The daughter b. the friend c. the daughterǿs friend

18. The maid of/next to the wizard who killed the dragon is holding a candle / was told to bring a candle / was from India.
Who killed the dragon? ( )
a. The maiden b. the wizard c. both

19. The pupil of/next to the monk who talks endlessly is cleaning the yard / was told to have lunch /is a Buddhist.
Who talks endlessly? ( )
a. The monk b. the pupil c. both

20. The friend of/next to the aunt who just came back from the UK is wearing a fur coat / was invited to a concert / is a scientist.
Who just came back from the UK? ( )
a. The friend b. the aunt c. the fur dealer

21. The child of/next to the researcher who is studying this topic wants to play football / was asked to play football / was a math genius.
Who is studying this topic? ( )
a. The kid b. the researcher c. High school teacher

22. The client of/next to the businessman who looks pale just came out of the hospital / was told to pay a big amount of money / was the CEO of the competing company.
Who looks pale? ( )
a. The client b. The businessman c. nobody

23. The nurse of/next to the patient who wears a pink wig is coming down the corridor / was called by the surgeon / is a very nice person.
Who wears a pink wig? ( )
a. The nurse b. The patient c. a doctor

24. The landlord of/next to the writer who has a nasty nose went to the seashore this morning / was asked to go to the seashore / is a very annoying person

Who grows a nasty nose? ( )
a. The landlord b. the writer c. both

A. Active Condition[75]

1. 刚才跳舞的女歌手（旁边）的姐姐正在角落里喝橙汁。
2. 会抽烟的女孩（旁边）的同伴准备去玩卡兵车。
3. 喜爱吃中餐的教授（身边）的妻子报名参加了厨艺培训。
4. 常常说谎的女孩子（旁边）的同学决定这次一句话也不说。
5. 爱在背地里说人坏话的女人（身边）的男人刚刚宰了一头猪。
6. 为贫困山区捐赠巨额财产的艺术家（身边）的父亲喜欢养花种树。
7. 待人刻薄的老板（身边）的老婆正在街上买东西。
8. 做了蠢事的女人（旁边）的孩子正在路边哭泣。
9. 看似高贵却蛮不讲理的阔太太（身边）的管家得了流感。
10. 有种族歧视的政治家（身边）的秘书正在起草大选文件。
11. 留着长头发的女孩（身边）的男朋友正在画了一幅的肖像画。
12. 刚抛弃了男朋友的女演员（身边）的经纪人决定辞去工作。
13. 即将离去的女儿（身边）的丈夫正在打包行李。
14. 烧得一手好菜的男孩（身边）的爷爷参加过硫黄岛战役。
15. 刚才还很兴奋的学生（身边）的家长们都回家了。
16. 曾经去过西藏的旅行家（身边）的助手将要出版一本游记。
17. 在超市工作的朋友（身边）的女儿刚刚高中毕业。
18. 杀死了恶龙的巫师（身边）的女仆点燃了一根蜡烛。
19. 一直说个不停的和尚（身边）的徒弟正在清扫庭院。
20. 刚从英国回来的姨妈（身边）的朋友穿了件草皮大衣。
21. 研究这个课题的研究员（身边）的孩子想去踢足球。
22. 看上去很苍白的商人（身边）的客户刚刚出院。
23. 带着粉红色假发的病人（身边）的护士从起身走了。
24. 长着奇怪鼻子的作家（旁边）的房东今早去了海边。

B. Passive condition

25. 刚才跳舞的女歌手（身边）的姐姐被一个帅气的年轻人恭维了一番。
26. 会抽烟的女孩（旁边）的同伴被警察以贩毒为由扣押。
27. 喜爱吃中餐的教授（身边）的妻子被邀请去参加舞会。
28. 常常说谎的女孩子（旁边）的同学被人发现躲在阁楼里面。
29. 爱在背地里说人坏话的女人（身边）的男人被人当成小偷抓住。
30. 为贫困山区捐赠巨额财产的艺术家（身边）的父亲被选为村代表。
31. 待人刻薄的老板（身边）的老婆被所有员工憎恨。
32. 做了蠢事的女人（旁边）的孩子刚被人从水里救起。
33. 看似高贵却蛮不讲理的阔太太（旁边）的管家昨天被打劫了。
34. 有种族歧视的政治家（身边）的秘书被炒了鱿鱼。
35. 留着长头发的女孩（身边）的男朋友被选送到法国去学习艺术。
36. 刚抛弃了男朋友的女演员（身边）的经纪人被一个百万富翁追求。
37. 即将离去的女儿（身边）的丈夫被感动得泪流满面。
38. 烧得一手好菜的男孩（身边）的爷爷曾被强迫服兵役打仗。

[75] The Chinese material is the direct translation from English in Appendix B.

39. 刚才还很兴奋的学生（身边）的家长们都被吓到了。
40. 曾经去过西藏的旅行家（身边）的助手被提升为一个当地杂志的编辑。
41. 在超市工作的朋友（身边）的女儿被保送到了一所大学。
42. 杀死了恶龙的巫师（身边）的女仆被释放自由了。
43. 一直说个不停的和尚（身边）的徒弟被叫去吃午饭了。
44. 刚从英国回来的姨妈（身边）的朋友被邀请去听音乐会。
45. 研究这个课题的研究员（身边）的孩子被人叫去踢足球了。
46. 看上去很苍白的商人（身边）的客户被告知要付一大笔钱。
47. 带着粉红色假发的病人（身边）的护士被急诊室叫去了。
48. 长着奇怪鼻子的作家（旁边）的房东今早被人叫去海边。

