

L2 Sentence Processing Strategies of  
Late Learners and Heritage Speakers: The Evidence from Mandarin-English Bilinguals

A Thesis Submitted to the  
College of Graduate and Postdoctoral Studies  
In Partial Fulfillment of the Requirements  
For the Degree of Master  
In the Department of Linguistics  
University of Saskatchewan  
Saskatoon

By

Yubin Xing

© Copyright Yubin Xing, July 12, 2021. All rights reserved.

Unless otherwise noted, copyright of the material in this thesis belongs to the author.

## Permission to Use

In presenting this thesis/dissertation in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis/dissertation in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis/dissertation work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis/dissertation or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis/dissertation.

Requests for permission to copy or to make other uses of materials in this thesis/dissertation in whole or part should be addressed to:

Head of the Department of Linguistics,  
Arts Building, 9 Campus Drive,  
University of Saskatchewan  
Saskatoon, Saskatchewan S7N 5A5 Canada

OR

Dean  
College of Graduate and Postdoctoral Studies  
University of Saskatchewan  
116 Thorvaldson Building, 110 Science Place  
Saskatoon, Saskatchewan S7N 5C9 Canada

## **Abstract**

This study investigates different processing strategies of L1 Mandarin speakers processing complex sentences in L2 English. Heritage speakers, L2 learners, and native English speakers as a control group were compared. This study tested the Shallow Structure Hypothesis (SSH) which argues that second language processing differs from first language processing. Age of acquisition (AoA) and L2 dominance were factored in to determine the difference between L2 learners and heritage speakers. In an online grammatical-Maze experiment, 72 participants completed a processing task on three sentential items: relative clause modifying subject (RCS), relative clause modifying object (RCO), and adverb phrase modifying predicate (ADVP). The results show that with an early AoA and using L2 English dominantly, heritage speakers show native speaker-like processing patterns. While late L2 learners have an opposite pattern of processing RCS and RCO items, they show a native speaker-like pattern on the ADVP item. This suggests that AoA is a critical predictor of processing relative clause attachment but not for ADVP attachment. L2 dominance does not predict attachment preference on the RCS and ADVP items in the statistical results. However, it predicts that bilingual participants process RCO sentences comparable to native speakers. Based on the SSH, these predictions from AoA and L2 dominance are to be attributed to different weightings of the two pathways (i.e., syntactic and heuristic) in the sentence parser of L2 learners. When parsing complex sentences, the weightings of both the syntactic and heuristic pathways are not only affected by AoA but also significantly affected by L2 dominance. Also, different syntactic structures of processing are likely taking variant effects from AoA and L2 dominance. Overall, this study provides evidence of AoA and L2 dominance critically affecting L2 processing which needs to be taken into account as part of the SSH.

## **Acknowledgments**

First, I would like to thank my supervisor Dr. Bettina Spreng who has always been guiding me in the path of linguistics research and providing plenty of valuable advice for planning and executing this project and writing my thesis. The same gratitude goes to the members of my thesis committee - Dr. Jesse Stewart, Dr. Katie Collins, and Dr. Zhi Li, and the committee chair Dr. Kohlberger. Their comments and ideas have been very helpful to my thesis writing. Meanwhile, I would like to express my gratitude to the Department of Linguistics (at the University of Saskatchewan) who gives me the chance of taking this MA program for studying this research topic. I certainly can not forget my friends and all the participants for their unconditional support during data collection. At last, I want to thank my family members, without whose support and encouragement I could never even start my study for this project.

## Table of Contents

Permission to Use .....	i
Abstract.....	ii
Acknowledgments.....	iii
Table of Contents .....	iv
List of Abbreviations .....	vii
List of Tables .....	viii
List of Figures.....	x
<b>Chapter 1 Introduction .....</b>	<b>1</b>
<b>1. Background .....</b>	<b>2</b>
<b>2. Models of Sentence Processing .....</b>	<b>5</b>
<b>2.1 Modular Processing: The Garden-path Model.....</b>	<b>6</b>
2.1.1 Minimal Attachment.....	8
2.1.2 Late Closure .....	9
2.1.3 Processing Modifiers: Relative Clause & Adverb Phrase .....	10
2.1.3.1 Relative Clause Modifier.....	11
2.1.3.2 Adverb Phrase Modifier .....	12
2.1.4 Previous Findings Based on the Garden-path Model .....	14
<b>2.2 Parallel Processing: Constraint-based Models .....</b>	<b>19</b>
2.2.1 Locality: Recency Preference & Predicate Proximity .....	20
2.2.2 Previous Findings Based on CBMs .....	23
<b>2.3 Processing by Non-native Speakers: Combining GPM and CBMs .....</b>	<b>27</b>
<b>2.4 Shallow Structure Hypothesis .....</b>	<b>33</b>
2.4.1 L2 Processing is Shallower than L1 Processing.....	35
2.4.1.1 The Self-paced Reading Approach.....	35
2.4.1.2 The Eye-Tracking Approach.....	37
2.4.1.3 Other Approaches .....	39
2.4.2 L2 processing patterns with L1 processing.....	40
<b>3. Heritage Speakers .....</b>	<b>43</b>
3.1 Critical Period Hypothesis (CPH).....	44
3.2 Age of Acquisition versus Language Dominance.....	45

3.2.1	Age of Acquisition .....	45
3.2.2	Language Dominance .....	48
3.3	L2 Processing by Early Bilinguals & Heritage Speakers.....	50
4.	Research Gaps and Questions.....	52
4.1	Rationale.....	52
4.2	Research Questions and Predictions.....	53
Chapter 2	Methods & Results .....	59
5.	Methods.....	60
5.1	Experimental Design .....	60
5.2	Equipment.....	62
5.3	Participant Recruitment .....	63
5.4	Procedures.....	63
5.5	Material .....	66
6.	Results.....	68
6.1	Accuracy and L2 Dominance Rates .....	68
6.1.1	Accuracy .....	68
6.1.2	L2 Dominance Rates .....	69
6.2	Predict from AoA and L2 Dominance .....	70
6.2.1	Predict from AoA .....	73
6.2.1.1	English Native Speakers.....	76
6.2.1.2	English Late Learners .....	78
6.2.1.3	Heritage Mandarin Speakers .....	80
6.2.2	Predict from L2 Dominance .....	82
6.2.3	Predict from L2 Dominance with AoA.....	84
6.3	Summary of the Results .....	86
Chapter 3	Discussion & Conclusion .....	88
7	Discussion .....	89
7.1	The AoA Effect .....	89
7.1.1	English Native Speakers .....	89
7.1.2	English Late Learners .....	92
7.1.3	Heritage Mandarin Speakers .....	95
7.1.4	RQ1 & RQ2 .....	96
7.2	The L2 Dominance Effect .....	98
7.2.1	Discuss L2 Dominance .....	98
7.2.2	Address RQ3.....	100

<b>8</b>	<b>General Discussion and Conclusion .....</b>	<b>102</b>
	<b>Appendix Items Presented in the Experiment.....</b>	<b>104</b>
	<b>References.....</b>	<b>109</b>

## List of Abbreviations

ADVP (AdvP)	adverb phrase
AoA	age of onset of acquisition
CBM	constraint-based models
CPH	Critical Period Hypothesis
ELL	English late learner
ENS	English native speaker
GPM	Garden-path model
HMCS	heritage Mandarin Chinese speaker
L1	first language
L2	second language
LC	Late Closure
MA	Minimal Attachment
NP	noun phrase
RCO	relative clause modifying object
RCS	relative clause modifying subject
VP	verb phrase
PP	prepositional phrase
RT	reading time
tRT	total reading time of a trial

## List of Tables

Table 1.1 Predicted RT Patterns of ENS (or HMCS) .....	56
Table 1.2 Predicted RT Patterns of ELL Based on J. Witzel et al. (2012) .....	57
Table 2.1 Three Items .....	62
Table 2.2 Mean Accuracy Scores of The Three Groups.....	68
Table 2.3 Rating Scale for L2 English Use in Self-Report Survey .....	69
Table 2.4 Descriptive Statistics of Self-Reported L2 Dominance Rates .....	69
Table 2.5 Observe the Total RTs across the Three Populations .....	72
Table 2.6 Observe the Total RTs across the Three Items .....	72
Table 2.7 Observe the Critical-Region RTs across the Three Items.....	72
Table 2.8 Inferential Statistics on Regions of Interest in The Three Items (Native Speakers) .....	77
Table 2.9 Inferential Statistics on Regions of Interest in The Three Items (English Late Learners) .....	79
Table 2.10 Inferential Statistics on Regions of Interest in The Three Items (Heritage Mandarin Speakers) .....	81
Table 2.11 Inferential Statistics of Critical Region RTs of RCO (Mandarin-English Bilinguals).....	84
Table 2.12 Inferential Statistics of Critical Region RTs of RCO (A Triple Interaction Model).....	85
Table 3.1 Hypothesis and Results (English Native Speakers).....	90
Table 3.2 Hypothesis and Results (English Late Learners).....	93
Table 3.3 Hypothesis and Results (Heritage Mandarin Speakers) .....	96

Table 3.4 Summary of Processing Patterns under the Effect of L2 Dominance ..... 98

Table 3.5 Summary of Processing Patterns under the Effect of L2 Dominance and AoA... 99

## List of Figures

Figure 1.1 Phrase Marker of the Sentence Example (1) (Frazier, 1978, p. 27) . . . . .	7
Figure 1.2 Phrase Markers of the Sentence Example (2a) and (2b) in Frazier & Rayner (1982, pp. 180–181).....	8
Figure 1.3 Phrase Markers of the Sentence Example (3a) and (3b) (Frazier 1978, p. 76) ...	10
Figure 1.4 Phrase Marker of ‘the son who bought the house’ in Sentence Example (4a)....	11
Figure 1.5 High (left) and Low (right) Attachment Versions of Phrase Marker for Example (4b).....	12
Figure 1.6 Low & Hight Attachment Conditions: Phrase Markers of Sentence Example (5a) and (5b).....	13
Figure 1.7 High & Low Attachment Conditions: Phrase Markers of Sentence Example (6a) and (6b).....	14
Figure 2.1 The G-Maze Paradigm .....	60
Figure 2.2 Regression Graphing on Total RTs & Critical RTs Among Three Groups .....	73
Figure 2.3 Boxplot of the RTs of the Critical Regions for Three Items (Three Populations) .....	74
Figure 2.4 Boxplot of the Total RTs of the Three Items (Three Populations).....	75
Figure 2.5 Interaction between Critical Region RT and L2 Dominance Level (All L2 Speakers).....	82

## **Chapter 1 Introduction**

## 1. Background

Sentence processing is an important component of language processing. One of the reasons is that a sentence is a structure that is considered both the smallest group of words semantically and the largest independent unit syntactically (Van Gompel, 2013). Models of sentence processing usually try to reveal the time course by which sentences are parsed online and how grammatical and non-grammatical information is represented during parsing a sentence (Van Gompel, 2013).

The Shallow Structure Hypothesis (SSH) (Clahsen & Felser, 2006b, 2006a, 2018) makes a set of predictions of second language (L2) processing. Specifically, it argues that L2 sentence processing strategies are different from first language (L1) processing strategies.

While many studies support the SSH with empirical findings (Clahsen & Veríssimo, 2016; Felser et al., 2003, 2009; Felser & Cunnings, 2012; Felser & Roberts, 2007; Marinis et al., 2005; Roberts et al., 2008), other studies disclose no significant differences between L1 speakers and L2 learners (Dussias, 2001; Frenck-Mestre, 2002; Frenck-Mestre & Pynte, 1997; Juffs & Harrington, 1996; J. Witzel et al., 2012).

According to the Critical Period Hypothesis (Johnson & Newport, 1989, 1991), L2 late learners are very unlikely to acquire a complete grammar like native speakers or heritage speakers. The reason is that age critically influences L2 attainment. As early bilinguals, heritage speakers usually have early ages of acquisition and speak L2 as their dominant language (Polinsky, 2018). Thus, dominance and early acquisition provide an advantage to form native-like grammatical representations.

This study tests the effects of age of onset of acquisition (AoA) and language dominance on L2 sentence processing. Specifically, sentence processing strategies of heritage speakers and

late L2 learners are compared with those of native speakers. Although the updated Shallow Structure Hypothesis (Clahsen & Felser, 2018) includes a discussion of AoA factoring into L2 morphological processing (Clahsen & Veríssimo, 2016; Veríssimo et al., 2018), early bilinguals are not discussed for sentence processing.

This study hypothesizes that with L2 acquired successfully before puberty and L2 being dominant in early life, heritage speakers can make use of native-like representations of grammatical structure when processing complex sentences. On the other hand, acquiring the L2 after puberty and having low language dominance levels in the L2 will affect L2 sentence processing in a variety of ways and grammatical information could be a less critical constraint than for native speakers. Moreover, L2 dominance will play an essential role in predicting L2 sentence processing patterns of bilinguals.

Chapter 1 reviews previous literature on relevant models of language processing (e.g., the Garden-path theory and constraint-based models), followed by a discussion of the SSH. After that, a comparative review of L1/L2 sentence processing differences is introduced, and the characteristics of heritage speakers will be discussed compared to late L2 learners. In this chapter, the focused variables of this study, (i.e., AoA and language dominance based on exposure quantities), are also introduced with respect to their effects on L2 acquisition and processing (Birdsong, 2006; Hernandez & Li, 2007; Weber-Fox & Neville, 1996). The hypothesis of this research study ends this chapter.

Chapter 2 reports the process and methodology for an online maze experiment, which was based on J. Witzel et al. (2012) and Felser et al. (2003). It also includes the results of the present study comparing three different speaker groups for the factors influencing processing strategies.

Chapter 3 firstly discusses the study results. Then it compares the results with the predicted patterns of these populations reported in previous literature (i.e. native speakers and L2 learners). Finally, the insights to be drawn from the results is discussed.

## 2. Models of Sentence Processing

Models of syntactic processing can be divided into two basic approaches, i.e. the modular approach and the constraint-based approach. The modular approach often takes a bottom-up type idea in that phrases are processed and assembled in small chunks for processing. The most common type of model is the Garden-path model (GPM), which assumes syntactic assembly without any information processing of semantics or phonology (Frazier, 1978, 1987; Frazier & Fodor, 1978; Rayner et al., 1983). Support for this idea comes from so-called garden path sentences that are initially ambiguous with respect to the attachment of complements or adjuncts.

Constraint-based models (CBMs) describe processing as an algorithm that uses grammatical structure, semantic cues, pragmatic bias, and other information for processing sentence structure (Gibson & Pearlmutter, 1998; MacDonald et al., 1994; Pearlmutter & Gibson, 2001). Different frequencies of sentential construction or word use are also argued to affect the final interpretation of a sentence. So, more commonly used structures such as active sentences would be processed faster than passive sentences. Structural information thus does not occupy a unique position in sentence processing.

Many studies have argued that there are differences between L1 and L2 processing (Clahsen & Felser, 2006b, 2006a, 2018; Clahsen & Veríssimo, 2016; Dussias, 2003; Felser et al., 2003; Marinis et al., 2005; Papadopoulou & Clahsen, 2003). The Shallow Structure Hypothesis (SSH) (Clahsen & Felser, 2006b, 2006a, 2018) argues that L1 and L2 processing could be performed in gradient different ways.<sup>1</sup> In contrast to L1 processing, L2 processing proceeds with

---

<sup>1</sup> Clahsen and Felser (2018) restate that this difference is not simply quantitative or qualitative, while it is gradient and could be quantified with some methodology like computational modelling.

less recourse to structure and more heuristics to disambiguate syntactic structure relations. Processing, in essence, follows a more linear “shallow” path than L1 processing.

## 2.1 Modular Processing: The Garden-path Model

The Garden-path theory is a modular-based processing model (Garfield, 1991). Frazier (1987) argued that the input for language processing includes different subsystems (i.e. phonological, semantic, and syntactic subsystems). Different types of information are encapsulated in various modules that belong to the subsystems mentioned above. Phonological, semantic, and syntactic information will thus be processed autonomously. The Garden-path model argues that online sentence processing is autonomous and involves the syntactic modules only.

The Garden-path theory is based on its method of testing. This model accounts for human parsing strategies to solve ambiguities during the processing of garden-path sentences.<sup>2</sup> A garden-path sentence is a sentence that appears to have one structure until it reveals itself to require a different structure at some point. At this point, the garden-path model predicts that such sentences show a slowing down in processing when the structure needs to be reassessed.

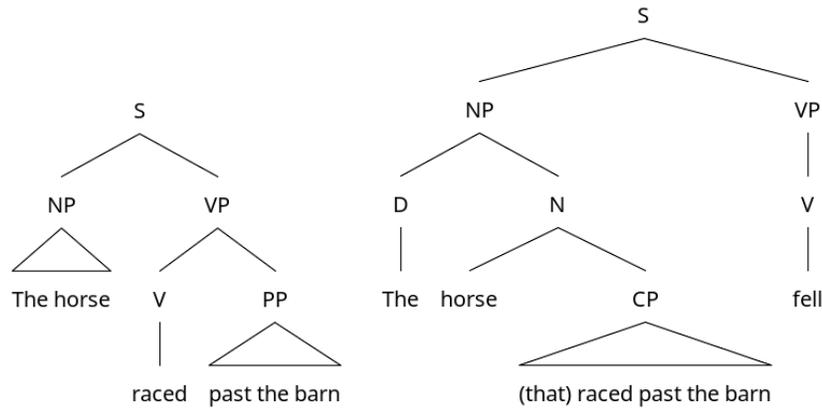
In item (1), the sentence initially would be structured with *past the barn* as part of the verb phrase (VP) *raised*. With the appearance of *fell*, that structure needs to be revised (see Figure 1.1). The parser is led down a ‘garden path’ due to the dropping of the relative pronoun *that*. Were it present, it would have signalled earlier that the VP *raced past the barn* is a relative clause (CP) modifying the noun phrase (NP) *the horse*.

(1) *The horse (that) raced past the barn fell.*<sup>3</sup>

---

<sup>2</sup> Encountering ambiguity during sentence comprehension reveals the absence or delay of necessary information for parsing a natural language by the human brain (Frazier, 1978).

<sup>3</sup> PP: prepositional phrase; D: determiner; CP: complementizer phrase, i.e. embedded clause; S: sentence.



*Figure 1.1 Phrase Marker of the Sentence Example (1) (Frazier, 1978, p. 27)<sup>4</sup>*

The garden path model assumes that syntactic processing happens in two stages (Frazier & Fodor, 1978). Stage 1 is the analysis of syntactic categories such as verb, noun, etc. Stage 2 is the assembly of these categories in syntactic structures. In the above example, while *raced* was correctly identified as a verb, the assembly was incorrect. It does not form the VP of the main clause.

Meaning and meaning relations (i.e., semantic information) is processed subsequently based on syntactic structure. In the parsing process, two strategies of sentence processing are applied: Minimal Attachment (MA) (see § 2.1.1) and Late Closure (LC) (see § 2.1.2) (Frazier, 1978; Frazier & Fodor, 1978). The third principle, Relativized Relevance (see § 2.1.3) (Frazier, 1990; Traxler & Frazier, 2008), takes into account discourse or pragmatic factors and may apply

---

<sup>4</sup> The syntactic structure representations of the example (1) were simplified.

after structural processing is completed. It may cause a reanalysis of sentence structure and change the final interpretation of a sentence based on context.<sup>5</sup>

### 2.1.1 Minimal Attachment

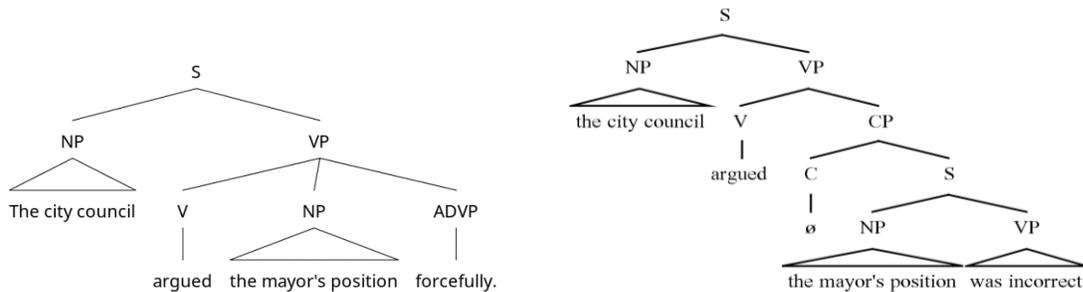
The Minimal attachment (MA) principle means that when the parser encounters an incoming item, this item will be attached to the syntactic structure in a way that requires the least nodes (or minimum nodes). Minimal Attachment is claimed as a strategy to minimize the working memory<sup>6</sup> cost during online sentence processing (Frazier, 1978).

When encountering the noun phrase (NP) *the mayor's position* in (2a), it will cost less working memory for the parser to attach the NP to the VP *argued* in (2a) than in (2b) because sentence (2a) has fewer syntactic nodes than (2b) (as shown in Figure 2).

(2a) *The city council argued the mayor's position forcefully.* (Minimum Attachment)

(2b) *The city council argued the mayor's position was incorrect.* (Non-minimum Attachment)

(Frazier & Rayner, 1982, pp. 180–181)



**Figure 1.2** *Phrase Markers of the Sentence Example (2a) and (2b) in Frazier & Rayner (1982, pp. 180–181)*

<sup>5</sup> Frazier (1990, p. 321), refers to *Relativized Relevance* as the preference of associating a phrase to the main assertion of the sentence being currently processed. To make this principle take effect, all other optional interpretations should be grammatically legitimate and inviolate to the discourse. Since this study is focused on the LC effect, details of *Relativized Relevance* will not be discussed here.

<sup>6</sup> Working memory is developed from the concept of short memory that refers to the very limited capacity of information temporarily stored in human memory system for people's retrieval during the process of cognition (Chai et al., 2018; Eysenck & Keane, 2010, Chapter 6).

In other words, predicted by the MA strategy, one should process *the mayor's position* easier and faster in (2a) than in (2b)<sup>7</sup>.

Frazier & Rayner (1982) propose the MA principle to account for their empirical findings from an eye-tracking experiment. Processing time (i.e. eye fixation durations) is longer when a sentence is set to disrupt the MA strategy. As shown in Figure 1.2, item (2a) has five nodes (i.e. NP, VP, V, NP, ADVP) while (2b) has eight nodes (NP, VP, V, CP, C, S, NP, VP) in the syntactic structure. Processing the NP *the mayor's position* in item (2b) will violate the MA principle, adding a higher cost for working memory and cause difficulty in online comprehension.

### 2.1.2 Late Closure

Late Closure (LC) refers to the parser's principle of associating a lexical item to the left material (Frazier 1978, p. 76). In example (3a), the noun phrase *the sock* is associated with the left material *While Mary was mending* as an object.

(3a) *While Mary was mending the sock fell off her lap.*

(3b) *While Mary was mending the sock it fell off her lap.*

(Frazier 1978, p. 76)

When the association of the incoming item requires building more syntactic nodes to complete a sentence (as shown by the syntactic structure in Figure 1.3), it will cause processing difficulty and reanalysis of the left materials (as predicted by the MA principle). Meanwhile, the LC principle predicts that in (3a), encountering *fell off her lap* makes the parser try to associate it to its left item *the sock*. Then the parser realizes that the first parse has failed. The verb phrase

---

<sup>7</sup> The syntactic structure representation of (2b) was modified to represent the embedded sentence structure more in line with current structure representations.

predicate must be combined with a subject *the sock*, and thus the parser is required to build a new clause structure to process *fell off her lap* as the predicate of *the sock*. As predicted by LC, the processing of (3b) should be easier than (3a). Item (3b) does not cause processing difficulty upon encountering the NP *fell off her lap* since it has a different subject from *the sock*. It is not required to reanalyze the former section *while Mary was mending the sock* (see Figure 1.3). The parser will just incrementally continue to process *it fell off her lap* as the main clause of the whole sentence.

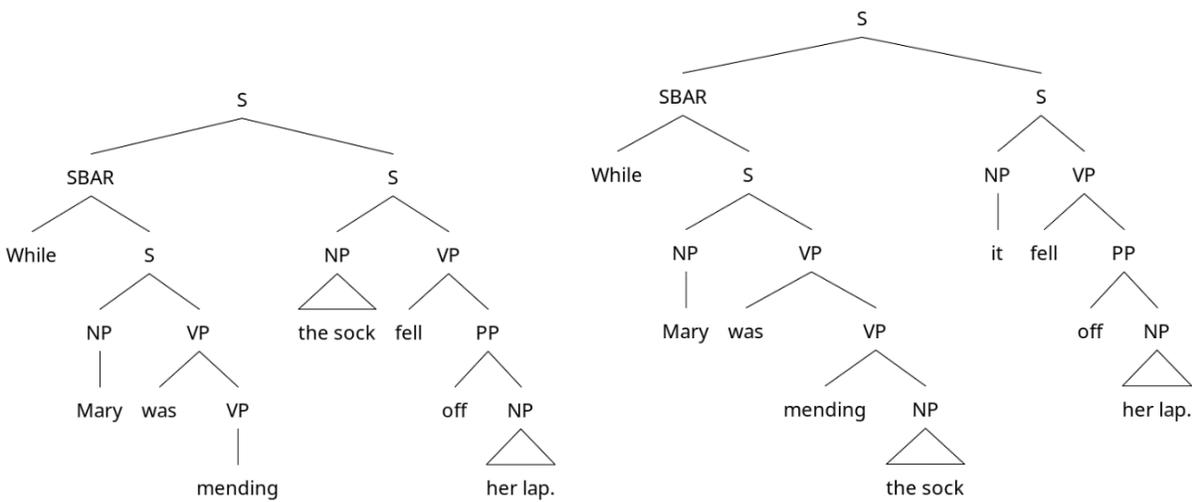


Figure 1.3 Phrase Markers of the Sentence Example (3a) and (3b) (Frazier 1978, p. 76)

MA & LC interact with each other (Frazier & Fodor, 1978) in that the MA will regulate the parser to build as few nodes as possible, while the LC requires early left attachment. As for whether these principles are universal, there is considerable dispute (see the constraint-based models introduced in § 2.2).

### 2.1.3 Processing Modifiers: Relative Clause & Adverb Phrase

Modifier processing is often used in experimental research to test whether a reader follows the LC parsing principle and whether LC is the only principle to guide the parser (Altmann et al., 1998;

Cuetos et al., 1996; Cuetos & Mitchell, 1988; De Vincenzi & Job, 1993; Gibson et al., 1996). One reason to choose a modifier as the experimental item is that a modifier can cause ambiguities in that it can attach at different sites.

### 2.1.3.1 Relative Clause Modifier

A **relative clause** (RC) is a modifier clause of an NP, and this NP is called the head of the relative clause. In (4a), the NP *the son* is the only possible head of the RC *who bought the house on the corner*, and the pronoun *who* is the relative pronoun. The position of the NP *the son* that is modified is the **attachment site** of the RC.

