The role of language proficiency and statistical learning in on-line comprehension of

syntax among bilingual adult readers

A thesis submitted for the degree of Master of Philosophy in

Research School of Psychology

The Australian National University

November 2017

By

(Keri) Yuk Ling CHUI

Abstract

Statistical learning (SL) is the ability to identify co-occurring regularities from the

environment, and has been implicated in learning across a range of skills, including

language. This research project investigated whether there are associations between

SL and on-line sentence processing in L1 Chinese L2 English bilinguals, and sought

to examine whether second language proficiency mediated the relationship between

visual SL and L2 language processing. To this end, two studies were conducted. In

Study 1, sixty Chinese-English bilinguals completed a self-paced reading task in

Mandarin and English, which tested participants' on-line processing of subject and

object relative clauses (RCs). They also completed a nonlinguistic visual SL task and

a battery of additional measures measuring L2 English proficiency and general

cognitive abilities. The results revealed that only nonverbal intelligence predicted L1

Chinese RCs processing, and neither visual SL capacity nor L2 proficiency predicted

L2 English RCs processing. One possible explanation is that SL is partially

modality-specific. Therefore, an auditory SL task was employed in addition to visual

SL task in Study 2. In Study 2, fifty-two native Mandarin-speaking adults completed

tests of visual and auditory SL, a self-paced reading task measuring the online

processing of Mandarin relative clauses, and measures of general cognitive abilities.

The results showed that auditory SL capacity independently predicted reading times

in the self-paced reading task. Visual SL was also related to language processing,

although the effect was marginal. The findings from Study 2 suggest that individual

differences in adults' capacity for SL are associated with on-line processing of

Chinese.

SL IN SENTENCE PROCESSING

Table of Contents

List of Tables

Study 1

Study 2

List of Figures

Chapter 1

Overview

Language comprehension is generally regarded as a rapid, on-line process that requires readers and listeners to construct an incremental representation of components of language and rapidly map it to meaning. While early theories of parsing did not assume the presence of individual differences, recent work suggests a range of sources of individual variability. One recent skill that has been linked to individual differences in parsing is statistical learning (SL): the human ability to identify the co-occurring regularities from the environment (Kaufman, DeYoung, Gray, Jiménez, Brown, & Mackintosh, 2010; Perruchet & Pacton, 2006). A growing number of studies have investigated the links between statistical learning (SL) and language. For instance, SL has been shown to be important for language acquisition in several domains, including from phonological learning (e.g., Speciale, Ellis & Bywater, 2004), word segmentation (e.g., Saffran, Newport, & Aslin, 1996; Evans, Saffran, & Robe-Torres, 2009), syntactic acquisition (Kidd, 2012; Kidd & Arciuli, 2016), and reading proficiency (e.g. Arciuli & Simpson, 2012).

While most early studies with adults concentrated on the role of SL in linguistic performance at the group level, a growing body of research indicates that SL is also skill subject to individual variation in mature language users. Along with this

framework, studies that have assessed individual differences in SL via artificial

grammar learning (AGL) tasks have reported associations with adult L1 language

processing (Misyak, Christiansen, & Tomblin, 2010; Misyak & Christiansen, 2012).

Although these data are no doubt important, most of the subjects in these studies

were monolingual participants and the linguistic tasks were conducted in languages

that use alphabetic scripts. Accordingly, relatively few empirical studies have

investigated the role of SL in second language (L2) acquisition or processing. Frost,

Siegelman, Narkiss, and Afek (2013) and Kaufman, DeYoung, Gray, Jiménez,

Brown, and Mackintosh (2010) are examples of such studies that have linked SL to

L2 literacy proficiency (specifically, literacy and general abilities measured via

secondary school exams). Other studies have shown that SL predicts L2 morphology

acquisition (Brooks & Kempe, 2013; Brooks, Kwoka, & Kempe, 2017; Granena,

2013; McDonough & Trofimovich, 2016). Still, no study has demonstrated a link

between SL and on-line syntactic processing in L2 populations.

To fill these research gaps, this thesis aims to investigate whether there is a

direct link between SL and language processing in the first (L1) and second (L2)

language of Chinese-English bilinguals and seeks to examine whether second

language proficiency mediates the relationship between visual SL and L2 language

processing. This thesis is divided into 5 chapters.

Chapter 2 is the Literature review. In the first section, I survey studies that have

investigated the online processing of syntax, with a focus on one particular structure

- relative clauses. These structures are interesting for several reasons. Firstly, they are

the most commonly used structures in past studies of individual differences, and

secondly, they are structurally very different in the two languages investigated in this

thesis – Mandarin and English. The first section of Chapter 2 largely focuses on the

key findings of studies of Mandarin relative clause processing. In the second section

of Chapter 2, I address the individual difference in relative clauses processing, and

review several theoretical accounts that predict different sources of individual

differences. The following section introduces SL capacity as a potential variable that

could explain individual differences in sentence processing among L1 and L2

learners. I define SL, and review some research findings on the link between SL

capacity and L1 adult language processing. The subsequent section reviews

empirical evidence for an association between SL capacity and L2 morphology

learning, along with the role of L2 language proficiency in L2 sentence processing.

In Chapter 3, I present the results of Study 1, which investigated the role of SL

in the L1 (Mandarin) and L2 (English) processing of RCs. Chapter 4 presents the

results of Study 2, which aimed to follow up the largely inconclusive results of

Study 1. Finally, Chapter 5, the General Discussion, draws the results of Study 1 and

2 together, discusses their limitations, and suggests future avenues of research.

Chapter 2

Literature review

One major goal of psycholinguistics is to discover the cognitive mechanisms that support human language. One method for testing the role of potential cognitive mechanisms is the study of individual differences (e.g., Kidd & Arciuli, 2016; Misyak, Christiansen, & Tomblin, 2010; Misyak & Christiansen, 2012), and this is the approach taken in this thesis. This chapter reviews studies of on-line syntactic processing, with a main focus on individual differences. A common feature of the past research, and one that I also followed in my own studies, is a focus on one structural type that elicits large individual differences in speakers – subject and object RCs. I next introduce RCs and review the different theoretical explanation of their processing difficulty.

Relative clauses

RCs are subordinate clauses that modify a noun phrase (NP). Although several different types of RCs exist, the field of sentence processing has almost exclusively concentrated on two types of restrictive RCs, which I will refer to here as subject and object RCs. Consider (1a) and (1b) below, (taken from Just & Carpenter, 1992).

(1a) The reporter [that_ attacked the senator] admitted the error.

(1b) The reporter [that the senator attacked_] admitted the error.

Sentence (1a) is a subject RC; the head noun *the reporter* serves as the subject of the RC, as indicated by the underscore gap. Sentence (1b) is an object RC, with the head noun serving as the grammatical object position of the embedded verb *attacked*.

An extensive literature has reported that object RCs (ORC) are typically more difficult to process than subject RCs (SRC). This result is most clearly observed in European languages such as English (e.g., Gibson, Desmet, Grodner, Watson, & Ko, 2005; Gordon, Hendrick, Johnson, & Lee, 2006; King & Just, 1991; Traxler, Morris, & Seely, 2002), French (e.g., Cohen & Mehler, 1996), German (Mecklinger, Schriefers, Steinhauer, & Friederici, 1995), and Dutch (e.g., Mak, Vonk, & Schriefers, 2002), but has also been observed in more typologically diverse languages such as Korean (e.g., Kwon, Gordon, Lee, Kluender, & Polinsky, 2010) and Japanese (e.g., Miyamoto & Nakamura, 2003; Ueno & Garnsey, 2008). This has come to be known as the *subject-object asymmetry*.

There are several proposed explanations of the *subject-object asymmetry*. This thesis focuses on reviewing the following accounts: (i) the word order / experience based account (Bever, 1970; MacDonald & Christiansen, 2002), (ii) the distance-based dependency locality theory (DLT) (Gibson, 1998, 2000) and (iii) the universal parsing account (Lin & Bever, 2006). Each is described below.

Word order / experience account

The word order account explains the preference for subject RCs with reference to the general preference to build structure according to the frequent, or canonical word order of a language (Bever, 1970; MacDonald & Christiansen, 2002). In English the canonical word order is subject-verb-object (SVO). SRCs follow this pattern, whereas ORCs do not, instead of having the marked object-subject-verb (OSV) word order. The suggestion is therefore that the difficulty associated with ORCs is due to the relative infrequency of the OSV order in the language (see Wells et al., 2009). In this sense the account is experience-based, since it argues that an individual's experience with a given word order influences syntactic processing.

Distance-based dependency locality theory (DLT)

An alternative account is Gibson's (1998, 2000) the distance-based dependency locality theory (DLT). The underlying assumption of DLT is that there are working memory limits on sentence processing, which can be quantified via two concepts: storage and integration costs. Specifically, storage cost refers to resource cost associated with storing partially-built structure, whereas integration cost refers to the integration of the current word into the structure. These concepts are best explained by way of example. In English, the DLT predicts greater difficulty for ORCs because the greater distance between the head noun and the gap within the RC results in

higher storage and integration cost when compared to SRCs (compare 1a to 1b).

Specifically, ORCs result in greater storage cost because two unintegrated NP

arguments must be held in memory before the identification of the gap. For example,

in sentence (1b) both *the reporter* and *the senator* must be held in memory, whereas

for subject RCs such as (1a) only the head noun is activated before the identification

of the gap (i.e., *the reporter*). ORCs also result in greater integration cost at the RC

verb. Here the two NPs must be integrated into the structure of the sentence at this

crucial point. In comparison for a SRC, there is no cost because the head noun and

the RC object occur either side of the verb, in their canonical order, and can be

integrated into the parse as soon as they are encountered.

Universal parsing account

Under the universal parsing account, Lin and Bever (2006) propose an

incremental minimalist parser which builds syntactic structure according to

hypothesised universal syntactic principles. On this account, the parser is argued to

build a hierarchical structure of a sentence. For structures containing RCs, the parser

begins searching for a gap to fill once it encounters the head noun, and will predict

gaps at potentially legitimate sites and attempt to fill them as quickly as possible (the

'Active Filler Strategy', Frazier, 1987). Since the underlying representational nature

of syntactic structure is assumed to be universal across languages, a gap situated at a

higher hierarchical syntactic structure gets filled earlier than one situated at a lower

syntactic position. In this sense, a SRC preference is predicted universally (i.e.,

across all languages of the world containing RCs), as the gap in a SRC is located

higher than the gap in an ORC. Therefore, whereas accounts like the DLT (Gibson,

2000) explain syntactic complexity effects as due to the *linear distance* between a

head noun and the hypothesized gap, the universal parsing account explains

complexity via *hierarchical* distance. Therefore, according to this universal parsing

account, the difficulty in processing an object RC is based on the longer structural

distance between the relativizer and the object gap than that between the relativizer

and the subject gap in a SRC. The gap and relativizer locations for the SRC and

ORC in the hierarchical syntactic structure are illustrated in (2a) and (2b)

respectively.