C. *Be* Condition

49. 刚才跳舞的女歌手（身边）的姐姐十个漂亮女孩。
50. 会抽烟的女孩（旁边）的同伴曾是个孤儿。
51. 喜爱吃中餐的教授（身边）的妻子大约 50 岁左右。
52. 常常说谎的女孩子（旁边）的同学是个游泳健将。
53. 爱在背地里说人坏话的女人（身边）的男人是个屠夫。
54. 为贫困山区捐赠巨额财产的艺术家（身边）的父亲是村代表。
55. 待人刻薄的老板（身边）的老婆是个赌徒。
56. 做了蠢事的女人（身边）的孩子是当地的象棋冠军。
57. 看似高贵却蛮不讲理的阔太太（身边）的管家是这家的私生子。
58. 有种族歧视的政治家（身边）的秘书来自德国。
59. 留着长头发的女孩（身边）的男朋友是一位艺术家。
60. 刚抛弃了男朋友的女演员（身边）的经纪人是一个女强人。
61. 即将离去的女儿（身边）的丈夫是一个战士。
62. 烧得一手好菜的男孩（身边）的爷爷曾是战地厨师。
63. 刚才还很兴奋的学生（身边）的家长们都是家长学校成员。
64. 曾经去过西藏的旅行家（身边）的助手是这本书的作者。
65. 在超市工作的朋友（身边）的女儿是一个高中毕业生。
66. 杀死了恶龙的巫师（身边）的女仆来自于印度某地。
67. 一直说个不停的和尚（身边）的徒弟相信佛教。
68. 刚从英国回来的姨妈（身边）的朋友是一位科学家。
69. 研究这个课题的研究员（身边）的孩子是一个小天才。
70. 看上去很苍白的商人（身边）的客户是竞争公司的总裁。
71. 带着粉红色假发的病人（身边）的护士是一个很好的人。
72. 长着奇怪鼻子的作家（旁边）的房东是一个很讨厌的人。

Appendix C: Instruction and Material for Study 3

Instruction and Consent

TITLE OF PROJECT: Crosslinguistic Study of Topicalized RC Attachment Preference: Position Effect in English and Chinese Relative Clause

PURPOSE OF STUDY: The purpose of this study is to investigate the sentence processing strategy. There have been many studies on English language, but more details needs to be explored. The current study is a component of the overall English RC processing study, on the resolution of RC ambiguity resolution at preverbal position.

PROCEDURE: You will be asked to read a sentence on the screen, and then press the ROUND button to carry on reading another page when you finished the previous one. After some of the sentences, you will be asked some yes/no questions. You are expected to answer these by pressing the RIGHT button to the ñYesò according to the content of the text, and LEFT to the ñNoò answers. Between sentences, there will be a black dot in the centre of the screen. You should focus on the dot, and press the ROUND button to continue.

The experiment will be confidential and will only be used for the purpose of this sentence processing experiment. The questionnaire will be destroyed when the study is over. Your record will be saved as an .EDF file without your name or other identifying characteristics.

REMUNERATION: Participants can receive £2 or half of a credit after the experiments.

POTENTIAL RISKS AND ETHICAL CONSIDERATIONS: The main risk associated with the experiment is the possibility that some participants will feel slightly bored. The helmet might make some people feel uncomfortable. No other risks are known by now.

BENEFITS: The direct benefit of this experiment would be to know how this language is processed, which is of particular interest since Chinese and English has the largest native speaker base of all world languages. Moreover, it is another chance to explore those special and unique Chinese and English psycholinguistic features, and develop the human language processing models so that they could benefit related fields and applications as language learning, language processing in a computer environment, etc., which may benefit everyone.

CONFIDENTIALITY: The information you give which is recorded will be kept strictly confidential, except as may be required by the law or professional guidelines for psychologists. All information will be identified by an identification code, not your name. Any form that requires your name (e.g., this consent form) will be stored separately from the other material. Your name or other identifying information will never be associated with any research reports or publications that use the results of your questionnaires or interviews.

WITHDRAWAL/PREMATURE COMPLETION: Your participation in the evaluation is voluntary, and you may discontinue at any time, without prejudice.

INVITATION TO ASK FURTHER QUESTIONS: You should ask any questions you have concerning this study before you sign this consent form.

CONSENT: I give my informed consent to participate in this study of **RC attachment preference in Chinese and English at preverbal position**. I have read and understand the consent form. Upon signing below, I will receive a copy of the consent form from the study investigator.

Materials

Version I

1. The male servant of the actress who was on the balcony with the guests was shot to death.
2. The designer of the pop star who is now having lunch was rewarded as ñthe best of the yearò.
3. The parents of the students who had been waiting in the yard were now led to the hall to rest.
4. The cover of the book that portrays the countryside was torn off by somebody.
5. I saw the driver of the boss who was eating in the restaurant yesterday.
6. The stone hit the tail of the green snake that was moving just now near a hole.
7. The villagers respect the son of the village-head who called on for fighting against flood.
8. Everybody was talking about the parents of the boy who only did things for money all the time
9. The boyfriend of the female student who covered her fault was driven out of the school.
10. The taste of the radish that mixed with other spices was savoured by the food inspector.
11. The grandmother of the senior student who always forbade kids partying was called by the police.
12. The assistant of the photographer who helps in the routines of the studio was offered an editorôs position in a magazine.
13. The voters were not happy with the opening speech of the conference that the chairman gave that day.
14. Readers are discussing the prologue of the photo album that a famous artist wrote before he died.
15. A reporter interviewed the teacher of the students who was preparing for the next lecture.
16. All of us now recognized the father of the young lady who was with a very young mistress in the street.
17. The wife of the mayor who just undertook his duty was robbed by a group of gangs.
18. The song of the choirs that has performed around the country was recorded in the tape.
19. The suspect of the detective who was just hired in this case was killed by the mafia.
20. The nice smell of the beef that has been on the stove for more than three hours was soon covered by some stinks.

21. People here all liked the granddaughter of the scientist who is studying this topic recently.
22. A guest found the box of the instrument that played the music in the concert.
23. Another company has copied the package design of the food that was baked according to traditional recipe.
24. I met the old classmate of my uncleôs who was the youngest son in motherôs family from my momôs side.

Version II

1. The male servant of the actress who was on the balcony with his wife was shot to death.
2. The designer of the pop star who is dressing up the performers was rewarded as ñthe best of the yearò.
3. The parents of the students who had been waiting for their children were now led to the hall to rest.
4. The cover of the book that has the authorôs signature on was torn off by somebody.
5. I saw the driver of the boss whose job was transporting the client for the company.
6. The stone hit the tail of the green snake that swung with the body near the hole.
7. The villagers respect the son of the village-head who came to visit his family from the city last week.
8. Everybody was talking about the parents of the boy who abandoned the kid mercilessly.
9. The boyfriend of the female student who became pregnant was driven out of the school.
10. The taste of the radish that the chef just took from the field was savoured by the food inspector.
11. The grandmother of the senior student who always studies hard was called by the police.
12. The assistant of the photographer who had many great exhibitions was offered an editorôs position in a magazine.
13. The voters were not happy with the opening speech of the conference that more than 10,000 people attended that day.
14. Readers are discussing the prologue of the photo album that will soon be published next year.
15. A reporter interviewed the teacher of the students who were guaranteed a position in the university.
16. All of us/ now recognized the father of the young lady who won the beauty competition this summer.
17. The wife of the mayor who was sitting in the carriage was robbed by a group of gangs.
18. The song of the choirs that is very popular here was recorded in the tape.
19. The suspect of the detective who wore black sunglasses was killed by the mafia.
20. The nice smell of the beef that he liked very much was soon covered by some stink.
21. People here all liked the granddaughter of the scientist who had long white beard in the department.
22. A guest found the box of the instrument that someone placed at the corner of the room.
23. Another company has copied the package design of the food that originated in Scotland in 1878.
24. I met the old classmate of my uncle who is teaching in the college on history.