The verb *bought* in the clause has a long-distance dependency (Sag et al., 2003, Chapter 14) with the NP *the son* (in 4a) or *the son of the actor* (in 4b). This dependency requires the parser to hold the attachment site(s) temporarily stored in the working memory waiting for retrieval when encountering the RC as shown in Figure 1.4. When multiple attachment sites emerge in a sentence, it will cause ambiguity.

(4a) *I talked to **the son** who **bought** the house on the corner.*

(4b) *I talked to **the son of the actor** who **bought** the house on the corner.*

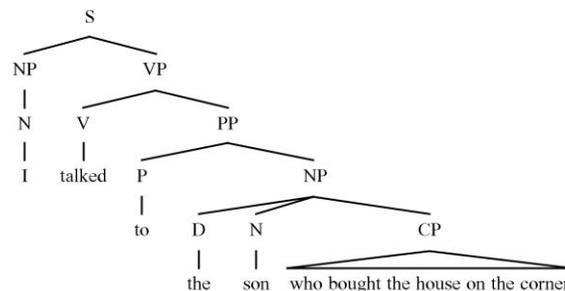


Figure 1.4 Phrase Marker of ‘the son who bought the house’ in Sentence Example (4a)

Ambiguous sentences such as (4b) have two potential attachment sites since the RC may modify the higher attachment site NP *the son* or the lower one *the actor* (See Figure 1.5). As predicted by the LC principle, people should prefer the lower attachment site *the actor* being modified by the RC. In other words, if one prefers attaching the modifier to the high attachment site, which is the NP *the son*, it is a violation of the LC parsing principle.

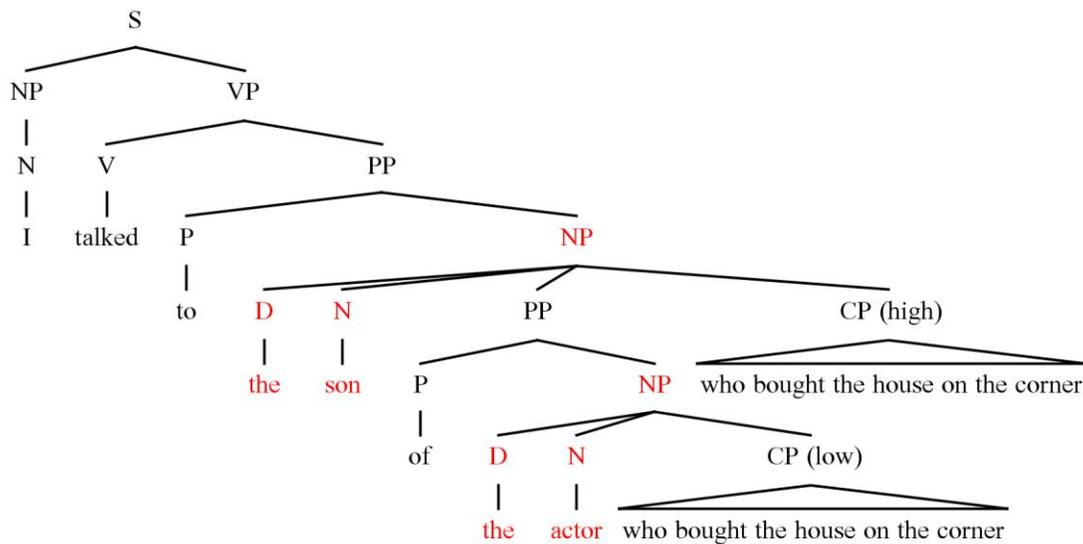


Figure 1.5 High (left) and Low (right) Attachment Versions of Phrase Marker for Example (4b)

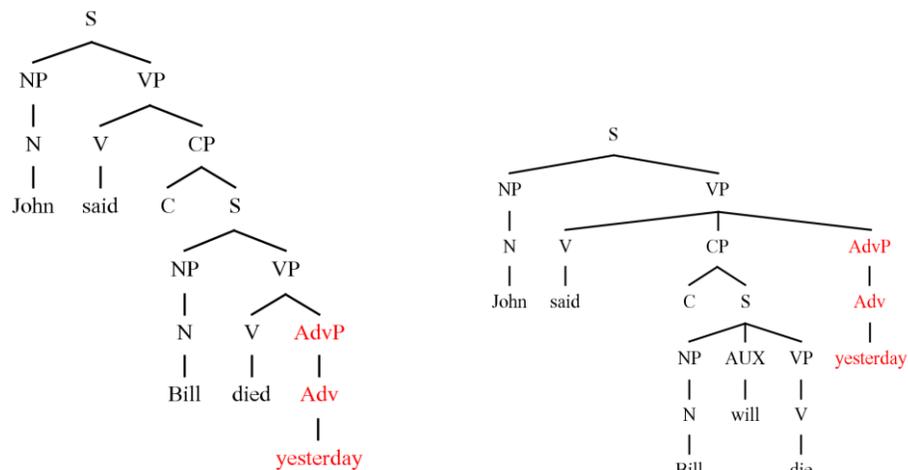
### 2.1.3.2 Adverb Phrase Modifier

**Adverb phrases** (AdvP) also serve as a testing ground for high and low attachment to test the parsing principles of the Garden path model. As a verb modifier, an AdvP can cause ambiguity of attachment preference in sentences with two or more attachment sites for selection. For example, when processing the sentences (5a) and (5b), an English native speaker should prefer the low attachment *died/will die* predicted by the LC principle. However, if the tense of the preferred

attachment site and the adverb meaning conflict, processing becomes difficult since the AdvP has to be attached high (5b). In other words, it is easier to process (5a) than (5b) (see Figure 1.6 for representation of the syntax structures of the low and high attachment condition).

(5a) *John said Bill died yesterday.* (Low Attachment Condition)

(5b) *John said Bill will die yesterday.* (High Attachment Condition)



**Figure 1.6 Low & High Attachment Conditions: Phrase Markers of Sentence Example (5a) and (5b)**

On the other hand, if the AdvP is changed to *tomorrow*, then a native English speaker would prefer to attach it to *will die* and (5b) would be easier to process.

A high attachment preference predicts that a speaker prefers to attach the AdvP *yesterday* or *tomorrow* to the verb *say*. As shown in example (6a) and (6b) (see Figure 1.7 for representation of syntactic structures of the high attachment condition), if someone has a high attachment preference, which violates the LC principle, one should have difficulty processing (6b) since there would be a conflict between the verb tense of the preferred attachment site and the adverb meaning,

while there is no such conflict between the verb tense of the preferred attachment site and the adverb meaning in (6a).

(6a) *John will say Bill died tomorrow.* (High Attachment Preference)

(6b) *John will say Bill died yesterday.* (Low Attachment Preference)

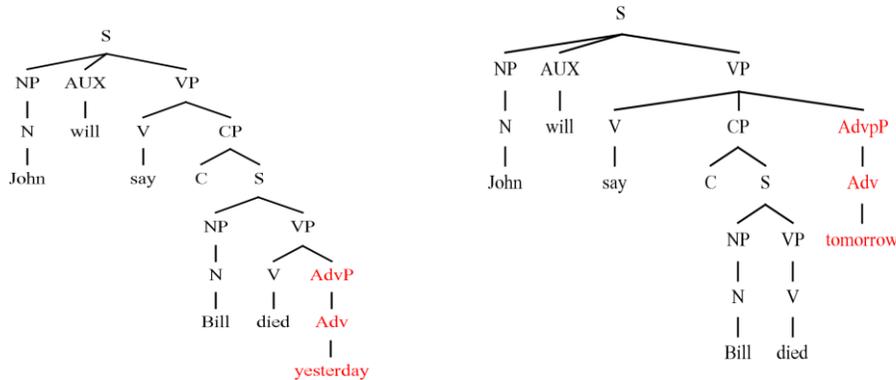


Figure 1.7 High & Low Attachment Conditions: Phrase Markers of Sentence Example (6a) and (6b)

## 2.1.4 Previous Findings Based on the Garden-path Model

Studies of English processing often show support for the LC principle with a low attachment preference for processing RC modifying multiple NPs. Following the study of Cuetos and Mitchell (1988) (as reviewed below), Carreiras and Clifton (1999) use eye-tracking experiments to test how English speakers process temporarily ambiguous sentences such as (7a) and (7b). Results support the low attachment prediction by the LC parsing principle. In (7a), two attachment sites evoke the ambiguity of whether the RC modifies NP1 *the sister* or NP2 *the handyman*. This ambiguity is then resolved by gender agreement between one of the NPs (i.e. *the sister* in (7a) or *the nursemaid* in (7b)) and the RC discourse *gave birth to twins*. The result of first-fixation durations and total

reading time indicates a significantly shorter processing duration of the disambiguity section in the low attachment trials (7b) than in the high attachment condition (7a).

(7a) *The police came from headquarters early this morning and arrested the sister of the handyman who recently gave birth to twins in the hospital.*

(7b) *The police came from headquarters early this morning and arrested the brother of the nursemaid who recently gave birth to twins in the hospital.*

(Carreiras & Clifton, 1999, p. 828)

Cuetos and Mitchell (1988) argue that LC is not a universal strategy of sentence processing. They report an early closure effect on Spanish relative clause processing. Using an offline questionnaire with sentences such as (8), Cuetos and Mitchell (1988) find that Spanish native speakers prefer interpreting the relative clause *que tuvo el accidente* ‘who had had the accident’ as the modifier of *la hija* ‘the daughter’, rather than *coronel* ‘colonel’.

(8) *El periodista entrevistó a la hija del coronel que tuvo el accidente.*

*‘The journalist interviewed the daughter of the colonel who had had the accident.’*

(Cuetos & Mitchell, 1988, p. 77)

This preference contradicts the prediction of the LC principle, which, in this case, would guide an English speaker to attach the relative clause to the second lower noun phrase ‘the colonel’, but not the first higher noun phrase ‘the daughter’. Also, Cuetos and Mitchell (1988) show in online reading experiments that Spanish speakers process a late closure sentence with more time than processing an early closure sentence. Thus, they argue that LC is not the only effect of sentence processing across languages.

Frazier (1990) argues that the different findings for Spanish result from a reanalysis process. This process happens once the structure-processing mechanism can not fix the final interpretation. It is the *Relativized Relevance* principle that causes the reanalysis to a high attachment preference. Therefore, processing a relative clause with two NP attachment sites in Spanish is predicted in the following order. First, the structure-processing mechanism (LC principle) will guide the Spanish parser to the low attachment site *the pretty nurse*, as shown in (9a) and (9b).

(9a) *The doctor called in the son of the pretty nurse who hurt herself.*

(9b) *The doctor called in the son of the pretty nurse who hurt himself.*

(Frazier, 1990, p. 323)

Second, regulated by the *Relativized Relevance* principle, a reanalysis process happens and causes the Spanish parser to attach the relative clause to the high attachment site *the son*. Third, the reanalysis process is affected by the discourse and changes the structure in the final interpretation (from a low site to a high site). Although this reanalysis process is reasonably fast, it is already beyond the two-stage parsing process proposed by Frazier and Fodor (1978). The reason for the discourse behind the reanalysis lies in grammar differences between English and Spanish. As argued by Frazier (1990), an English structure like *the son of the pretty nurse* can be alternated by *the pretty nurse's son*, while it can not be replaced in Spanish because a genitive form *the pretty nurse's son* does not exist in Spanish. Thus, the whole NP *the son of the pretty nurse* is the main assertion to be modified in a Spanish sentence, while *the pretty nurse* itself is not the main assertion to be modified.

De Vincenzi and Job (1993) support the reanalysis hypothesis with findings from an experiment with Italian sentences of the same structure as items (9a) and (9b) with Italian native speakers. A late closure effect (low attachment preference) is found in an online processing task,

while the offline task results suggest that the participants prefer to interpret the relative clause as a modifier of the higher NP (i.e. in the case of item (9b), *the son*).

Although many studies support the LC principle as a universal parsing strategy, some literature reports high attachment preferences in languages other than English, including Spanish (Carreiras & Clifton, 1999; Cuetos & Mitchell, 1988). In a self-paced reading experiment, Brysbaert and Mitchell (1996) also find a high attachment preference when native speakers process Dutch. A high attachment effect is also reported in an eye-tracking experiment of French relative clause processing by Zagar et al. (1997).

According to the discussion of Gibson et al. (1996, p. 25), Spanish speakers do apply the LC principle for low attachment preference when processing sentences like the example (10). As shown in the English version, the adverb phrase '*yesterday*' is preferable when attached to '*Bill died*' because readers encounter processing difficulty when the predicate is replaced by '*will die*', which will violate the time agreement between the verb phrase and the AdvP '*yesterday*'.

(10) *Juan dijo que Bill se murió (# morira) ayer.*

*'John said Bill died (# will die) yesterday.'*

(Gibson et al., 1996, p. 25)

This effect of low site preference for AdvP attachment has also been found in English sentence processing. In an eye-tracking experiment by Altmann et al. (1998), English speakers show a low attachment preference even though the experimental items are manipulated with contextual bias supporting a high attachment interpretation. As shown in examples (11a) and (11b), Altmann et al. (1998) create two conditions of context, respectively supporting high and low AdvP attachment preferences in one of their experiments.

(11a) Low attachment supporting context + low attachment preference:

*Tom's got two young dogs and they like playing in the fields.*

*Tom washed one of the dogs yesterday but the other one last week.*

*He'll brush the dog he washed yesterday to make its fur shine again.*

(11b) High attachment supporting context + high attachment preference:

*Tom's got two young dogs and they like playing in the fields.*

*Tom washed one of the dogs but did not want to bother with the other dog.*

*He'll brush the dog he washed tomorrow to make its fur shine again.*

(Altmann et al., 1998, p. 462)

With a low attachment supporting context in (11a), an English speaker will prefer to attach *yesterday* to *he washed*, but not *he'll brush* as predicted by Late Closure. In (11b), the context supports the reader to attach the AdvP *tomorrow* to *He'll brush*, which is on the high attachment site. Thus, item (11b) is designed against the parsing principle of Late Closure. However, eye movement data indicates significantly more first pass regressions in the critical region *tomorrow* when processing the high attachment condition (11b) than the low one (i.e. *yesterday* of 11a). These results are interpreted as signaling difficulty when processing the high attachment condition. Compared to the low attachment trials, longer reading and rereading times are also observed on the adverbial phrase in the high attachment condition, even when the context supports it.

The authors state that even with a contextual bias, English-speaking participants still show a preference for attaching the modifier to the low attachment site. This result validates the LC strategy with native English speakers for AdvP attachment (but see the results of other experiments in Altmann et al., 1998 reviewed in § 2.2.2).

In summary of Section 2.1, the modular-based model indicates that the two-stage parsing of a sentence is quite automatic without the interference of any lexical-semantic information. After structure parsing is completed, reanalysis may happen if other alternatives for interpretation exist due to non-grammatical factors (e.g. semantic, discourse, or pragmatic bias). Many empirical findings in psycholinguistic studies support the modular-based model. Meanwhile, literature in neurolinguistics also presents findings that support the modular-based processing model using neuroimaging methodologies (Friederici, 1995, 2002; Hahne & Friederici, 1999). Details will not be reviewed here since the current study follows a behavioral approach in psycholinguistics, which does not always echo the neuroimaging data.

To summarize Section 2.1, previous findings for native speakers' processing find support for and against the Garden-path model. Studies on L2 processing usually combine this modular-based approach with other approaches under the framework of constraint-based models. These studies will be reviewed in Section 2.3.

## **2.2 Parallel Processing: Constraint-based Models**

Parallel processing models like the constraint-based models (CBM) emphasize the interaction between different types of information during processing (Tanenhaus & Lucas, 1987). The term *constraint* refers to factors influencing processing such as syntactic bias, lexical meanings, word categories, word probabilities, phrase-based structure probabilities, common knowledge, discourse, pragmatic information, intonation, and other bias that take effect both intra-sententially and extra-sententially (see McRae & Matsuki, 2013, p. 3 for an overview; see also MacDonald et al., 1994 and Gibson & Pearlmutter, 1998 for early proposals of constraint-based models). For instance, Gibson and Pearlmutter (1998) argue that lexical frequency affects online sentence interpretation. When one reads the sentence, *The spy saw the cop with the binoculars*, two possible interpretations

can be made depending on how often a prepositional phrase (PP) is used to modify a verb phrase (VP) or a noun phrase (NP). If the frequency of modifying a verb phrase is relatively higher, one is more likely to attach the PP *with the binoculars* to the VP *saw the cop* → *saw the cop with binoculars*, instead of to the NP *the cop* → *the cop with binoculars*.<sup>8</sup>

### 2.2.1 Locality: Recency Preference & Predicate Proximity

Garden-path model (GPM) accounts take LC as a universal principle of syntactic processing to resolve ambiguity upon processing modifiers like RC and AdvP. As reviewed in Section 2.1, Cuetos and Mitchell (1988) argue that Spanish speakers have a high attachment preference when processing a relative clause preceded by two NPs. Gibson et al. (1996) propose the *Recency Preference* and *Predicate Proximity* principles as phrase-based factors that are weighted differently across languages to account for crosslinguistic variation in attachment preferences. Gibson et al. (1996) argue that *Predicate Proximity* hides the Late Closure (LC) effect in Spanish sentence processing and thus causes a high attachment preference. *Predicate Proximity* refers to “attach as close as possible to the head of a predicate phrase” (Gibson et al., 1996, p. 41).

Meanwhile, the authors refer to *Recency Preference* as a variant of the LC principle.<sup>9</sup> The *Recency Preference* principle means the parser prefers to attach an incoming lexical item to the most recently constructed structure, and it interacts with other principles to achieve various patterns of attachment preference. Guided by this principle, the parser will choose the most recent

---

<sup>8</sup> Except for psychological experiments, computational modeling has also been used to simulate the interaction of different variables to achieve a parsing result (e.g. accuracy of ambiguity resolution) that can be compared with human parsing performance (McRae & Matsuki, 2013; Smolensky et al., 2014). This approach also inspires the argument of the SSH that grammatical and non-grammatical (heuristics) pathways exist in parallel (Clahsen & Felser, 2018).

<sup>9</sup> *Recency Preference* predicts preference of the most recent attachment site or lexical items. However, LC only predicts choosing one attachment site over other sites in the first stage parsing. After that LC will no longer function during the reanalysis against possible processing difficulties (Gibson et al., 1996). Also, Recency is not restricted to structural preference but can be applied to discourse-based preference as well (Pearlmutter & Gibson, 2001).

structure as the preferred attachment, while the *Predicate Proximity* principle will guide the parser to choose the closest site to the head of a predicate phrase. For an item like *I talked to the son of the actor who bought the house on the corner* (4b), the *Recency Preference* principle will guide the parser to attach the relative clause to the lower site *the actor* because it is the most recent structure to the relative clause *who bought the house on the corner*. But the *Predicate Proximity* principle will predict high attachment preference (*the son*) because it is the closest site to the predicate phrase *talked to* in the main clause.

Gibson et al. (1996) thus claim that *Predicate Proximity* will interact with *Recency Preference* during the processing of ambiguous RC attachments. In the very early stage of parsing, the *Recency Preference* principle will first take effect, while the *Predicate Proximity* principle weighs differently across languages to interact with the *Recency Preference*. Therefore, in a Spanish sentence with a relative clause modifying two NP attachment sites, the high attachment is preferred (Cuetos & Mitchell, 1988) because the weighting of *Predicate Proximity* is stronger in Spanish, which is different from English sentence processing.

To test the existence of the *Predicate Proximity* effect, Gibson et al. (1996) manipulate an experimental item (i.e., a temporarily ambiguous RC structure) into three conditions in both Spanish and English. In each condition (12a, 12b, 12c), the relative clause has three attachment sites (NP1 *la(s) lampara(s)* ‘*the lamp(s)*’, NP2 *la(s) pintura(s)* ‘*the painting(s)*’, and NP3 *la(s) casa(s)* ‘*the house(s)*’). The RC modifying ambiguity is resolved using number agreement between the predicate (*fue dañada* ‘*was damaged*’) and a potential NP attachment site (either NP1, NP2 or NP3). Attachment preference will then be forced by the singular verb in the relative clauses and the singular form of either NP1(12c), NP2(12b), or NP3(12a). Using the self-paced reading

experiment, Gibson et al. (1996) find a common effect of non-monotonicity preference in Spanish and English sentence processing.

(12a) *las lámparas*<sub>NP1</sub> *cerca de* *las pinturas*<sub>NP2</sub> *de* ***la casa***<sub>NP3</sub> *que fue dañada en la inundación*  
*'the lamps*      *near*      *the paintings*      *of* ***the house*** *that was damaged in the flood'*

(12b) *las lámparas*<sub>NP1</sub> *cerca de* ***la pintura***<sub>NP2</sub> *de* *las casas*<sub>NP3</sub> *que fue dañada en la inundación*  
*'the lamps*      *near*      ***the painting***      *of* *the houses* *that was damaged in the flood'*

(12c) ***la lámpara***<sub>NP1</sub> *cerca de* *las pinturas*<sub>NP2</sub> *de* *las casas*<sub>NP3</sub> *que fue dañada en la inundación*  
***'the lamp***      *near*      *the paintings*      *of* *the houses* *that was damaged in the flood'*

(Gibson et al., 1996, p. 27)

Reading time results suggest that both Spanish and English speakers prefer the lowest attachment site (**NP3** *las casa(s)* '*the house(s)*') more than the other two attachment sites (**NP1** *la(s) lámpara(s)* '*the lamp(s)*' and **NP2** *la(s) pintura(s)* '*the painting(s)*'). Meanwhile, the preference effect is stronger for NP1 than for NP2. In other words, item (12a) is the easiest structure to parse for the participants, while item (12b) is the hardest one. If only *Recency Preference* would take effect here, the participants should have processed item (12a) with a faster duration than the other two items while process item (12c) with the longest duration. This hypothesis is not supported by the results. Gibson et al. (1996) argue that this non-monotonicity result suggests *Predicate Proximity* interacts with the *Recency Preference* effect because the *Recency Preference* or *Predicate Proximity* principle alone could only predict a monotonicity order of preference, which should be either NP3-NP2-NP1 or NP1-NP2-NP3.

By combining these two principles, Gibson and Pearlmutter (1998) propose a constraint-based model that includes three other categories of constraint: lexical constraints, contextual constraints, and phrase-level contingent frequency constraints. Since the focus of this project is on

the effect of age of acquisition and L2 dominance factors on attachment preference in L2, these other three categories of constraints will not be discussed in detail. The respective weightings of *Predicate Proximity* and *Recency Preference* could cause either high or low attachment in RC ambiguous sentences or sentences with AdvP attachment. However, due to cross-linguistic variability, this syntactic factor (i.e., *Predicate Proximity & Recency Preference*) still needs further exploration, as discussed in Gibson and Pearlmutter (1998) and remains unresolved in constraint-based models.

### 2.2.2 Previous Findings Based on CBMs <sup>10</sup>

It is noticeable that complex structures with RC attachment are the first item used to argue against the Garden-path model. As mentioned in the above section, Gibson et al. (1996) have already proposed *Recency & Predicate Proximity* to complement the Late Closure principle. Results of processing the structure *NI of NP2 of NP3 + RC* in English and Spanish suggest the weighting of *Predicate Proximity* is different across languages. Furthermore, Gibson & Pearlmutter (1998) claim *Recency & Predicate Proximity* as a constraint of the computational cost of the parser. Other than this cost hypothesis, researchers also explore lexical/discourse constraints that interact with the *Recency Principle & Predicate Proximity*.

Rohde et al. (2011) found a verb effect that can reverse the low attachment preference in English sentence processing. It is shown that implicit causality verbs like *detest* will crucially influence the online processing of a sentence like (13a) and (13b) (Rohde et al., 2011, pp. 8–9).

(13a) *John detests the children of the musician who lives in La Jolla.* [low]

---

<sup>10</sup> Considering this project only discusses the syntactic processing principles through the Garden-path model and constraint-based models, section 2.2.2 will briefly review the literature that explores other types of constraint but investigate the details that are related to the *Recency Preference* and *Predicate Proximity* principles.

(13b) *John detests the children of the musician who are arrogant and rude.* [high]

(14a) *John babysits the children of the musician who lives in La Jolla.* [low]

(14b) *John babysits the children of the musician who are students at a private school.* [high]

(Rohde et al., 2011, p. 26)

Implicit causality refers to verbs that implicitly give inferences of causing some result. The verb *detest* can cause an inference that the subject dislikes someone or something for some reason. The authors argue that implicit causality verbs are associated with discourse processing. In a self-paced reading (SPR) experiment, they report significantly shorter reading times of high attachment cases (13b) than that of reading low attachment cases (13a). For (14a) and (14b), the verb *babysit* does not have the property of implicit causality. The reading times of these items did not show any processing advantage for the high attachment trials, and the online processing result was consistent with the low attachment preference predicted by the *Recency Preference* principle. Thus, due to the implicit causality property of the verb *detest*, (13b) is more plausible for the participants than (13a), while the verb *babysit* (a non-implicit causality verb) does not elicit such an effect on processing (14a) and (14b). Therefore, Rohde et al. (2011) argue that discourse processing interacts with syntactic processing when an implicit causality verb such as *admire*, *adore*, *blame*, etc. occurs in a sentence (see Rohde et al., 2011, p. 26 for the experimental materials).

Similarly, Altmann et al. (1998) incorporate the discourse constraint into their study on attachment preference of verb modifiers. Findings in their study challenge the Garden-path model as a universal parsing model when processing the AdvP attachment. The authors find a Late Closure effect on AdvP attachment processing during their first experiment, as mentioned in Section 2.1.4. However, when the context activates some prediction effect in the parser, attachment

preferences could be overridden by a bias other than the syntactic modules. For instance, (15a) has a contextual bias that the AdvP *next week* is preferred to be attached to the high attachment site *she'll* but not the low one *she proposed*.

(15a) *When will Fiona implement the plan she proposed?*

*She'll implement the plan she proposed next week (last week), of course.*

(15b) *She'll implement the plan she proposed tomorrow (yesterday), they hope.*

(Altmann et al., 1998, p. 467)

On the other hand, (15b) has no such context. An eye-tracking experiment by Altmann et al. (1998) was conducted to compare the processing patterns between these two conditions. The results show significant processing difficulties for the adverb *tomorrow* in the high attachment condition (15b), while this effect is reversed for (15a). The authors argue that contextual bias overrides the Late Closure strategy during processing structures with AdvP attachment, which means grammatical structure is not the only bias by which a sentence parser is constrained.