SL IN SENTENCE PROCESSING

(2a) syntactic tree for English SRC
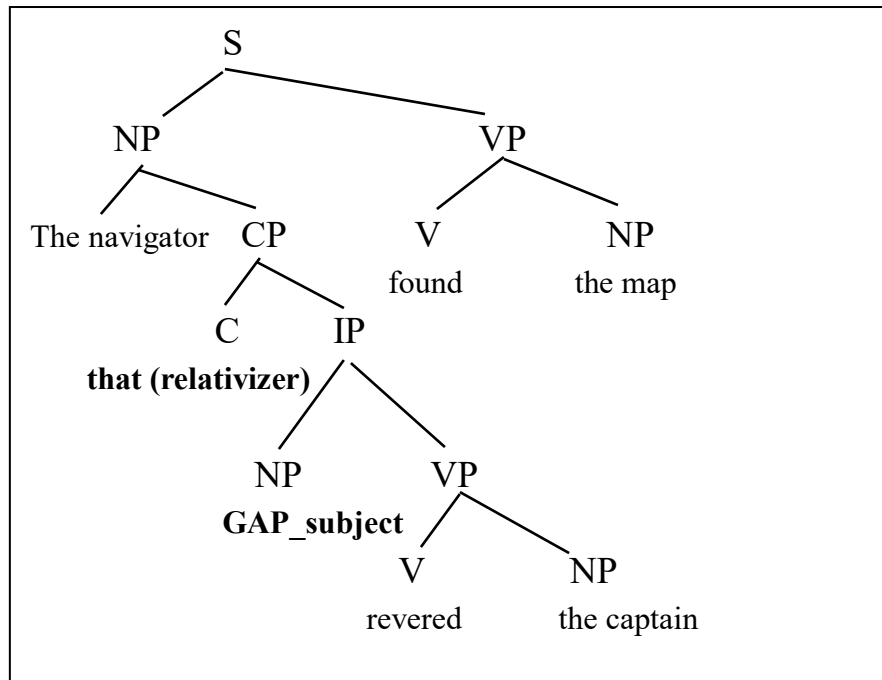
(2b) syntactic tree for English ORC
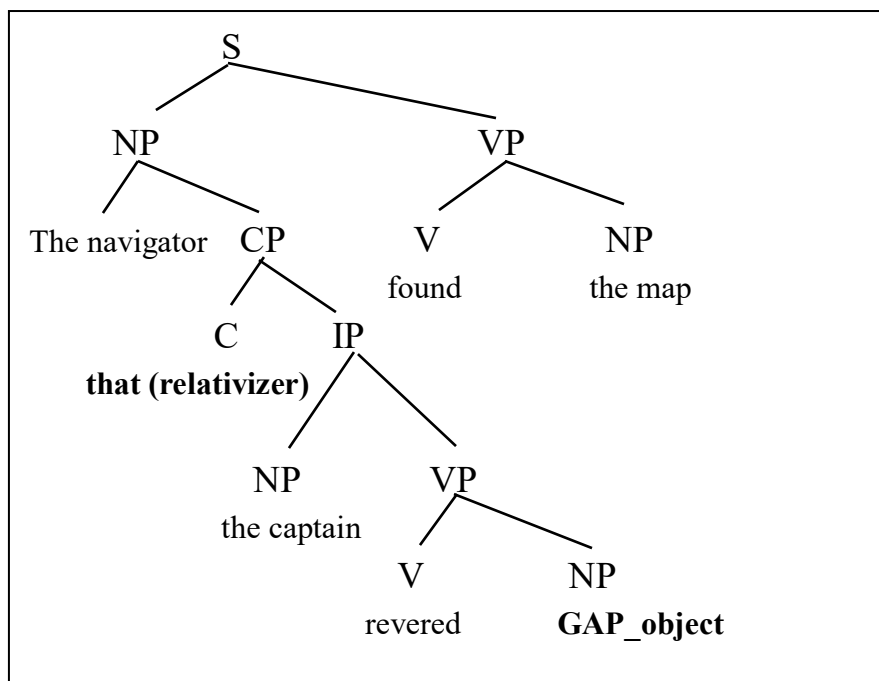
Note: S = sentence (the root node); NP = noun phrase; VP = verb phrase; CP = complementizer phrase; IP = Inflection phrase; C = complementizer; V= verb

Mandarin RCs: An interesting test case.

Although the SRC preference has been widely regarded to be universal across

both head-initial (e.g., for English, see Gibson, et al., 2005; for Dutch, see Mak,

Vonk, & Schriefers, 2002; and for French, see Cohen & Mehler, 1996), and head-

final languages (e.g., for Korean, see Kwon, et al., 2010; for Japanese, see Miyamoto

& Nakamura, 2003), results from studies of Mandarin RC processing have produced

a range of puzzling findings. Interestingly, different theories make differing

predictions regarding RC difficulty in Mandarin. In contrast to those theories that

predict a universal subject preference (e.g., Lin & Bever, 2006), both the word order

account (MacDonald & Christiansen, 2002) and the DLT (Gibson, 1998,2000)

predict an object advantage. This is because Mandarin has the typologically rare

combination of SVO word order and head-final RCs (Dryer, 2005). As such, whereas

SRCs like (3a) have a non-canonical VOS word order and a large distance between

the head noun (*athlete*) and the gap, object RCs like (3b) follow the canonical SVO

word order and have a shorter distance between the head noun (*athlete*) and the gap.

(3a) Mandarin SRC

崇拜　艺术家　的　　运动员　享用了　一顿　美味的　　晚餐。

[**Chongbai** yishujia　de] **yundongyuan** xiangyongle yidun meiweide wancan

[__$_i$ **adored**　the artist　that] the **athlete**$_i$　enjoyed　a　delicious　dinner

'The athlete that adored the artist enjoyed a delicious dinner.'

(3b) Mandarin ORC

艺术家 **崇拜** 的 **运动员** 享用了 一顿 美味的 晚餐。

[yishujia **chongbai** de] **yundongyuan** xiangyoungle yidun meiweide wancan

[The artist adored __*i* that] the **athlete***i* enjoyed a delicious dinner.

'The athlete that the artist adored enjoyed a delicious dinner.'

Therefore, Mandarin, like other Chinese languages such as Cantonese (see also Chan, Yang, Chang, & Kidd, 2017), potentially serves as a key language in debates concerning RC processing and its universality. Accordingly, there have been numerous recent studies focusing on the language, which have notably yielded inconsistent results. In the next section I summarise these studies.

Mandarin RCs processing

In line with those studies that support the universal SRC preference, Lin and Bever (2006) investigated the processing of two types of RCs in Mandarin (i) SRCs vs. ORCs, and (ii) relative clauses modifying the main clause subject vs. the main clause object. Examples of the two types of RCs are shown in (4)

(4) a. Subject-modifying SRC

勾引 院長 的 少女 撞到 了 議員。

gouyin yuanzhang de shaonyu zhuangdao le yiyuan
V1 N1 DE N2 V2 N3

[__*i* **seduce** dean DE] **young lady** *i* bump into ASP congressman

'The young lady that seduced the dean bumped into the congressman.'

b. Subject-modifying ORC

院長　　勾引 的 少女　撞到　　了議員。

yuanzhang gouyin de shaonyu zhuangdao le yiyuan

　　N1　　　V1　　DE　N2　　　V2　　　　N3

　　[dean　　seduce__ *i*　DE]**young lady** *i* bump into ASP congressman

'The young lady that the dean seduced bumped into the congressman.'

c. Object-modifying SRC

議員　　撞到　　了 勾引　院長　　的 少女。

yiyuan zhuangdao le gouyin yuanzhang de shaonyu

　N1　　　V1　　　　V2　　　N2　　　DE　N3

congressman bump into ASP[__ *i* seduce dean DE] **young lady** *i*

'The congressman bumped into the young lady that seduced the dean.'

d. Object-modifying ORC

議員　　撞到　　了 院長　　勾引 的 少女。

yiyuan zhuangdao le yuanzhang gouyin de shaonyu

　N1　　　V1　　　　N2　　　V2　DE　N3

congressman bump into ASP [dean seduce __ *i* DE] **young lady** *i*

'The congressman bumped into the young lady that the dean seduced.'

The results showed that the reading times on both relativizer (*de*) and the head

noun (*young lady*) like (4) were significantly shorter for SRCs than for ORCs,

regardless the RC modification. This result is consistent with the universal parsing

account.

Vasishth, Chen, Li, and Guo (2013) have also found evidence in favor of a SRC

advantage in Mandarin. In a meta-analysis of 13 prior studies that have used the self-

paced reading method, they found evidence for an overall subject advantage. They

also reported two of their own empirical studies that were consistent with the meta-

analytic evidence. Also using the self-paced reading methodology, they found a

significant subject advantage at the head noun in their Experiment 1, and a

significant subject advantage in the word region following the head noun in their

Experiment 2. In contrast to Lin and Bever (2006), however, they did not interpret

their data to be consistent with the universal parsing account. Instead, the authors

attributed the subject-relative advantage in Mandarin RC processing to structural

frequency. Specifically, they argued that, while Mandarin has canonical SVO word

order, which according to the word order account should favor ORCS, SRCs are

actually more frequent in Mandarin than ORCs, therefore potentially explaining the

effect. They reported corpus data consistent with this interpretation.

While those studies reviewed above found a significant SRC advantage, several

studies have found significant oRC advantages across a range of methodologies,

including eye-tracking while reading (e.g. Sung, Cha, Tu, Wu, & Lin, 2015; Sung,

Tu, Cha, & Wu, 2016), a maze task (i.e., a task that similar to self-paced reading, in

which the participants are required to make a choice between two alternatives when

reading every word of the sentence) (e.g. Qiao, Shen & Forster, 2012), and self-

paced reading (e.g. Chen, Ning, Bi, & Dunlap, 2008; Gibson & Wu, 2013; He, Xu,

& Ji, 2017). Thus, these data support the word order and DLT accounts. For instance,

Hsiao and Gibson (2003) found that, in singly embedded RCs, participants read the

first two words in ORCs faster than the first two words in SRCs, but the reading time

difference at the relativizer (*de*) and the head noun regions did not reach

significance. In a subsequent study, Gibson and Wu (2012) investigated RC

processing in discourse context. That is, participants read passages of text that

supported either an SRC or an ORC interpretation. Studies in English have found

that a supporting discourse context can remove the subject advantage (Grodner,

Gibson, & Watson, 2005). Regardless, their Mandarin results still suggested an

object advantage, with participants processing ORCs faster than SRCs at the head

noun. In line with the DLT, the authors argued that Chinese RCs processing is

constrained by working memory limitations, which causes the parser to favor

building the ORC structure.