Version III

1. The male servant of the actress who was on the balcony with her husband was shot to death.
2. The designer of the pop star who is performing in the stage was rewarded as ñthe best of the

yearò.

3. The parents of the students who just finished the big examination were now led to the hall to rest.
4. The cover of the book that tells the most boring story was torn off by somebody.
5. I saw the driver of the boss who chairs the board meeting in this company.
6. The stone hit the tail of the green snake that likes eating rats near a hole.
7. The villagers respect the son of the village-head whom they elected recently.
8. Everybody was talking about the parents of the boy who likes wearing skirts ever since young.
9. The boyfriend of the female student who never obeys regulations was driven out of the school.
10. The taste of the radish that everybody disliked was savoured by the food inspector.
11. The grandmother of the senior student who always uses rude languages was called by the police.
12. The assistant of the photographer who had been to Tibet was offered an editorôs position in a magazine.
13. The voters were not happy with the opening speech of the conference that had strong racist colour that day.
14. Readers are discussing the prologue of the photo album that contains ironic meanings towards the dictator.
15. A reporter interviewed the teacher of the students who won many prizes in the past two years.
16. All of us now recognized the father of the young lady who bought many things in that shopping mall.
17. The wife of the mayor who always wears heavy makeup was robbed by a group of gangs.
18. The song of the choirs that was very pure and smooth was recorded in the tape.
19. The suspect of the detective who was preparing to escape was killed by the mafia.
20. The nice smell of the beef that filled the room just now was soon covered by some stink.
21. People here all liked the granddaughter of the scientist who just went to elementary school in the autumn.
22. A guest found the box of the instrument that had been open and empty for a while.
23. Another company has copied the package design of the food that this advertisement agency has proposed recently.
24. I met the old classmate of my uncle who married his sister 10 years ago.

Version IV

1. Someone shot the male servant of the actress who was on the balcony with the guests that day.
2. He fell in love with the designer of the pop star who is having lunch now.
3. The announced result disappointed the parents of the students who had been waiting in the yard.
4. Somebody has torn the cover of the book that portrays the countryside of Italy.
5. The driver of the boss who was eating in the restaurant was fired today.
6. The tail of the green snake that was moving just now was hit by a stone.
7. The son of the village-head who called on for fighting against flood was arrested for no reason.
8. The parents of the boy who only did things for money were questioned by the others.
9. The school has expelled the boyfriend of the female student who covered her fault in that incident.
10. The food inspector savoured the taste of the radish that mixed with other spices at the dinner.
11. The police called the grandmother of the senior student who always forbade kids partying at home.

12. A magazine offered a position to the assistant of the photographer who helps in the routines of the studio.
13. The opening speech of the conference that the chairman gave was interrupted by the audience.
14. The prologue of the photo album that will soon be published was now translated to many languages.
15. The teacher of the students who was preparing for the next lecture was interviewed by a journalist.
16. The father of the young lady who was with a very young mistress was pointed to be the CEO of the company.
17. A group of gangs robbed the wife of the mayor who just undertook his duty last month.
18. I recorded the song of the choirs that has performed around the country.
19. The mafia has killed the suspect of the detective who was just hired in this case.
20. The stink soon covered the nice smell of the beef that has been on the stove for more than three hours.
21. The granddaughter of the scientist who is studying this topic was chosen to be the representative of the group.
22. The box of the instrument that played the music in the concert was stained with some cake
23. The package design of the food that was baked according to traditional recipe was illegally copied by another company.
24. The old classmate of my uncleȸs who was the youngest son in motherȸs family has been promoted to be a professor.

Version V

1. Someone shot the male servant of the actress who was on the balcony with his wife that day.
2. He fell in love with the designer of the pop star who is dressing up the performers now.
3. The announced result disappointed the parents of the students who had been waiting for their children for the whole afternoon.
4. Somebody has torn the cover of the book that has the authorȸs signature on according to the librarian.
5. The driver of the boss whose job was transporting the clients was fired today.
6. The tail of the green snake that swung with the body was hit by a stone.
7. The son of the village-head who came to visit his family from the city was arrested for no reason.
8. The parents of the boy who abandoned the kid mercilessly were questioned by the others.
9. The school has expelled the boyfriend of the female student who became pregnant 3 months ago.
10. The food inspector savoured the taste of the radish that the chef just took from the field.
11. The police called the grandmother of the senior student who studies very hard constantly.
12. A magazine offered a position to the assistant of the photographer who had many great exhibitions during the past 10 years.
13. The opening speech of the conference that more than 10,000 people attended was interrupted by the audience.
14. The prologue of the photo album that a famous artist wrote was now translated to many languages.
15. The teacher of the students who were guaranteed a position in the university was interviewed by

a journalist.

16. The father of the young lady who won the beauty competition was pointed to be the CEO of the company.
17. A group of gangs robbed the wife of the mayor who was sitting in the carriage near the gate.
18. I recorded the song of the choirs that is very popular in this area.
19. The mafia has killed the suspect of the detective who wore black sunglasses all the time.
20. The stink covered the nice smell of the beef that he liked a lot in no time.
21. The granddaughter of the scientist who had long white beard was chosen to be the representative of the group.
22. The box of the instrument that someone placed at the corner of the room was stained with some cake.
23. The package design of the food that originated in Scotland was illegally copied by another company.
24. The old classmate of my uncleęs who is teaching in the college has been promoted to be a professor.