Native speakers of Mandarin Chinese are reported to have a low attachment preference when processing Mandarin Chinese relative clauses, but the parsing process is not purely syntactic for subject-relative clauses (Kwon et al., 2019; Shen, 2006). Notably, the Mandarin counterpart of a complex NP in English (NP1 of NP2) is connected by the genitive case marker *de*, which assigns case from left to right. For instance, in Mandarin, an object-relative clause could be *The journalist interviewed had the accident + de + the colonel de daughter*, whose English counterpart is *The journalist interviewed the daughter of the colonel who had had the accident*. In this case, a low attachment preference means Mandarin native speakers prefer to attach the relative clause *'had the accident'* to *'the colonel'*, rather than *'daughter'*. The NP *'daughter'* is at a high attachment site

of the relative clause (Kwon et al., 2019, p. 3). Kwon et al. (2019) argue that preference patterns of subject-relative clause attachment in Mandarin are not affected by syntactic structure alone. The animacy of the complex NP is a bias to interact with a syntactic preference of attachment site. When processing a sentence with a subject-relative clause like *receive subsidy + de + farmer de farm + main clause* 'the farm of the farmer that receives subsidy + main clause' (Kwon et al., 2019, p. 3), Mandarin speakers are reported to have a low attachment preference because the low site is an animate NP 'farmer'. However, when they read the sentence with *farm de farmer*, they prefer to rate the relative clause as a modifier of the high attachment site *farmer*, an animate NP. Also, in many other studies, an animacy effect of sentence subject is detected on relative clause processing which indicates that lexical constraints bias syntactic processing in an early stage of sentence parsing (Kidd et al., 2007; Mak et al., 2002; Weckerly & Kutas, 1999).

Hare et al. (2003) argue that lexical frequency (of verbs) plays a role in resolving ambiguities when processing garden path sentences. The study of Wells et al. (2009) reports a frequency effect of syntactic structure when processing object and subject relative clauses.

Both the GPM and CBMs accounts have found evidentiary support for different situations of low or high preference of RC and AdvP attachment when processing complex sentences in L1. The GPM predicts the LC effect occurs at an early stage of parsing while discourse factors will cause a reanalysis of syntactic structure (De Vincenzi & Job, 1993; Frazier, 1978, 1990; Frazier & Fodor, 1978), accounting for crosslinguistic variation in attachment preferences. Compared to the GPM, the CBMs predict that a non-grammatical bias interacts with phrase-based (grammatical) bias during the very early processing of a sentence. Specifically, it is the interaction between the locality factor and other cues (predicate proximity, lexical frequencies, contextual cues) that results in different interpretations when processing modifier attachment in different languages (Altmann

et al., 1998; Desmet et al., 2006; Gibson & Pearlmutter, 1998; Hare et al., 2003; Kidd et al., 2007; Kwon et al., 2019; Mak et al., 2002; Rohde et al., 2011; Weckerly & Kutas, 1999; Wells et al., 2009).

### **2.3 Processing by Non-native Speakers: Combining GPM and CBMs**

Many studies on L2 processing of RC and AdvP attachment are approached combining the LC principle of the Garden-path model and *Recency/Predicate Proximity* of the constraint-based models. Felser et al. (2003) argue that L2 processing is qualitatively different from native language processing. In a self-paced reading (SPR) experiment, Felser et al. (2003) tested English learners of both German and Greek with two types of items (16a and 16b). In (16a), the ambiguous relative clause is preceded by a complex genitive antecedent with two noun phrases (i.e. NP1 *the secretary*, and NP2 *the professor*) linked by the preposition *of*. However, in (16b), the same NP1 and NP2 are linked by the thematic preposition *with*.

(16a) *The dean liked the secretary of the professor who was reading a letter.*

(16b) *The dean liked the professor with the secretary who was reading a letter.*

(Felser et al., 2003, p. 462)

Subjects were requested to read the items in a self-controlled manner. The online reading times of critical regions of the items were recorded for statistical analysis. The data shows that German and Greek speakers do not have any significant preference for either high attachment (NP1) or low attachment (NP2) for the “genitive case assigner” items, while native English-speaking children have a significant preference for low attachment meaning *the professor* is reading a letter. Instead,

the L2 learner participants show a clear low attachment preference during processing the ambiguous relative clauses with a thematic preposition *with* in the antecedent.<sup>11</sup>

As for (16a), when the genitive antecedent with *of* provides no or fewer lexical cues for L2 learners, they tend to have no preference for either the high or the low attachment site. The result is interpreted such that the L2 learners in this study hardly constructed full representations of a grammatical structure as native English speakers do. This kind of performance in L2 processing is inconsistent either with predictions from the Late Closure (LC) principle (Frazier, 1978) or the predictions from the *Recency & Predicate Proximity* principles (Gibson et al., 1996; Gibson & Pearlmuter, 1998; Pearlmuter & Gibson, 2001). Thus, it is suggested that L2 speakers/learners do not process like L1 or native speakers (Felser et al., 2003).

Similarly, Papadopoulou and Clahsen (2003) show that Greek language learners fail to apply the LC principle during parsing an ambiguous sentence with a relative clause. Two groups of adult Greek learners were recruited as participants whose L1s are respectively Spanish and German. The authors concentrated on the critical region of a Greek sentence with a temporarily ambiguous relative clause. The experimental stimulus was manipulated with gender agreement (as shown in example 17) between the complex NP *ton fititi me tin kathigitria* ‘the student of the teacher’ and the VP *apoghoitevmenos* ‘disappointed’.

---

<sup>11</sup> Based on the construal theory (Frazier & Clifton, 1997), Felser et al. (2003) argue that German and Greek speakers are more affected by lexical information when processing ambiguous L2 relative clauses since the preposition “with” leaves the NP1 out of the thematic domain of the relative clause and makes the high attachment (i.e. NP1) too costly to be preferred (De Vincenzi & Job, 1993). Thus, NP2 is their preference as being modified by the ambiguous relative clause in item 16b).

(17) *Enas kirios fonakse ton fititi me tin kathighitria pu itan apoghoitevmenos apo to neo ekpedheftiko sistima.*

*'A man called the student<sub>masculine</sub> of the teacher<sub>feminine</sub> who seemed disappointed<sub>masculine</sub> by the new educational system.'*

(Papadopoulou & Clahsen, 2003, p. 511)

However, there is no significant effect on reading times of the VP when Greek learners process the sentence with a genitive antecedent *me 'of'*. Namely, L2 learners have no significant difference in reading times between the sentences of high and low attachment designs.

This result patterns with the results in Felser et al. (2003), in which the same methodology is applied for processing tasks of English complex sentences with relative clauses.

Spanish learners are also argued to have no significant attachment preference when processing ambiguous RC attachment in their second language (Dussias, 2003). The study of Dussias (2003) is focused on RC attachment resolution by highly proficient L2 learners. In a questionnaire experiment, Dussias recruited two groups of proficient L2 speakers: L1 Spanish – L2 English (BSE), and L1 English – L2 Spanish (BES). Another two groups of monolinguals speaking English (ME) and Spanish (MS) respectively were recruited as control. All participants were asked to answer questions on how to attach a relative clause to the head of an NP1 + *of* + NP2 structure. The selection was to be made to choose the RC as the modifier of either NP1 (the high attachment site) or NP2 (the low attachment site). All bilingual participants were asked to complete both English and Spanish versions of the questionnaire (18a and 18b). The monolingual groups were asked to complete the tasks in their native language only.

(18a) Peter fell in love with the daughter of the psychologist who studied in California.

*Who studied in California?*

a) *The daughter studied in California.*

b) *The psychologist studied in California.*

(18b) Pedro se enamoró de la hija del psicólogo que estudió en California.

‘Peter fell in love with the daughter of the psychologist who studied in California.’

*¿Quién estudió en California? ‘Who studied in California?’*

a) *La hija estudió en California. ‘The daughter studied in California.’*

b) *El psicólogo estudio´ en California. ‘The psychologist studied in California.’*

(Dussias, 2003, pp. 541–542)

In addition, an online reading experiment was conducted with Spanish items such as in (19a) and (19b). Three groups of speakers were asked to participate (i.e. BSE, BES, and MS). The ambiguous sentence is resolved by the NP *su esposo* ‘her husband’ in the relative clause. So (19a) and (19b) represent low attachment and high attachment conditions, respectively.

(19a) *El perro mordió al cuñado de la maestra / que vivió en Chile / con su esposo.*

(low attachment)

*‘The dog bit the brother-in-law of the teacher<sub>FEM</sub> who lived in Chile with her husband.’*

(19b) *El perro mordió a la cuñada del maestro / que vivió en Chile / con su esposo.*

(high attachment)

*‘The dog bit the sister-in-law of the teacher<sub>MAS</sub> who lived in Chile with her husband.’*

(Dussias, 2003, p. 547)

All English speakers had a low attachment preference when they answered the questions in the offline task. For instance, they were more likely to choose b) *The psychologist studied in*

*California* as the answer to (18a). However, when BES participants took the online task for processing a Spanish sentence with ambiguities, they had nonsignificant differences between the processing times of low and high attachment trials (19a) and (19b). Monolingual Spanish speakers processed items in the high attachment condition faster than items in the low attachment condition. These results suggest that L2 Spanish learners do not use syntactic processing strategies like native speakers when they process a sentence in their second language, even when they have a high level of L2 proficiency.

Following this approach, non-native speakers are also tested on processing AdvP attachment items in some studies. Findings by J. Witzel et al. (2012) indicate a similar pattern on sentence ambiguity reaction between proficient English learners of native Chinese speakers and native English speakers in eye movement data. Three types of structure (with RC high/low attachment, adverbial high/low attachment, and NP/sentence coordination, respectively) were presented as the stimuli to the two groups of subjects in an eye-tracking experiment (20a, 20b, and 20c for reference). The first two items were used to detect the late closure or locality strategy in the parser of L2 learners, while the third item is used to determine whether a minimal attachment strategy is applied.

(20a) RC attachment:

*The son of the actress who shot herself/himself on the set was under investigation.*

(20b) Adverb phrase attachment:

*Anne will serve the apples she picked yesterday/tomorrow, but she won't serve the plums.*

(20c) NP versus Sentence coordination:

*The nurse examined the mother(,) and the child played quietly in the corner.*

(J. Witzel et al., 2012, pp. 424–427)

The eye movement data of high proficient L2 learners suggest some interesting patterns of attachment preference during their online processing.<sup>12</sup> The critical regions of (20a) and (20b) elicited significant effects of attachment preference. Specifically, the results show a high attachment preference for (20a) and a low attachment preference for (20b). In the third pair of conditions (i.e., noun phrase versus sentence coordination), the sentence coordination item (*the mother, and the child*) was easier processed than the noun phrase structure (*the mother and the child*) for English learners. Since native English speakers were reported having low attachment preference when processing these structures, the authors concluded that highly proficient L2 learners follow native-like principles such as the Minimal Attachment/Late Closure (Frazier, 1978) or the *Recency/Predicate Proximity* Principle (Gibson et al., 1996; Gibson & Pearlmutter, 1998).

Another study using offline questionnaires found no particular preference for low or high attachment with Cantonese-English bilinguals (Bai, 2018). Both high and low proficient English L2 learners were asked to read a sentence like *The director who attends the meeting frequently amends the agenda*. After that, the subjects needed to answer a question on which verb the adverb *frequently* modifies. The results did not show any significant difference between the answer rates for the verb *amends* and the verb *attends*. In this case, even highly proficient English learners of Cantonese speakers do not follow the Late Closure principle.

In sum, the findings from studies using self-paced reading experiments suggest that late L2 learners do not build native-like representations of grammatical knowledge for modifier processing.

---

<sup>12</sup> Five categories of eye movement measures are recorded in Witzel et al. (2012): first-pass reading time, go-past reading time (i.e. regression-path reading time), total reading time, first-pass regression proportion, and total sentence reading time.

However, eye-tracking experiments, which are claimed as the more sensitive tools, show results that support native-like processing of highly proficient L2 learners.

## **2.4 Shallow Structure Hypothesis**

Taking different models of L1 sentence processing as the foundation, the SSH predicts that L2 processing strategies are different from L1 processing strategies. Clahsen & Felser (2018) argue that sentence processing relies on a dual-path mechanism for L1 speakers and L2 learners. One pathway is the grammatical route, and the other one is more heuristic and recruits other constraints (e.g. lexical, contextual, or other non-grammatical information) to drive the parser. Both paths run parallel during sentence processing. When processing a sentence in the first language, the parser relies more on the grammatical route and less on the heuristically driven pathway, while late L2 learners rely more on the heuristic pathway. The SSH predicts that a late L2 learner will be unlikely to follow the syntactic parsing principles proposed in the Garden-path theory.

Clahsen and Felser (2006b, 2006a, 2018) comprehensively review the literature on L2 processing. Most studies are focused on comparing native speaker adults, native speaker children, and late L2 learners in psycholinguistics and neurolinguistics. As for similarities and differences between L1 and L2 sentence processing, the literature falls into two broad categories:

- 1) behavioral studies of ambiguity resolution on relative clause attachment, filler-gap dependencies
  - Support for the SSH (Dussias, 2003; Dussias & Cramer Scaltz, 2008; Felser et al., 2003, 2009; Felser & Cunnings, 2012; Havik et al., 2009; Marinis et al., 2005; Papadopoulou & Clahsen, 2003; Roberts et al., 2006, 2008)

- Evidence against the SSH (Frenck-Mestre, 2002; Juffs, 2004; Williams, 2006; J. Witzel et al., 2012)

2) Neuroimaging studies on Event-related potential (ERP) components or other types of neural activities that are activated by syntactic anomalies in the experimental stimuli

- Support for the SSH: (Hahne & Friederici, 2001; Kubota et al., 2003; Perani & Abutalebi, 2005; Weber-Fox & Neville, 1996)
- Evidence against the SSH: (Covey, 2018; van Hell & Tokowicz, 2010).

An update for the SSH framework is proposed in Clahsen and Felser (2018). Instead of saying “fundamental and qualitative differences between L1 and L2 processing” (Clahsen & Felser, 2018, p. 695), they clarify that there is a gradient difference between L2 and L1 processing. In their dual-pathway model, the grammatical representation is less weighted than that of the heuristic resource for the parser of late L2 learners. This account is partly inspired by computational modeling approaches such as Smolensky et al. (2014) which assume that linguistic constraints are weighted in the language parser. This approach to language processing can quantify the grammatical and non-grammatical constraints of L2 processing and thus develop the SSH in more detail. However, the exact weighting and time-course of various constraints in the L2 processing mechanism remain unknown (Clahsen & Felser, 2018). Since this question is beyond the scope of this project, it will not be discussed further.

However, the SSH framework underspecifies how AoA and language dominance may predict the behavioral patterns of sentence processing of late and early L2 learners. The current study fills this gap.

## 2.4.1 L2 Processing is Shallower than L1 Processing

### 2.4.1.1 The Self-paced Reading Approach

L2 parsing difficulties have been shown for a long time.<sup>13</sup> As discussed in Section 2.3, some studies using the self-paced reading (SPR) experiment have revealed different processing strategies between L1 speakers and late L2 learners when processing modifiers (i.e. RC and AdvP) (Dussias, 2003; Felser et al., 2003; Papadopoulou & Clahsen, 2003). Literature addressing other sentence structures also provides support to this hypothesis.

For example, Marinis et al. (2005) tested L2 processing of *wh*-dependencies in English, and the final result supports the Shallow Structure Hypothesis. English learner participants (Chinese, Japanese, German, and Greek as their L1) do not rely on the presence of complementizers (e.g. *that* in 21a) in the same way native English-speaking participants rely on to help processing *wh*-dependency constructions.

(21a) *The manager who the secretary claimed **that**<sub>(intermediate gap)</sub> the new proposal had pleased<sub>(gap)</sub> will hire five workers tomorrow.*

(21b) *The manager who the secretary's claim about the new proposal had pleased<sub>(gap)</sub> will hire five workers tomorrow.*

(Marinis et al., 2005, p. 75)

Items like (21a) and (21b) are used in an SPR experiment to test whether L1 and L2 speakers could benefit from some kind of comprehension facilitation by the complementizer *that* (i.e. before an intermediate gap) which should reactivate the filler *who* temporarily stored in the working memory

---

<sup>13</sup> See Juffs & Harrington (1995) for an early study of online processing by late L2 learners of Chinese speakers.

(Clifton & Frazier, 1989; Gibson & Warren, 2004). Assuming learners have the same processing mechanism as native speakers, L2 learners recruited by Marinis et al. (2005) were expected to spend shorter reading times on the VP *claimed* in (21a) than the VP *claim about* in (21b), which is the processing result for native speakers. However, the final results reveal that processing by English learners is not facilitated by this complementizer. Meanwhile, both native speakers and learners show a filler-gap effect when reading the position *had pleased*, which is lexically driven by verb-subcategorization (Pickering & Barry, 1991). This result is interpreted as a case of incomplete representations of grammatical structure in the parser of L2 learners. Thus, their L2 learners rely more on lexical-semantic bias than on grammatical representations of *wh*-dependency processing.

Following the study of Marinis et al. (2005), these results were tested and supported by Felser and Roberts (2007) using cross-modal picture priming. In a cross-modal experiment, the participants are asked to judge the animacy of a picture when reading the sentences from the computer screen and listening to the pre-recorded sentences. Items with long-distance *wh*-dependencies like *John saw the peacock to which the small penguin gave the nice birthday present in the garden last weekend* are used in the experiment. The participants are Greek L1 speakers learning English as L2. Results from a previous study on native English-speaking adults and children (Roberts et al., 2006) were used as control. Pictures were inserted into critical positions of the experimental sentences. The authors ruled out the working memory factor because the reading span test result<sup>14</sup> has no significant interaction with the reaction times. However, the between-subjects analysis indicated that the results of high proficient L2 learners are inconsistent

---

<sup>14</sup> The reading span task is a type of experiment for working memory comparison among the participants (Daneman & Carpenter, 1980).

with that of native speakers in processing gap-filler dependencies. By failing to find any effect of the intermediate gap on L2 learners' processing, the authors provide some support to the SSH.

Ambiguity resolution of subject versus object is another sentential construction to test how native-like or non-native-like L2 learners respond to the "Garden Path". Roberts and Felser (2011) weighted semantic cues (plausibility effect) in processing L2 English sentences with weak and strong "Garden Paths" while Havik et al. (2009) examined L2 ambiguity resolution with a German-Dutch combination. Both these two studies elicited different processing strategies for L1 and L2 during the reanalysis of the Garden-path sentences. In Dussias and Cramer Scaltz (2008), L2 learners use all possible information from their L1 and L2 to interpret Garden-path sentences, while native speakers construct full grammatical representations when trying to resolve complex ambiguities.

#### **2.4.1.2 The Eye-Tracking Approach**

The eye-tracking methodology is also used to compare L1 and L2 sentence processing. Unlike most of the studies using SPR, which result in a significant difference between L1 and L2 grammatical processing, the results of many studies with the eye-tracking experiment are controversial.

Felser et al. (2009) argue that there is a notable difference between L1 and L2 processing of binding principle A in English based on an eye-tracking experiment. Binding-principle A processing has been studied in L1 processing literature (Sturt, 2003).<sup>15</sup> L1 speakers should encounter an early processing difficulty when they read a sentence with a mismatch between the

---

<sup>15</sup> The Binding Principle can be divided into two principles: the binding principle A and binding principle B. Specifically, binding principle A means "A reflexive pronoun must be bound by a preceding argument of the same verb" (Sag et al., 2003, p. 205). For example, in the sentence *Susan likes herself*, the reflexive pronoun *herself* and the preceding argument *Susan* must be bound by the same verb (in this case, *like*).

gender of an anaphoric reflexive (i.e., herself) and the stereotypical gender of the binding antecedent. A stereotypical gender could be posed on some nouns such as *surgeon* in the following context: *Jonathan (Jennifer) was pretty worried at the City Hospital. He (or 'She') remembered that the surgeon had pricked herself with a used syringe needle.* In this case, for native speakers, the gender of the binding-inaccessible antecedent (i.e., *Jonathan* or *Jennifer*) should have no significant effect on the first fixation data of eye movement in the reflexives *herself*. This is interpreted as native speakers being sensitive to binding principle A in a very early stage of online sentence processing (Sturt, 2003). Felser et al. (2009) use a similar item to test whether Japanese speakers learning English could be affected by other factors beyond grammatical information. In their experiment, participants are asked to process inaccessible antecedents and reflexives in different contexts. Sentences such as *John noticed that Richard had cut himself with a very sharp knife* and *Jane noticed that Richard had cut himself with a very sharp knife* are presented as stimuli of an eye-tracking experiment. There is a distinguishable difference in the first-pass durations<sup>16</sup> of the reflexive region (*himself*) between these two contexts for L2 learners. Native speakers seem to have no such effect that is imposed by discourse. Thus, binding-**inaccessible** antecedent to a reflexive has elicited a significant discourse effect on the L2 parser. While native speakers adhere to the *Recency/Predicate Proximity* principle (Gibson et al., 1996; Gibson & Pearlmutter, 1998), L2 learners are more strongly influenced by semantic or pragmatic cues (which refers to the conventional gender of names *Jane* and *John* in this case). Binding principle A is a weaker constraint of sentence processing for L2 speakers/learners than native speakers (Felser et al., 2009).

---

<sup>16</sup> In eye movement data, first pass duration and first fixation duration are usually used to reveal very early parsing tendencies of a reading participant.

Felser and Cunnings (2012) conducted two sessions of eye-tracking experiments on Japanese and German speakers learning English. Similar to Felser and Roberts (2007), Felser and Cunnings find a discourse effect that delays the processing durations for L2 learners but not for native English speakers. Both groups of L2 learners (Japanese and German speakers) had the same patterns of violation to binding principle A. These results also rule out L1 transfer since German is not a long-distance binding language like Japanese. In other words, German speakers were not fully constrained by grammatical representations in this experiment. L1 transfer does not provide any facilitation for German speakers. The longer processing time reflects a behavioral performance with less grammatical representation in the parsers of L2 learners.

Roberts et al. (2008) have also found distinct patterns in eye-tracking experiments between native Dutch speakers and their L2 learners on processing complex sentences. The researchers argue that L2 Dutch learners relied more on the discourse-pragmatic information than the grammatical constraints to perform the online processing of syntactically ambiguous sentences.

#### **2.4.1.3 Other Approaches**

It should be noted that electrophysiological and neuroimaging measures have become very powerful tools for the research of neural circuits and their activities during language processing. For instance, using event-related functional magnetic resonance imaging (fMRI), Newman et al. (2001) present findings of distinctive spatial activations in the human cerebrum caused by syntactic and semantic violations. According to Newman et al. (2001), the frontal lobes are more likely to be activated for syntactic violations, while temporal-parietal regions light up more during semantic violations of sentence processing. Also, unlike L1 speakers, L2 learners (especially late learners) tend to explore other areas of the cortex to process L2 languages as a kind of compensation that interacts with factors of AoA, language exposure, and language proficiency (Perani & Abutalebi,

2005). Meanwhile, proposing a memory extraction model, Ullman (2001) argues that late L2 learners use more declarative but less procedural memory to extract grammatical information compared to L1 speakers. It is suggested that age of first exposure and length of experience are strong predictors of how much effort L2 learners make to compute grammar with a declarative memory system instead of the procedural memory system. Many studies on neural activities support significant differences between L1 and L2 grammatical processing (Hahne & Friederici, 2001, 2001; Kubota et al., 2003; Weber-Fox & Neville, 1996).

#### **2.4.2 L2 processing patterns with L1 processing**

While many studies support the SSH (Clahsen & Felser, 2006b, 2006a, 2018), there is also evidence of native-like processing patterns for L2 learners, especially for highly proficient learners. Some studies using the SPR approach show that L2 processing shares similar patterns with L1 processing. For instance, Juffs (2004) tested learners with garden-path sentences like *After the children cleaned the house looked very neat and tidy* in an SPR experiment and no significant difference effect was observed in the critical regions (i.e., verbs like *cleaned* and *looked*). In other words, English L2 learner participants did not behave significantly differently from native English speakers. A native-like processing pattern of L2 learners is also reported by Williams (2006), where Chinese-English and Roman-English bilinguals were asked to process *wh*-dependency constructions in SPR tasks for English.

Results seem more mixed for eye-tracking studies. Section 2.4.2 reviewed eye fixation studies revealing considerable differences between L1 and L2 syntactic processing. However, Frenck-Mestre and Pynte (1997) report that both French-English and English-French bilingual participants in their study have a garden-path effect in common. Moreover, the ambiguity

resolution strategies of learners are found to be similar to what native speakers perform in an eye-tracking experiment.

Notably, as summarized in the overview of eye-tracking studies in Frenck-Mestre (2002), language exposure plays an important role in proficiency, which could influence the L2 processing critically. As discussed in Section 2.3, J. Witzel et al. (2012) also report a native-like pattern of sentence processing when testing highly proficient Chinese speakers learning English.

Some ERP studies have shown evidence of neural activities of highly proficient learners similar to those of L1 speakers when processing sentences (See van Hell & Tokowicz, 2010 for an overview). Covey (2018) tests *wh*-dependency processing with native English speakers and Chinese-English bilinguals. It is shown that gap prediction is guided by grammatical knowledge for L1 speakers. Meanwhile, highly proficient L2 learners also benefit from the gap prediction effect for sentence processing with *wh*-movement. Native-like processing is also reported in some other ERP studies (Bowden et al., 2013; Foucart & Frenck-Mestre, 2012). Since the current study does not intend to detect the neural activities of grammatical processing, these studies will not be discussed further.

As predicted by the SSH, late L2 learners in many studies use grammatical knowledge less than native speakers do to parse complex sentence structures. However, studies have also shown that with high L2 proficiencies or strong L1/L2 similarities, some late learners may follow the same or similar processing patterns as native speakers. These findings are found especially in neural potentials in some EEG experiments.

Overall, a very rare chance of native-like processing can be detected for late L2 learners when a high level of L2 proficiency is reported. However, L2 proficiency is not the only factor

affecting syntactic processing. It is age and length of exposure that critically influence L2 processing and/or acquisition (Ullman, 2001; Weber-Fox & Neville, 1996).

### **3. Heritage Speakers**

Heritage speakers (HS) are defined as bilinguals whose weaker or non-dominant language corresponds to a minority language (Polinsky, 2018). Most often, the minority language is the first language, while the dominant language is the second language (in the case of sequential bilinguals) (Polinsky, 2018, p. 9). Heritage speakers have been referred to by many different terms such as semispeakers, incomplete acquirers, and early bilinguals (Polinsky & Kagan, 2007). Meanwhile, as HS are usually early bilinguals, the Shallow Structure Hypothesis (Clahsen & Felser, 2006b, 2006a, 2018) does not describe L2 processing patterns of this population. Therefore, this study expands on the applicability of SSH on the L2 processing of HS. Since the SSH covers late L2 learners whose dominant language is the L1, this study adds dominance and early acquisition as factors that may influence processing strategies.