In a more recent series of eye-tracking studies, Sung et al. (2015) and Sung et

al. (2016) found an ORC advantage at the head noun using multiple indices of

reading, including gaze duration, regression path duration, and regression rate. The

authors concluded that the ORC preference in processing Chinese RCs is attributable

to the unique typological properties of Mandarin. That is, because Mandarin has both

SVO word order and head-final RCs, ORCs follow canonical SVO word order and

have a shorter distance between the gap and the head noun, making them easier to

process.

A more recent self-paced reading study reported similar results. He et al. (2017)

used self-paced reading task to investigate SRCs and ORCs processing in young

(aged 19 –25 years) and older adults (aged 60-65 years). They showed that ORCs

were easier to process than SRCs in both young and older adults. Specifically, the

authors reported that the older adults had difficulty processing SRCs compared to

ORCs from the head noun onwards. He et al. interpreted their findings to suggest

that a range of factors contribute to the difficulty of Chinese RCs processing, such as

age, language specificity, and linguistic experience.

Finally, individual differences research suggests that, as in more commonly

studied languages like English, the processing of Chinese RCs is also subject to

individual differences. Chen et al. (2008) examined the role of working memory

capacity in Mandarin-speaking participants' processing of centre-embedded RCs.

The participants completed a self-paced reading task and an independent measure of

working memory (WM). The participants were then divided into high and low WM

capacity groups. They found that the first two words in ORCs were processed

significantly faster than in SRCs among low working memory capacity group.

However, no significant difference was found in high working capacity group,

presumably because they possessed the WM resources to process both structural

variants (especially SRCs) without significant cost.

In summary, although there have been several studies investigating the

processing of SRCs and ORCs in Mandarin, the results have been mixed One feature

of the data that any theory must explain is the pervasive presence of individual

differences in syntactic processing, and particularly in RCs. There has been a long

history of individual differences research that have focused on RCs (e.g. Just &

Carpenter, 1992) in addition to other syntactic phenomena (e.g., Farmer et al., 2017;

Pearlmutter & MacDonald, 1995). The potential sources of these varied abilities

have been revealed by examining the parsing difficulty during on-line sentence

comprehension, which have then been linked to independently measured cognitive

skills. I discuss this research next.

Individual differences in on-line English RCs processing

Although traditional linguistic approaches to language such as the Universal

Parsing account (Lin & Bever, 2006) do not predict significant and systematic

individual differences in syntactic processing, more psycholinguistically-oriented

theories such as the DLT (Gibson, 2000) and Experience-based account (MacDonald

& Christensen, 2002) postulate different mechanisms that can be linked to individual

differences research.

The DLT (Gibson, 2000) assumes that WM capacity constrains structure

building. Since WM capacity varies in the population, variation in reading times

should be directly related to variation in WM. Accordingly, WM-based accounts

predict that complex sentence structures impose cognitive burden on parsing because

they consume significant WM resources, which results in longer reading times (e.g.

Gibson, 1998; Just & Carpenter, 1992; King & Just, 1991). In contrast, experience-

based theories predict that the ease of parsing a particular syntactic structure depends

on its frequency within the language and the person's experience with that structure

(e.g., MacDonald & Christiansen, 2002; Wells, Christiansen, Race, Acheson, &

MacDonald, 2009; Gennari, Mirkovic, & MacDonald, 2012; see also Kidd, Brandt,

Lieven, & Tomasello, 2007). Structures that are less frequent will hinder

comprehension, which in turn slows processing speed in comparison to more

frequent structures.

Individual differences in working memory capacity have been empirically

linked to the ease with which individuals process RCs. King and Just (1991) tested

46 English speaking participants' verbal working memory capacity using the

Daneman and Carpenters (1980) reading span task. Participants were divided into

two groups according to their WM capacity (high versus low span, via a median

split). Participants also completed a self-paced reading task, where the focus was on

the processing of English SRCs and ORCs. The results showed that participants with

low working memory capacity processed ORCs more slowly than participants with

high working memory capacity, which was indexed by the increased reading time on

the main verb (a critical region that indicated the syntactic complexity). While

shorter reading times were reported on the main verb for SRC than for ORC among

high-span participants and low-span participants, no significant correlation was

found between reading span scores and SRC. Thus, the authors argued that SRC

were easier to process than ORC as the word order of ORC was more demanding on

the working memory resources. The authors concluded that variability in WM

capacity constrained the on-line processing of syntactic information, as measured via

reading times and comprehension-question accuracies.

Just and Carpenter (1992) used a computational model of reading, CC

READER (Just & Carpenter, 1992, p.140), to simulate the reading time patterns

reported by King and Just (1991) and suggested that, in comparison to the high WM

span adults, the low WM span adults did not have sufficient processing resources to

comprehend the more difficult ORCs. Just and Carpenter further investigated several

studies comparing participants with high and low working memory spans, which

demonstrated that complex syntactic processing is mediated by a single WM

capacity, which interacts with the comprehension cues at syntactic, lexical,

contextual, semantic, and pragmatic levels as a whole. Under this capacity constraint

account, individual differences in verbal working memory capacity are implicated in

the online comprehension of syntactically complex sentences, and are specifically

constrained by a domain-general WM capacity.

Working from the assumption that sentence difficulty is directly related to an

individual's experience with language rather than capacity limits per se, MacDonald

and Christiansen (2002) simulated individual differences in RC processing using a

connectionist modelling. In the simulations, the models used simple recurrent

networks (SRNs, Elman, 1991) to acquire and parse grammar. Notably, the models

are sensitive to frequency: the more frequent the structure, the more quickly it will

be acquired the more easily it will be processed. The authors trained 10 connectionist

networks to predict the upcoming word in syntactically simple and complex

sentences. The sentences in the training corpus were generated from a probabilistic

context-free grammar, including subject noun-verb agreement, present and past tense

verbs that varied in terms of their argument structure (transitive/intransitive), and

embedded subject and object relative clauses. Specifically, the occurrence

probability of SRCs and ORCs was evenly distributed; that is, the two constructions

occurred with equal frequency (2.5% the training corpus).

To examine the role of experience with the frequency distribution in the input,

the networks' learning performance was assessed in terms of grammatical prediction

error scores (GPE, Christiansen, 2001; Christiansen & Chater, 1999) across training

epochs. The results showed that, after the first training cycle, the networks

encountered processing difficulty on the critical main verb region in ORCs, as

indicated by higher error scores, but not in SRCs. However, the difference in

processing difficulty between SRCs and ORCs diminished after the networks

receiving the second and the third training cycles. As such, the training effect for

ORCs, which have non-canonical and therefore irregular word order, was larger than

that for the *regular* canonical SRCs. As the training corpus contained relatively

larger number of simple transitive sentences, which follow English canonical SVO

word order, SRC processing was facilitated by the exposure to other sentences

following the same word order pattern. In contrast, there are no other sentences in

English that follow the irregular OSV pattern (except for object clefts, e.g., *It was the*

apple that the boy ate, which did not occur in the training corpus). Hence, the

processing of irregular ORCs relied heavily on the exposure to ORCs themselves.

MacDonald and Christiansen call this effect the *frequency x regularity interaction*.

The idea is that, because in English SRCs have the regular SVO word order, their

acquisition and processing should not be affected by their frequency. However,

because ORCs do not have a common word order, their acquisition and processing is

wholly dependent on the frequency with which they occur in the language. Because

ORCs are not frequent, this therefore explains their difficulty relative to SRCs.

The frequency by regularity interaction hypothesis therefore makes the

prediction that increases in exposure should improve an individual's processing of

ORCs but not SRCs. To test this hypothesis, Wells, Christiansen, Race, Acheson, and

MacDonald (2009) conducted a training study. Participants completed a reading span

task (Daneman & Carpenter, 1980) before they were assigned to either an

experimental or a control group. In the experimental group, participants were

exposed to SRCs and ORCs over several sessions, thereby increasing their

experience with the structures. In contrast, the control group was exposed to a

different set of structures, such as sentential complements (e.g., *The organizers*

estimated that more than 1000000 people attended the peace rally last year) and

conjoined clauses (e.g., *The amateur golfer had beaten many of the pros and even*

won the celebrated championship). Participants' online processing proficiency of

SRCs and ORCs was assessed using self-pace reading task pre- and post-training. In

the pre-test, two groups of participants showed longer reading times for ORCs than SRC at the main verb (a critical region that indicated the syntactic complexity). In the post-test, reading times for both RCs had reduced across groups, while the decrease was larger for ORCs than SRCs in the experience group when compared to that with the control group. Thus, the findings were consistent with the predictions of the *frequency ✕ regularity* interaction. That is, exposure to RC structures increases processing efficiency, but does so more for the non-canonical object RCs. The authors concluded that linguistic exposure plays a significant role in syntactic processing, thus supporting the experience-based account. Moreover, they argue against a WM account, since the two groups were equated on WM capacity.

Wells et al. suggested that the additional RC experience may have altered participants' knowledge about distributional frequency of head noun type and RC type co-occurrences. That is, the greater number of object RCs in particular may have changed the experimental group's expectation at ambiguous points in the structure, namely, the relative pronoun (*the student that ...*), such that an ORC analysis was more likely than prior to training. Taken together, MacDonald and Christiansen (2002) and Wells et al. (2009) suggest language experience play significant role in RC comprehension difficulty independent of any effect of WM.

The findings of Wells et al. also suggest that differences in experience can lead to different patterns of results.

Overall, the experience-based approach argues that individual differences in the difficulty of parsing a particular syntactic structure depends on distributional properties within the language and the person's experience with that structure. Yet, the biological mechanism that individuals adopt to detect the sensitivity to statistical structure and the extraction of these distributional regularities remains debatable. Recently, a growing number of studies have suggested individual differences in statistical learning ability is one possible mechanism (e.g., Brooks & Kempe, 2013; Brooks, Kwoka, & Kempe, 2017; Granena, 2013; Kidd, 2012; Kidd & Arciuli, 2016; McDonough & Trofimovich, 2016; Misyak & Christiansen, 2012; Misyak, Christiansen, & Tomblin, 2010a, 2010b). I consider this possibility in the next section.

What is SL?

This section begins with a general definition of statistical learning (SL), focusing on how individual differences in SL are related to language learning. I then briefly review some empirical findings on the link between SL capacity and L1 adult language processing.

Statistical learning has been broadly defined as the sensitivity to detect the co-occurring regularities from the environment (Kaufman, DeYoung, Gray, Jiménez, Brown, & Mackintosh, 2010; Perruchet & Pacton, 2006). As such, SL has been proposed to be implicated in numerous human abilities, including language learning, perception, categorization, segmentation of continuous input, prediction and generalization via different sensory modalities (Frost, Armstrong, Siegelman, & Christiansen, 2015). Over the past two decades, the role of SL in language acquisition has received a lot of attention. In their seminal study, Saffran et al. (1996) showed that 8-month-old infants were capable of segmenting continuous speech based on co-occurrence probabilities between adjacent syllables. The task began with a 2-minute familiarization phase, in which twenty-four 8- month-old infants were exposed to a continuous auditory input sequence that consisted of four three-syllable words (*bidakupadotigolabubidaku*). The input word strings followed a statistical distribution in which words had high transitional probabilities (TP) within syllables (TP = 1.0, e.g. *bida*) but low transitional probabilities at word boundaries (TP = 0.33, e.g. *kupa*). Following familiarization, the infants were required to discriminate the target words from the extra foil nonwords (TP = 0), which were from the same pool of syllables as the input words. The result demonstrated that infants exhibited a significantly longer listening times for the foil words when

compared to target words, suggesting that they discriminated between the trained and untrained words.