Version VI

1. Someone shot the male servant of the actress who was on the balcony with her husband that day.
2. He fell in love with the designer of the pop star who is performing in the stage right now.
3. The announced result disappointed the parents of the students who just finished the big examination five minutes ago.
4. Somebody has torn the cover of the book that tells the most boring story in the world.
5. The driver of the boss who chairs the board meeting was fired today.
6. The tail of the green snake that likes eating rats was hit by a stone.
7. The son of the village-head who they just elected was arrested for no reason.
8. The parents of the boy who likes wearing skirts were questioned by the others.
9. The school has expelled the boyfriend of the female student who never obeys regulations as an example to the others.
10. The food inspector savoured the taste of the radish that no one liked in the restaurant.
11. The police called the grandmother of the senior student who always uses rude languages to other people.
12. A magazine offered a position to the assistant of the photographer who had been to Tibet many times.
13. The opening speech of the conference that had strong racist colour was interrupted by the audience.
14. The prologue of the photo album that contains ironic meanings was now translated to many languages.
15. The teacher of the students who won many prizes was interviewed by a journalist.
16. The father of the young lady who bought many things in that shopping mall was pointed to be the CEO of the company.
17. A group of gangs robbed the wife of the mayor who always wears heavy makeup near the gate.
18. I recorded the song of the choirs that was very pure and smooth in the concert.
19. The mafia has killed the suspect of the detective who was preparing to escape to another country.
20. The stink soon covered the nice smell of the beef that filled the room just now.

21. The granddaughter of the scientist who just went to elementary school was chosen to be the representative of the group.
22. The box of the instrument that had been open and empty for a while was stained with some cake.
23. The package design of the food that this advertisement agency has proposed was illegally copied by another company.
24. The old classmate of my uncleǿs who married his sister has been promoted to be a professor.

Version I[76]

1. 昨天晚上，和宾客们一同站在阳台上的女演员的男仆被人开枪打死了。
2. 工作室里，正在吃午餐的歌星的化妆师被授予“年度最佳”称号。
3. 在操场上等候良久的学生们的父母被人领到小礼堂休息。
4. 图书员说描绘乡村风情的小说的封面被人撕掉了。
5. 昨天中午我看见正在餐厅吃饭的老板的司机穿了件红衬衫。
6. 那石头砸到刚才还在游动的绿蛇的尾巴把蛇惊吓了。
7. 村民们非常尊敬号召大家抗洪的村长的儿子因为他是硕士。
8. 所有人都在谈论那钻在钱眼里的小男孩的父母在法庭上所言。
9. 替女友顶罪的那个女生的男朋友被学校开除了。
10. 和其他佐料混在一起的萝卜的味道被美食家品出来了。
11. 那禁止孩子们哄闹的学长的祖母被警察叫去了。
12. 负责照相馆日常事务的摄影师的助手刚被聘为一家杂志社的编辑。
13. 选民们非常不满主席发表的招待会的开幕词因为内容太陈旧了。
14. 参观者正在讨论这著名艺术家撰写的影集的前言并且哀悼他的逝世。
15. 记者采访了正在课备课的学生们的老师因此拖延了他的工作时间。
16. 大家认出了那和年轻新妻在一起的女孩子的父亲并且欢呼起来。
17. 刚刚上任的市长的夫人被一伙歹徒打劫了。
18. 刚刚在全国巡回演出的合唱团的歌曲被录进了磁带中。
19. 刚被聘来的侦探的嫌疑人被黑社会分子暗杀了。
20. 被炖了三个多小时的牛肉的香味一下子被这股恶臭覆盖了。
21. 这里的人都很喜欢近来研究这个课题的科学家的孙女所以这里气氛很融洽。
22. 一位听众发现演奏会上用来表演的乐器的盒子在大厅角落里。
23. 另一家公司仿造了这家根据传统工艺烘烤的食品的包装设计并牟取暴利。
24. 我在学校时常碰到我们家里排行最小的舅舅的老同学因为他是我们的老师。

Version II

1. 昨天晚上和妻子一同站在阳台上的女演员的男仆被人开枪打死了。
2. 工作室里，正在帮演员化装的歌星的化妆师被评为“年度最佳”。
3. 在操场上等待孩子的学生们的父母被人领到小礼堂休息。
4. 图书员说那作者亲笔签了名的小说的封面被人撕掉了。
5. 昨天中午我看见负责运送公司客户的老板的司机穿了件红衬衫。
6. 那石头砸到刚才还在随身体摆动的绿蛇的尾巴把蛇惊吓了。
7. 村民们非常尊敬上礼拜从城里回来探亲的村长的儿子因为他是硕士。
8. 所有人都在谈论那狠心抛弃亲身骨肉的小男孩的父母在法庭上所言。

[76] This is a direct translation from the English material above in Appendix C.

9. 已证实怀孕了的那个女生的男朋友被学校开除了。
10. 那厨师刚刚烹饪的萝卜的味道被美食家品出来了。
11. 那学习一直很努力的学长的祖母被警察叫去了。
12. 举办了多次个人影展的摄影师的助手刚被聘为一家杂志社的编辑。
13. 选民们非常不满那一万多人参加的招待会的开幕词因为内容太陈旧了。
14. 参观者正在讨论明年将要出版的影集的前言并且哀悼作者的逝世。
15. 记者采访了被保送去上大学的学生们的老师因此拖延了他的工作时间。
16. 大家认出了那赢得选美比赛大奖的女孩子的父亲并且欢呼起来。
17. 坐在马车里的市长的夫人被一伙歹徒打劫了。
18. 在这一带非常受欢迎的合唱团的歌曲被录进了磁带中。
19. 带着黑色太阳镜的侦探的嫌疑人被黑社会分子暗杀了。
20. 他非常喜欢的牛肉的香味一下子被这股恶臭覆盖了。
21. 这里的人都很喜欢长着白胡子的科学家的孙女所以这里气氛很融洽。
22. 一位听众发现某人放起来的乐器的盒子在大厅角落里。
23. 另一家公司仿造了这家 1878 年起源于苏格兰的食品的包装设计并牟取暴利。
24. 我在学校时常碰到教历史课的舅舅的老同学因为他是我们的老师。