Many psycholinguistic studies on HS are focused on the transfer effect from the L2 (dominant language) to the L1 (heritage language), heritage language acquisition, and heritage language attainment (see Bolger & Zapata, 2011 for an overview). There is very little work on L2 syntactic processing of HS populations and how L2 dominance affects processing (but see Ge et al., 2019).

In the current project, late L2 learners refer to late bilinguals who arrive in a country after maturation or the Critical Period (see § 3.1). Except at a very young age before puberty, most HS speak their L2 dominantly compared to late learners whose dominant language mostly remains the L1 or becomes roughly on par with the L2.

### **3.1 Critical Period Hypothesis (CPH)**

The Critical Period of first language acquisition refers to a period before L1 language development starts to cease (Lenneberg, 1967). This period starts from infancy and ends around puberty. It is argued in the literature that acquiring a second language will encounter the same challenge as L1 development faces when acquisition starts after the critical period (Gathercole & Thomas, 2009; Johnson & Newport, 1989; Long, 1990; Newport, 1990, 1991; Weber-Fox & Neville, 1996). When people are exposed to a second language, there is a performance decline of L2 attainment as the age of onset of acquisition increases. Specifically, compared to early learners who start the acquisition of their L2 before 8 years old, a decline of grammatical attainment is reported for the population that starts learning an L2 after the age of 8-10 years, i.e., the end of the critical period (Johnson & Newport, 1989; Newport, 1990). Language competence will not have significant improvement and will be very unlikely to reach native-like levels if a second language starts to be acquired after the critical period. However, much work also suggests that L2 improvement or L2 native-like proficiency can be achieved even if L2 acquisition starts after puberty during adulthood (Bialystok & Hakuta, 1999; Birdsong, 1992; Birdsong & Molis, 2001; Flege, 1999).

As reviewed in Section 2.4, the Shallow Structure Hypothesis predicts a dual-pathway model of language processing where the L2 parser does not rely heavily on the grammatical pathway but relies more on the heuristic one (Clahsen & Felser, 2018). Mental grammars are the basic resource of sentential representations in the grammatical pathway of a parser of L2 learners. Therefore, given the controversial views on the CPH mentioned above (i.e., whether a critical period exists or not for L2 ultimate attainment), some late L2 learners may possess enough grammatical knowledge to use the grammatical pathway for processing even when they start immersion in the L2 context after or at the end of the critical period.

Meanwhile, the Critical Period Hypothesis does not specify what role language dominance plays in L2 acquisition and attainment. For some HS, L2 is acquired after the completion of L1 acquisition, but L2 is the dominant language in their life (i.e., sequential bilinguals). In an English-dominant community, early Welsh-English bilinguals are found with no challenge in acquiring their dominant L2 English (Gathercole & Thomas, 2009). Thus, adult HS might have acquired systematic L2 grammars for L2 sentence processing. Their high levels of L2 dominance could be one of the reasons to facilitate and benefit L2 acquisition.

This study predicts that L2 dominance at least partially overrides the properties of heuristic L2 processing for late learners and early learners. Specifically, this project will test if L2 dominance can be used as a predictor of native-like patterns of sentence processing for both HS and late L2 learners.

### **3.2 Age of Acquisition versus Language Dominance**

Several predictors of processing strategies are to be distinguished: Age of Acquisition (AoA), Age of (First) Exposure (AoE), and Length of Residence (LoR). AoA refers to the age at which the learners start to immerse themselves in the L2 context as immigrants, while AoE is the age when classroom teaching of an L2 begins, the learner visits the L2 country, or other instances when the learner is first exposed to the L2. LoR refers to the duration for which the learner has been continuously immersed in the L2 context (Birdsong, 2006).

#### **3.2.1 Age of Acquisition**

Some work argues that there is a maturational constraint of L2 acquisition (Gathercole & Thomas, 2009; Johnson & Newport, 1989; Long, 1990; Newport, 1990, 1991; Weber-Fox & Neville, 1996). Other work shows this age hypothesis is not absolute (Bialystok & Hakuta, 1999; Birdsong, 1992;

Birdsong & Molis, 2001; Flege, 1999; Friederici et al., 2002). Some research suggests that many other factors like formal education in the L2 context, socioeconomic status, and L2 use rates will affect L2 proficiency (Birdsong & Molis, 2001; Flege et al., 1999; Hakuta et al., 2003).

Johnson and Newport (1989) extend the Critical Period Hypothesis of Lenneberg (1967) to L2 acquisition. Groups of Korean and Chinese early and late arrivals in the U.S were compared based on a grammaticality judgment task of twelve morphosyntactic rules. The groups were compared based on the age of arrival, which ranged from 3 to 39 years. Early arrivals have an AoA ranging from 3 to 15 years, while late arrivals have an AoA that ranged from 17 to 39 years. Two situations of AoE are identified for late arrivals. If the late arrivals took English lessons in their native country before immigrating to the US, their AoE is earlier than AoA. If the late arrivals had no exposure in their native country, they are identified as having the same AoE and AoA. Two hypotheses were tested.

Firstly, the exercise hypothesis predicts that the superior capacity of first language acquisition will remain intact after puberty, under the condition that the language learning capacity for L2 acquisition is exercised before maturation. If it is not exercised, then the subjects will lose this capacity after maturation. Secondly, the maturational state hypothesis predicts that maturation signifies the decline or disappearance of the superior capacity for L2 acquisition no matter how the language learning capacity has been exercised. The results show no significant effect on the scores of late arrivals by an earlier AoE in their native country. In other words, taking English lessons in the home country does not seem to be a significant advantage for L2 attainment. However, AoA plays a significant role in the final score of early and late arrivals compared to the native control group. Notably, the age of 7 as an AoA is a prominent line that distinguishes the ultimate performance. For all L2 learners, the scores show a significant decline starting from an

AoA of 7 years old. This is interpreted as a critical age limit for the onset of L2 acquisition to achieve a native level of proficiency. Also, the results support the maturational state hypothesis that the language learning capacity will not remain intact along with maturation (i.e., the age of 16). Therefore, the Critical Period Hypothesis accounts for L1 Acquisition (L1A) and L2 ultimate attainment (Johnson & Newport, 1989). Thus, no significant benefits are reported for late arrivals who arrive after 7 years old with an early AoE of taking English classes in their home countries. The age of arrival in the L2 context is the best measure of the age of onset of acquisition (AoA). In this case, AoA of 7 is a critical point to determine whether L2 attainment of grammatical knowledge can develop to a native level.

For this study, this means that AoA after 7 should significantly influence processing strategies. If native-like attainment and processing strategy are parallel, then L2 speakers with an AoA after seven years should process heuristically according to the SSH.

Birdsong (1992), on the other hand, reports results disputing the maturational state hypothesis in a study on French learners who speak English as their first language. The bilingual participants had started acquiring their L2 French at or post-puberty. However, some bilingual individuals performed like French native speakers in interpretation and grammaticality judgment tasks.

Flege et al. (1999) do not find any significant effect of AoA on scores of grammaticality judgment tasks in a group of Korean-English bilinguals. Instead, their study results suggest that the rate of English use significantly predicts lexically based morphosyntactic acquisition. Thus, language dominance is a factor in native-like attainment and processing, assuming the rate of use illustrates dominance.

Birdsong and Molis (2001) replicate the study of Johnson & Newport (1989) with a larger sample size of Spanish-English bilinguals. The cut-off age is set at 17 as the maturational age. The results indicate that early bilinguals with AoA less than 8 show consistent AoA effects on L2 attainment. However, instead of a dispersed data of late arrivals in Johnson and Newport (1989), Birdsong and Molis (2001) find a linear decline negative to the increase of AoA after maturation. It is notable that some late learners of this study even have similarly high scores as native speakers do. The authors argue that early exposure may cause a higher score. However, the percentage of L2 English use is a significant predictor for the performance of late learners with an AoA after 17.

A census study by Hakuta et al. (2003) suggests a decline of L2 English proficiency as AoA increases, but again, the linear trend is not flattened after the age of 15 or even 20, contradicting the Critical Period Hypothesis prediction on L2 acquisition. Also, socioeconomic status (especially the formal education of immigrants) positively affects the variance in L2 attainment. Longer formal education in the immigrant country is reported to promote higher levels of L2 proficiency (Hakuta et al., 2003).

The evidence from most studies suggests that L2 ultimate attainment declines as AoA increases and the critical period ends (Birdsong & Vanhove, 2016). Thus, Birdsong (2006) argues that AoA is the most prominent factor for L2 ultimate attainment. However, language use and thus language dominance is also a predictor of L2 performance for English learners (Birdsong, 2006).

### **3.2.2 Language Dominance**

According to Birdsong (2014), language dominance affects linguistic competence, language production, and processing. It is shown in “fluency of speech, lexical diversity, morphosyntactic knowledge” (linguistic competence), “length of utterances” (for production), and “parsing speed and accuracy” (for processing) (Birdsong, 2014, p. 375).

Domains of language dominance may differ for different contexts for language use such as “conversations with elder relatives, child-directed speech, watching TV news, and interactions in the workplace”. (Birdsong, 2014, p. 2) Thus, it needs to be assessed for different contexts compared to another language of the speaker (Birdsong, 2014).

AoA and L2 dominance are independent of each other with respect to domains of language dominance (Birdsong, 2014). In other words, a bilingual person acquiring an L2 at a later age than that of L1 does not mean the L1 must be the dominant language. For example, heritage speakers may simultaneously acquire the L1 and the L2 or acquire the L2 later than the L1. Nevertheless, they are L2 dominant speakers with the L2 used primarily in most domains. Therefore, language dominance and AoA should not be considered as always being significantly dependent on each other (Birdsong, 2014).

On the other hand, some studies report that AoA and L2 dominance are closely related to different dimensions and domains of language dominance. Flege et al. (2002) assess the language dominance of ninety Italian-English bilinguals with subjective (i.e., self-rating on use amount of L1 and L2) and objective (i.e. L1 and L2 production, L1 and L2 accent evaluation, L2 to L1 translation) methods. They found that the earlier a bilingual speaker arrives in the L2 context and the less the speaker uses the L1 in the new context, the more likely they consider the L2 to be the dominant language (and the more proficient they rate their L2). The subjective measure results align with the ones of objective measures.

This current study predicts that language dominance and AoA are different variables that may independently affect L2 processing. It tests the correlation between dominance measurement and age factors and addresses the processing dimension when considering the percentage of L2 use of different bilinguals (i.e., late and early arrivals). Based on the assessment method in Flege

et al. (2002), this current study evaluates the L2 dominance of all English learners through a subjective report in a survey. Along with AoA (of either less than 8 or higher than 18), the self-rated level of language dominance is used as an independent variable to predict the processing cost of reading RC and ADVP constructions.

### **3.3 L2 Processing by Early Bilinguals & Heritage Speakers**

Findings of neuroimaging studies on early bilinguals support the Critical Period Hypothesis of L2 acquisition and syntactic processing (Weber-Fox & Neville, 1996). In a study on early and late Italian-German bilinguals, Wartenburger et al. (2003) show that AoA has more impact on cortical representation when executing L2 grammatical judgment tasks than doing L2 semantic judgment tasks. Thus, compared to late learners, HS populations are more likely to perform native-like processing of complex sentences in their early acquired L2. In addition, the grammatical processing of HS in their minority language is argued to be more native-like because of an early AoA, compared to late L2 learners of that language (Montrul, 2006; Montrul et al., 2008, 2014). HS speak L2 as their dominant language. However, very few psycholinguistic studies on heritage speakers are focused on L2 dominance and AoA for L2 online processing.

In a study comparing Korean heritage speakers (HS) with late arrivals in the US, no significant difference in offline reading strategy is reported (Ha, 2005). Korean-English bilinguals with AoA before the age of 8 have non-significant differences in their parsing strategy compared to that of the English control participants. An off-line questionnaire was applied to ask a question about which noun phrase should be attached to the relative clauses in sentential items like *Someone shot the dog of the actress who was on the balcony*. Specifically, participants should decide whether the RC *who was on the balcony* is modifying NP1 *the dog* or NP2 *the actress*. The results reveal that with an early AoA before 8, the heritage speakers are more likely to attach the RC *who*

*was on the balcony* to the low site *the actress* than other bilinguals with AoA older than 8. Also, the data of all late bilinguals (having high attachment preference) is significantly different from the pattern of English native controls. Ha (2005) thus shows that HS and English native speakers apply the same parsing strategy.

On the other hand, late bilinguals have a high attachment preference when reading ambiguous RC attachments in their offline tasks. To some extent, this high attachment preference is accounted for by constraint-based models of language processing (Gibson & Pearlmutter, 1998). Thus, both HS and late bilinguals apply native-like processing strategies.

Ge et al. (2019) argue that, unlike native English speakers, Cantonese heritage speakers have difficulty integrating prosodic information into the semantic processing of sentences in their dominant L2 English. Subjects participated in a “look and listen” task with the visual world paradigm. When listening to different audio-recorded sentence trials, a picture depicting the meaning of a critical phrase in each trial was shown to the participants. The eye movement of participants on the pictures was recorded to convey the processing patterns. Different patterns were discovered between the HS participants and the native controls.

Therefore, it is still controversial whether HS can incorporate native-like strategies in processing their dominant L2. This study will add more evidence for the processing strategies found for HS and late English learners.

## 4. Research Gaps and Questions

### 4.1 Rationale

Some of the previous work suggests that L2 learners do not show a clear pattern of high or low attachment preference when processing a sentence like *The dean liked the secretary of the professor who was reading a letter* (Felser et al., 2003). This absence of a clear pattern is interpreted as shallow processing by Clahsen & Felser (2006a, 2006b). However, other studies provide empirical findings of native-like processing on highly proficient L2 learners (e.g. (Frenck-Mestre, 2002; J. Witzel et al., 2012)). For instance, J. Witzel et al. (2012) demonstrate a high attachment effect on processing *NI of N2 + RC* and a low attachment effect on processing *ADVP* modifiers by highly proficient learners of Chinese speakers residing in the US. Thus, J. Witzel et al. (2012) argue that the SSH of Clahsen & Felser (2006b, 2006a) can not fully account for L2 syntactic processing of constructions with *RC* and *ADVP* attachment in English, especially when considering highly proficient later L2 learners.

According to the language processing literature on native speakers (Gibson et al., 1996; Gibson & Pearlmutter, 1998), L1 sentence processing is guided by the parsing principles of *Recency & Predicate Proximity* (constraint-based models). Both J. Witzel et al. (2012) and Clahsen & Felser (2006b, 2006a) treat this principle as similar to the Late Closure principle of the Garden-path model (Frazier, 1978; Frazier & Fodor, 1978; Frazier & Rayner, 1982).

The current study does not argue for a modular or a parallel processing model. Instead, this study hypothesizes that *RC* and *AdvP* attachment constructions are parsed in different patterns for native speakers of English and English learners who speak Mandarin as their first language. The contributing factors for different patterns of L2 learners are *AoA* and *L2 dominance levels*.

This study follows the studies of J. Witzel et al. (2012) and Felser et al. (2003) to test the Shallow Structure Hypothesis but includes AoA and L2 dominance as additional independent variables. The SSH (Clahsen & Felser, 2006b, 2006a, 2018) has not yet been tested under the inclusion of age factors on L2 syntactic processing. Therefore, it makes no statement about how populations parse L2 sentences with different AoAs. On the other hand, the Critical Period Hypothesis (Johnson & Newport, 1989, 1991; Long, 1990; Newport, 1990) does not discuss L2 dominance for its influence on L2 ultimate attainment. Therefore, it does not predict how L2 dominance ratings interact with L2 grammatical attainment and, thus, processing strategies.

The SSH makes no prediction for AoA, and the CPH makes no prediction for L2 dominance of how they will affect L2 sentence processing. To fill these gaps, this study takes two steps.

- 1) It looks at how AoA individually predicts L2 processing patterns.
- 2) It tests how L2 dominance affects processing patterns including and excluding AoA.

Specifically, this study tests whether early and late L2 learners (with different L2 dominance ratings) follow the Late Closure principle or use *Recency & Predicate Proximity* strategies during processing complex sentences with RC and ADVP attachment. If L2-speaking subjects follow the parsing principles of native speakers, then the SSH hypothesis must be revised to include AoA and language dominance. This project thus provides insight into how L2 sentence processing is affected by language dominance in heritage speakers.

## **4.2 Research Questions and Predictions**

The following research questions (RQs) are investigated.

**RQ1.** Do late English learners (Mandarin L1) perform shallow or native-like processing?

If they perform native English-like processing, there should be a low attachment preference pattern when processing RCO, RCS, or ADVP attachment structures. In that case, the reading times of critical regions in the low attachment conditions should be significantly shorter than those in the high attachment condition. If they do not apply any native English-like strategies, they should be very unlikely to have any significant preference of high or low attachment on processing any type of these complex sentences.

**RQ2.** Do heritage speakers (L1 Mandarin) with high dominance and early AoA have the same processing patterns as native English speakers?

If the answer is yes, they should have low attachment preference when processing RCO, RCS, and ADVP items in their dominant L2 (English). In other words, the reading times of critical regions in trials of the low attachment condition should be shorter than those in the high attachment condition.

**RQ3.** Does L2 dominance level individually predict native English-like processing strategies for Mandarin-English bilingual speakers?

If it does, then the bilinguals with higher L2 dominance scores should have significant preference patterns of native English speakers while low L2 dominant bilinguals should not, regardless of their AoA.

The Shallow Structure Hypothesis is tested with heritage Mandarin Chinese Speakers (HMCS) and English Late Learners (ELL) who all speak Mandarin Chinese as their L1. A group of English Native Speakers (ENS) is also tested as the control group to be compared as showing L1 parsing strategies. Since the former two populations have experienced different AoA, their L2 attainment of grammatical knowledge will be distinctively comparable, based on the Critical Period Hypothesis (Johnson & Newport, 1989, 1991; Long, 1990; Newport, 1990; Weber-Fox &

Neville, 1996). Furthermore, as predicted by the Shallow Structure Hypothesis, L2 syntactic processing should vary depending on how native-like the L2 parser constructs the grammatical representation. Therefore, the HMCS subjects are predicted to perform native-like processing since they start their L2 acquisition before or during the Critical Period. Most importantly, HMCS speak their L2 dominantly starting from an early age. Under this circumstance, they should have a low attachment preference like native English speakers (Felser et al., 2003; Frazier, 1978; J. Witzel et al., 2012; N. Witzel et al., 2012).

The findings reported in J. Witzel et al. (2012) suggest that ELL speakers are expected to have a high attachment preference on processing sentences with the construction of NP1 *of* NP2 + RC. A low attachment preference on processing verb modifiers (ADVP) is also predicted. However, if the existing SSH framework is considered, the ELL subjects should not perform either native English-like or HMCS-like attachment preference when processing these items. The following list summarizes the previous findings of ADVP and RC attachment processing patterns in English learner populations.

- AdvP Attachment Preference: Low preference - J. Witzel et al. (2012)
- RCS attachment Preference: High Preference - J. Witzel et al. (2012)
- RCO attachment Preference: None - The SSH Framework <sup>17</sup>

Many SPR studies do not find any native-like pattern of sentence processing, as argued by J. Witzel et al. (2012). Previous findings support that data from Maze experiments (Forster et al., 2009) will indicate the processing costs of reading ambiguous sentences with RC and AdvP attachment (N. Witzel et al., 2012). N. Witzel et al. (2012) compare effect sizes of attachment

---

<sup>17</sup> RCO refers to the relative clause modifying object; RCS refers to the relative clause modifying subject.

preference elicited in self-paced reading (SPR), eye-tracking, and Maze paradigms. They argue that the grammatical Maze (G-Maze) experiment provides similar effect sizes of grammatical processing as the eye-tracking methodology. The current study follows this approach.

*Table 1.1 Predicted RT Patterns of ENS (or HMCS)*

RCS Predictions				
RCS	Region 5 (critical)	Region 6	Region 7	Total RT
RCS_Low	<b>herself</b>	in the theater	was...	---
RCS_High	<b>herself</b>	in the theater	was...	---
Low_RT pattern	<b>Shorter</b> <sup>18</sup>	Shorter	Shorter	Shorter
High_RT pattern	<b>Longer</b> <sup>19</sup>	Longer	Longer	Longer
RCO Predictions				
RCO	Region 6 (critical)	Region 7	Region 8	Total RT
RCO_Low	<b>were</b>	busy	in the office.	---
RCO_High	<b>was</b>	busy	in the office.	---
Low_RT pattern	<b>Shorter</b>	Shorter	Shorter	Shorter
High_RT pattern	<b>Longer</b>	Longer	Longer	Longer
ADVP Predictions				
ADVP	Region 6 (critical)	Region 7	Region 8	Total RT
ADVP_Low	<b>yesterday,</b>	but she won't serve	the plums.	---
ADVP_High	<b>tomorrow,</b>	but she won't serve	the plums.	---
Low_RT pattern	<b>Shorter</b>	Shorter	Shorter	Shorter
High_RT pattern	<b>Longer</b>	Longer	Longer	Longer

To explain how a complex sentence with modifiers could elicit a processing response, the following predictions are listed, based on the G-Maze data patterns of native speakers in the study of N. Witzel et al. (2012). The average response times (RTs) of ELL participants for each sentential

<sup>18</sup> “Longer” means the reading time in this region is significantly longer than its counterpart in a high or low condition.

<sup>19</sup> “Shorter” means the reading time in this region is significantly shorter than its counterpart in a high or low condition.

type are predicted to be significantly longer than that of HMCS and ENS participants. ENS participants are predicted to have the shortest RT among the three groups since English is their first and dominant language.

1) **English Native Speakers.** ENS participants are predicted to have a low preference for RC and AdvP attachment sites. All items with a low-attachment design will be processed easier than the ones with a high-attachment design. The RT patterns are predicted, as shown in Table 1.1. The critical regions are the most important data points for final statistical analysis.

*Table 1.2 Predicted RT Patterns of ELL Based on J. Witzel et al. (2012)*

RCS Predictions				
RCS	<b>Region 5 (critical)</b>	Region 6	Region 7	Total RT
RCS_Low	<b>herself</b>	in the theater	was...	---
RCS_High	<b>herself</b>	in the theater	was...	---
Low_RT pattern	<b>Shorter</b>	Longer	Longer	Longer
High_RT pattern	<b>Longer</b>	Shorter	Shorter	Shorter
RCO Predictions				
RCO	<b>Region 6 (critical)</b>	Region 7	Region 8	Total RT
RCO_Low	<b>were</b>	busy	in the office.	---
RCO_High	<b>was</b>	busy	in the office.	---
Low_RT pattern	<b>Shorter</b>	Longer	Longer	Longer
High_RT pattern	<b>Longer</b>	Shorter	Shorter	Shorter
ADVP Predictions				
ADVP	<b>Region 6 (critical)</b>	Region 7	Region 8	Total RT
ADVP_Low	<b>yesterday,</b>	but she won't serve	the plums.	---
ADVP_High	<b>tomorrow,</b>	but she won't serve	the plums.	---
Low_RT pattern	<b>Shorter</b>	Shorter	Shorter	Shorter
High_RT pattern	<b>Longer</b>	Longer	Longer	Longer

2) **Heritage Mandarin Chinese Speakers.** Since HMCS participants are argued to have acquired native-like grammars because of the early AoA and speaking L2 dominantly, their processing should have similar patterns to the ENS patterns (Table 1.1).

3) **English Late Learners.** As predicted by the SSH and Felser et al. (2003), ELL will not show any significant effect of attachment preference (e.g. lower or higher) when processing modifiers. However, based on J. Witzel et al. (2012) findings, it is predicted that ELL participants in the current study will have a high preference for RC attachment sites and a low preference for AdvP attachment sites. Table 1.2 lists the predictions for ELL processing patterns under this condition.

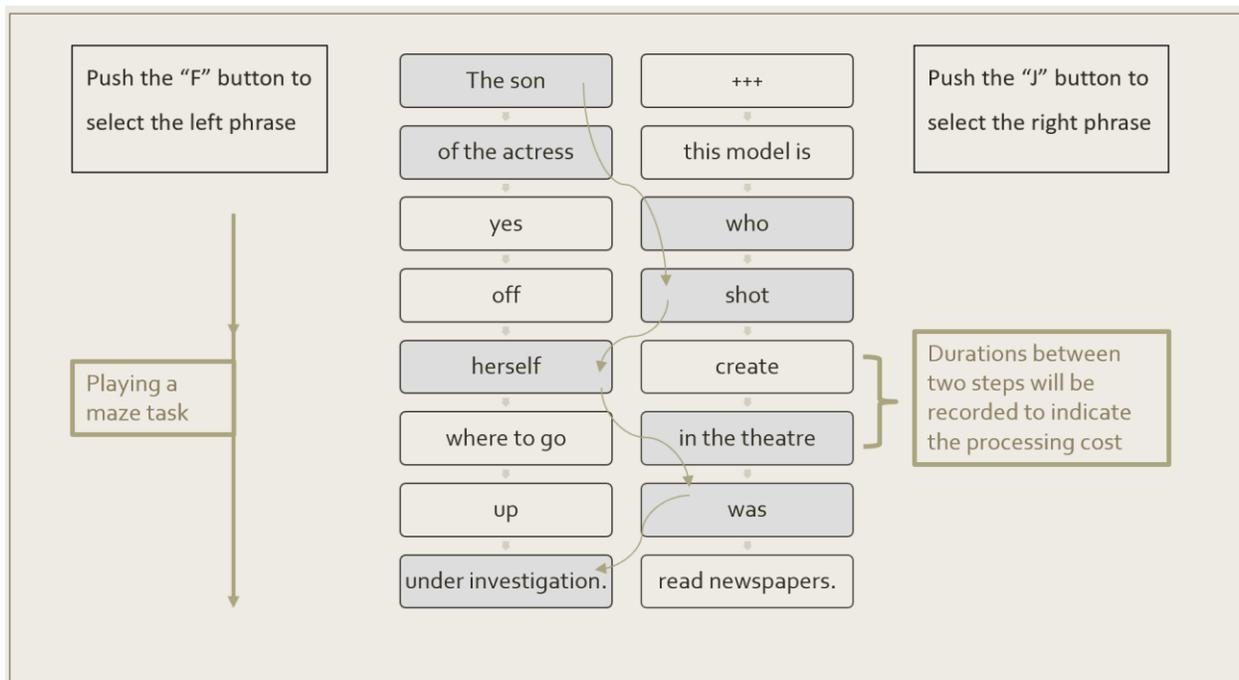
According to both accounts of GPM (Frazier, 1978; Frazier & Fodor, 1978) and CBMs (Gibson et al., 1996; Gibson & Pearlmutter, 1998; Pearlmutter & Gibson, 2001), if the ELL participants have no significant preference for low or high attachment processing RC- or AdvP-modifying sentences, they do not process those sentences with native-like parsing strategies. Then the Shallow Structure Hypothesis would be supported for this group. On the other hand, if the data of ELL conforms with the prediction listed in Table 1.2, the SSH needs to be adjusted.