Similarly, Saffran, Newport and Aslin (1996) constructed an artificial language of six three-syllable words (*babupu, bupada, dutaba, patubi, pidabu, and tutibu*) to investigate adults' segmentation of speech and detection of word boundaries. The six words were then used to construct a continuous auditory input sequence. Specifically, the input word strings followed a statistical distribution, where words had varied transitional probabilities (TP) within syllables (TP range from 0.31 to 1.0), and low transitional probabilities at word boundaries (TP between 0.1 and 0.2). After training for 21 minutes, participants were required to discriminate the target word from the foil, which was either a nonword or a part-word foil in a two-alternative forced choice task. Participants' performance on identifying the target word was significantly above chance. The authors concluded that adults are also able to detect word units by using transitional probability cues identified from the input. That is, an SL mechanism that might be important for acquisition is operational in adults.

Taken together, the above studies suggest that infants and adults are sensitive to and can identify statistical regularities present in auditory sequences. Based on these findings, subsequent research has linked SL to phonological acquisition (e.g., Maye,

Werker, & Gerken, 2002), vocabulary learning (e.g., Hay, Pelucchi, Graf Estes, &

Saffran, 2011), syntax acquisition (e.g., Kidd, 2012; Kidd & Arciuli, 2016) and even

specifically to the processing of RC structures (e.g., Misyak et al., 2010).

Importantly, recent has shown that, just like WM, SL ability systematically varies in

the population and is therefore subject to individual differences.

Statistical Learning and Adult Language Processing

Several studies have identified correlations between independent tests of SL and

first language processing in adult populations. For instance, Conway,

Bauernschmidt, Huang, and Pisoni (2010) tested SL in two modalities (visual and

auditory), and showed that performance on both tasks was related to language. In the

visual sequence learning task, participants were required to view a sequence of

colored squares (red, blue, yellow, green), which were displayed one at a time in one

of four different positions on a computer screen (upper left, upper right, lower left,

lower right) and which followed a probabilistic artificial grammar. The participants

were then required to reproduce sequences that either followed or did not follow the

artificial grammar. Performance on this task was related to an auditory language

processing task in which participants were required to identify words in predictable

or unpredictable contexts (e.g. *I've got a cold and a sore **throat** versus David knows*

*long **wheels**)*.

In a second study, Conway et al. (2010) investigated whether auditory SL was related to performance on a task measuring participants' ability to predict upcoming words in distorted speech while watching a person speaking. In addition, the participants' language ability was assessed by two subtests of the Test of Adolescent and Adult Language (TOAL-3; Hammill, Brown, Larsen, & Wiederholt, 1994); namely, the Reading/Vocabulary test and Reading/Grammar test. IQ was also measured. The results once again demonstrated a significant relationship between SL and language processing, a relationship which held after controlling for language proficiency and IQ. In a final study, Conway et al. again showed a strong relationship between visual statistical learning and language processing, this time controlling for working memory, inhibition (a measure of Executive Function), and IQ. Overall, this research provides convincing evidence that SL is implicated in online language processing. However, the results are limited in that the language processing task only measured participants' ability to identify words in context. More recent studies have investigated the role of SL in the processing of a central feature of language – grammar.

Misyak, Christiansen and Tomblin (2010) investigated whether individual differences in SL predicted online processing of SRC and ORC structures. They measured the L1 English adults' SL using sequences of nonwords (*dak,pel,vot,jic,rud*

tood) that followed an artificial grammar (AG) containing both adjacent and non-adjacent dependencies between items. For instance, for the sequence *dak-pel-vot* there is a local relationship between adjacent non-words (an *adjacent dependency*), but a non-adjacent and long-distance relationship between *dak* and *vot* (a *non-adjacent dependency*). The results showed that individual differences in nonadjacent SL correlated with the processing of RC sentences that contained non-adjacent dependencies, suggesting that individual differences in SL predict variation in grammatical processing.

In an additional study, Misyak and Christiansen (2012) investigated whether adjacent SL and nonadjacent SL tasks in AG was related to processing of adjacent and non-adjacent dependencies in natural language. Participants were tested on an independent measure of language comprehension, where the sentences contained adjacent and non-adjacent dependencies. Three syntactic structures were examined, namely, subject (6) and object (7) relative clauses, animate (8) and inanimate (9) reduced relative clauses, and sentences containing lexical ambiguities (10 - 11).

(6)   The reporter that attacked the senator admitted the error.

(7)   The reporter that the senator attacked admitted the error.

(8)   The defendant examined by the lawyer turned out to be unreliable.

(9)   The evidence examined by the lawyer turned out to be unreliable.

(10) Chris and Ben are glad that the bird perches [seem easy to install] /

[comfortably in the cage].

(11) The teacher told the principal that the student needs [were not being met] /

[to be more focused].

In addition, they controlled for a set of six linguistic, cognitive and personality

variables. The results indicated that SL, verbal working memory, and comprehension

of all three sentence types were correlated. After controlling for the effect of all the

above six predictors, only adjacent SL predicted the comprehension accuracy of

sentences containing lexical ambiguities, whereas only nonadjacent SL predicted the

comprehension accuracy of subject-object RCs. The authors concluded that

individual differences in SL of adjacent and nonadjacent tasks in the artificial

grammars predicted parsing of local and long-distance relationships in sentences. It

is important to note, however, that while significant effects were observed, their

sample was relatively small $(n = 30)$, and as such their regression analyses may have

been unstable given the large number of independent variables they included in their

models.

Clear relationships between SL and grammatical comprehension have also been

found in developmental populations. Kidd and Arciuli (2016) developed a battery of

tests for 68 six-to-eight-year-old children. The children's SL capacity was tested

using a visual SL task developed by Arciuli and Simpson (2012). In this task,

children were exposed to a series of 12 "alien" cartoon characters displayed one at a

time on the computer screen. Unbeknown to the children, the visual presentation of

aliens followed a probabilistic pattern. Specifically, the aliens appeared in four

triplets in which three aliens always occurred together. Therefore, there were high

transitional probabilities (TP) between aliens within triplets and low TP between

triplets. Following familiarization, a surprise test phase began. In the test, children

were required to identify triplets from foils in a two-alternative forced choice task.

The children's grammatical comprehension was tested using a picture pointing task.

Four syntactic structure were measured, including simple active (e.g. *which mouse is*

kissing the chicken?), passives (*e.g. which mouse is being kissed by the chicken?*),

subject RCs (e.g. *where is the mouse that is kissing the chicken?*), and object RCs

(e.g. where is the mouse that the chicken is kissing?). The children were also tested

on three covariate measures. General verbal and non-verbal ability was measured via

the Peabody Picture Vocabulary Test IV (PPVT-4, Dunn& Dunn, 2007) and Raven's

Colored Progressive Matrices (RCPM, Raven, Raven & Court, 1998) respectively.

WM was assessed by the Listening Span task (Gathercole & Pickering, 2001). The

results showed that SL predicted comprehension of passives and object RCs over and

above the influence of the three covariate measures. The authors concluded that a

domain-general capacity for SL is implicated in the acquisition of syntactic structure.

A recent event-related potential (ERP) study conducted by Daltrozzo, Emerson,

Deocampo, Singh, Freggens, Lee and Conway (2017) has also reported correlations

between visual SL of nonlinguistic stimuli and grammatical ability in L1 English

speaking adults. They investigated the association between visual SL and language

performance by employing three standardized measures. Seventeen native English

adults' SL capacity was measured by visual SL task developed by Jost, Conway,

Purdy, Walk, and Hendricks (2015). In the task, participants viewed a sequence of

six colored circles on a computer screen one at a time. Unbeknown to the

participants, a set of statistical regularities were created among the 6 colored circles.

For each participant, a *standard* stimulus, a *high predictability* (HP) predictor, a *low*

predictability (LP) predictor, and the *target* stimulus were randomly assigned to one

of six coloured circles. In each trial, the *standard* stimulus repeatedly occurred and

appeared prior to either HP or LP predictors (i.e. each with a 50% probability of

appearance). The *target* and the *standard* followed the HP predictor on 90% and 10%

of the trials respectively, whereas the *target* and the *standard* followed the LP

predictor formed 20% and 80% of the trials respectively. After the exposure of

predictor-target probabilities, adults were asked to identify the *target* stimulus from a

continuous stream of colored circles by hitting the button on the computer. The

adults' language ability was assessed on three standardized measures: final words

prediction and grammatical ability was tested using the Sentence Completion subtest

and the Grammaticality Judgement subtest of the Comprehensive Assessment of

Spoken Language (CASL, Carrow-Woolfolk, 1999), respectively. Receptive

vocabulary was measured via PPTV-4 (Dunn & Dunn, 2007). The results showed

that visual SL predicted the grammatical ability and receptive vocabulary, but there

was no association with sentence completion. Noting that the role of attention in SL

tasks is controversial, Daltrozzo et al. further tested adults' general selective

attention as a covariate measure. The authors reported that the association between

SL and grammatical performance held irrespective of attention, whereas the

correlation of SL and receptive vocabulary did not. These findings provide additional

evidence for the link between SL and grammar ability.

Overall, the above studies on L1 populations demonstrate that individual

differences in SL are linked to language processing in children and adults. To the

extent that this link is attested *across* modalities (e.g., visual SL predicting spoken

language), the SL mechanism supporting language processing may be domain-

general. Since second language (L2) learning may also rely on the acquisition of

statistical regularities and transitional probabilities of a particular linguistic

environment, which reflect a general capacity of SL (Onnis, 2012; Frost, 2012; Frost,

Siegelman, Narkiss, & Afek, 2015), it is worthwhile examining whether SL has a

role to play in L2 processing. In the next section I discuss past research that has

investigated the link between SL and L2 acquisition

Statistical Learning and Second Language Acquisition

Relatively few studies have investigated the role of SL in second language (L2)

acquisition or processing. In one study, Frost, Siegelman, Narkiss, and Afek (2013)

investigated the role of SL in predicting L2 reading proficiency of Hebrew by native

speakers of English. They tested participants' visual SL via their ability to detect

adjacent dependencies in the continuous stream of abstract shapes (Turk-Browne,

Jungé, & Scholl, 2005; see also Glicksohn & Cohen, 2011; Kim, Seitz, Feenstra, &

Shams, 2009). In this task, participants saw the stream of 24 abstract shapes

presented one at a time on a computer screen. Similar to the task used by Kidd and

Arciuli (2016), the shapes followed a probabilistic pattern, but contained 8 triplets.