Version III

1. 昨天晚上和丈夫一起站在阳台上的女演员的男仆被人开枪打死了。
2. 工作室里，将要上台表演的歌星的化妆师被评为“年度最佳”。
3. 刚刚考完重要考试的学生们的父母被人领到小礼堂休息。
4. 图书员说那本内容非常冗长的小说的封面被人撕掉了。
5. 昨天中午我看见主持董事会议的老板的司机穿了件红衬衫。
6. 那石头砸到喜欢吃老鼠的绿蛇的尾巴把蛇惊吓了。
7. 村民们非常尊敬他们最近刚推举的村长的儿子因为他是硕士。
8. 所有人都在谈论那喜欢玩电脑游戏的小男孩的父母在法庭上所言。
9. 从来不遵守学校规章制度的那个女生的男朋友被学校开除了。
10. 没有人喜欢的萝卜的味道被美食家品出来了。
11. 言语粗俗不讲理的学长的祖母被警察叫去了。
12. 曾经去过西藏的摄影师的助手刚被聘为一家杂志社的编辑。
13. 选民们非常不满宣传民族主义的招待会的开幕词因为内容太陈旧了。
14. 参观者正在讨论这大胆讽刺独裁者的影集的前言并且哀悼作者的逝世。
15. 记者采访了多次获得大奖的学生们的老师因此拖延了他的工作时间。
16. 大家认出了那在街上买了许多东西的女孩子的父亲并且欢呼起来。
17. 一天到晚浓妆艳抹的市长的夫人被一伙歹徒打劫了。
18. 听上去音色纯美的合唱团的歌曲被录进了磁带中。
19. 正准备逃跑的侦探的嫌疑人被黑社会分子暗杀了。
20. 充斥在屋子里的牛肉的香味一下子被这股恶臭覆盖了。
21. 这里的人都很喜欢近来刚上小学的科学家的孙女所以这里气氛很融洽。
22. 一位听众发现敞开着空荡荡的乐器的盒子在大厅角落里。
23. 另一家公司仿造了这家广告公司刚提议的食品的包装设计并牟取暴利。
24. 我在学校时常碰到娶了姨妈的舅舅的老同学因为他是我们的老师。

Version IV

1. 昨天晚上某人开枪打死了和宾客们一同站在阳台上的女演员的男仆然后逃跑了。
2. 他爱上了正在吃午餐的歌星的化妆师并准备求婚。
3. 公布的结果伤害了在操场上等候良久的学生们的父母。
4. 某人撕掉了描绘意大利乡村风情的小说的封面很不道德。
5. 在餐厅吃饭的老板的司机今天被开除了。
6. 刚才还在游动的绿蛇的尾巴被石头砸到了。
7. 号召大家抗洪的村长的儿子被无缘无故的逮捕了。
8. 那钻在钱眼里的小男孩的父母被公众们盘问起来。
9. 学校开除了替女友顶罪的那个女生的男朋友并且全校通报。
10. 美食家品尝了和其他佐料混在一起的萝卜的味道并表示赞赏。
11. 警方请去了学习一直非常努力的学长的祖母说明了情况。
12. 一家杂志社聘请了负责照相馆日常事务的摄影师的助手担任编辑。
13. 主席发表的招待会的开幕词被听众打断了。
14. 著名艺术家撰写的影集的前言将要被翻译成多国文字。
15. 正在课备课的学生们的老师被一位记者采访了。
16. 和年轻新妻在一起的女孩子的父亲被认定为公司的 CEO。
17. 一伙歹徒打劫了刚刚上任的市长的夫人并且刺伤了她。
18. 我录下了刚刚在全国巡回演出的合唱团的歌曲并且反复听。
19. 黑社会分子暗杀了刚被聘来的侦探的嫌疑人使案件失去线索。
20. 这股恶臭一下子覆盖了被炖了三个多小时的牛肉的香味非常恶心。
21. 近来研究这个课题的科学家的孙女被选为小组代表出席会议。
22. 演奏会上用来表演的乐器的盒子被人发现在大厅角落里。
23. 这家根据传统工艺烘烤的食品的包装设计另一家公司仿造了。
24. 我们家里排行最小的舅舅的老同学被学校聘为教授。

Version V

1. 昨天晚上某人开枪打死了和妻子一同站在阳台上的女演员的男仆然后逃跑了。
2. 他爱上了正在为演员化装的歌星的化妆师并准备求婚。
3. 公布的结果伤害了在操场上等候孩子良久的学生们的父母。
4. 某人撕掉了写有作者亲笔签名的小说的封面很不道德。
5. 负责接送客户的老板的司机今天被开除了。
6. 刚才还随身体摆动的绿蛇的尾巴被石头砸到了。
7. 从城里回来探亲的村长的儿子被无缘无故的逮捕了。
8. 那狠心抛弃亲身骨肉的小男孩的父母被公众们盘问起来。
9. 学校开除了怀孕三个月的那个女生的男朋友并且全校通报。
10. 美食家品尝了由著名厨师刚刚烹饪的萝卜的味道并表示赞赏。
11. 警方请去了学习一直非常努力的学长的祖母说明了情况。。
12. 一家杂志社聘请了开办过许多个人影展的摄影师的助手担任编辑。
13. 有一万多人参加的的招待会的开幕词被听众打断了。
14. 将要出版的影集的前言将要被翻译成多国文字。
15. 被保送去上大学的学生们的老师被一位记者采访了。
16. 赢得选美比赛冠军的女孩子的父亲被认定为公司的 CEO。
17. 一伙歹徒打劫了坐在马车里的市长的夫人并且刺伤了她。
18. 我录下了在这一带非常受欢迎的合唱团的歌曲并且反复听。
19. 黑社会分子暗杀了带着黑色太阳镜的侦探的嫌疑人使案件失去线索。

20. 这股恶臭一下子覆盖了他很喜欢的牛肉的香味非常恶心。
21. 长着白胡子的科学家的孙女被选为小组代表出席会议。
22. 说是收拾好了的乐器的盒子被人发现在大厅角落里。
23. 这家根据传统工艺烘烤的食品的包装设计另一家公司仿造了。
24. 在学校里教历史的舅舅的老同学被学校聘为教授。