## **Chapter 2    Methods & Results**

## 5. Methods

### 5.1 Experimental Design

An online experiment of the G-Maze paradigm was designed to measure the time course of processing ambiguous RC attachment and AdvP attachment of three populations (i.e., heritage Mandarin speakers, English late learners with L1 Mandarin, and English native speakers). During a trial, subjects are asked to serially put together an eight-region sentence like *The son, of the actress, who, shot, herself, in the theatre, was, under investigation* through selection between the correct phrase and a distractor (See Figure 2.1). Keyboard-press durations were recorded as processing costs and temporarily saved on the server of Pavlovia<sup>20</sup>. After completion, the data was downloaded and saved on the server at the University of Saskatchewan.



**Figure 2.1** *The G-Maze Paradigm*

<sup>20</sup> <https://www.pavlovia.org>

In the G-Maze experiment, the participants are presented with sentences with three types of modifiers: 1) RC modifying subjects (RCS); 2) RC modifying objects (RCO); 3) AdvP modifying VPs. Each item is divided into eight regions as shown in Table 2.1. The bolded texts in Table 2.1 are the critical regions for each sentential item since they are the ones to disambiguate the sentences. In detail, the participants see two words/phrases at a time (or per page in the window screen of the experiment). This means each trial has eight pages (or eight steps) to complete. One of the presented words (or phrases) grammatically fits the sentence (see the grey-highlighted boxes in Figure 2.1). The other one is a distractor word or phrase (see the non-highlighted boxes in Figure 2.1) which is presented on either the right or left side of that grammatically acceptable word or phrase. At each step or each page, the participants are asked to make their selections by pressing one of two buttons: the "F" on the keyboard to select the word or phrase on the left; the "J" on the keyboard to select the word or phrase on the right. The first region of each trial has no word-type or phrase-type distractor. Instead, there is a "+++" symbol which comes as a distractor. The "+++" symbol also tells the participants that this page is the 1<sup>st</sup> page of a trial (i.e., the start of a new trial). The participants need to select the grammatically acceptable word or phrase as fast as they can. Once a selection is made by pressing either the "J" or the "F" button, the screen will automatically jump to the next page (which will be the next region of the same sentence or a new trial if the previous page is the last region of a trial).

Reading times (RTs) of the critical regions are extracted to signal the processing cost. Post-critical regions are also considered to detect any spillover effect if present. The spillover effect refers to an extension of processing difficulties into the next few words after an ambiguity region. Take the trial *The son of the actress who shot himself in the theatre...* as an example, a native speaker will encounter processing difficulty in the reflexive *himself* because they prefer to interpret

the RC as a modifier of *the actress*, but not *the son*. Therefore, they will slow down when encountering *himself*. This effect of slowing down will spill over onto the next few words, in this case, *in the theatre*. On the other hand, if they have no processing difficulty with *himself*, this critical region is unlikely to elicit an effect of spillover in post-critical regions.

*Table 2.1 Three Items*

Region 1	Region 2	Region 3	Region 4	<b>Region 5</b>	Region 6	Region 7	Region 8
The son	of the actress	who	shot	<b>herself/ himself</b>	in the theater	was	under investigation.

RCS

Region 1	Region 2	Region 3	Region 4	Region 5	<b>Region 6</b>	Region 7	Region 8
The doctor	phoned	the boss	of the secretaries	who	<b>were/ was</b>	busy	in the office.

RCO

Region 1	Region 2	Region 3	Region 4	Region 5	<b>Region 6</b>	Region 7	Region 8
Anne	will serve	the apples	she	picked	<b>yesterday/ tomorrow,</b>	but she won't serve	the plums.

AdvP

Since the first region of each trial has no distractors, the variance of RTs in this region could potentially affect the results. Therefore, only RTs of region 2 ~ 8 are summed to calculate the total RT (tRT) of a trial.

## 5.2 Equipment

Using Psychopy<sup>21</sup> as the builder, a G-Maze experiment of reading ambiguous sentences with RC attachment and AdvP attachment was developed through a laptop within the Windows-10

<sup>21</sup> <https://www.psychopy.org/about/index.html>

operating system. Using the same laptop, the G-Maze experiment was then published online through Pavlovia.org. Each participant used their own computer or laptop at home or any other preferred location with internet access. An access link was generated during the publication to Pavlovia. The experiment was accessible through a web browser by clicking that link.

### **5.3 Participant Recruitment**

Recruitment was done through PAWS (an internal forum of the University of Saskatchewan), linguistlist.org (the online platform of the international organization for linguists), Facebook, and Twitter. Three population groups were targeted. Potential HMCS participants were expected to be living in Canada, the US, or another English-speaking country. The ELL participants were expected to be living in an English-speaking country or China. Participants are required to speak Mandarin Chinese as their L1 and as their most dominant language before the age of 18. Monolinguals were expected to be living in Canada, the US, or another English-speaking country.

Data was collected through Pavlovia.<sup>22</sup> Following the current version of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TPCS 2 2018), the University of Saskatchewan Behavioural Research Ethics Board has reviewed and approved this project with a certificate of ethics approval.

### **5.4 Procedures**

The procedures follow the grammatical-Maze method used by N. Witzel et al. (2012). In the experiment, a survey is inserted before the Maze task starts. The survey includes a list of questions about the linguistic background and demographic information of the participants. Due to the anonymous nature of the recruitment process, all participant groups were identified through their

---

<sup>22</sup> Pavlovia website: <http://pavlovia.org>; The project repository: <https://gitlab.pavlovia.org/yux580/mazeusask>

answers to the demographic questions. Each participant had to answer the questions to be able to continue to the maze task.

Demographics were asked in this survey. There are also four critical questions about L2 use in daily life: 1) How much do you use English per week with your family; 2) How much do you use English per week with your friends; 3) How much do you use English (per week) with other people (e.g. Facebook friends); 4) How much do you use English per week at school or work? Each question is followed by 11 options from level 0 (0%) to level 10 (100%). For instance, if a participant self-reports 10% of English use at home, 100% at school/workplace, 90% with friends, and 90% in other places (like Facebook), the final score will be 29 (i.e.  $1+10+9+9=29$ ). The level of 29 is referred to as the L2 dominance level of this participant. For the AoA questions in the survey, bilingual participants were asked at which age they started to learn English and how many years of education they had taken with English as the instructional language.

After completing the survey, the participants were brought to the maze task interface. Each sentence had eight separate regions in the maze task (as introduced in § 5.1). Each region has both a grammatically legal answer and a distractor. To succeed in a trial, participants need to choose the grammatically legal option instead of a distractor. Only the trials with all eight regions being selected correctly were considered accurate. Before the formal task, they were presented with 2 to 3 warmup sentences to familiarize themselves with the formal procedure. The entire process of the maze task took 20 to 30 minutes to complete. All trials were played in random order with fillers in between them.<sup>23</sup> The formal maze task was designed to present at least one filler immediately after two trials were played adjacently. These fillers prevented the participants from playing too many

---

<sup>23</sup> But see Section 6 for an ORDER effect.

trials continuously and becoming aware of the sentential patterns. The idea was to drive the readers to parse each trial unconsciously. For each sentence, participants were asked to assemble its eight regions into a complete sentence as fast as possible. The paradigm and fragmentized sentence examples are shown in Figure 2.1 and Table 2.1. There are a total of 90 sentences presented in the task, which includes 60 trials and 30 fillers. After playing half of the trials, a break page was shown to the participants to let them take a rest. The duration of this break was controlled by the participants. An instruction of pressing the “SPACE” key was presented on the break-page to continue the task.

In the end, a total number of 29 ELLs, 25 ENSs, and 18 HMCSs participated in this online experiment. All of them reported having completed or receiving post-secondary education. All of the ELL participants have reported an AoA which is more than 18, while each HMCS participant reported an AoA below 8. According to the results of the survey, each bilingual participant speaks Mandarin as their first language and English as their second language. Only two late learners reported their length of residence (in an English-speaking country) as being not less than 10 years. Meanwhile, 17 ELLs reported their length of residence as 0 which means they had resided in an English-speaking country for less than one year. Other ELLs reported different lengths of residence from 1 to 9 years. Up to the date of participation, all HMCS participants had been living in an English-speaking country for more than 10 years. All HMCSs reported that their education in an English-speaking country started from elementary school with English as the instruction of language since then. Most of the ELLs reported that they had not received any formal education with English as the instruction of language before taking their post-secondary education.

Three ELL and one HMCS participants were removed from the data due to their low accuracy rates. They failed most of the trials (only 20-30% of trials were done successfully). That

means they either did not understand the rules of the maze task or just did not pay attention. Therefore, the final data for analysis was composed of records from 17 HMCS, 25 ENS, and 26 ELL participants.

## 5.5 Material

All formal trials were taken either from Felser et al. (2003) or J. Witzel et al. (2012). By recording and analyzing every region's processing time in a sentence like (22a), (22b), or (22c), different patterns of attachment preference can be determined for the three groups of participants. Since the results of Felser et al. (2003) and J. Witzel et al. (2012) are qualitatively different, the RCO items are picked from the experimental material of Felser et al. (2003), who did not apply either the RCS or AdvP constructions. The RCS and AdvP items were taken from the experimental materials in the study of J. Witzel et al. (2012), who did not test their hypothesis on the RCO constructions. Each item has 10 cases respectively for low attachment and high attachment constructions. In sum, 60 trials were applied for these three items. Thirty random fillers were mixed into the trials.

(22a) RCS: The son of the actress who shot herself/himself in the theater was under investigation. (Low/High attachment)

(22b) RCO: The reporter phoned the boss of the secretaries who were/was busy in the office. (Low/High attachment)

(22c) ADVP: Anne will serve the apples she picked yesterday/tomorrow, but she won't serve the plums. (Low/High attachment)

Each item has two counterbalanced conditions that could cause either low or high attachment site preference to be modified by the RC or AdvP. As indicated by item (22a), (22b), and (22c), a whole sentence is disambiguated when encountering a preferred reflexive (e.g.

*herself/himself*), auxiliary *be* (e.g. *was/were*), or adverb phrase (e.g. *yesterday/tomorrow*) (See the **Appendix** for the material presented in the experiment).

## 6. Results

### 6.1 Accuracy and L2 Dominance Rates

Data was collected and converted into a CSV file, which was then processed in R Version 3.6.2<sup>24</sup> for statistical analysis. The inferential statistical models are introduced in Section 6.2.

#### 6.1.1 Accuracy

Heritage Chinese Mandarin Speakers (HMCS) have the top mean accuracies (RCS 89.41%, RCO 90.29%, ADVP 85.88%) among the three groups. Followed is the English Native Speakers (ENS) group which has higher mean scores in all three items (RCS 86.73%, RCO 84.04%, ADVP 83.08%) compared to those of English Late Learners (ELL) (RCS 85.93%, RCO 79.44%, ADVP 73.89%). Mean scores of the three subject groups are summarized in Table 2.2.

*Table 2.2 Mean Accuracy Scores of The Three Groups*

Subject Groups	Item Types	Accuracy Mean (%)	SD (%)
ENS	RCS	86.73	12.80
	RCO	84.04	14.21
	ADVP	83.08	16.68
ELL	RCS	85.93	12.17
	RCO	79.44	20.21
	ADVP	73.89	18.41
HMCS	RCS	89.41	11.84
	RCO	90.29	11.11
	ADVP	85.88	16.13

Meanwhile, as shown in Table 2.2, with the exception of the RCS item (SD: 12.17), the ELL group has the highest SD values, which were 20.21 of RCO and 18.41 of ADVP. The HMCS participants have the lowest SDs for all these three items.

---

<sup>24</sup> <https://www.r-project.org/>

These descriptive statistics in Table 2.2 suggest that both HMCS and ENS processed these items more successfully than the ELL participants. Moreover, the distribution of the ELL data shows greater variation than the other two groups indicating this group of late learners had the largest variance with respect to accuracy in their maze task when processing the experimental items.

### 6.1.2 L2 Dominance Rates

The dominance levels of both bilingual groups were evaluated according to self-reported L2 use. Each ELL and HMCS subject was asked in the survey to answer four questions. The participants needed to choose one option to rate their L2 English use under four circumstances, as shown in Table 2.3. L2 English use in the four conditions is summed as the total L2 dominance score (or level).

*Table 2.3 Rating Scale for L2 English Use in Self-Report Survey*

L2 Use AT HOME	L2 Use AT SCHOOL/WORKPLACE	L2 Use WITH FRIENDS	L2 Use OTHERS	L2 Use Score IN TOTAL
0~10	0~10	0~10	0~10	SUM (of Four Conditions)

Only 9 ELLs reported their L2 dominance levels higher than 20, while most HMCSs rated themselves on a level higher than 30. The mean L2 dominance levels of the two bilingual groups are reported in Table 2.4.

*Table 2.4 Descriptive Statistics of Self-Reported L2 Dominance Rates*

Subject Groups	Min	Max	Median	Mean	SD
ELL	0.00	40.00	3	11.00	14.12
HMCS	21.00	40.00	36	34.94	4.94

## 6.2 Predict from AoA and L2 Dominance

After the CSV file was imported into R, graphing was done using the `ggplot2` library (Wickham, 2016). Three generalized linear mixed-effect regression (LMER) models were created to estimate the variance of reading times dependent on different variables. The LMER models were built with the `lmer` function of the `lme4` package (Bates et al., 2015). The `confint` function was used to calculate the 95% confidence intervals. Each model included item number and speaker as random effects. Trials with tRT above 30 seconds and critical RT above 7.5 seconds were removed from the data. This is to remove the anomalous data that surprisingly deviated from a normal distribution. In other words, if a participant spent more than 30 seconds in total on a trial or spent more than 7.5 seconds on the critical region, the trial is considered as being anomalous and failed. Furthermore, the RT values in milliseconds were converted to log values to normalize the data for the inferential statistics which are reported in Sections 6.2.1 and 6.2.2.

To test the AoA factor, an interaction between *population* (ENS, ELL, and HMCS) and *Item\_Attachment* (RCS\_Low, RCS\_High, RCO\_Low, RCO\_High, ADVP\_Low, and ADVP\_High) was considered in two LMER models as the independent variables. Respectively, these two models had RT of the critical region (*RT\_Cr*) and RT of the whole trial (*RT\_TOTAL*) as the dependent variables. Meanwhile, the variables of *SPEAKER*, *Item\_No*, *L2Dominance\_score* (L2 dominance scores), and *LoR* (length of residence) are set in the model as random variables.

- $model.RT\_Cr = (RT\_Cr\_Log \sim Population * Item\_Attachment + (1|Item\_No) + (1|SPEAKER) + (1|L2Dominance\_score) + (1|LoR), data = DATA)$
- $model.RT\_TOTAL = (RT\_TOTAL\_Log \sim Population * Item\_Attachment + (1|Item\_No) + (1|SPEAKER) + (1|L2Dominance\_score) + (1|LoR), data = DATA)$

To test the L2 dominance factor, an interaction between *L2Dominance* (ENS, L2LESS, and L2MORE)<sup>25</sup> and *Item\_Attachment* (RCS\_Low, RCS\_High, RCO\_Low, RCO\_High, ADVP\_Low, and ADVP\_High) was considered in another LMER model. The RT of the critical region was considered as the independent variable. The variables of *SPEAKER*, *Item\_No*, *AoA* (age of acquisition), and *LoR* (length of residence) are set in the model as random variables.

- $model.RT\_Cr = (RT\_Cr\_Log \sim L2Dominance * Item\_Attachment + (1|Item\_No) + (1|SPEAKER) + (1|AoA) + (1|LoR), data = DATA)$

To rule out the effect of AoA (i.e., to compare the effects between AoA and L2 dominance), a model with a triple interaction of *L2Dominance\*Population\*Item\_Attachment* as the fixed variables was created to predict sentence processing patterns. In this model, *SPEAKER*, *Item\_No*, and *LoR* (length of residence) are the random variables.

- $model.RT\_Cr = (RT\_Cr\_Log \sim L2Dominance * Item\_Attachment * Population + (1|Item\_No) + (1|SPEAKER) + (1|AoA) + (1|LoR), data = DATA)$

By rotating the intercepts in these models, processing patterns (Low VS. High) of each population or group were predicted in the statistical results. For instance, if the intercepts include the log value of RTs of ELL and RCS\_Low, it was then compared to the estimated result of RCS\_High, accompanied with the relevant t-values, p-values, and confidence intervals.

A linear mixed-effect regression model was run to just observe the total reading times across the three groups of participants. As Table 2.5 shows, the ELL group has the longest tRT with an average duration of 15.3 seconds (s) (intercept). This value is significantly longer than the

---

<sup>25</sup> To further test the effect of L2 dominance, all Mandarin speakers (i.e., ELL and HMCS) were firstly treated as a single group. To find out the dominance level in which these L2 learners could be reaching in a native-like processing pattern, this group was then divided into two subgroups with a cut-off dominance score. Therefore, like the age of acquisition (AoA), the L2 dominance factor includes three categorical variables: ENS (i.e., English native speakers), L2MORE (i.e., higher L2 dominant speakers), and L2LESS (i.e., lower L2 dominant speakers).

average durations of both the ENS (11.3 s) and HMCS (11.6 s) groups. The ENS participants spent the shortest durations when processing these trials.

*Table 2.5 Observe the Total RTs across the Three Populations*

Three groups_Total RT	Estimate	Std. Error	2.50%	97.50%	t value	Pr(> t )
(Intercept)	15386.38	337.66	14729.96	16042.74	45.57	< 2e-16 ***
Population_ENS	-4065.88	481.53	-5001.97	-3129.80	-8.44	3.90e-12 ***
Population_HS	-3761.68	541.87	-4814.94	-2708.11	-6.94	1.99e-09 ***
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1						

*Table 2.6 Observe the Total RTs across the Three Items*

Item Comparison in Total RTs	Estimate	Std. Error	2.50%	97.50%	t value	Pr(> t )
(Intercept)(i.e., ADVP)	13666.95	346.86	12987.66	14346.37	39.40	0.00000***
RCO	-1004.04	252.17	-1497.32	-510.69	-3.98	0.000197 ***
RCS	-1099.82	252.18	-1593.10	-606.41	-4.36	0.000055 ***
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1						

*Table 2.7 Observe the Critical-Region RTs across the Three Items*

Item Comparison in Critical Region RTs	Estimate	Std. Error	2.50%	97.50%	t value	Pr(> t )
(Intercept)(i.e., ADVP)	2001.97	63.65	1878.08	2125.88	31.45	0.00000***
RCO	-242.48	76.30	-391.30	-93.68	-3.18	0.0024 **
RCS	-355.07	76.30	-503.91	-206.27	-4.36	0.00002 ***
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1						

Also, Table 2.6 and Table 2.7 present the comparison of tRTs and the critical-region RTs across the three items. Specifically, a linear mixed-effect regression model predicts that participants read the ADVP with the longest duration (13.7s) than the other two items (see Table 2.6). Then the RCO item seems to have taken a longer duration to process than the RCS item (with a difference of 95 ms in tRTs between them). This is also reflected by the critical region RTs as reported in Table 2.7 (with a difference of 113 ms in critical-region RTs between them). Even

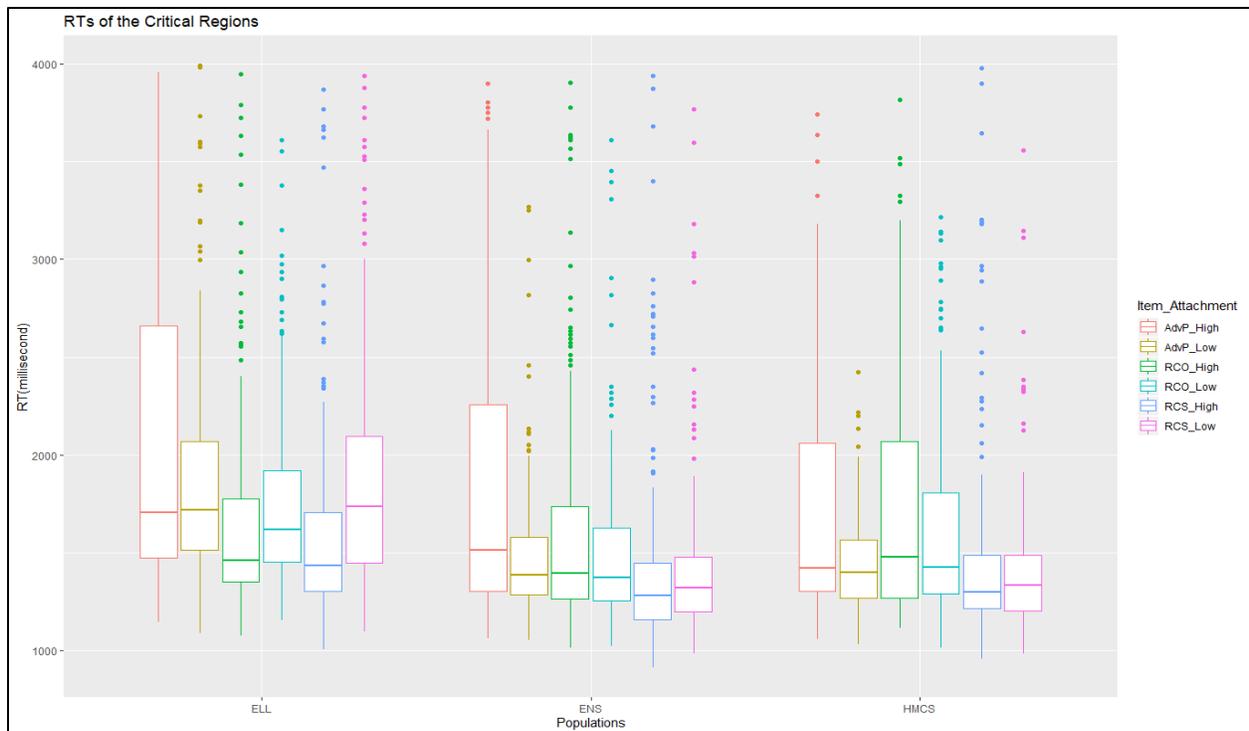
though, RCO should be the most challenging item among the three constructions in this experiment, which is further discussed at the end of Section 7.1.1.

### 6.2.1 Predict from AoA

By graphing the raw data in Figure 2.2, it appears the distribution of the critical region RTs against the tRTs shows a clear distinction between ELL and the other two groups. Meanwhile, the HMCS participants do not seem to differ sharply from the native speakers. Overall, this figure indicates it costs the ELL participants more to process these complex sentences, compared to both native speakers and heritage speakers. The HMCS group parallels with the ENS participants with respect to the processing cost.



*Figure 2.2 Regression Graphing on Total RTs & Critical RTs Among Three Groups*

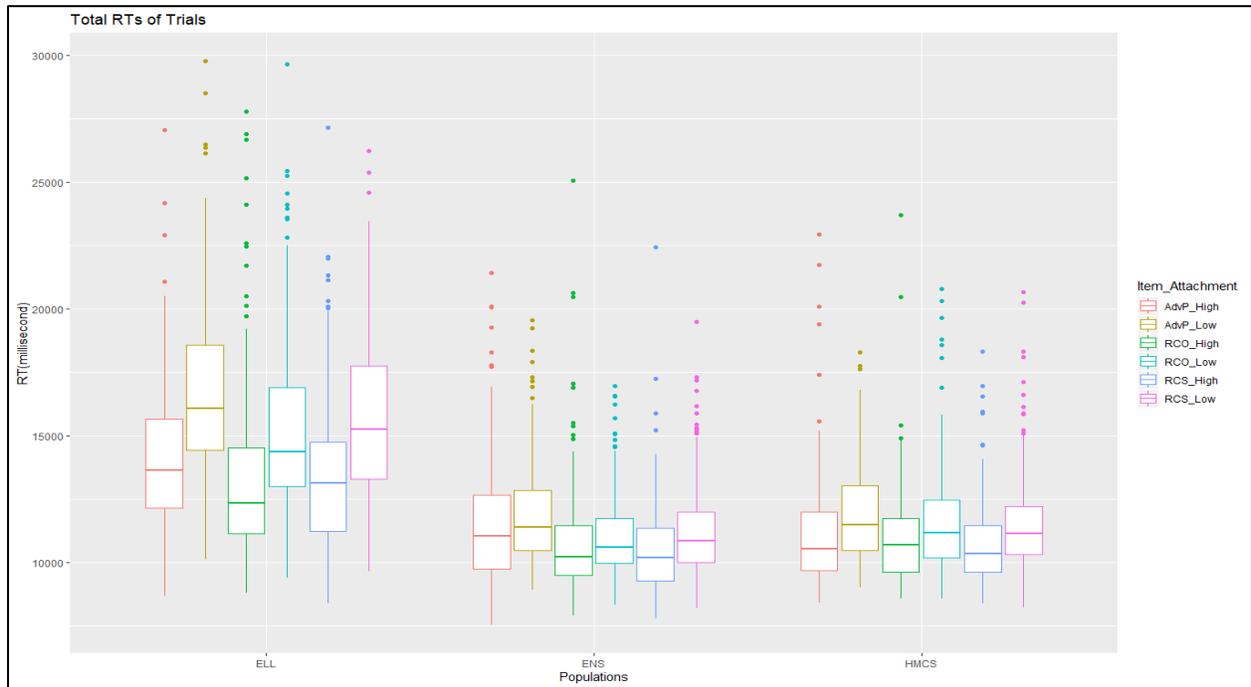


**Figure 2.3** Boxplot of the RTs of the Critical Regions for Three Items (Three Populations)

Furthermore, the RTs of the critical regions in the three items show different patterns across the three groups as indicated by Figure 2.3. English native controls seem to have shorter RTs in the low attachment conditions than in the high attachment conditions when reading the ADVP and RCO items, while for the RCS item, the processing times of the critical region in this graph are not distinguishable between the low and high attachment conditions. The heritage speakers show similar patterns with the native controls, which agrees with Figure 2.3. While the late English learners also have a similar pattern on the ADVP and RCS items with those of native speakers and heritage speakers, their RT pattern of the critical region in the RCO item is the opposite.

However, Figure 2.4 shows that all these three groups tend to read the high attachment trials faster than doing the low attachment trials, through plotting the total reading times (tRTs) of

each item. Also, the difference between low and high attachment trials of the ELL group looks more distinguishable than the patterns shown for ENS and HMCS.