Following familiarization, participants were required to discriminate the triplets from

the false triplets in a two-alternative forced choice task. Three tasks were used to

measure participants' Hebrew reading proficiency, one measuring non-word

decoding, another measuring real word reading, and another measuring

morphological knowledge via a priming task. The adults' working memory capacity

and general cognitive abilities were also measured. The results revealed a significant

correlation between SL and Hebrew (L2) reading proficiency in all three reading

tasks, irrespective of general cognitive factors. The authors concluded that SL

predicted second-language literacy acquisition.

Kaufman, DeYoung, Gray, Jiménez, Brown, and Mackintosh (2010) examined

the relationship between implicit learning in 16 to 18-year-old English students and

L2 exam performance. They used a Serial Reaction Time (SRT) task to test

participants' probabilistic sequence learning. In the task, participants viewed an

abstract shape (an asterisk) occur in one of four positions on a computer screen, and

the participants were required to press the button on the keyboard that corresponded

to the position in which the shape appeared as quickly as possible. Unbeknown to the

participants, the order of the digit sequences was based on a specific probabilistic

distribution (Schvaneveldt & Gomez, 1998). The participants' SL was assessed via a

reduction in reaction time across trials in patterned blocks versus random blocks.

The participants' L2 performance was assessed using the L2 (i.e. French / German)

General Certificate of Secondary Education (GCSE) exam scores (UK Year 10

equivalent). The results revealed a significant correlation between SL and L2 GCSE

exam score. While suggestive, the result is not very specific, as it is unclear to which

aspect of L2 learning SL is related.

Similarly, Granena (2013) investigated the role of SL ability in early and late

Chinese-Spanish bilinguals' L2 morphosyntactic acquisition. SL was measured via

two tasks. The first investigated participants' ability to extract phonotactic statistical

information (i.e., internal statistical structure) from a series of non-words (Speciale,

Ellis, & Bywater, 2004). The second tested participants' ability to learn statistical

sequences of visual patterns in a serial reaction time task (SRT task). The

participants' knowledge of several aspects of Spanish was tested, including (i) noun-

adjective gender agreement, (ii) subject-verb agreement, (iii) noun-adjective number

agreement, (iv) subjunctive mood, (v)perfective/imperfective aspect, and (vi)

passives. A significant association was found between auditory sequence learning

ability and knowledge of Spanish morphological agreement in early L2 learners (i.e.,

those who had begun learning Spanish before puberty). In contrast, performance on

the visual probabilistic SRT task was significantly associated with grammatical

sensitivity to agreement relations in late L2 learners (i.e., those who began learning

Spanish post-puberty). Thus, these findings indicated that individual difference in

sequence learning predicted L2 language proficiency scores.

Recently, McDonough and Trofimovich (2016) examined the role of SL and

WM in 140 Thai-English bilinguals' acquisition of two grammatical constructions.

The bilingual adults completed a SL task in which they learned non-adjacent

dependencies (following Misyak et al., 2010). In addition, the bilinguals' WM was

measured using the backwards digit span task from the Wechsler Adult Intelligence

Scale, WAIS-III (Psychological Corporation, 1997). Two L2 constructions were

tested: the novel Esperanto transitive (e.g *tauro batas cevalon*, i.e. *bull hits horse*),

and English double-object dative construction (e.g., *John built the table a leg*). In the

exposure phase, participants heard the Esperanto transitive alternation, where word

order because the suffix *-n* is used to mark the object in the sentence (i.e. SVO, e.g.

tauro batas cevalon and OVS, e.g. *cevalon batas tauro [bull hits horse]. T*he English

double-object dative construction varied in terms of prototypical sentence with

human recipient (12) and nonprototypical sentence with inanimate noun phrases as

object and recipient (13).

> (12) Mr. Smith enjoyed teaching students about chemistry. But they had a lot of
>
>   problems with the last exam, so he told them the answer.

> (13) John's children broke a table while they were playing. It was his favorite
>
>   table, so John built the table a leg.

Following exposure, the bilinguals' L2 grammatical structure learning was

assessed using a forced-choice picture selection task. The results showed that SL

predicted the learning of novel L2 Esperanto pattern and the nonprototypical English

double-object dative construction, whereas WM was not related to learning. Hence,

these findings suggest that SL supports the learning of L2 morphosyntactic patterns.

With respect to the frequency distribution of L2 input structure, Brooks, Kwoka

and Kempe (2017) investigated the role of SL and general cognitive abilities in 54

English adults' learning of L2 Russian inflectional morphology. The participants' SL

was assessed using auditory sequences of 10 monosyllabic nonwords (*hep, tam, biv,*

dupp, jux, lum, meep, sig,zoet,rauk) that followed an artificial grammar containing

adjacent-dependencies. After training, the participants were required to discriminate

sequences that followed the rules in a two-alternative forced choice task. L2 learning

was assessed through the participants' learning of Russian noun cases. The

frequencies of each noun were exposed to the participants in a balanced or skewed

distribution (i.e. the specific nouns occurred more frequent than other nouns).

Participants' L2 Russian morphology learning was tested across three sessions, in

which session 2 and 3 were composed of training and testing phases, whereas

session 1 merely included training phase.

The results showed that SL and nonverbal intelligence predicted L2 Russian

morphology learning. In addition, the authors reported a significant correlation

between nonverbal intelligence and the generalization of L2 case-marking suffixes to

new vocabulary. With respect to the frequency distribution of L2 input structure, they

showed that balanced frequency input supported generalization in L2 learning.

Furthermore, phonological short-term memory and verbal working memory were not

related to L2 learning. Brooks et al. (2017) interpreted their findings to suggest that

individual difference in SL and nonverbal intelligence have a prevailing role over

frequency distribution of input in L2 morphology acquisition.

However, one study by Robinson (2005) did not find a relationship between SL

and L2 grammar. This study investigated the correlation between SL and novel L2

Samoan grammar learning in 37 Japanese-English adults. In the SL task, the author

tested the proficient L2 learners (i.e., English Majors) using letters strings that

followed a statistical pattern, developed by Knowlton and Squire (1996), whereas the

novel L2 Samoan learning was assessed by the grammatical judgment task. There

was no significant link between SL task and the novel L2 Samoan grammaticality

performance. In contrast, a significant link was found between intelligence (IQ) and

the above two tasks, while WM was correlated with the grammatical judgement task.

Therefore, although the weight of evidence suggests that SL is implicated in L2

learning, null results have been reported.

The Present Research

This review of the literature has shown that infants and adults are able to detect

statistical regularities via input across auditory and visual modalities, and they were

sensitive to distributional properties of their own language. However, SL appears to

be a skill that is subject to significant and meaningful individual differences.

Accordingly, SL has been linked to individual differences in how children acquire

and how adults process language (e.g., in children: Kid & Arciuli, 2016; in adults:

Misyak & Christiansen, 2012). Most of the research has investigated first language

acquisition and processing; however, there is now a small literature looking at L2

learning and processing. Most published studies have reported a significant role for

SL in L2 learning, but despite these suggestive results many questions remain. For

instance, is SL similarly implicated in an individual's L1 and L2? Is SL differentially

implicated in L2 learning depending on an individual's level of proficiency, given

that proficiency plays a crucial role in L2 processing (Hopp, 2006; Jackson, 2008;

Rah & Adone, 2010)? Is the SL-language link evident or even different in a non-

alphabetic language? This thesis explores these questions.

Chapter 3

Study 1: The Role of Visual SL and L2 Language Proficiency in Chinese and English Sentence Processing

Study 1 aimed to investigate the relationship between the capacity of visual statistical learning (VSL) and on-line sentence processing in bilingual speakers who speak Mandarin as their L1 and English as their L2, and sought to examine whether second language proficiency mediates the relationship between VSL and L2 language processing. Given that SL has been reported to predict English language processing among L1 populations, and that a role for SL in L2 language learning has been identified (Frost, Siegelman, Narkiss, & Afek, 2013; Mcdonough & Trofimovich, 2016; Brooks, Kwoka, & Kempe, 2017), it was hypothesized that participants' VSL ability would be independently associated with L1 (Chinese) and L2 (English) syntactic processing. Specifically, we expected to find that Chinese-English bilinguals with higher VSL scores would show better performance in both L1 and L2 sentence comprehension tasks than those with lower VSL scores, as reflected by faster reading times, even when other linguistic and cognitive variables were considered.

Method

Participants

Sixty (40 females, 20 males; age: M = 22.3 years, range: 18-34) Chinese-

English bilingual undergraduate and postgraduate students were recruited via the

Psychology Research Participation Scheme at The Australian National University.

All participants provided their informed consent and were paid $30 as compensation

for their participation across two testing sessions. The participants were all native

speakers of Mandarin, with no reported history of uncorrected visual, auditory or

neurological impairments. The data of participants with less than 80% accuracy on

the language comprehension tasks was removed from the analysis. Based on the

inclusion criteria, 40 participants were selected to be valid in English sample (11

males, 29 females; age: M = 22.2 years, SD = 3.65, range = 18-34) and 55

participants (16 males, 39 females) with a mean age of 22.27 (SD = 3.26, range = 18-

34) were considered in Chinese sample.

Participants were asked to complete a L2 language history questionnaire

adapted from Li, Zhang, Tsai and Puls (2015) about (i) their age at which they

started learning English (L2), (ii) self-rated proficiency of English in terms of

listening, speaking, reading, and writing, using 7-point Likert scale (from 1= *very*

poor to 7 = *native-like*), and (iii) a rating of how good the participant's L2 learning

skill was in comparison to their friends (7-point scale, from 1 = *Very Poor* to 7 =

Excellent).The forty valid participants started using English as L2 at the mean age of

7.4 years (*SD* = 5.3 years, range = 3-18). Table 1 summarizes the participants' self-

rated proficiency and learning skill of English.

Table 1

Self-rated Proficiency and Learning skill of English (*n* = 40)

Proficiency measure	Mean	SD	Minimum	Maximum
Listening Ability	5.63	1.10	1	7
Speaking Ability	5.05	1.18	2	7
Reading Ability	5.53	1.01	3	7
Writing Ability	5.03	0.83	3	7
Learning Ability	4.48	0.85	3	7
Sum of rating score (1-35)	25.7	3.84	14	35

On average, participants reported good command of English in the domains of

listening, reading, speaking and writing. However, more than half (52.5 %) of the

participants rated their English learning ability as limited to average, indicated by

self-rating between 3 and 4. As the participants were students of The Australian

National University, they all had achieved a minimum academic IELTS score of 6.5

to meet English Language admission requirements. Therefore, while some self-

ratings were low, we can be confident that they had reasonable English-language

proficiency.