Version VI

1. 昨天晚上某人开枪打死了和丈夫一同站在阳台上的女演员的男仆然后逃跑了。
2. 他爱上了正在舞台表演的歌星的化妆师并准备求婚。
3. 公布的结果伤害了刚刚考完重要考试的学生们的父母。
4. 某人撕掉了内容非常冗长的小说的封面很不道德。
5. 主持董事会议的老板的司机今天被开除了。
6. 那喜欢吃老鼠的绿蛇的尾巴被石头砸到了。
7. 村民们刚刚选举的村长的儿子被无缘无故的逮捕了。
8. 那个喜欢打电脑游戏的小男孩的父母被公众们盘问起来。
9. 学校开除了从不遵守规章制度的那个女生的男朋友并且全校通报。
10. 美食家品尝了大家都不喜欢的萝卜的味道并表示赞赏。
11. 警方请去了语言粗俗不讲理的学长的祖母说明了情况。
12. 一家杂志社聘请了曾经去过西藏的摄影师的助手担任编辑。
13. 宣传民族主义的招待会的开幕词被听众打断了。
14. 大胆讽刺独裁者的影集的前言将要被翻译成多国文字。
15. 多次获得大奖的学生们的老师被一位记者采访了。
16. 刚才在街上买了许多东西的女孩子的父亲被认定为公司的 CEO。
17. 一伙歹徒打劫了一天到晚浓妆艳抹的市长的夫人并且刺伤了她。
18. 我录下了听上去音色纯美的合唱团的歌曲并且反复听。
19. 黑社会分子暗杀了准备逃跑的侦探的嫌疑人使案件失去线索。
20. 这股恶臭一下子覆盖了刚才充斥在屋子里的牛肉的香味非常恶心。
21. 刚上小学的科学家的孙女被选为小组代表出席会议。
22. 敞开着空荡荡的乐器的盒子被人发现在大厅角落里。
23. 这家广告公司刚提议的食品的包装设计另一家公司仿造了。
24. 娶了姨妈的舅舅的老同学被学校聘为教授。

Appendix D: Instruction and Material for Study 4 & 5

Instruction and Consent

TITLE OF PROJECT: Crosslinguistic Study of Language Processing

PROCEDURE: This study will include two parts. The first part is an eye-tracking task. You will be introduced to a computer screen, where there will be a brief verbal introduction about the experiment and instructions for how to continue. Then you can acquaint yourself with the experiment through practice sentences to give you an idea of how the experiment is conducted. You will be introduced to the experiment part as you go on. The second part is a questionnaire study asking you to rate the sentences you just read in the first part.

Eye-tracking: You will be asked to read a sentence in the screen, and then press a **ROUND** button to carry on reading the next page as soon as you finished the previous one. After some of the sentences, you will be asked some **YES/NO** questions. You are expected to answer these by pressing a button according to the content of the text. By responding to this question, you do not need to press the **ROUND** button any more.

REMUNERATION: Participants can receive 0.5 credit or £3 after the experiments.

POTENTIAL RISKS AND ETHICAL CONSIDERATIONS: The main risk associated with the experiment is the possibility that some participants will feel slightly bored. No other risks are known by now.

CONFIDENTIALITY: The information you give which is recorded will be kept strictly confidential, except as may be required by the law or professional guidelines for psychologists. All information will be identified by an identification code, not your name. Any form that requires your name (e.g., this consent form) will be stored separately from the other material. Your name or other identifying information will never be associated with any research reports or publications that use the results of your questionnaires or interviews.

WITHDRAWAL/PREMATURE COMPLETION: Your participation in the evaluation is voluntary, and you may discontinue at any time, without prejudice.

INVITATION TO ASK FURTHER QUESTIONS: You should ask any questions you have concerning this study before you sign this consent form.

CONSENT: I give my informed consent to participate in this study of **Crosslinguistic Study of Language Processing**. I have read and understand the consent form. Upon signing below, I will receive a copy of the consent form from the study investigator.

Materials[77]

1. Mary enjoyed watching the two birds fighting for food on the branch. But she did not notice that her biscuits were all eaten by her brother / by the biscuits (expt4) at her brother (expt5) / her brother during that time.
   马丽开心地看着树上两只鸟儿为了争夺食物而打架。但她却没有注意到自己的饼干已经被她的弟弟/被她的饼干在她的弟弟/她的弟弟吃完了。
2. Mr. Wang is about to retire. His position will be taken by his assistant / by his book at his assistant / his assistant in a week.
   王先生即将退休。他的职位将在一周内被他的秘书/被他的书在他的秘书/他的秘书所取代。
3. The cat never saw a real mouse in its modern home. It was attacked by the big rat / by a game at the big rat / the big rat in the pet hospital.
   那只猫从来没有在自己时髦的家中见过真的老鼠。于是它在宠物医院被大老鼠/被游戏在大老鼠/大老鼠攻击了。
4. Cinderella was looking forward to dancing with the prince again. She was extremely flattered by the prince / by the abuse at the prince / the prince the other day.
   灰姑娘一心希望可以和王子再跳一次舞。那天她被王子/被侮辱在王子/王子奉承了很久。
5. The chef could cook chicken in more than one hundred different ways. All the food was eaten up by hungry guests / by empty bowls at hungry guests / hungry guests at the table.
   那位厨师用一百多种方法烹饪鸡。所有的食物都被饥饿的客人/被空空的碗在飢餓的客人/饥饿的客人吃完了。
6. The website was such a mess that no one could find any useful information there. The layout was designed by an arrogant artist / by an elegant keyboard at an arrogant artist / an arrogant artist who did not know much about IT.
   那个网页糟透了，在那里找不到任何有效信息。排版是被一位高傲的艺术家/被一个精美的键盘在一位高傲的艺术家/一位高傲的艺术家设计的。
7. The wine was made from some local plant that does not grow anywhere else. The entire stock was bought by a wine taster / by a big tree at a wine taster / wine taster after the exhibition.
   这种酒取材于当地的一种植物，这种植物在别处无法生存。展销会后，所有的储存都被一个品酒师/被一棵大树在一个品酒师/一个品酒师买下了，并且立刻下了订单。
8. The fashion model was late for the show because of the traffic. But she was sharply criticised by the media / by the infection at the media / the media for ruining the whole show.
   这位时装模特因为交通堵塞而迟到。之后她却被媒体/被传染病在媒体/媒体强烈批评为整个时装秀的破坏者。
9. The kid insisted that he did not kill the dog. He was seen by a neighbour / by the teeth at a neighbour / a neighbour sneaking out of its ownerõs yard.
   那个孩子坚持说他没有杀死那只狗。但是他被一个邻居/被几颗牙齿在一个邻居/一个邻居看到鬼鬼祟祟地从狗的主人的后院出来。
10. The author realized his work had become so popular that he eventually became a celebrity. He was stopped by many strangers / by many bags at many strangers / many strangers in the street for autographs.
    那位作家终于意识到自己的小说这么受欢迎，他已经是一位名人了。走在街上，他常常被

[77] The red texts are the normal sentences and the green texts represent the anomalous ones. Texts in yellow background are the modified semantic anomaly used in Study 5.