*Figure 2.4 Boxplot of the Total RTs of the Three Items (Three Populations)*

Therefore, before presenting the LMER model results, a note is mentioned here. As reported below in this Section 6.2, models on tRTs yielded a universal pattern for all three constructions and all three populations. For each sentential construction (either RCS, RCO, or ADVP), the results show that the tRTs of high attachment items are significantly shorter than that of the low attachment condition.<sup>26</sup> This result is slightly surprising for native speakers and heritage speakers (who were expected to have shorter tRTs of reading low attachment trials than their high attachment trials, according to previous findings). It is believed to be caused by the experimental

<sup>26</sup> Also, being outside the scope of this study, results for pre-critical regions are not reported in this thesis. There is a universal pattern for all the three groups of participants that high attachment trials were read significantly faster in these regions, which is identical to the tRT patterns.

design. The actual experiment had been designed that all high attachment items were played after their low counterparts were completed. There is a short break between the low and high attachment sections. Even so, it seems this design imperfection (hereafter referred to as the ORDER effect) has eased the processing of high attachment trials and caused universally shorter tRTs of these trials. Participants could have become familiar with the high attachment counterparts after they finished the low attachment trials. Detailed statistics are reported in the following subsections. This issue will be discussed in detail in Section 7.

Nevertheless, for the ENS participants, the statistical results of the critical regions in RCS, RCO, and ADVP items do suggest similar patterns to those that are reported in previous literature (Felser et al., 2003; J. Witzel et al., 2012; N. Witzel et al., 2012).

#### **6.2.1.1 English Native Speakers**

Table 2.8 lists the inferential statistical results of the ENS group. For the RCS item, ambiguity is resolved by choosing the critical region *herself/himself* (i.e. region 5) in the relative clause to modify a preferred noun phrase *the son* or *the actress*. An effect of low attachment preference is found in this region by native English speakers in the SPR experiment of Felser et al. (2003). In the current study, the LMER model result of this region shows a shorter RT (7.24 log ms or 1402ms) of the low attachment condition for the ENS participants. But this result is not significant. Moreover, the result of total RTs (tRTs) is significantly shorter in the high attachment condition (9.27 log ms or 10615 ms) than that of its low counterparts (9.34 log ms or 11338 ms). These results of critical region RTs and tRTs could be attributed to the above-mentioned ORDER effect. It seems the language parser has generated a particular pattern when participants start to read the high attachment trials.

**Table 2.8 Inferential Statistics on Regions of Interest in The Three Items (Native Speakers)**

Examples	RCS_Critical Region RTs	Estimate (Log Value)	Std. Error	2.50%	97.50%	t value	Pr(> t )	Estimate (ms)
himself herself	(Intercept)	7.28	0.04	7.21	7.35	199.14	0.0000***	1454
	Low Attachment	-0.04	0.04	-0.11	0.04	-0.90	0.36828	1402
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1								
The son of the actress who shot herself/himself in the theater was under investigation.	RCS_Total RTs	Estimate (Log Value)	Std. Error	2.50%	97.50%	t value	Pr(> t )	
	(Intercept)	9.27	0.03	9.22	9.33	323.30	0.0000***	10615
	Low Attachment	0.07	0.02	0.03	0.10	3.31	0.00121 **	11338
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1								
was were	RCO_Critical Region RTs	Estimate (Log Value)	Std. Error	2.50%	97.50%	t value	Pr(> t )	
	(Intercept)	7.41	0.04	7.34	7.48	202.26	0.0000***	1645
	Low Attachment	-0.08	0.04	-0.16	0.00	-2.01	0.046989 *	1515
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1								
The reporter phoned the boss of the secretaries who were/was busy in the office.	RCO_Total RTs	Estimate (Log Value)	Std. Error	2.50%	97.50%	t value	Pr(> t )	
	(Intercept)	9.30	0.03	9.25	9.36	324.23	0.0000***	10960
	Low Attachment	0.01	0.02	-0.03	0.05	0.61	0.54187	11094
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1								
tomorrow yesterday	ADVP_Critical Region RTs	Estimate (Log Value)	Std. Error	2.50%	97.50%	t value	Pr(> t )	
	(Intercept)	7.55	0.04	7.48	7.62	205.75	0.0000***	1906
	Low Attachment	-0.26	0.04	-0.34	-0.18	-6.29	0.00000 ***	1472
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1								
Anne will serve the apples she picked yesterday/tomorrow, but she won't serve the plums.	ADVP_Total RTs	Estimate (Log Value)	Std. Error	2.50%	97.50%	t value	Pr(> t )	
	(Intercept)	9.36	0.03	9.31	9.42	326.00	0.0000***	11660
	Low Attachment	0.03	0.02	0.00	0.07	1.72	0.08825 .	12068
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1								

For the RCO item, an effect of low attachment preference is found in this region by native English speakers in the SPR experiment of Felser et al. (2003). Notably, as Table 2.8 indicates, the critical region RT (for *were/was*) in the low attachment condition (7.33 log ms or 1515 ms) is significantly lower than its high attachment counterpart (7.41 log ms or 1645 ms). LMER model of RTs for post-critical region-1 shows a spillover effect that low attachment RTs were

significantly shorter than the high attachment counterparts. However, post-critical region-2 has an opposite pattern compared to post-critical region-1.<sup>27</sup> This pattern continues to exist in the tRTs (Low: 9.31 log ms or 11094 ms; High: 9.30 log ms or 10960 ms), which is very similar to the result of the tRTs of the RCS construction, although this RCO result is not significant. These behavioral patterns of processing RCO indicate that native speakers have a low attachment preference. However, the low attachment preference is shaded by the ORDER effect in the RTs of post-critical region-2 and the tRTs.

Ambiguity in the ADVP item is resolved by the adverb phrases (e.g., *yesterday/tomorrow*) in those trials (see the examples in Table 2.8). The processing pattern of ADVP aligns with that of RCO. The effect of low attachment preference on the critical region is significant (Low: 7.29 log ms or 1472 ms; High: 7.55 log ms or 1906 ms). However, the ORDER effect seems also significant in tRTs which has caused a pseudo-preference of the high attachment (High: 9.36 log ms or 11660 ms; Low: 9.39 log ms or 12068 ms).

### **6.2.1.2 English Late Learners**

As shown in Figure 2.3, the RT patterns of the critical region in the RCS item and RCO item processed by the ELL participants are just opposite to those of native speakers. Subsequently, the LMER model (see Table 2.9) estimates that RT of the critical region in the low attachment condition (7.57 log ms or 1950 ms) is significantly longer than the one in the high attachment condition (7.34 log ms or 1544 ms) for the RCS item. When looking at the statistics of the critical region (e.g. *were/was*) in the RCO item, it appears that low attachment trials were read in a longer duration (estimated as 7.49 log ms or 1794 ms) than the estimate of the high attachment trials (7.38

---

<sup>27</sup> Post-critical region statistics are not presented since they are not the most interesting data in previous literature.

log ms or 1611 ms). Therefore, both the RCS and RCO items were processed with a high attachment preference by ELL.

**Table 2.9 Inferential Statistics on Regions of Interest in The Three Items (English Late Learners)**

Examples	RCS_Critical Region RTs	Estimate (Log Value)	Std. Error	2.50%	97.50%	t value	Pr(> t )	Estimate (ms)
	himself	(Intercept)	7.34	0.04	7.27	7.41	201.33	
herself	Low Attachment	0.23	0.04	0.15	0.31	5.69	0.00000***	1950
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1								
The son of the actress who shot herself/himself in the theater was under investigation.	RCS_Total RTs	Estimate (Log Value)	Std. Error	2.50%	97.50%	t value	Pr(> t )	13004
	(Intercept)	9.47	0.03	9.22	9.33	330.84	0.00000***	
	Low Attachment	0.19	0.02	0.03	0.10	9.24	0.00000***	
	Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1							
was	RCO_Critical Region RTs	Estimate (Log Value)	Std. Error	2.50%	97.50%	t value	Pr(> t )	1611
	(Intercept)	7.38	0.04	7.31	7.45	202.21	0.00000***	
	Low Attachment	0.11	0.04	0.03	0.19	2.63	0.009799**	
	Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1							
were	RCO_Total RTs	Estimate (Log Value)	Std. Error	2.50%	97.50%	t value	Pr(> t )	13108
	(Intercept)	9.48	0.03	9.25	9.36	330.89	0.00000***	
	Low Attachment	0.16	0.02	-0.03	0.05	8.14	0.00000***	
	Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1							
tomorrow	ADVP_Critical Region RTs	Estimate (Log Value)	Std. Error	2.50%	97.50%	t value	Pr(> t )	2319
	(Intercept)	7.75	0.04	7.68	7.82	210.17	0.00000***	
	Low Attachment	-0.19	0.04	-0.27	-0.11	-4.67	0.00000***	
	Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1							
yesterday	ADVP_Total RTs	Estimate (Log Value)	Std. Error	2.50%	97.50%	t value	Pr(> t )	14779
	(Intercept)	9.60	0.03	9.31	9.42	333.66	0.00000***	
	Low Attachment	0.14	0.02	0.00	0.07	6.73	0.00000***	
	Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1							
Anne will serve the apples she picked yesterday/tomorrow, but she won't serve the plums.	ADVP_Critical Region RTs	Estimate (Log Value)	Std. Error	2.50%	97.50%	t value	Pr(> t )	16934
	(Intercept)	9.60	0.03	9.31	9.42	333.66	0.00000***	
	Low Attachment	0.14	0.02	0.00	0.07	6.73	0.00000***	
	Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1							

As for the ADVP item, the critical region in the high attachment condition is processed significantly longer (7.75 log ms or 2319 ms) than the one in the low attachment condition (7.56 log ms or 1912 ms) which is in alignment with the ADVP pattern of the ENS group. This indicates a low attachment preference of the ELL participants on processing this item.

Other than that, post-critical regions of all these three items show significantly longer reading times in the low attachment conditions. This also happened to the results of total reading times (tRTs) as indicated by Table 2.9. Like for the ENS group, it appears that the ORDER effect even overrode an effect of low attachment preference for the ADVP item, which was supposed to be reflected by a shorter tRT in its low attachment condition.

### **6.2.1.3 Heritage Mandarin Speakers**

As shown in Table 2.10, the HMCS patterns of RTs of the critical regions are similar to the results of native speakers. There is a non-significant result by comparing the low (7.26 log ms or 1422 ms) and high attachment trials (7.29 log ms or 1458 ms) with respect to the critical regions in the RCS item. Looking at the tRTs of the RCS item in Table 2.10, there is a significantly longer duration of the low attachment trials (9.37 log ms or 11706 ms) than the high attachment ones (9.29 log ms or 10797 ms). This reverse situation between critical regions and total reading times also happened to the RCO item as indicated by the statistical results in Table 2.10. Meanwhile, a low attachment preference in the ADVP item is reflected by a significantly shorter duration in the low attachment condition (7.29 log ms or 1460 ms) compared to the high attachment one (7.49 log ms or 1786 ms). Parallel with the RCS and RCO items, the ADVP tRTs have also been affected by the ORDER effect which provided some significant facilitation on reading the high attachment trials.

Other than that, the post-critical regions of all these three items have a significant pattern showing that longer durations were spent on reading the low attachment trials than doing the high counterparts.

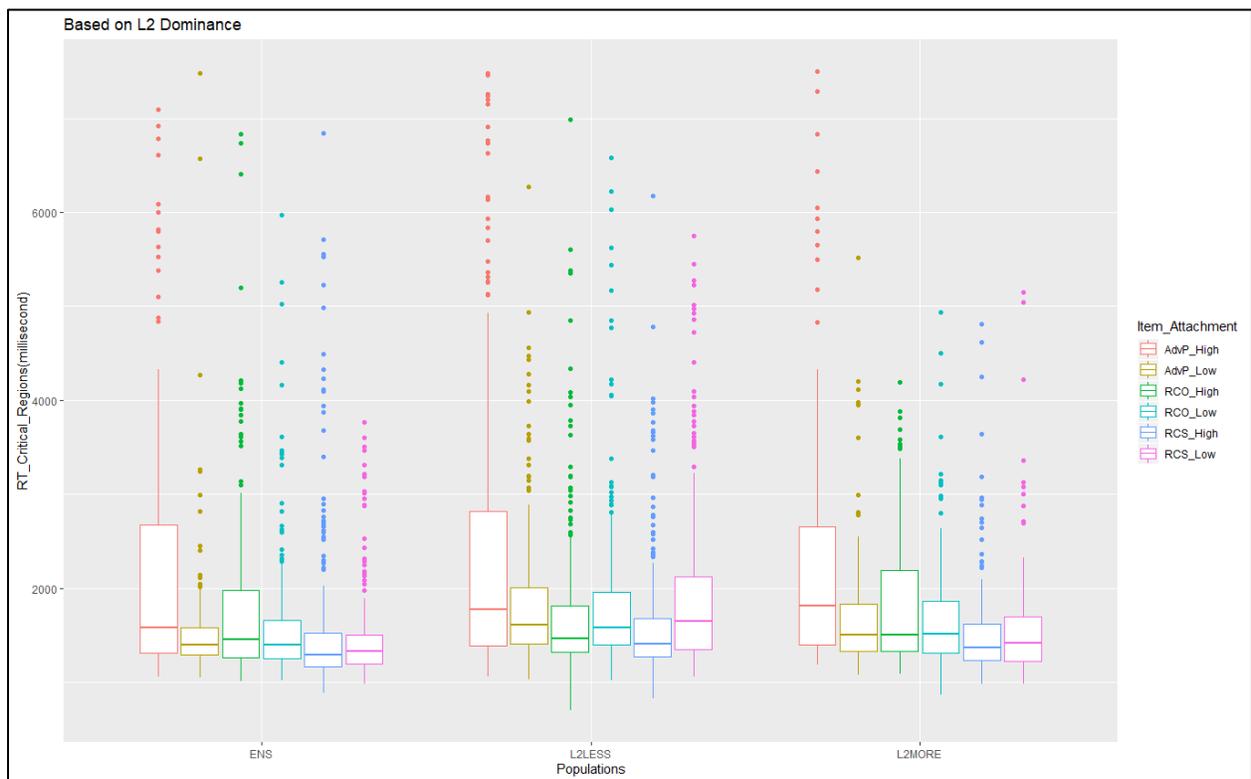
**Table 2.10 Inferential Statistics on Regions of Interest in The Three Items (Heritage Mandarin Speakers)**

Examples	RCS_Critical Region RTs	Estimate (Log Value)	Std. Error	2.50 %	97.50 %	t value	Pr(> t )	Estimate (ms)
	himself herself	(Intercept)	7.29	0.04	7.20	7.37	172.17	
	Low Attachment	-0.03	0.05	-0.11	0.06	-0.55	0.581308	1422
Signif. codes: 0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 0.1								
The son of the actress who shot herself/himself in the theater was under investigation.	RCS_Total RTs	Estimate (Log Value)	Std. Error	2.50 %	97.50 %	t value	Pr(> t )	10797 11706
	(Intercept)	9.29	0.03	9.22	9.35	267.79	0.00000**	
	Low Attachment	0.08	0.02	0.04	0.12	3.59	0.000416***	
	Signif. codes: 0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 0.1							
was were	RCO_Critical Region RTs	Estimate (Log Value)	Std. Error	2.50 %	97.50 %	t value	Pr(> t )	1721 1618
	(Intercept)	7.45	0.04	7.37	7.53	175.91	0.00000**	
	Low Attachment	-0.06	0.05	0.03	0.19	-1.36	0.176855	
	Signif. codes: 0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 0.1							
The reporter phoned the boss of the secretaries who were/was busy in the office.	RCO_Total RTs	Estimate (Log Value)	Std. Error	2.50 %	97.50 %	t value	Pr(> t )	11070 11779
	(Intercept)	9.31	0.03	9.25	9.38	268.40	0.00000**	
	Low Attachment	0.06	0.02	0.02	0.11	2.76	0.006384**	
	Signif. codes: 0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 0.1							
tomorrow yesterday	ADVP_Critical Region RTs	Estimate (Log Value)	Std. Error	2.50 %	97.50 %	t value	Pr(> t )	1786 1460
	(Intercept)	7.49	0.04	7.41	7.57	176.79	0.00000**	
	Low Attachment	-0.20	0.05	-0.29	-0.11	-4.42	0.00002***	
	Signif. codes: 0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 0.1							
Anne will serve the apples she picked yesterday/tomorrow, but she won't serve the plums.	ADVP_Total RTs	Estimate (Log Value)	Std. Error	2.50 %	97.50 %	t value	Pr(> t )	11350 12224
	(Intercept)	9.34	0.03	9.27	9.40	269.12	0.00000**	
	Low Attachment	0.07	0.02	0.03	0.12	3.29	0.001197**	
	Signif. codes: 0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 0.1							

In summary, data of the HMCS participants indicate a low attachment preference in each of these three items. However, there was no spill-over effect in post-critical regions or it was just overridden by the ORDER effect. The tRTs are fully covered by the ORDER effect which generated an opposite pattern to that of the critical regions.

## 6.2.2 Predict from L2 Dominance

Reading times of the critical regions are plotted in Figure 2.5 across the control group and the other two bilingual groups who have different L2 dominance scores. The cut-off score of these two subgroups (i.e., L2LESS and L2MORE) in this graph was set at 21. Before that, plotting was done on cut-off scores of 10 and 15. Neither of them showed any differences in the raw-data plots between higher and lower L2-dominance speakers.



*Figure 2.5 Interaction between Critical Region RT and L2 Dominance Level (All L2 Speakers)*

When it was set at 21, as shown in Figure 2.5, the plotted L2LESS patterns of these three items are all in alignment with the patterns of the ELL results reported in Section 6.2.1.2. However, there is a slightly reversed pattern of the RCO item for the L2MORE participants who have higher dominance scores than 21. This pattern looks like a low attachment preference which is consistent

with the RCO processing results of the native speakers reported in Section 6.2.1.1. However, the LMER model did not predict any significant result for RCO attachment preference for the L2MORE participants. As for the RCS and ADVP items, model results showed the L2MORE group has similar preference patterns as those of the ELL participants. Not surprisingly, the statistical results of all three items for the L2LESS group are consistent with the results of the ELL group. This means the L2 dominance level of 21 has no significant effect on the processing patterns of these items. But the pattern of high attachment preference in RCO (which happens to both the L2LESS group and ELL group) disappears from the results of the bilinguals who have a higher L2 dominance level than 21.

Therefore, to further test the predictor of L2 dominance, cut-off levels were respectively raised to 30, 35, and 37. This was to determine whether the RCO processing pattern could be reversed to a low attachment preference as that of English native speakers.

No significant effect was shown at the point of 30 and 35. When the cut-off score was set at 37 (i.e. each of the L2MORE participants has a dominance score of no less than 37), the pattern of the critical-region RTs of the RCO (see Table 2.11) shows an eased processing behavior on the low attachment trials (Low: 7.41 log ms or 1656 ms; High: 7.46 log ms or 1742 ms). Although the p-value (0.176855) is not significant enough to indicate a strong effect, this pattern can be claimed as a low attachment preference, considering the ORDER effect that had been imposed on almost all regions of the trials and neutralized some effects of low attachment preference. Meanwhile, the RCO pattern of a significant high-attachment preference (High: 7.39 log ms or 1617 ms; Low: 7.47 log ms or 1747ms) of the L2LESS group is still in alignment with the plot in Figure 2.5.

On the other hand, the ADVP pattern of low attachment preference did not change for both the L2LESS and L2MORE groups. Other than that, the L2LESS group still showed a high

attachment preference for the RCS item, while no significant preference on the RCS item was found for the L2MORE group. This means the L2 dominance level of 37 does not impose any significant effect on the processing patterns of the ADVP and RCS items. Nevertheless, the RCS processing pattern of high attachment preference vanished when the L2 dominance level was raised to 37.

**Table 2.11 Inferential Statistics of Critical Region RTs of RCO (Mandarin-English Bilinguals)**

<i>The reporter phoned the boss of the secretaries who were/was busy in the office. (L2Dominance cut-off score: 37)</i>								
	<b>L2MORE_RCO_Critical-Region RTs</b>	Estimate (Log Value)	Std. Error	2.50%	97.50%	t value	Pr(> t )	<b>Estimate (ms)</b>
<i>was</i>	<b>(Intercept)</b>	7.46	0.05	7.37	7.56	149.78	0.00000***	1742
<i>were</i>	<b>Low Attachment</b>	-0.05	0.05	0.03	0.19	-0.97	0.176855	1656
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1								
	<b>L2LESS_RCO_Critical-Region RTs</b>	Estimate (Log Value)	Std. Error	2.50%	97.50%	t value	Pr(> t )	
<i>was</i>	<b>(Intercept)</b>	7.39	0.04	7.32	7.46	202.36	0.00000***	1617
<i>were</i>	<b>Low Attachment</b>	0.08	0.04	0.00	0.15	1.96	0.05296 .	1747
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1								

### 6.2.3 Predict from L2 Dominance with AoA

The model with a triple interaction (*L2Dominance\*Item\_Attachment\*Population*) tells us how processing patterns could be affected by both AoA and L2 dominance simultaneously. This model yields results that hold three specific variables in the intercept. For instance, in Table 2.12, the intercept of *ELL: L2LESS: RCO\_Critical Region RTs* indicates an estimate of the critical RTs of (RCO) high attachment condition processed by ELL participants who have L2 dominance scores lower than 37.

**Table 2.12 Inferential Statistics of Critical Region RTs of RCO (A Triple Interaction Model)**

The reporter phoned the boss of the secretaries who were/was busy in the office.								
was were	<b>ELL: L2LESS: RCO_Critical Region RTs</b>	Estimate (Log Value)	Std. Error	2.50%	97.50%	t value	Pr(> t )	<b>Estimate (ms)</b>
	<b>(Intercept)</b>	7.37	0.11	7.30	7.45	68.37	0.00000***	1592
	<b>Low Attachment</b>	0.12	0.04	0.04	0.20	2.76	0.00659**	<b>1794</b>
	Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1							
was were	<b>ELL: L2MORE: RCO_Critical Region RTs</b>	Estimate (Log Value)	Std. Error	2.50%	97.50%	t value	Pr(> t )	<b>Estimate (ms)</b>
	<b>(Intercept)</b>	7.43	0.12	7.29	7.57	60.22	0.00000***	1689
	<b>Low Attachment</b>	0.06	0.07	-0.08	0.20	0.83	0.405129	<b>1792</b>
	Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1							
was were	<b>HMCS: L2LESS: RCO_Critical Region RTs</b>	Estimate (Log Value)	Std. Error	2.50%	97.50%	t value	Pr(> t )	<b>Estimate (ms)</b>
	<b>(Intercept)</b>	7.42	0.10	7.31	7.52	74.14	1	1667
	<b>Low Attachment</b>	-0.02	0.06	-0.12	0.09	0.83	0.76612	<b>1640</b>
	Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1							
was were	<b>HMCS: L2MORE: RCO_Critical Region RTs</b>	Estimate (Log Value)	Std. Error	2.50%	97.50%	t value	Pr(> t )	<b>Estimate (ms)</b>
	<b>(Intercept)</b>	7.49	0.12	7.38	7.60	64.57	0.00000***	1783
	<b>Low Attachment</b>	-0.11	0.06	-0.23	0.00	-1.92	0.05566 .	<b>1593</b>
	Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1							

Unlike the results excluding AoA in Section 6.2.2, no dramatic effect from L2 dominance is indicated by the results of this model, especially for the ADVP and RCS items. In other words, either the ELL and HMCS group did not show any significant difference of processing patterns between low and high L2 dominant speakers on processing the ADVP and RCS items. However, there is a tendency that the effect of AoA on processing RCO is gradually overridden by the L2 dominance factor as the L2 dominance score increases. As seen in Table 2.12, lower L2 dominant speakers (L2 dominance score < 37) who have AoA higher than 18 (i.e., the *ELL: L2LESS* group) show a significant high preference in RCO processing ( $p = 0.00659$ ). This effect disappears for the ELL participants who have a higher level of L2 dominance score (the *ELL: L2MORE* group) (with a p-value of 0.40513). Meanwhile, a similar situation of override is shown by HMCS

participants who have much earlier AoAs than the ELL participants. Specifically, the low attachment preference of the RCO processing by the *HMCS: L2LESS* group is not significant (see Table 2.12). However, as the dominance score increases to a level as high as 37 (i.e., the *HMCS: L2LESS* group), low attachment preference becomes significant, even under an ORDER effect <sup>28</sup>.

### 6.3 Summary of the Results

As introduced in Section 4, this study tested the Shallow Structure Hypothesis (Clahsen & Felser, 2006b, 2006a, 2018) with Heritage Mandarin Chinese speakers (HMCS) and very late English learners who speak Mandarin as their L1 (ELL). The research questions are:

- **RQ1.** Do late English learners of L1 Mandarin speakers process sentences like native speakers?
- **RQ2.** Do heritage Mandarin speakers process sentences like native speakers?
- **RQ3.** Is L2 dominance a critical variable to predict the processing patterns of L2 learners?

In the maze experiment reported in **Chapter 2**, three groups of participants (ENS, ELL, and HMCS) were asked to read three items as stimuli. The three items are the sentential constructions with relative clause modifying a subject (RCS), relative clause modifying an object (RCO), and an adverb phrase modifier (ADVP). Previous literature suggests native English speakers have low attachment preference when reading sentences of these constructions (Felser et al., 2003; J. Witzel et al., 2012; N. Witzel et al., 2012). This low attachment preference can be accounted for by not only modular-based language processing models (Frazier, 1978, 1990; Frazier & Fodor, 1978; Frazier & Rayner, 1982) but also constraint-based models (Gibson et al.,

---

<sup>28</sup> Under an ORDER effect, this effect of low attachment preference is not significant when AoA is the only predictor in the models reported in Section 6.2.1.3.

1996; Gibson & Pearlmutter, 1998). L2 learners should not have any preference for low or high attachment according to the SSH (Clahsen & Felser, 2006b, 2006a, 2018; Felser et al., 2003), while another study reports that highly proficient L2 learners have certain particular attachment preferences when processing the RCS and ADVP constructions (J. Witzel et al., 2012).

The results of the current study on the RCS and ADVP items are consistent with findings of the latter (J. Witzel et al., 2012) to some degree. However, the ELL results of processing the RCO item do not conform with findings reported in some SSH literature (Felser et al., 2003). Compared to the native speakers, the HMCS group has a very similar pattern of attachment preference (i.e., low attachment preference) when reading all these three items. Other than that, a large amount of L2 use (with a very high L2 dominance level, i.e., above 36) is a factor that has an overriding effect on AoA when a triple interaction (*Population\*L2Dominance\*Item\_Attachment*) was considered as a predictor for statistics of the reading times in critical regions of the RCO item. All these results will be discussed in Chapter 3.

## **Chapter 3 Discussion & Conclusion**

## 7 Discussion

The following Section 7.1 will discuss how the age of onset of L2 acquisition (AoA) predicts L2 processing patterns of the two bilingual groups, focusing on the processing patterns of two bilingual groups with considerably different AoAs. This discussion will address the first and second research questions (**RQs**) of whether these L2 learners process native-like or shallowly. After that, the results of data analysis with L2 dominance (as an independent variable) are discussed in Section 7.2 to address the third **RQ**.