Materials and Procedures

The participants were tested on five tasks across two testing sessions. The

duration of the experiment was approximately 2 hours. As is customary in individual

differences studies, each participant was tested on the tasks in the same order, which

aims to reduce any unwanted error variance associated with variation in testing

procedures. An on-line Chinese (L1) sentence comprehension task, a non-verbal IQ

test, and the L2 English Language proficiency measures were included in the first

session. For the L2 English Language proficiency measures, the participants were

asked to complete (i) the L2 language history questionnaire (as described

previously), and (ii) the Peabody Picture Vocabulary Test (PPVT, Dunn & Dunn,

2007). An on-line English (L2) sentence comprehension task, a measure of working

memory, and a visual SL task were conducted in the second session. Each is

described below.

Sentence Comprehension Task

Two sets of stimulus items were developed to test participants' Chinese (L1)

and English (L2) sentence comprehension. The target sentence structures were

subject and object RCs. Each stimulus set (i.e., Chinese and English) contained 160

test sentences, composing 64 Relative Clause (RC) sentences (32 subject RCs and 32

Object RCs) and 96 filler sentences.

Mandarin Self-paced Reading Task

Thirty-Two Chinese singly-embedded RC sentence pairs were constructed, with

each pair containing a subject- and object-extracted version of the sentence. Thirteen

pairs were taken from Sung, Cha, Tu, Wu and Lin (2015), and 15 pairs of Chinese

singly-embedded RCs were adapted from Hsiao and Gibson (2003). The remaining

four pairs were created anew. All noun phrases were animate, since past studies have

shown that animacy moderates the complexity associated with ORCs in Mandarin

(e.g., Wu, Kaiser, & Andersen, 2012). The following are examples for the two types

of Chinese RCs.

(14 a) Subject-extracted RC (SRC)

崇拜　　　艺术家　的　　　运动员　　享用了　　一顿　美味的　　晚餐。
[Chongbai yishujia　　de] yundongyuan xiangyongle yidun meiweide wancan
[__*i* adored　　the artist　that]　the **athlete***i*　enjoyed　　a　　delicious　　dinner
　　　V1　　N1　　de　　N2　　V2　　DET　　Adj.　　N3
"The athlete that adored the artist enjoyed a delicious dinner."

(14 b) Object-extracted RC (ORC)
艺术家　崇拜　　　的　　运动员　　　享用了　　一顿　美味的　　晚餐。
[yishujia　chongbai de]　yundongyuan　xiangyoungle yidun　meiweide wancan
[The artist adored __*i*　that] the **athlete***i*　enjoyed　　a　　delicious　　dinner.
N1　　　V1　　de　　N2　　V2　　DET　　Adj.　　N3
"The athlete that the artist adored enjoyed a delicious dinner."

In sentence (14a), the head noun *the athlete* occupies the subject position within the

RC, as denoted by the underscore gap. In contrast, in sentence (14b), the head noun

occupies the object position. The test sentences varied in length, from 11 to 19 words

(mean length = 15 words). The length of the RCs was five to seven characters. All

the target sentences are listed in the Appendix.

The Chinese RCs were interspersed within 96 filler sentences, which contained

three different types of structures: (i) Chinese pivotal sentences[1] as shown in (15),

(ii) passive sentences (BEI structure with an agent) as shown in (16), and (iii)

declarative sentence, as shown in (17).

(15) 教授 鼓励 小陈 进行 癌症 研究。
    jiaoshou guli [xiaochen jinxing aizheng yanjiu]
    N1   V1   N2   V2   NP3
"Professor encouraged Mr. Chen to conduct cancer research."

(16) 这些 议员 的 提案 全 被 财政 部长 否决 了。
    Zhexie yiyuan de tian quan bei caizheng buzhang foujue le
"The parliamentarians' proposals were rejected by the finance minister."

(17) 芬兰 是 观察 北极 光 的 最佳 地点 之一。
    Fenlan shi guancha beiji guang de zuijia didian zhi yi
"Finland is one of the best places to observe the northern lights."

There were 32 tokens of each filler type. All target structures and fillers were

presented in simplified characters. After reading the complete sentence on a

computer screen, participants were required to answer a Yes/No comprehension

question. For Chinese RC sentences, half of the comprehension questions asked

about the main clause and the other half asked about the relative clause. For each

---

[1] Here, Chinese Pivotal Sentence refers to the generalized structure of NP1 + VP1+NP2+VP2, in which NP2 (the pivot Noun) is the object of VP1 and the subject of VP2 (pivot verb) (Dong, Qiu and Chen ,2013). Semantically, the event indicated by the first verb is the trigger of the action designated by the pivot verb. The Chinese Pivotal Sentence can be transformed into Chinese RC structure, see Li and Thompson (1981) and Peng (2016) for more details.

category of filler sentences, eight comprehension questions were generated.

English Self-paced Reading Task

Thirty-Two English RC sentence pairs were adapted from Hutton and Kidd

(2011). Once again, each pair had a subject-relative and an object-relative version.

Examples are shown in sentences (18a) and (18b).

(18a) Subject-extracted RC (SRC)

The **navigator [**that __revered the captain] found the map without any trouble.

(18b) Object-extracted RC (ORC)

The **navigator [**that the captain revered __] found the map without any trouble.

Sentence (18a) is a subject RC; the head noun *navigator* serves as the subject of

the RC, as indicated by the underscore gap. Sentence (18b) is an object RC, with the

head noun serving as the grammatical object position of the embedded verb *revered*.

All English RC sentences were controlled for length in words (mean length= 12.5

words, Range = 10 : 15). The length within RCs consisted of 4 words. As in the

Chinese version of the task, all NPs were animate, since the animacy of the head

noun also modulates object RC difficulty in English-speaking L1 adults (Traxler,

Morris, & Seely, 2002) and in L2 English learners (e.g., Baek, 2012). All the target

sentences are listed in the Appendix.

The English RCs were interspersed within 96 fillers, which were adapted from

Hutton and Kidd (2011). Thirty-two fillers were sentential complements (19), 32

were conjoined clause sentences (20), and the remaining 32 fillers were simple

declarative sentences (21)

(19) The instructor implied that the student would not understand the answer.

(20) The translator reread the word and then referred to his dictionary for an

exact definition.

(21) The eagle led the hunter.

After reading the complete sentence on a computer screen, participants were

required to answer a Yes/No comprehension question. As in the Chinese version of

the task, half of the comprehension questions asked about the main clause of the

English RCs and the other half asked about the relative clause. For each category of

filler sentences, eight comprehension questions were constructed.

Procedure

Eight experimental lists were created for both the Chinese and English self-

paced reading tasks. Each trial incorporated six practice sentences to allow

participants to become familiar with the self-paced presentation format of the

experiment. The test RC structures were pseudo-randomly interspersed among filler

items, such that no RC structures occurred consecutively. Thus, participants were

individually tested on 32 RCs (16 SRCs and 16 ORCs) and 96 fillers in total. In each

list, yes-no comprehension questions were approximately asked after every fourth

sentence. A total of sixteen comprehension questions were asked for the target

structures and 24 for the filler sentences in Chinese and English script, respectively.

The self-paced moving window method was used (Just, Carpenter, & Woolley, 1982), presented via E-prime 2.0 (Schneider, Eschman, & Zuccolotto, 2002). At the beginning of each trial, participants saw a series of dashes indicating the length of the sentence. Each dash corresponded to a word or phrase. Participants pressed the space bar to reveal subsequent segments of the sentence. When the new word segment was presented, the previous one would become hidden from view. After displaying the final word segment of each sentence, the participant was required to either answer a comprehension question, or proceed to the next sentence at their own pace following the presentation of "Ready for next sentence?"

Visual Statistical Learning Task

SL was measured using Siegelman, Bogaerts and Frost's (2017) visual SL task

(VSL), because past studies have reported that individual differences in VSL are

predictive of linguistic proficiency (e.g., in L1: Arciuli & Simpson, 2012; Kidd &

Arciuli, 2016; in L2: Frost, Siegelman, Narkiss, & Afek, 2013). The task was chosen

because it has good to excellent psychometric properties (test-retest reliability = .68,

Cronbach's α = .88). The VSL task tests participants' ability to detect adjacent

dependencies in a continuous stream of meaningless shapes. The task began with a

familiarization phase, in which participants were exposed to a series of novel shapes,

one at a time, on a computer screen. Unbeknown to participants, the 16 shapes were

used to construct eight triplets with different transitional probabilities (TP), including

four triplets that were constructed from four shapes, and which had internal

transitional probabilities (TPs) of .33, and another four triplets made from the

remaining 12 shapes with internal TPs of 1. These eight triplets were repeatedly

displayed to the participants, one at a time, during the familiarization phase, with

each triplet presented 24 times. The order of the triplets was randomly presented to

the participants with a constraint that no repeated triplets appeared consecutively.

Each shape was displayed on a computer screen for 800ms, with a 200-ms pause

between shapes. The familiarization stream lasted approximately 10 minutes. The 16

shapes adopted in the current VSL task are presented in Figure 1.
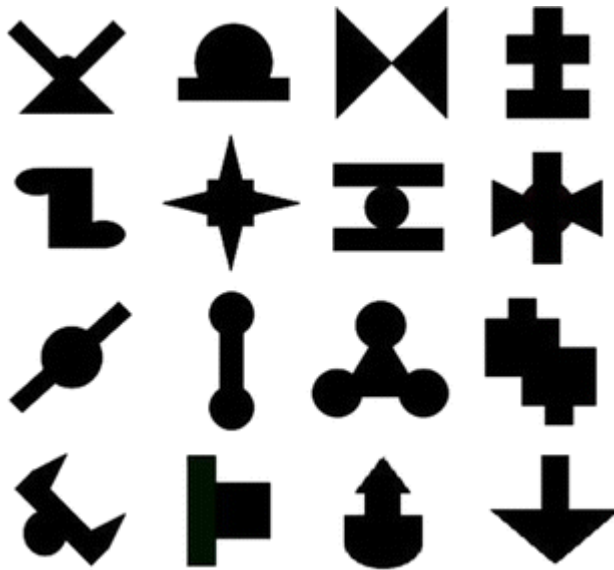
Figure 1. The 16 shapes used in the visual SL task

In the 42-item test phase, participants were required to identify targets from

foils in 34 pattern recognition items test and eight pattern completion items. For the

pattern recognition test, participants were required to select the familiar pair or triplet

(from a foil or multiple foils), whereas in pattern completion test they were asked to

identify a missing shape. The number of alternative choices in each trial varied from

two to four. The foils were constructed into pairs or triplets with internal TPs ranging

from 0 to 0.5. In the pattern completion test, participants were required to identify a

missing shape in a pattern, in which the pattern could be a triplet or a pair. The

statistical learning score of each participant was determined by how many target

items they could correctly identify and complete. The total score of the test was 42.