不认识的人/被不认识的包在不认识的人/不认识的人叫住寻求签名。

11. The politician regretted having said such a stupid thing. Now he has been sacked by the government / by the soap at the government / the government for being a racist.
那位政客深深为自己的愚蠢言行而后悔。目前，因为种族歧视他已经被政府/被肥皂在政府/政府开除了。
12. The band did not care about what the media said and went on with their concert plan. Their fans were stopped by the police / by the clothes at the police / the police in front of the concert hall.
乐队不理会媒体所言，径自按照他们的音乐会安排行事。在演奏厅前，歌迷们被警方/被衣服在警方/警方拦住了。
13. The old prefect decided to resign because of health problems. However, his letter of resignation was not accepted by the administrator / by the chopsticks at the administrator / the administrator in the meeting.
老长官因为健康原因决定辞职。可是他的辞职信没有被行政执行官/被那双筷子在行政执行官/行政执行官批准。
14. The secretary felt it wrong to forge the report, even though her boss insisted she should do it. She was caught by the police / by a holiday at the police / the police later although the boss managed to escape.
那秘书觉得写假报告是错误的，虽然她的老板坚持要她这么做。她后来被警方/被假日在警方/警方抓住，然而老板却逃脱了。
15. The two big banks were negotiating about merging for a long time. In the end, one was acquired by the other / by the clock at the other / the other at a very low price.
两家银行因为合并事宜磋商了很久。最后其中一家银行被另一家/被一座钟在另一家/另一家廉价收购了。
16. Kevin helps with watering his neighbourôs room plants when they are away. His pet is taken great care of by his neighbour / by modern history at his neighbour / his neighbour when he is away.
凯文在邻居外出期间帮他们家浇花。而他不在的时候，他的宠物被邻居一家/被现当代史在邻居一家/邻居一家照顾得很周到。
17. As spring has come, walking along the riverside has become very popular. Kiki was invited by her friends / by some oil at her friend / her friends to have a picnic there this weekend.
春天来了，许多人喜欢到河边散步。这个周末琦琦被朋友们/被一些油在朋友們/朋友们邀请去那里野餐。
18. My uncleôs best friend used to get up in the middle of the night to watch his fish in a big water tank. He was scolded by his wife / by his problem at his wife / his wife for his abnormal behaviour.
我叔叔的好朋友曾经半夜起床观看鱼缸里的鱼。为此，他被他的妻子/被他的难题在他的妻子/他的妻子骂了好几顿。
19. He finally got the job. He was interviewed by the CEO / by his nose at the CEO / the CEO even though he did not know it during his interview.
他最终得到了那份工作。他被公司总裁/被他的鼻子在公司总裁/公司总裁面试的，虽然面试当时他并不知道这一点。
20. The big lake divides the city into two parts. The two parts have been connected by an old bridge / by an old card at an old bridge / an old bridge for hundreds of years.
一片大湖将这座城市分成两部分。这两部分被一座古桥/被一张卡片在一座古桥/一座古桥联接了几百年。
21. A painting in the museum was stolen recently. The security department of the museum was

criticised by the government / by the chocolate at the government / the government in the meeting.

博物馆里的一幅油画最近被盗。在会上，馆内的安全部门被政府/被巧克力在政府/政府严厉地批评了一番。

22. The battle ended gravely with many more soldiers killed. The wounded were sent to hospital by helicopters at helicopters / by candles / helicopters as quickly as possible.

战役艰难的结束了，更多的士兵为此牺牲了。伤员们被直升飞机/被蜡烛在直升飛機/直升飞机尽快运往医院

23. The company has introduced a new game console to the mediaǿs acclaim. In the morning of the launch day, the shops were surrounded by eager gamers / by a wooden door at eager gamers / eager gamers before the opening time.

公司推出了新一款合乎大众口味的游戏机。在销售日当天早上，还没有到开门时间，商店都被焦急的玩家/被一木质的门在焦急的玩家/焦急的玩家围满了。

24. John broke his ankle when he fell off a ladder. He was carried by his friend / by his cake at his friend / his friend all the way to the hospital.

约翰从梯子上面摔下来，摔坏了脚踝。他被他的朋友/被他的蛋糕在他的朋友/他的朋友一路抬到医院。

Appendix E: The Matlab Programme for Eye-tracking data cleaning

```
function [results] = time_c

r=dlmread('D:\matlabmatrix\mat5.txt');
d=dlmread('c:\e.txt');



ccc=0;
ccc2=0;

counter=1;
% 3 versions of 24 sentences
for v=1:3
    for s=1:24
        % take relevant part of data matrix (for a sentence-version pair) and regions
        data_vs=d(d(:,1)==v & d(:,2)==s,:);
        region=r(r(:,1)==v & r(:,2)==s,:);

        % identify subjects
        subjects=unique(data_vs(:,11));

        % do for each subject
        for subq=1:length(subjects)
            sub=subjects(subq);
            % choose part of data that relates to current subject
            partdata=data_vs(data_vs(:,11)==sub,:);

            % go through each of three regions
            for reg=1:3
                % find region borders
```

```
start=region(2*reg+1);
stop=region(2*reg+2);

% state: 0 - not yet in region, 1 - in region 1st time, 2 -
% regression pass, 3 - beyond region (right)
state=0;
% first-pass time
fptime=0;
% regression-path time
rptime=0;
% total time
ttime=0;

firstout=0;
firstout_dir=0;
firstfix=0;
prev=0;
regdist=0;
regind=0;
old_regdist=0;

% first fixation of a regression
first_regfix=0;

% difference of y coordinates for some fixations
y_coord_diff=0;

% go through partial data line by line
for i=1:size(partdata,1)
    % check current region against borders
    cur=partdata(i,4);
    curfocus=partdata(i,3);
    prevfix=partdata(i,10);
```

```
nextfix_dir=partdata(i,8);

% update total time
if cur>=start && cur<=stop
    ttime = ttime + curfocus;
end

% find direction of first fixation exiting the region
if state==1 && (cur<start || cur>stop)
    firstout_dir=pf;
    firstout=cur;
    % implicit assertion i-1 > 1 (since two state changes happened)
    y_coord_diff=partdata(i,6)-partdata(i-1,6);
end
% find first fixation of a regression
if state==2 && first_regfix==0
    first_regfix=curfocus;
end

% state transitions
if state==0 && cur>=start && cur<=stop
    % entering region for the first time
    state=1;
elseif (state==1 || state==2) && cur>stop
    % passed beyond region
    state=3;
elseif (state==1 || state==2) && cur==0 && prev>=start && prev<=stop
&& prevfix==1
    state=3;
elseif state==1 & cur<start
%           if cur==0
 %              fprintf(1,'Warning: including zero region in s %d v %d
(regression originated from region %d); time %d\n',s,v,reg,curfocus);
                %continue;
```

```
                %                       end
            state=2;
        end

        % update first-pass time
        if state==1
            fptime = fptime + curfocus;
            if firstfix==0
                firstfix=curfocus;
            end
        end

        if state==2 && regdist==0
            regdist=old_regdist;
            regind=cur;
        end

        % update regression path time
        if state==1 || state==2
            rptime = rptime + curfocus;
        end
        prev=cur;
        pf=nextfix_dir;
        old_regdist=partdata(i,7);
    end
    results(counter,:)=[v s reg fptime rptime ttime firstout firstout_dir firstfix regdist regind first_regfix y_coord_diff sub];
    counter=counter+1;
   end
  end
 end
end
```