In detail, attachment preference patterns will be discussed to say if these bilingual populations have followed the Late Closure principle (Frazier, 1978; Frazier & Fodor, 1978; Frazier & Rayner, 1982) or the *Recency/Predicate Proximity* Principles (Gibson et al., 1996; Gibson & Pearlmutter, 1998) which were claimed as the native processing strategies. In the Shallow Structure Hypothesis (Clahsen & Felser, 2006b, 2006a, 2018), these processing principles are argued that they hardly apply to late L2 learners.

### 7.1 The AoA Effect

#### 7.1.1 English Native Speakers

The comparison between the predictions and results of the ENS participants is reported in Table 3.1. On the right side, i.e., in the “Results” section of the table, the word “High” refers to a pattern of high attachment effect which means the RTs in the high attachment condition are shorter than their counterparts in the low attachment condition. The “Low” refers to an opposite effect of the “High”. For instance, the result of “Low” in section 5 of the RCO item means the participants read the low-attachment trials non-significantly shorter than reading the high-attachment ones. The

result of “High” in section 6 means a non-significant effect of high attachment preference (with a p-value above 0.1).

For the native-speaker participants, results in the critical regions are consistent with the findings in the literature (Felser et al., 2003; J. Witzel et al., 2012; N. Witzel et al., 2012). As summarized in Table 3.1, they were found to have a significant preference for low attachment when processing the RCO and ADVP items. The low attachment preference of RCS is not significant. Other than that, the post-critical region-1 of the RCO trials has elicited a significant spill-over effect.

*Table 3.1 Hypothesis and Results (English Native Speakers)*

RCS Predictions					Results			
RCS	Region5 (critical)	Region6	Region7	Total RT	Region5 (critical)	Region6	Region7	Total RT
RCS_Low	herself	in the theater	was...	---	Low	High	High	High***
RCS_High	himself	in the theater	was...	---				
Preference	Low	Spillover_Low	Spillover_Low	Low				
RCO Predictions					Results			
RCO	Region6 (critical)	Region7	Region8	Total RT	Region6 (critical)	Region7	Region8	Total RT
RCO_Low	were	busy	in the office.	---	Low*	Low*	Low	High
RCO_High	was	busy	in the office.	---				
Preference	Low	Spillover_Low	Spillover_Low	Low				
ADVP Predictions					Results			
ADVP	Region6 (critical)	Region7	Region8	Total RT	Region6 (critical)	Region7	Region8	Total RT
ADVP_Low	yesterday,	but she won't serve	the plums.	---	Low***	High*	High**	High .
ADVP_High	tomorrow,	but she won't serve	the plums.	---				
Preference	Low	Spillover_Low	Spillover_Low	Low				
<b>Notes:</b> <b>Low:</b> Low Attachment effect (i.e. shorter RT in the low attachment condition) <b>High:</b> High Attachment effect (i.e. shorter RT in the high attachment condition) <b>Signif. codes:</b> 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1								

The tRT (i.e. Total RT) results are opposite to those reported in previous studies. Participants tend to have an advantage when processing post-critical regions of high attachment trials in the RCS and ADVP items. For the RCO item, lower tRTs of high attachment trails could be attributed to the significantly shorter RTs in the regions before the critical region. These results form a contradiction with the results in critical regions, which was not expected.

A reason for the contradiction between the critical region and tRT results could be the ORDER effect which is already mentioned in Section 6.2.1. High attachment trials were all played after the participants completed the low attachment trials. The participants had gotten familiar somewhat with the repeated phrases in post-critical regions of the high attachment trials. This familiarity may have made them read these regions faster than their counterparts in the low attachment condition. This could also be the cause of the situation that the low attachment effect of RCS processing on critical regions is not significant.

It is notable that for the RCO item, the post-critical regions were not largely affected by the ORDER effect. One reason could be the RCO structure like *The reporter phoned the boss of the secretaries who was/were busy in the office* was more costly to process, compared to the other two constructions. According to the account of working-memory demand (Frazier & Fodor, 1978), the processor probably needed much storage “space” of the working memory when processing the post-critical regions (region 7 and 8) of an RCO trial. Participants must hold a subject *the reporter*, a predicate *phoned*, and an object *the boss of the secretaries who was/were* in the working memory to proceed to the post-critical regions. However, for the RCS item such as *the son of the actress who shot herself/himself in the theatre was under investigation*, the processor only needs to keep a subject chunk *the son of the actress who shot herself/himself in the theatre* in the working memory before encountering the post-critical regions. Similarly, for the ADVP item such as *Anne*

*will serve the apples she picked yesterday, but she won't serve the plums*, the processor was temporarily paused by a comma after parsing a subject, a predicate, and an object. Therefore, the post-critical regions of RCS and ADVP were unlikely to pose any spillover difficulty to the parser or processor under the ORDER effect.

In summary, native speakers have a low attachment preference when processing a complex sentence with modifiers like RCs or ADVPs. The ORDER effect may have overridden the effect of low attachment preference in the post-critical regions of the RCS and ADVP items. However, it was not strong enough to cover the low preference effect in the critical regions of all three items and the post-critical regions of the RCO item.

### **7.1.2 English Late Learners**

As for late learners, RCO and RCS/ADVP items are discussed separately. The findings of J. Witzel et al. (2012) suggest that late learners have a high and low preference for the RCS and ADVP items, respectively, while Felser et al. (2003) argue that L2 learners have no significant preference for either high or low attachment upon processing the RCO item.

As shown in Table 3.2, the results of RCS and ADVP are consistent with the hypothesis based on J. Witzel et al. (2012). The critical regions of the RCS item elicited a significant effect of high attachment preference for the ELL participants. Meanwhile, a low preference effect also reaches a significant level in the critical region of the ADVP item.

It appears that the post-critical regions and tRT of the ADVP item have the opposite effect to the result of its critical region. This situation parallels the ENS group. Meanwhile, for both the RCS and ADVP items, the difference gradient of the LMER-estimated tRTs between low and high attachment conditions for ELL participants is much more inclined than the ENS group. This

stronger incline of the ELL group is likely caused by a combination of the attachment preference effects and the ORDER effect. For the RCS trials, the ELL participants received an effect of high attachment preference plus the ORDER (which would enhance the processing advantage on high attachment trials). However, the ENS group was affected by a low attachment preference effect plus the ORDER effect (which would override the processing advantage on low attachment trials). Therefore, considering all factors that could have influenced the data, I argue that the results support that those late L2 learners have a high attachment preference processing the RCS item (opposite to native speakers) while having a low preference on the ADVP item.

*Table 3.2 Hypothesis and Results (English Late Learners)*

RCS Hypothesis					Results			
RCS	Region5 (critical)	Region6	Region7	Total RT	Region5 (critical)	Region6	Region7	Total RT
RCS_Low	herself	in the theater	was...	---				
RCS_High	himself	in the theater	was...	---	High***	High***	High***	High***
Preference	High	Spillover_High	Spillover_High	Spillover_High				
RCO Predictions					Results			
RCO	Region6 (critical)	Region7	Region8	Total RT	Region6 (critical)	Region7	Region8	Total RT
RCO_Low	were	busy	in the office.	---				
RCO_High	was	busy	in the office.	---	High**	High***	High***	High***
Preference	High	Spillover_High	Spillover_High	Spillover_High				
ADVP Predictions					Results			
ADVP	Region6 (critical)	Region7	Region8	Total RT	Region6 (critical)	Region7	Region8	Total RT
ADVP_Low	yesterday,	but she won't serve	the plums.	---				
ADVP_High	tomorrow,	but she won't serve	the plums.	---	Low***	High***	High***	High***
Preference	Low	Spillover_Low	Spillover_Low	Low				
Notes:								
Low: Low Preference effect (i.e. shorter RT in the low attachment condition)								
High: High Preference effect (i.e. shorter RT in the high attachment condition)								
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1								

The RTs in the RCO critical region imply a significantly shorter duration in the high attachment condition than in its counterpart in the low condition. Even considering the ORDER effect has caused a processing advantage on high attachment trials, this result is still convincing enough to be interpreted as a high attachment preference of late learners. The results of the ELL participants processing RCO are not consistent with the previous findings on L2 learners reported in Felser et al. (2003). On the other hand, it is also opposite to the low attachment preference of native English speakers in the literature (Felser et al., 2003; Frazier, 1990; J. Witzel et al., 2012; N. Witzel et al., 2012).

Notably, in the literature of Mandarin Chinese processing, a low preference is reported on processing relative clause attachment, while syntactic structure is not the only constraint for this pattern (Kwon et al., 2019; Shen, 2006). Therefore, the high attachment preference in L2 English is unlikely to be attributed to an effect of L1 transfer.

It is unclear whether the opposite preference of L2 English to English native speakers should be interpreted as a native-like syntactic processing strategy. The Garden-path model account (Frazier, 1990) claims high attachment preference in relative clause constructions should be caused by reanalysis and discourse factors, while the account of constraint-based models (Gibson et al., 1996; Gibson & Pearlmutter, 1998) takes the high attachment preference as a recency effect (see § 2.2.1) syntactically. Considering the dual-pathway model in the updated Shallow Structure Hypothesis (Clahsen & Felser, 2018), it is reasonable to say that the ELL participants have acquired enough grammatical knowledge to use the grammatical pathway for ambiguity resolution of RCS, RCO, and ADVP constructions. On the other hand, during this maze experiment, the heuristics (e.g. a discourse effect argued by Frazier, 1990) of the late learners must have taken up more weighting to parse the RCS and RCO trials compared to that of a native parser.

In Mandarin Chinese, no case marker exists and functions like the one *of* in English. Therefore, for ELL participants, it is their main assertion of attaching a relative clause to NP1 of a complex phrase with an NP1+ *of* +NP2 construction. This discourse factor drives them to complete a rapid reanalysis and finally have a high attachment preference because subconsciously, they believe it is illegal to attach the relative clause to the low site (i.e. NP2). The weighting of the discourse heuristics was heavy enough to reverse the preference effect. I argue that this has caused an opposite preference of RC attachment site between the ELL group and the native participants.

Overall, the ELL participants have applied native-like strategies in processing the ADVP trials. However, for the RCS and RCO items, they tend to rely more on the heuristic pathway of the parser compared to the ENS participants.

### **7.1.3 Heritage Mandarin Speakers**

In this study, each of the heritage Mandarin speakers has an AoA of 8 at the most. They were predicted to process more native-like than the late learners in processing complex English sentences. The results indicate that the hypothesis is supported. As compared in Table 3.3, results in critical regions all have shown a low preference effect even though it is not significant for the RCS and RCO items.

It appears that the low preference effect was competing with the ORDER effect, which had already been found on the other two groups of participants. Finally, the ORDER won the competition in most post-critical regions and elicited a processing advantage on the high attachment trials. This could be why the low preference effect is not significant enough in critical regions of the RCS and RCO items.

**Table 3.3 Hypothesis and Results (Heritage Mandarin Speakers)**

RCS Predictions					Results			
RCS	Region5 (critical)	Region6	Region7	Total RT	Region5 (critical)	Region6	Region7	Total RT
RCS_Low	herself	in the theater	was...	---	Low	High	High*	High***
RCS_High	himself	in the theater	was...	---				
Preference	Low	Spillover_Low	Spillover_Low	Low				
RCO Predictions					Results			
RCO	Region6 (critical)	Region7	Region8	Total RT	Region6 (critical)	Region7	Region8	Total RT
RCO_Low	were	busy	in the office.	---	Low	High	High*	High**
RCO_High	was	busy	in the office.	---				
Preference	Low	Spillover_Low	Spillover_Low	Low				
ADVP Predictions					Results			
ADVP	Region6 (critical)	Region7	Region8	Total RT	Region6 (critical)	Region7	Region8	Total RT
ADVP_Low	yesterday,	but she won't serve	the plums.	---	Low***	High***	High***	High**
ADVP_High	tomorrow,	but she won't serve	the plums.	---				
Preference	Low	Spillover_Low	Spillover_Low	Low				
<b>Notes:</b>								
Low: Low Preference effect (i.e. shorter RT in the low attachment condition)								
High: High Preference effect (i.e. shorter RT in the high attachment condition)								
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1								

Due to the ORDER effect, the statistical results of the tRTs and non-critical region RTs are no longer convincing enough to argue for the original hypothesis in the current study. The ORDER effect could have overridden the effect of attachment preference in these data. However, the ORDER issue could be combined with the critical regions, and together they could be discussed to gain some insight into the online parsing mechanisms. After all, the results of the critical regions are of the most interest (N. Witzel et al., 2012).

### 7.1.4 RQ1 & RQ2

In this experiment, the results of the ENS group pattern with findings in the literature. As very late learners with AoAs no less than 18 years old, the ELL participants have applied native-like

strategies when processing complex sentences with adverb phrase modifiers. However, they tend to be constrained by the late AoAs when processing the RCS and RCO items. The late AoAs even have resulted in a reversal effect of native processing patterns (low preference), as for the RC attachment constructions. On the other hand, since the HMCS participants have much earlier AoAs than the ELL group, they performed native-like processing strategies on all those three items. The results suggest that the age of acquisition is a critical predictor of whether L2 learner participants use native-like strategies when processing complex sentences with RCO and RCS attachment. However, for the ADVP item, the AoA of an L2 learner does not seem like a significant constraint.

In summary, Section 7.1 addressed the first two research questions of the current study. Whether late L2 learners perform native-like or shallow processing (**RQ1**), the answer is not an absolute yes or no. Native speaker participants were found to have low attachment preference on modifier processing, consistent with previous studies on English sentence processing. The late-learner results of the maze experiment on processing the ADVP item indicate a native-like processing pattern, while this is not the case for processing the RCO and RCS items. L2 late learner participants of L1 Mandarin speakers tend to have a high attachment preference when processing the RCO and RCS items.

For **RQ2** (Do heritage speakers with high dominance and early AoA have the same processing patterns as native English speakers?), the answer is yes. An effect of low attachment preference was found for all the three items processed by the group of heritage speakers who have much earlier AoAs than the late learners.

## 7.2 The L2 Dominance Effect

### 7.2.1 Discuss L2 Dominance

As reported in Section 6.2.2, an interaction of L2 dominance \* Item\_Attachment was considered as the independent variable in a generalized linear mixed-effect regression model. Table 3.4 summarizes the different processing patterns of the three items across two sub-groups. Specifically, two cut-off scores (i.e., 21 and 37) were found most interesting since there were some evident effects of L2 dominance at these two points. Firstly, it is unlikely to find any effect of L2 dominance on ADVP processing since all the three populations have shown an identical preference for low attachment upon this item. However, for RCO processing, English learners with L2 dominance scores of more than 21 have no significant preference patterns of attachment, which is not identical to those with L2 dominance levels not higher than 21.

*Table 3.4 Summary of Processing Patterns under the Effect of L2 Dominance*

<b>Cut-off Scores of L2 Dominance</b>	<b>Sub-groups</b>	<b>RCO</b>	<b>RCS</b>	<b>ADVP</b>
21	lower L2 dominant speakers	High	High	Low
	higher L2 dominant speakers	None	High	Low
37	lower L2 dominant speakers	High	High	Low
	higher L2 dominant speakers	Low	None	Low

With the L2 dominance levels reaching 37 and higher, L2 English speakers tend to have a low attachment preference on RCO, like native speakers. Furthermore, the cut-off score of 21 has no impact on RCS processing, while the high attachment preference disappeared from the statistical results of these participants who have L2 dominance scores higher than 36. Considering

the HMCS participants who have shown a low attachment preference (i.e., like the native speaker participants) on RCS even under the situation of an ORDER effect, these results suggest that the higher the L2 dominance levels are for the bilingual participants in this experiment, the more likely they apply English native-like strategies to process the RCO and RCS items.

*Table 3.5 Summary of Processing Patterns under the Effect of L2 Dominance and AoA*

AoA(Population): L2Dominance: Item_Attachment			Results of Preference Patterns	
ELL(AoA>=18)	L2Dominance_More	RCS	High	significant
ELL(AoA>=18)	L2Dominance_More	RCO	None	
ELL(AoA>=18)	L2Dominance_More	ADVP	Low	significant
ELL(AoA>=18)	L2Dominance_Less	RCS	High	significant
ELL(AoA>=18)	L2Dominance_Less	RCO	High	significant
ELL(AoA>=18)	L2Dominance_Less	ADVP	Low	significant
HMCS(AoA<=8)	L2Dominance_More	RCS	Low	nearly significant
HMCS(AoA<=8)	L2Dominance_More	RCO	Low	significant
HMCS(AoA<=8)	L2Dominance_More	ADVP	Low	significant
HMCS(AoA<=8)	L2Dominance_Less	RCS	Low	nearly significant
HMCS(AoA<=8)	L2Dominance_Less	RCO	Low	non-significant
HMCS(AoA<=8)	L2Dominance_Less	ADVP	Low	significant

However, results of the generalized LMER model with a triple interaction (*Population\*Item\_Attachment\*L2Dominance*) indicate no effect of the *AoA\*L2Dominance* on the RCS and ADVP items as suggested by the processing patterns listed in Table 3.5. Even for the RCO item, the *L2Dominance* effect is not as strong as the one predicted from the LMER model with a double interaction of *L2Dominance\*Item\_Attachment*. Looking at the highlighted rows in Table 3.5, it is noticed that high *L2Dominance* neutralizes or partially overrides the effect of a higher AoA (i.e., for ELL participants). Meanwhile, it reinforces the significant effect of being native-like (i.e., a low attachment preference) for heritage speakers during processing the RCO

item. After all, it could not reverse the high attachment preference of the late learner participants to a low preference pattern which is a native-like processing pattern.

Therefore, and to summarize, when L2 dominance was considered in the statistical analysis for the Mandarin-English bilingual population, an effect of low attachment preference was found for highly L2 dominant bilinguals in the critical regions of the RCO item. This effect distinguishes them from the high attachment preference of low L2 dominant speakers. It appears that the RCO processing patterns have been affected by the L2 dominance factor significantly. However, considering the results of the triple interaction in an LMER model, this *L2Dominance* variable was not strong enough to reverse the effect of AoA which predicts a high attachment preference for late English learners through this maze experiment.

### **7.2.2 Address RQ3**

Overall, for **RQ3** of this current study (Does L2 dominance level individually predict native-like processing strategies for bilingual speakers?), the answer is no, telling from the collected data. To address this question, L2 use was first measured in a self-reporting survey and combined with the item attachment conditions (i.e., low and high) as an interaction to predict the reading time patterns in a generalized LMER model. The results suggest that the more dominantly the learners speak their L2 English, the more native-like they will behave when processing complex English sentences with modifiers like the relative clause. However, to rule out the effect of AoA, a triple interaction was then created in a generalized LMER model. Although the results show that a higher L2 dominance level tends to impose some effect on processing patterns of RCO, it could not fully override the effect of AoA.

Other than that, there is a variance in the extent to which these different constructions are affected by AoA and L2 use. For L2 learners, AoA is a critical factor when predicting native-like or non-native-like processing on the RCS and RCO items, but not on the ADVP item (as discussed in § 7.1). On the other hand, L2 dominance is a less critical variable than AoA for processing the RCS and RCO items. L2 dominance does override some effect of AoA when processing the RCO item, but not completely. With AoAs older than 18, which has passed the critical period (Johnson & Newport, 1989, 1991; Long, 1990; Newport, 1990, 1991), learners with a very high level of L2 use have shown no significant preference pattern of the RCO item. This is important to be discussed because a lower level of L2Dominance of late learners predicts a high attachment preference. One possibility here is higher L2 dominant speakers with late AoAs tend to abandon the high preference and start to change their processing behaviors. A larger dataset with more variant dominance scores and AoAs could reveal more insight in future research.

## 8 General Discussion and Conclusion

This study does not challenge the SSH framework. However, the results provide more empirical evidence to be used for updating the existing SSH (Clahsen & Felser, 2018), according to which different populations might have different weightings of input in a dual-pathway parser. With final L2 grammatical attainment being more native-like, learners tend to have the capability to construct more native-like representations in the grammatical pathway of their sentence parser, which results in a more native-like processing pattern. This was shown in this experiment with the early learners and the late learners with very high L2 dominance levels. Instead, the heuristic pathway is more weighted to recruit heuristic information when processing complex sentences if speakers have higher AoAs and/or less L2 dominance levels.

Specifically, early learners use a native-like parser when reading complex sentences with RC and ADVP attachments. Meanwhile, it is obvious for late L2 learners to have a high weighting of the grammatical pathway when processing the ADVP construction. For parsing the RCS and RCO constructions, without an early AoA, high L2 use alone is not enough for an English learner to build up highly native-like representations in the grammatical pathway (for a low attachment preference). After all, being contradictory on the preference of attachment site does not mean having a processing mechanism in common, as per the dispute between the GPM (Frazier, 1990) and CBM account (Gibson & Pearlmutter, 1998, 1998). Notably, late L2 learners who mostly use L2 in their daily life are likely to be different from these late learners with a low rate of L2 use to interpret the RCO construction. This difference is predicted to be a tendency for building more native-like grammatical representations in the parser, although it does not completely override the effect of AoA. On the other hand, the Critical Period Hypothesis can not fully predict from AoA how native-like L2 sentence processing proceeds, at least for the ADVP construction.

A limitation of the current study is the issue of non-randomness between low and high attachment conditions. This could have caused the processing advantage of reading trials in the high attachment condition. However, it did not damage the effect elicited by the critical region of an ambiguous sentence. Meanwhile, it seems that a bigger sample size of L2 speakers with more variance on AoA and L2 dominance conditions could help to produce more convincing results from the statistical analysis when comparing the effects between AoA and L2 dominance. Moreover, the effect of L1 transfer is not fully covered in this current study, which could be studied in future research combining the effect of AoA and L2 dominance. A typical method for such a study is using counterbalanced items in both L1 and L2 languages of highly L2 dominant bilinguals as processing stimuli. This could yield a specific result of how L1-L2 transfer will affect L2 sentence processing.

Nevertheless, the current study provides more support for the effect of AoA and L2 dominance on bilingual sentence processing. The age of acquisition is a critical factor for L2 grammatical processing, but L2 dominance should not be negligible, especially for processing high cost-taking structures like the RCO item in this current study.

## Appendix Items Presented in the Experiment

### *The Low Attachment Condition:*

- RCS1 The son of the actress who shot herself in the theater was under investigation.
- RCS2 The brother of the bride who embarrassed herself at the wedding felt ashamed.
- RCS3 The mother of the waiter who hurt himself in the kitchen was shocked by the accident.
- RCS4 The son of the lady who introduced herself on the platform was popular at the party.
- RCS5 The brother of the schoolgirl who burned herself with a lighter was very upset.
- RCS6 The grandmother of the man who killed himself last summer was sent to hospital.
- RCS7 The grandma of the fireman who criticized himself far too often was anxious.
- RCS8 The mother of the schoolboy who hurt himself in the school was under custody.
- RCS9 The daughter of the man who complimented himself in public was very beautiful.
- RCS10 The sister of the prince who injured himself in swimming pool was still sad.
- RCO1 The reporter phoned the boss of the secretaries who were busy in the office.
- RCO2 The cleaning lady saw the kids of the player who was working very late.
- RCO3 The man trusted the teacher of the students who were ready to go home.
- RCO4 The principal smiled at the babies of the care giver who was eating her lunch.
- RCO5 The inspector called the assistant of the policemen who were watching TV.
- RCO6 The journalist hated the soldiers of the colonel who was sitting in the chair.
- RCO7 The movie star noticed the kids of the doctor who was wearing a green dress.
- RCO8 The director congratulated the father of the actors who were winning the prize.
- RCO9 The doctor recognized the nurse of the patients who were feeling very tired.
- RCO10 The cameraman spoke to the employer of the cleaners who were standing on the platform.
- ADVP1 Anne will serve the apples she picked yesterday, but she won't serve the plums.

ADVP2 Robert will meet the friend he phoned yesterday, but he doesn't want to do that.

ADVP3 David caught the fish he will cook tomorrow, but it is not his favorite kind.

ADVP4 Sue insulted the candidate she will debate tomorrow, but she wishes she hadn't.

ADVP5 Jane prepared the lecture she will give next week, but she still needs to review it.

ADVP6 Joseph brewed the beer he will serve next week, but it is not very tasty.

ADVP7 Tom will plant the tree he bought last week, but he isn't sure where to put it.

ADVP8 Lisa will change the plans she made last week, but she won't cancel any of them.

ADVP9 Jeff planned the party he will hold next month, but he hasn't sent invitations.

ADVP10 Dan wrote the speech he will deliver next month, but he hasn't practiced it yet.

ADVP11 Paul will marry the woman he just met last month, but the wedding will be small.

ADVP12 Amy will visit the man she worked with last month, but she is nervous about it.

ADVP13 Mary called the applicant she will interview tomorrow, but there was no answer.

ADVP14 Mike watered the flower he will sell tomorrow, but he forgot to water the bush.

ADVP15 Susan bought the wine she will drink next week, but she didn't buy any cheese.

*The High Attachment Condition:*

RCS11 The son of the actress who shot himself in the theater was under investigation.

RCS12 The brother of the bride who embarrassed himself at the wedding felt ashamed.

RCS13 The mother of the waiter who hurt herself in the kitchen was shocked by the accident.

RCS14 The son of the lady who introduced himself on the platform was popular at the party.

RCS15 The brother of the schoolgirl who burned himself with a lighter was very upset.

RCS16 The grandmother of the man who killed herself last summer was sent to hospital.

RCS17 The grandma of the fireman who criticized herself far too often was anxious.

RCS18 The mother of the schoolboy who hurt herself in the school was under custody.

RCS19 The daughter of the man who complimented herself in public was very beautiful.

RCS20 The sister of the prince who injured herself in swimming pool was still sad.

RCO11 The reporter phoned the boss of the secretaries who was busy in the office.

RCO12 The cleaning lady saw the kids of the player who were working very late.

RCO13 The man trusted the teacher of the students who was ready to go home.

RCO14 The principal smiled at the babies of the care giver who were eating lunch.

RCO15 The inspector called the assistant of the policemen who was watching TV.

RCO16 The journalist hated the soldiers of the colonel who were sitting in the chair.

RCO17 The movie star noticed the kids of the doctor who were wearing a green dress.

RCO18 The director congratulated the father of the actors who was winning the prize.

RCO19 The doctor recognized the nurse of the patients who was feeling very tired.

RCO20 The cameraman spoke to the employer of the cleaners who was standing on the platform.

ADVP16 Anne will serve the apples she picked tomorrow, but she won't serve the plums.

ADVP17 Robert will meet the friend he phoned tomorrow, but he doesn't want to do that.

ADVP18 David caught the fish he will cook yesterday, but it is not his favorite kind.

ADVP19 Sue insulted the candidate she will debate yesterday, but she wishes she hadn't.

ADVP20 Jane prepared the lecture she will give last week, but she still needs to review it.

ADVP21 Joseph brewed the beer he will serve last week, but it is not very tasty.