Participants' raw scores were then computed to generate a proportion correct score

for analysis.

Verbal Working memory

Individual differences in verbal working memory (WM) are correlated with

online measures of sentence comprehension (Misyak & Christiansen, 2012; Jeon &

Yamashita, 2014; Chen, Ning, Bi and Dunlap, 2008). The sentence-span task from

Lewandowsky, Oberauer, Yang and Ecker (2010) for adults was used to measure

participants' working memory. The task has good psychometric properties

(*Cronbach's α* =.76). The sentence-span task was administered individually in the

participants' L1 (Chinese) in order to avoid the potential influence of differences in

L2 proficiency (Mackey, Adams, Stafford, & Winke, 2010; McDonough &

Trofimovich, 2016). In the task, participants read a series of Chinese sentences via

Power-Point slides. The entire task consisted of 60 sentences, which were further

divided into 5 blocks. Each block consisted of two, three, four, five or six sentences

that were 15 to 18 words in length (Mean length = 16.5 words). The sentences were

presented one at a time on the screen for 7 seconds. In each trial, participants had to

evaluate the truth value of the sentence (True/False) (e.g. False sentence: 我没带钱

出门,*不幸遇到老友才没丢脸。[*I went out without taking any money, but**

***unfortunately** I ran into an old friend who helped me out.*]) and remember the final

word (e.g.丢脸) of each sentence for later recall at the end of the block. The number

of sentences and the number of target words in each block increased from two to six

as the participants proceeded through the task, with three trials administered at each

level. Participants were required to complete three trials within each level before

progressing to the next highest sentence level. A question mark '?' was displayed at

the end of each trial in order to signal the participants to recall all of the target words.

Two practice blocks were included at the beginning of the task in order to familiarize

the participants with the procedure.

Participants received one point for every correct word recalled. This differs

from other scoring methods of the test, where participants are given a score that

corresponds to the highest level of the test at which they were successful (e.g., a

score of 3 would indicate that a participant was able to pass the level that contained

three sentences and therefore 3 words to recall). Friedman and Miyake (2005)

demonstrated that the total number of words recalled across all trials had good

reliability (three trials per level: test-retest reliability = .72) and relatively high

correlations with a separate reading task ($r = .51$) when compared to an absolute

span scoring method, which counted only the highest perfectly recalled trials (three

trials per level: $r = .42$) In scoring the performance, participants were given one

point for each correct word recalled and one point for each correct plausibility

judgment. The maximum score was 120. Participants' raw scores were used for

analysis.

Non-Verbal IQ

Non-verbal IQ was assessed to ensure that any association between SL and

language comprehension did not reflect general intellectual ability. Non-verbal IQ

was measured using Ravens Advanced Progressive Matrices (RAPM, Raven, Raven

& Court, 1998). The RAPM has good internal consistency and reliability in

university student samples (split-half reliabilities >.8, Paul, 1985; Alderton &

Larson, 1990, and the test-retest reliability, r = .83, Bors & Forrin, 1995). In this test,

participants were asked to identify complex visual patterns in abstract pictures and

were required to select the most appropriate pattern to complete the missing matrix

from an array of eight alternatives. The sets progressively increased in difficulty, raw

scores on each set (i.e., Set I and Set II) were summed to form a total non-verbal IQ

score for analysis. The total maximum score was 48.

L2 Vocabulary knowledge

A large body of research has reported significant positive correlations between

vocabulary knowledge and reading comprehension (e.g., Qian, 2002; Qian & Schedl,

2004). Specifically, L2 vocabulary knowledge has been found to be a significant

determinant in L2 reading comprehension (Jeon & Yamashita, 2014). Therefore,

participants' L2 (English) receptive vocabulary knowledge was measured using the

Peabody Picture Vocabulary Task (PPVT-4, Dunn & Dunn, 2007). The fourth edition

was selected because it is suitable for assessing the English vocabulary of non-

English-speaking adults (Dunn & Dunn, 2007). In addition, PPVT-4 has been used as

a means of assessing adults' L2 English vocabulary size (e.g., Hellman, 2011). The

PPVT4 has excellent psychometric properties (split-half reliability and test-retest

reliability >.90).

In the task, participants were required to select one picture out of four in

response to the examiner's verbal prompt. The test has two parallel forms (PPVT-4-

A and PPVT-4-B). Each form of the PPVT consisted of 228 single-word items,

divided into 19 item sets. The sets progressively increase in difficulty, and testing

terminated when participants made 8 or more errors in a set. Participants' vocabulary

score was calculated by subtracting the number of incorrect items from the number

of the last item administered. The total maximum score was 228, raw scores were

used for analysis.

Results

Table 2 shows the descriptive statistics for VSL task, WM, vocabulary and

nonverbal IQ performance among 55 participants in Chinese sample, and the

correlations between the measures are shown in Table 3.

SL IN SENTENCE PROCESSING

Table 2

Descriptive Statistics for Predictor Variables in Chinese Sample

	M	SD	Range	Skewness	Kurtosis
VSL	0.70	0.15	(0.48,1)	1.16	-1.10
WM	45.67	25.44	(17,110)	3.35	-0.64
Nonverbal IQ	38.35	5.61	(25, 48)	-1.74	0.10

Note.VSL: visual statistical learning; WM: working memory
VSL task: min score=0, max score=1
WM task: min score= 0, max score=120.
Vocabulary: min score=0, max score=228
Nonverbal IQ: min score= 0, max score= 48

Table 3

Simple Bivariate Pearson Correlations between Predictor Variables in Chinese Sample

	VSL	WM
Nonverbal IQ	-.215	.116
WM	-.090	

None of the predictor variables were significantly correlated. Group

performance on the VSL task was significantly above chance (one-sample t test

tested against 0.5 chance performance, t (54) = 9.84, p < .001, Cohen's d = 1.33).

One anomaly in the data is the particularly low performance of the participants on

the WM task. In fact, performance was so low (mean below 40%) that it is doubtful

that the task was adequately tapping the participants' true WM skills. Since the

reliability of the scores is in question, performance on the WM task was eliminated

from further analysis. This problem is rectified in Study 2.

I next present the results for the participants' processing of Mandarin RCs,

followed by the results for their processing of English RCs. Participants who scored

higher than 80% accuracy ($N = 55$) on all comprehension questions in Chinese

sentence comprehension task were included. Figure 2 shows the mean reading times

for each region of interest in the Mandarin RCs. RTs greater than 2000ms were

excluded from the analysis since the reading and response execution for each phrase

normally require less than 2000ms, and as such RTs > 2000ms cannot be considered

to be tapping into online syntactic processing. Analyses focused on the head noun

and at the region following head noun, since previous studies have shown that this is

where complexity effects are found in Mandarin using self-paced reading (e.g.,

Gibson & Wu, 2013; Lin & Bever, 2006; Vasishth, Chen, Li, & Guo, 2013).

At both the head noun and at the word following the head noun, the reading

time was shorter for ORCs than SRCs; however, this difference did not reach

significance. The mean reading times at the head noun were 519.41ms ($SD = 310.51$)

for SRCs, and 519.29ms ($SD = 291.87$) for ORCs. At the word following the head

noun, the mean reading times for SRC was 533.97ms ($SD = 287.95$) and 533.73ms

($SD = 273.21$) for ORC. No significant effects were found between SRC and ORC at

the first two words and the relativizer (*de*).
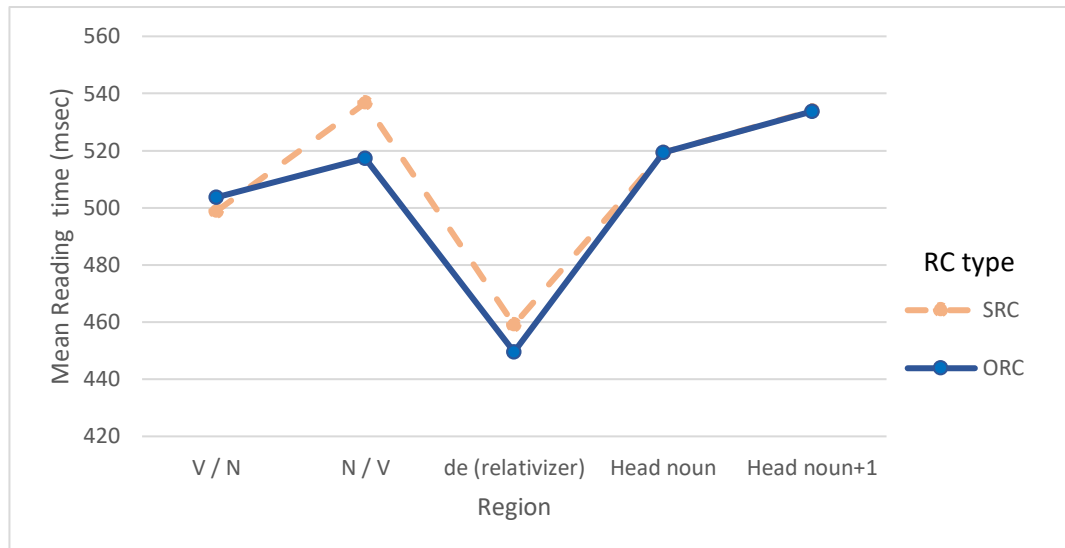
Figure 2. Mean raw reading times of each region of interest in Chinese RCs.

The data were analyzed in two steps. Firstly, I analyzed the self-paced reading data by themselves to determine whether there was a reliable subject-object asymmetry. I then analyzed whether individual participants' reading times could be predicted by the measured covariate variables. The raw reading times for the regions of interest as shown in Figure 2 were analyzed using log transformation since the residuals were strongly skewed. As mentioned earlier, the WM task was excluded. Thus, three variables, VSL, non-verbal IQ, and number of characters of the whole RC sentence (which acted as a control variable), were included as fixed effects in the individual differences analyses. The data were analyzed using the linear mixed-effects models in R, using the *lme4* package (Bates, Maechler & Bolker, 2012). All predictor variables were converted to z-scores in order to reduce the collinearity and make the coefficients more interpretable. Structure was um coded (i.e., SRC = -1,

ORC = +1) to allow ANOVA-like interpretations. Residuals of linear mixed effects

models were inspected to avoid large deviations from the normality. Random slopes

and intercepts were included to explain the variance of random effects for

participants and items. Models were specified with a maximal random effects

structure and were then compared to simplified models using AIC (Barr et al., 2013).

The model with the smallest AIC value (i.e. the best fitting model) was selected.

I first analyzed the group performance on the test structures. As such, the data

were entered in a linear mixed effect model with RC structure as fixed effect.

Following the above procedure to determine the random effects structure, the

random-effect structure thus included a by-participants random intercept, by-items

random intercept and by-participants random slopes for RC structure. The results for

the regions of interest are shown in Table 4.