Appendix F: *de* as a Postposition in Chinese

When we were discussing the Chinese syntax in Chapter 6, we mentioned about the categorization of the word *de* in Chinese. In most Chinese syntax books published in Mainland China (e.g., *Modern Chinese, Chinese Syntax*), *de* is not defined as a *postposition*, but 助词 *Zhuci* unique to mainland. *Zhuci*, translated directly into English, is *Auxiliary*. However, as we know, in English, *Auxiliary* only means a word related to an *Auxiliary verb*. In the case of relative clause, *de* is definitely not an *Auxiliary*, since it has nothing to do with verbs. We then look at some Chinese grammar books in English. To our surprise, the category of *de* is not explained clearly and the closest explanation is that *de* is a particle. Now what kind of particle could *de* possibly be?

Not being able to find a satisfactory explanation from Chinese grammar book, we then instead looked at some Japanese grammar books. We know that the two languages, although not related, have many common features, especially for historical reasons, i.e., the two languages have borrowed many words from each other and share the same meanings and forms. Though the pronunciations are slightly different (such as Japanese onyomi and Mandarin pronunciation), it does make sense to use Japanese for reference sometimes, since typically, as modern science developed faster in Japan than in China, the Japanese language is also well studied. It is really helpful to compare the two.

In the case of relative clause, Chinese complex NP has exactly the same structure as Japanese (recall that Chinese appears head-final in Noun Phrases), and the usage and position of Chinese *de* is identical with Japanese *no*. We found that the Japanese grammar looked on *no* as one kind of 助詞 *jyoshi* (*no* is *kaku-jyoshi*), which is identical with the Chinese characters used in the Chinese grammar books. Translated into English, 助詞 is ñparticleò. In Japanese grammar studies in English, *no* is considered a postposition.

In light of the match of Japanese *no* and Chinese *de*, we conclude that the Chinese particle *de* is actually a postposition. Later, we found the further prove of our conclusion from some Taiwanese linguistic WebPages (e.g. http://www.edutech.org.tw/Prospects/MLT-TMSS(3).htm#F) that one use of *de* is considered to be post-positional.

We present this course of finding the right category for *de*, in order to suggest that the Chinese researchers in Mainland China would be well advised to adopt terminology that is better suited for other languages, especially English, for the benefit of better international scientific communication.

Appendix G: List of Abbreviations

Abbre.	Full Term	Page Number
Theoretical		
UG	Universal Grammar	2, 4, 6, 18, 19, 21, 28, 30-32, 35, 60, 73, 79, 80, 86, 92, 95, 190, 222- 225, 236, 243
TG	Transformational Grammar	6, 18-20, 22-25, 39, 190
GB	Government & Binding	6, 19, 21-27, 29, 40, 78-80, 99, 100, 197,198
MP	Minimalist Programme	6, 19, 28-31, 39, 60, 100, 197, 198, 231, 232, 236, 243
EPP	Extended Projection Principle	6, 29, 30, 80, 81, 98
PF	Phonological Form	21, 22, 27, 28
LF	Logical Form	21, 22, 27, 28
GPT	Garden Path Theory	6, 37-42, 109-111, 114
RR	Relativized Relevance	6, 37, 39-41, 46, 83, 84, 86, 87, 109, 110, 114, 124, 159, 160, 162
GPTG	Garden Path Theory Group	37, 42, 43, 75, 81-83, 90, 91, 109, 162
RP/PP	Recency Preference / Predicate Proximity	4, 43, 45, 46, 86, 88, 109, 111, 114, 162
SOUL	Semantics-Oriented Unification-based Language	46, 47
AB	Attachment Binding	4, 46, 47, 86-88, 109, 111, 112, 114, 162
ETH	Exposure-based Tuning Hypothesis	3, 4, 43, 49-51, 60, 89, 91-93, 109, 112-114, 125, 162, 248
CIM	Competition Incremental Model	4, 89-91
IPH	Implicit Prosody Hypothesis	3, 4, 56-58, 61, 72, 95-97, 109, 113, 114, 116, 160, 162, 164
LAST	Late Assignment of Syntax Theory	2, 3, 5, 6, 16, 56, 58, 59-61, 100, 170, 192, 195, 205, 214, 217, 218, 223, 224, 227, 228, 231, 236, 241, 244-251
Linguistic		
NP	Noun Phrase	2, 6, 7, 16, 38, 40, 44, 46, 50, 52, 53, 58, 77, 78, 80-86, 91-95,99, 106-116, 118-126, 128,

		130-137, 139-149, 153-156, 158, 160, 166, 167, 171, 193, 194, 198, 199, 201, 216, 217, 219, 221, 223, 227, 233, 234, 237-240, 244
PP	Pre/post-positional Phrase	25, 35, 39, 40, 46, 58, 77, 86, 91, 94, 99, 106, 109, 111, 114, 126, 131-135, 162, 166, 167, 171-185, 187, 193, 194, 198, 199, 201-214, 216-219, 223, 233, 244
VP	Verb Phrase	27, 29, 38, 40, 44, 46, 57, 76, 77, 81, 84, 91, 94, 95, 106, 109, 110, 111, 165, 166, 193
RRC	Reduced Relative Clause	76, 77
RC	Relative Clause	2, 4, 5, 16, 45, 47, 49, 50, 58, 76, 77, 82-89, 91, 92, 94-96, 100, 101, 103, 104, 106, 108-115, 118-129, 131-135, 140, 144, 153, 157, 159-161, 164, 192, 195, 232-235, 243