ADVP22 Tom will plant the tree he bought next week, but he isn't sure where to put it.

ADVP23 Lisa will change the plans she made next week, but she won't cancel any of them.

ADVP24 Jeff planned the party he will hold last month, but he hasn't sent invitations.

ADVP25 Dan wrote the speech he will deliver last month, but he hasn't practiced it yet.

ADVP26 Paul will marry the woman he just met next month, but the wedding will be small.

ADVP27 Amy will visit the man she worked with next month, but she is nervous about it.

ADVP28 Mary called the applicant she will interview yesterday, but there was no answer.

ADVP29 Mike watered the flower he will sell yesterday, but he forgot to water the bush.

ADVP30 Susan bought the wine she will drink last week, but she didn't buy any cheese.

*Fillers:*

F1 The princess who scratched herself in public was awfully embarrassed.

F2 The policeman's brother who found himself in trouble called their father.

F3 The nephew and the maid who cut himself with a knife screamed at the dog.

F4 When her boyfriend treated her like his family, she was very satisfied and happy.

F5 The brother of the queen bought her a pet which was so cute and beautiful.

F6 The guards of the prison have released the innocent prisoners and they were very proud.

F7 The policewoman in New York City who found herself in trouble called her father.

F8 The maid's nephew cut himself with the knife which he bought yesterday.

F9 Her boyfriend was proud of her because she was a judge in the court.

F10 His daughter and the king got married after he left the country.

F11 The young girl helped the driver who were talking to an old woman.

F12 I watched the fans of the singer dancing about throughout the concert.

F13 The doctor contacted the nurses when the lawyer was talking on the phone.

F14 The photographer liked the models who were smiling at him all the time.

F15 That little girl envied the princess who was just eating chocolates.

F16 The solders just followed the general who was both a father and a husband.

- F17 Jack trained the horse of the cowboy who was playing with a pistol.
- F18 The teacher talked to a student's parents because he failed the final exam.
- F19 They were fond of music and arts, which was good for the community.
- F20 Yesterday the reporter criticized the politicians working with the president.
- F21 The daughter of Chris cleaned the bookcase for sale last week, but it is still very dusty.
- F22 Mark and his friend answered the email, but they don't know what to do in the next step.
- F23 The boy will not burn the wood, only because he wants to save some of it.
- F24 Jim painted the picture he displayed for sale, but he wasn't happy with it.
- F25 John hired the clerk and promoted him last month, but he fired another employee.
- F26 Lisa bought the flower she likes the most, but the color is not very vibrant.
- F27 The waitress is going to clean up the tables, and the bar owner will be happy.
- F28 The criminal broke a piece of glass on the police car, but he didn't get away from it.
- F29 The teacher insulted the girl who argued with me yesterday, but she wishes she hadn't.
- F30 Anna borrowed a book from the library last week, and she will return it next week.

## References

- Altmann, G. T. M., van Nice, K. Y., Garnham, A., & Henstra, J.-A. (1998). Late Closure in context. *Journal of Memory and Language*, 38(4), 459–484.  
<https://doi.org/10.1006/jmla.1997.2562>
- Bai, B. (2018). Syntactic ambiguity resolution of Cantonese-English bilinguals: Evidence from prepositional phrase and adverb. *Open Journal of Modern Linguistics*, 08(03), 39–47.  
<https://doi.org/10.4236/ojml.2018.83005>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1). <https://doi.org/10.18637/jss.v067.i01>
- Bialystok, E., & Hakuta, K. (1999). Confounded age: Linguistic and cognitive factors in age differences for second language acquisition. In *Second Language Acquisition and The Critical Period Hypothesis* (pp. x, 191). Lawrence Erlbaum Associates Publishers.
- Birdsong, D. (1992). Ultimate Attainment in Second Language Acquisition. *Language*, 68(4), 706. <https://doi.org/10.2307/416851>
- Birdsong, D. (2006). Age and second language acquisition and processing: A selective overview. *Language Learning*, 56, 9–49. <https://doi.org/10.1111/j.1467-9922.2006.00353.x>
- Birdsong, D. (2014). Dominance and Age in Bilingualism. *Applied Linguistics*, 35(4), 374–392.  
<https://doi.org/10.1093/applin/amu031>
- Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in Second-Language Acquisition. *Journal of Memory and Language*, 44(2), 235–249.  
<https://doi.org/10.1006/jmla.2000.2750>
- Birdsong, D., & Vanhove, J. (2016). Age of second language acquisition: Critical periods and social concerns. In E. Nicoladis & S. Montanari (Eds.), *Bilingualism across the lifespan:*

- Factors moderating language proficiency.* (pp. 163–181). American Psychological Association. <https://doi.org/10.1037/14939-010>
- Bolger, P. A., & Zapata, G. C. (2011). Psycholinguistic approaches to language processing in heritage speakers. *Heritage Language Journal*, 8(1), 1–29.
- Bowden, H. W., Steinhauer, K., Sanz, C., & Ullman, M. T. (2013). Native-like brain processing of syntax can be attained by university foreign language learners. *Neuropsychologia*, 51(13), 2492–2511. <https://doi.org/10.1016/j.neuropsychologia.2013.09.004>
- Brysbaert, M., & Mitchell, D. C. (1996). Modifier attachment in sentence parsing: Evidence from Dutch. *The Quarterly Journal of Experimental Psychology Section A*, 49(3), 664–695.
- Carreiras, M., & Clifton, C. (1999). Another word on parsing relative clauses: Eyetracking evidence from Spanish and English. *Memory & Cognition*, 27(5), 826–833. <https://doi.org/10.3758/BF03198535>
- Chai, W. J., Abd Hamid, A. I., & Abdullah, J. M. (2018). Working memory from the psychological and neurosciences perspectives: A review. *Frontiers in Psychology*, 9, 401. <https://doi.org/10.3389/fpsyg.2018.00401>
- Clahsen, H., & Felser, C. (2006a). Continuity and shallow structures in language processing. *Applied Psycholinguistics*, 27(1), 107–126. <https://doi.org/10.1017/S0142716406060206>
- Clahsen, H., & Felser, C. (2006b). Grammatical processing in language learners. *Applied Psycholinguistics*, 27(1), 3–42. <https://doi.org/10.1017/S0142716406060024>
- Clahsen, H., & Felser, C. (2018). Some notes on the Shallow Structure Hypothesis. *Studies in Second Language Acquisition*, 40(3), 693–706. <https://doi.org/10.1017/S0272263117000250>

- Clahsen, H., & Veríssimo, J. (2016). Investigating grammatical processing in bilinguals: The case of morphological priming. *Linguistic Approaches to Bilingualism*, 6(5), 685–698.  
<https://doi.org/10.1075/lab.15039.cla>
- Clifton, C., & Frazier, L. (1989). Comprehending sentences with long-distance dependencies. In G. N. Carlson & M. K. Tanenhaus (Eds.), *Linguistic Structure in Language Processing* (Vol. 7, pp. 273–317). Springer Netherlands. [https://doi.org/10.1007/978-94-009-2729-2\\_8](https://doi.org/10.1007/978-94-009-2729-2_8)
- Covey, L. (2018). *An ERP investigation of individual differences in the processing of wh-dependencies by native and non-native speakers* [PHD Dissertation, University of Kansas]. <http://hdl.handle.net/1808/27993>
- Cuetos, F., & Mitchell, D. C. (1988). Cross-linguistic differences in parsing: Restrictions on the use of the Late Closure strategy in Spanish. *Cognition*, 30(1), 73–105.  
[https://doi.org/10.1016/0010-0277\(88\)90004-2](https://doi.org/10.1016/0010-0277(88)90004-2)
- Cuetos, F., Mitchell, D. C., & Corley, M. M. B. (1996). Parsing in different languages. In *Language processing in Spanish*. (pp. 145–187). Lawrence Erlbaum Associates, Inc.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19(4), 450–466.  
[https://doi.org/10.1016/S0022-5371\(80\)90312-6](https://doi.org/10.1016/S0022-5371(80)90312-6)
- De Vincenzi, M., & Job, R. (1993). Some observations on the universality of the late-closure strategy. *Journal of Psycholinguistic Research*, 22(2), 189–206.  
<https://doi.org/10.1007/BF01067830>
- 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 Processes*, 21(4), 453–485. <https://doi.org/10.1080/01690960400023485>
- Dussias, P. E. (2001). Sentence Parsing in Fluent Spanish-English Bilinguals. In *One mind, two languages: Bilingual language processing*. (pp. 159–176). Blackwell Publishing.
- Dussias, P. E. (2003). Syntactic ambiguity resolution in L2 learning: Some effects of bilinguality on L1 and L2 processing strategies. *Studies in Second Language Acquisition*, 25(4), 529–557. <https://doi.org/10.1017/S0272263103000238>
- Dussias, P. E., & Cramer Scaltz, T. R. (2008). Spanish–English L2 speakers’ use of subcategorization bias information in the resolution of temporary ambiguity during second language reading. *Acta Psychologica*, 128(3), 501–513. <https://doi.org/10.1016/j.actpsy.2007.09.004>
- Eysenck, M. W., & Keane, M. T. (2010). *Cognitive psychology: A student’s handbook* (Sixth edition..).
- Felser, C., & Cunnings, I. (2012). Processing reflexives in a second language: The timing of structural and discourse-level constraints. *Applied Psycholinguistics*, 33(3), 571–603. <https://doi.org/10.1017/S0142716411000488>
- Felser, C., & Roberts, L. (2007). Processing wh-dependencies in a second language: A cross-modal priming study. *Second Language Research*, 23(1), 9–36. <https://doi.org/10.1177/0267658307071600>
- Felser, C., Roberts, L., Marinis, T., & Gross, R. (2003). The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics*, 24(3), 453–489. <https://doi.org/10.1017/S0142716403000237>

- Felser, C., Sato, M., & Bertenshaw, N. (2009). The on-line application of binding Principle A in English as a second language. *Bilingualism: Language and Cognition*, *12*(4), 485–502. <https://doi.org/10.1017/S1366728909990228>
- Flege, J. E. (1999). Age of learning and second language speech. *Second Language Acquisition and the Critical Period Hypothesis*, *101*, 131.
- Flege, J. E., Mackay, I. R. A., & Piske, T. (2002). Assessing bilingual dominance. *Applied Psycholinguistics*, *23*(4), 567–598. <https://doi.org/10.1017/S0142716402004046>
- Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language*, *41*(1), 78–104.
- Forster, K. I., Guerrero, C., & Elliot, L. (2009). The maze task: Measuring forced incremental sentence processing time. *Behavior Research Methods*, *41*(1), 163–171. <https://doi.org/10.3758/BRM.41.1.163>
- Foucart, A., & Frenck-Mestre, C. (2012). Can late L2 learners acquire new grammatical features? Evidence from ERPs and eye-tracking. *Journal of Memory and Language*, *66*(1), 226–248. <https://doi.org/10.1016/j.jml.2011.07.007>
- Frazier, L. (1978). On Comprehending Sentences: Syntactic Parsing Strategies. *ETD Collection for University of Connecticut*.
- Frazier, L. (1987). Theories of sentence processing. In J. L. Garfield (Ed.), *Modularity in Knowledge Representation and Natural-Language Understanding*. The MIT Press.
- Frazier, L. (1990). Parsing modifiers: Special purpose routines in the human sentence processing mechanism. *Comprehension Processes in Reading*, 303–330.

- Frazier, L., & Clifton, C. (1997). Construal: Overview, motivation, and some new evidence. *Journal of Psycholinguistic Research*, 26(3), 277–295.  
<https://doi.org/10.1023/A:1025024524133>
- Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, 6(4), 291–325. [https://doi.org/10.1016/0010-0277\(78\)90002-1](https://doi.org/10.1016/0010-0277(78)90002-1)
- Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, 14(2), 178–210. [https://doi.org/10.1016/0010-0285\(82\)90008-1](https://doi.org/10.1016/0010-0285(82)90008-1)
- Frenck-Mestre, C. (2002). An on-line look at sentence processing in the second language. In *Advances in Psychology* (Vol. 134, pp. 217–236). Elsevier.  
[https://doi.org/10.1016/S0166-4115\(02\)80012-7](https://doi.org/10.1016/S0166-4115(02)80012-7)
- Frenck-Mestre, C., & Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native languages. *The Quarterly Journal of Experimental Psychology A*, 50(1), 119–148. <https://doi.org/10.1080/027249897392251>
- Friederici, A. D. (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and Language*, 50(3), 259–281. <https://doi.org/10.1006/brln.1995.1048>
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, 6(2), 78–84. [https://doi.org/10.1016/S1364-6613\(00\)01839-8](https://doi.org/10.1016/S1364-6613(00)01839-8)
- Friederici, A. D., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing: Evidence challenging the critical period hypothesis. *Proceedings of the National Academy of Sciences*, 99(1), 529–534. <https://doi.org/10.1073/pnas.012611199>

- Garfield, J. L. (Ed.). (1991). *Modularity in Knowledge Representation and Natural-Language Understanding*. The MIT Press. <https://doi.org/10.7551/mitpress/4735.001.0001>
- Gathercole, V. C. M., & Thomas, E. M. (2009). Bilingual first-language development: Dominant language takeover, threatened minority language take-up. *Bilingualism: Language and Cognition*, 12(2), 213–237. <https://doi.org/10.1017/S1366728909004015>
- Ge, H., Matthews, S., & Yip, V. (2019). *Cantonese heritage speakers' online processing of English: Evidence from eye movements in the visual world paradigm*. <https://doi.org/10.13140/RG.2.2.12850.96960>
- Gibson, E., Pearlmutter, N., Canseco-Gonzalez, E., & Hickok, G. (1996). Recency preference in the human sentence processing mechanism. *Cognition*, 59(1), 23–59. [https://doi.org/10.1016/0010-0277\(95\)00687-7](https://doi.org/10.1016/0010-0277(95)00687-7)
- Gibson, E., & Pearlmutter, N. J. (1998). Constraints on sentence comprehension. *Trends in Cognitive Sciences*, 2(7), 262–268. [https://doi.org/10.1016/S1364-6613\(98\)01187-5](https://doi.org/10.1016/S1364-6613(98)01187-5)
- Gibson, E., & Warren, T. (2004). Reading-time evidence for intermediate linguistic structure in long-distance dependencies. *Syntax*, 7(1), 55–78. <https://doi.org/10.1111/j.1368-0005.2004.00065.x>
- Ha, J. A. (2005). Age-related effects on syntactic ambiguity resolution in first and second languages: Evidence from Korean-English bilinguals. *Proceedings of the 7th Generative Approaches to Second Language Acquisition Conference (GASLA 2004)*, 111–123. [www.lingref.com](http://www.lingref.com), document #1160
- Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience*, 11(2), 194–205. <https://doi.org/10.1162/089892999563328>

- Hahne, A., & Friederici, A. D. (2001). Processing a second language: Late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition*, 4(2), 123–141. <https://doi.org/10.1017/S1366728901000232>
- Hakuta, K., Bialystok, E., & Wiley, E. (2003). Critical evidence: A test of the Critical-Period Hypothesis for Second-Language Acquisition. *Psychological Science*, 14(1), 31–38. JSTOR.
- Hare, M., McRae, K., & Elman, J. L. (2003). Sense and structure: Meaning as a determinant of verb subcategorization preferences. *Journal of Memory and Language*, 48(2), 281–303. [https://doi.org/10.1016/S0749-596X\(02\)00516-8](https://doi.org/10.1016/S0749-596X(02)00516-8)
- Havik, E., Roberts, L., van Hout, R., Schreuder, R., & Haverkort, M. (2009). Processing Subject-Object Ambiguities in the L2: A Self-Paced Reading Study With German L2 Learners of Dutch. *Language Learning*, 59(1), 73–112. <https://doi.org/10.1111/j.1467-9922.2009.00501.x>
- Hernandez, A. E., & Li, P. (2007). Age of acquisition: Its neural and computational mechanisms. *Psychological Bulletin*, 133(4), 638–650. <https://doi.org/10.1037/0033-2909.133.4.638>
- Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21(1), 60–99. [https://doi.org/10.1016/0010-0285\(89\)90003-0](https://doi.org/10.1016/0010-0285(89)90003-0)
- Johnson, J. S., & Newport, E. L. (1991). Critical period effects on universal properties of language: The status of subadjacency in the acquisition of a second language. *Cognition*, 39(3), 215–258. [https://doi.org/10.1016/0010-0277\(91\)90054-8](https://doi.org/10.1016/0010-0277(91)90054-8)

- Juffs, A. (2004). Representation, Processing and Working Memory in a Second Language. *Transactions of the Philological Society*, 102(2), 199–225.  
<https://doi.org/10.1111/j.0079-1636.2004.00135.x>
- Juffs, A., & Harrington, M. (1995). PARSING EFFECTS IN SECOND LANGUAGE SENTENCE PROCESSING: Subject and Object Asymmetries in wh-Exaraction. *Studies in Second Language Acquisition*, 17(4), 483–516.
- Juffs, A., & Harrington, M. (1996). Garden path sentences and error data in second language sentence processing. *Language Learning*, 46(2), 283–323. <https://doi.org/10.1111/j.1467-1770.1996.tb01237.x>
- Kidd, E., Brandt, S., Lieven, E., & Tomasello, M. (2007). Object relatives made easy: A cross-linguistic comparison of the constraints influencing young children’s processing of relative clauses. *Language and Cognitive Processes*, 22(6), 860–897.  
<https://doi.org/10.1080/01690960601155284>
- Kubota, M., Ferrari, P., & Roberts, T. P. L. (2003). Magnetoencephalography detection of early syntactic processing in humans: Comparison between L1 speakers and L2 learners of English. *Neuroscience Letters*, 353(2), 107–110.  
<https://doi.org/10.1016/j.neulet.2003.09.019>
- Kwon, N., Ong, D., Chen, H., & Zhang, A. (2019). The Role of Animacy and Structural Information in Relative Clause Attachment: Evidence From Chinese. *Frontiers in Psychology*, 10, 1576. <https://doi.org/10.3389/fpsyg.2019.01576>
- Lenneberg, E. H. (1967). The biological foundations of language. *Hospital Practice*, 2(12), 59–67. <https://doi.org/10.1080/21548331.1967.11707799>

- Long, M. H. (1990). Maturational Constraints on Language Development. *Studies in Second Language Acquisition*, 12(3), 251–285. <https://doi.org/10.1017/S0272263100009165>
- MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101(4), 676–703. <https://doi.org/10.1037/0033-295X.101.4.676>
- Mak, W. M., Vonk, W., & Schriefers, H. (2002). The influence of animacy on relative clause processing. *Journal of Memory and Language*, 47(1), 50–68. <https://doi.org/10.1006/jmla.2001.2837>
- Marinis, T., Roberts, L., Felser, C., & Clahsen, H. (2005). GAPS IN SECOND LANGUAGE SENTENCE PROCESSING. *Studies in Second Language Acquisition*, 27(01). <https://doi.org/10.1017/S0272263105050035>
- McRae, K., & Matsuki, K. (2013). Constraint-based models of sentence processing. In *Sentence processing*. (pp. 51–77). Psychology Press.
- Montrul, S. (2006). On the bilingual competence of Spanish heritage speakers: Syntax, lexical-semantics and processing. *International Journal of Bilingualism*, 10(1), 37–69. <https://doi.org/10.1177/13670069060100010301>
- Montrul, S., Davidson, J., De La Fuente, I., & Foote, R. (2014). Early language experience facilitates the processing of gender agreement in Spanish heritage speakers. *Bilingualism: Language and Cognition*, 17(1), 118–138. <https://doi.org/10.1017/S1366728913000114>
- Montrul, S., Foote, R., & Perpiñán, S. (2008). Gender agreement in adult second language learners and Spanish Heritage Speakers: The effects of age and context of acquisition. *Language Learning*, 58(3), 503–553. <https://doi.org/10.1111/j.1467-9922.2008.00449.x>

- Newman, A. J., Pancheva, R., Ozawa, K., Neville, H. J., & Ullman, M. T. (2001). An event-related fMRI study of syntactic and semantic violations. *Journal of Psycholinguistic Research*, 30(3), 339–364. <https://doi.org/10.1023/A:1010499119393>
- Newport, E. (1990). Maturational Constraints on Language Learning. *Cognitive Science*, 14(1), 11–28. [https://doi.org/10.1207/s15516709cog1401\\_2](https://doi.org/10.1207/s15516709cog1401_2)
- Newport, E. (1991). Contrasting conceptions of the critical period for language. *The Epigenesis of Mind: Essays on Biology and Cognition*, 111–130.
- Papadopoulou, D., & Clahsen, H. (2003). Parsing strategies in L1 and L2 sentence processing: A study of relative clause attachment in Greek. *Studies in Second Language Acquisition*, 25(4), 501–528. <https://doi.org/10.1017/S0272263103000214>
- Pearlmutter, N. J., & Gibson, E. (2001). Recency in verb phrase attachment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(2), 574–590. <https://doi.org/10.1037/0278-7393.27.2.574>
- Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language processing. *Current Opinion in Neurobiology*, 15(2), 202–206. <https://doi.org/10.1016/j.conb.2005.03.007>
- Pickering, M., & Barry, G. (1991). Sentence processing without empty categories. *Language and Cognitive Processes*, 6(3), 229–259. <https://doi.org/10.1080/01690969108406944>
- Polinsky, M. (2018). *Heritage Languages and Their Speakers*. Cambridge University Press.
- Polinsky, M., & Kagan, O. (2007). Heritage Languages: In the ?Wild? and in the Classroom. *Language and Linguistics Compass*, 1(5), 368–395. <https://doi.org/10.1111/j.1749-818X.2007.00022.x>

- Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior*, 22(3), 358–374.  
[https://doi.org/10.1016/S0022-5371\(83\)90236-0](https://doi.org/10.1016/S0022-5371(83)90236-0)
- Roberts, L., & Felser, C. (2011). Plausibility and recovery from garden paths in second language sentence processing. *Applied Psycholinguistics*, 32(2), 299–331.  
<https://doi.org/10.1017/S0142716410000421>
- Roberts, L., Gullberg, M., & Indefrey, P. (2008). Online pronoun resolution in L2 discourse: L1 Influence and General Learner Effects. *Studies in Second Language Acquisition*, 30(3), 333–357. <https://doi.org/10.1017/S0272263108080480>
- Roberts, L., Marinis, T., Felser, C., & Clahsen, H. (2006). Antecedent priming at trace positions in children’s sentence processing. *Journal of Psycholinguistic Research*, 36, 175–188.  
<https://doi.org/10.1007/s10936-006-9038-3>
- Rohde, H., Levy, R., & Kehler, A. (2011). Anticipating explanations in relative clause processing. *Cognition*, 118(3), 339–358. <https://doi.org/10.1016/j.cognition.2010.10.016>
- Sag, I. A., Wasow, T., & Bender, E. M. (2003). *Syntactic Theory: A Formal Introduction* (2nd ed). Center for the Study of Language and Information.
- Shen, X. (2006). *Late assignment of syntax theory: Evidence from Chinese and English* [PHD Dissertation, University of Exeter]. <http://hdl.handle.net/10036/30017>
- Smolensky, P., Goldrick, M., & Mathis, D. (2014). Optimization and quantization in gradient symbol systems: A framework for integrating the continuous and the discrete in cognition. *Cognitive Science*, 38(6), 1102–1138. <https://doi.org/10.1111/cogs.12047>

- Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language*, 48(3), 542–562. [https://doi.org/10.1016/S0749-596X\(02\)00536-3](https://doi.org/10.1016/S0749-596X(02)00536-3)
- Tanenhaus, M. K., & Lucas, M. M. (1987). Context effects in lexical processing. *Cognition*, 25(1–2), 213–234. [https://doi.org/10.1016/0010-0277\(87\)90010-2](https://doi.org/10.1016/0010-0277(87)90010-2)
- Traxler, M. J., & Frazier, L. (2008). The role of pragmatic principles in resolving attachment ambiguities: Evidence from eye movements. *Memory & Cognition*, 36(2), 314–328. <https://doi.org/10.3758/MC.36.2.314>
- Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition*, 4(2), 105–122. <https://doi.org/10.1017/S1366728901000220>
- Van Gompel, R. P. G. (Ed.). (2013). *Sentence Processing*. Psychology Press.
- van Hell, J. G., & Tokowicz, N. (2010). Event-related brain potentials and second language learning: Syntactic processing in late L2 learners at different L2 proficiency levels. *Second Language Research*, 26(1), 43–74. <https://doi.org/10.1177/0267658309337637>
- Veríssimo, J., Heyer, V., Jacob, G., & Clahsen, H. (2018). Selective Effects of Age of Acquisition on Morphological Priming: Evidence for a Sensitive Period. *Language Acquisition*, 25(3), 315–326. <https://doi.org/10.1080/10489223.2017.1346104>
- Wartenburger, I., Heekeren, H. R., Abutalebi, J., Cappa, S. F., Villringer, A., & Perani, D. (2003). Early setting of grammatical processing in the bilingual brain. *Neuron*, 37(1), 159–170. [https://doi.org/10.1016/S0896-6273\(02\)01150-9](https://doi.org/10.1016/S0896-6273(02)01150-9)
- Weber-Fox, C. M., & Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual

- speakers. *Journal of Cognitive Neuroscience*, 8(3), 231–256.  
<https://doi.org/10.1162/jocn.1996.8.3.231>
- Weckerly, J., & Kutas, M. (1999). An electrophysiological analysis of animacy effects in the processing of object relative sentences. *Psychophysiology*, 36(5), 559–570.  
<https://doi.org/10.1111/1469-8986.3650559>
- Wells, J. B., Christiansen, M. H., Race, D. S., Acheson, D. J., & MacDonald, M. C. (2009). Experience and sentence processing: Statistical learning and relative clause comprehension. *Cognitive Psychology*, 58(2), 250–271.
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.  
<https://ggplot2.tidyverse.org>
- Williams, J. N. (2006). Incremental interpretation in second language sentence processing. *Bilingualism: Language and Cognition*, 9(1), 71–88.  
<https://doi.org/10.1017/S1366728905002385>
- Witzel, J., Witzel, N., & Nicol, J. (2012). Deeper than shallow: Evidence for structure-based parsing biases in second-language sentence processing. *Applied Psycholinguistics*, 33(2), 419–456. <https://doi.org/10.1017/S0142716411000427>
- Witzel, N., Witzel, J., & Forster, K. (2012). Comparisons of Online Reading Paradigms: Eye Tracking, Moving-Window, and Maze. *Journal of Psycholinguistic Research*, 41(2), 105–128. <https://doi.org/10.1007/s10936-011-9179-x>
- Zagar, D., Pynte, J., & Rativeau, S. (1997). Evidence for early closure attachment on first pass reading times in French. *The Quarterly Journal of Experimental Psychology Section A*, 50(2), 421–438. <https://doi.org/10.1080/713755715>