SL IN SENTENCE PROCESSING

Table 4

Main effect of RC structure by region of interest

Region		Coefficient	SE	P value
de (relativizer)	Intercept	6.02	0.07	<.001***
	RC structure	-0.01	0.01	0.44
Head noun	Intercept	6.05	0.09	<.001***
(HN)	RC structure	0.01	0.02	0.67
HN+1 region	Intercept	6.09	0.09	<.001***
	RC structure	-0.0003	0.01	0.99

Note: A negative sign on the estimated coefficient indicates an object preference;
RC=relative clause

***p<.001.

Table 4 demonstrates that there were no significant differences in the processing

of subject and object RCs across the region of interest.

I next examined whether SL and other cognitive factors predicted individual

differences in the data. The three variables VSL, non-verbal IQ and number of

characters (which acted as a control variable) were simultaneously entered as fixed

effects in the linear mixed-effects model with by participant random slopes for RC

structure, random intercepts for items and participants. The interaction between the

predictor variables and RC structure were also included. The statistical analyses of

the full models for each region of interest are shown in Table 5.

Table 5

Main effects of VSL, non-verbal IQ, RC structure and their interaction by region of interest.

Region		Coefficient	SE	p value
Relativizer **(*de*)**	Intercept	6.02	0.07	<.001***
	RC structure	-0.01	0.01	0.44
	VSL	0.01	0.03	0.79
	Non-verbal IQ	**-0.07**	**0.03**	**.03***
	No. of characters	0.002	0.004	0.66
	VSL x RC structure	0.01	0.01	0.35
	Non-verbal IQ x RC structure	-0.001	0.01	0.94
Head noun **(HN)**	Intercept	6.05	0.09	<.001***
	RC structure	0.01	0.02	0.67
	VSL	-0.003	0.04	0.94
	Non-verbal IQ	**-0.11**	**0.04**	**.006****
	No. of characters	0.01	0.01	0.24
	VSL x RC structure	-0.002	0.01	0.8
	Non-verbal IQ x RC structure	**0.02**	**0.01**	**.008****
HN+1 **region**	Intercept	6.09	0.09	<.001***
	RC structure	-0.0003	0.01	0.99
	VSL	-0.02	0.03	0.62
	Non-verbal IQ	**-0.07**	**0.03**	**.04***
	No. of characters	0.01	0.01	0.23
	VSL x RC structure	0.01	0.01	0.56
	Non-verbal IQ x RC structure	0.003	0.01	0.77

Note: The dependent variable is log-transformed reading time; VSL= visual statistical learning, RC=relative clauses.   *p<.05. **p< .01. ***p<.001.

In the current Chinese sample, reading times for each region of interest were not significantly predicted by the number of characters. As can be seen in Table 5, VSL did not significantly contribute to the models, and the interaction between VSL and RC structure was not significant for the relativizer (*de*), head noun and the word

following head noun region respectively. Non-verbal IQ, however, significantly

predicted RTs for both Chinese SRCs and ORCs at the relativizer (*de*) and the word

following head noun region, as indicated by the main effect of non-verbal IQ. The

negative coefficient shows that higher non-verbal IQ capacity is related to decreased

reading time in those regions.

There was also a significant interaction between non-verbal IQ and RC

structure at the head noun region. The interaction at the head noun suggests the

difference between the processing of SRCs and ORCs varied according to non-

verbal IQ. Figure 3 plots this interaction over the whole sample ($n = 55$) for: a) SRCs

and b) ORCs

a)

b)
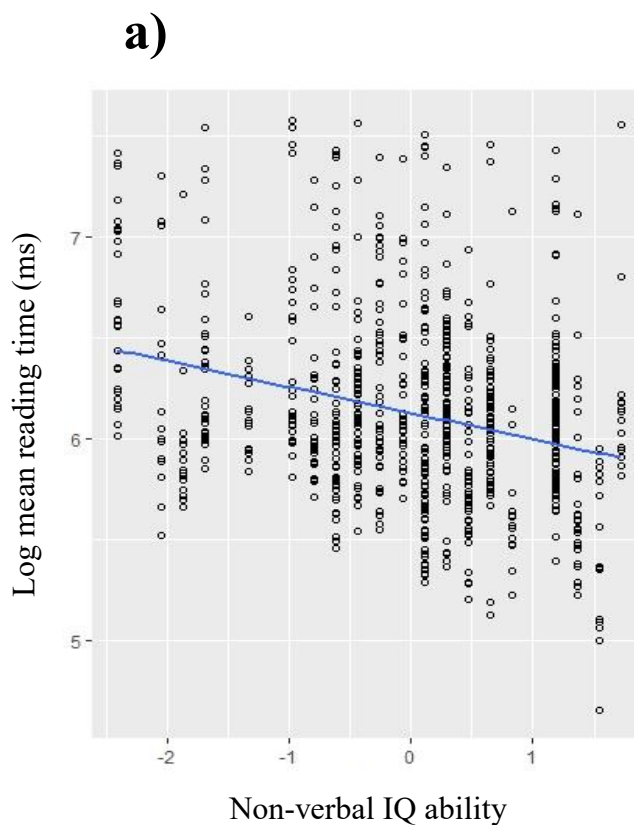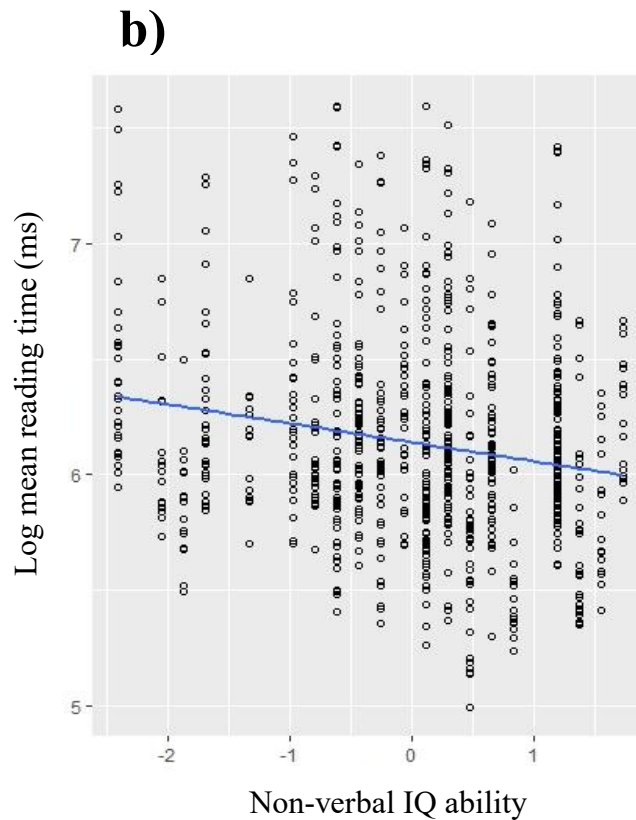
Figure 3. Correlation between the non-verbal IQ ability and RC reading times for:

a) SRCs and b) ORCs, over the whole sample.

As can be seen in Figure 3a) and 3b), it appears that the difference in reading

time between subject and object RCs at the head noun was lower for participants

with higher IQ versus those with low IQ.

To further examine the interaction between non-verbal IQ and the Chinese RC

structure, participants were divided into high IQ group and low IQ group using a

median split. The high IQ group consisted of 25 participants and the low IQ group

consisted of 30 participants. The effect of construction was examined for each group

separately using a linear mixed-effect model. The model included the fixed effects of

RC structure, while the random effect structure included the by participant intercept,

by item intercept and by participant random slopes for RC structure. Table 6 presents

the statistical analyses for the main effect of RC structure by IQ group.

Table 6

Main effect of RC structure by non-verbal IQ group

Non-verbal IQ group	Head noun region	Coefficient	SE	p value
High	Intercept	6.05	0.10	<.001***
	RC structure	0.02	0.02	0.13
Low	Intercept	6.07	0.11	<.001***
	RC structure	-0.005	0.02	0.80

Note: The dependent variable is log-transformed reading time at head noun as the residuals were strongly skewed before log transformation. A negative sign on the estimated coefficient indicates an object preference. RC=relative clauses. ***$p<.001$.

Table 6 shows that there was no significant RC structure effect in either high

and low IQ group, suggesting that SRCs were not processed significantly faster than

ORCs within each IQ group. Thus, although the interaction was significant in the

overall analyses, effects of structure were not evident *within* the groups, which is

likely due to the drop in statistical power associated with dividing the group via a

median split.

The results from the Mandarin self-paced reading task show that non-verbal IQ

was associated with L1 Chinese RCs processing at head noun, but that VSL was not

associated with language comprehension or any of the other tasks. The lack of

association between VSL and general intelligence is consistent with previous studies

showing that SL capacity is independent of intelligence. However, the finding that

there was no VSL-language link is inconsistent with studies that have found an

association between VSL and language (Arciuli & Simpson, 2012; Kidd & Arciuli,

2016). In the next exploratory analysis, I took a different tact. Following past

research on WM and on-line language processing (e.g., Caplan & Waters, 1999), I

examined whether VSL ability, as a dichotomized variable, predicted L1 language

processing.

Participants were divided into high VSL capacity group and low VSL capacity

group based on their VSL task scores. The current sample was distinguished as high

VSL capacity participant group from low VSL capacity participant group using a

median split. High VSL capacity group consisted of 29 participants, while low VSL

capacity group consisted of 26 participants. The data were analysed in a linear mixed

effect model with RC structure and VSL group as fixed effects (code as +1 for high

VSL capacity and -1 for low VSL capacity), and random intercepts for items and

participants.

Following the above analysis procedure, the non-verbal IQ capacity was also

dichotomized into high and low capacity IQ group to further investigate L1 RCs

processing times of the IQ group within each region of interest

Table 7 displays the statistical analyses for the main effect of VSL group, non-verbal

IQ group and Chinese RC structure and their interaction by region of interest.

SL IN SENTENCE PROCESSING

Table 7

The main effect of VSL group, non-verbal IQ group and Chinese RC structure and their interaction by region of interest.

Region		Coefficient	SE	p value
Relativizer (*de*)	Intercept	6.02	0.07	<.001***
	RC structure	-0.01	0.01	0.64
	VSL group	-0.01	0.03	0.72
	IQ group	-0.06	0.03	**0.04 ***
	VSL group x RC structure	0.01	0.007	0.13
	IQ group x RC structure	0.003	0.007	0.69
Head Noun (HN)	Intercept	6.05	0.09	<.001***
	RC structure	0.01	0.01	0.66
	VSL group	-.002	0.04	0.97
	IQ group	-0.09	0.04	**0.02 ***
	VSL group x RC structure	-0.004	0.009	0.67
	IQ group x RC structure	0.01	0.01	**0.09#**
HN+1 region	Intercept	6.09	0.09	<.001***
	RC structure	.0002	0.01	0.99
	VSL group	-0.02	0.03	0.51
	IQ group	-0.06	0.03	**0.06#**
	VSL group x RC structure	0.007	0.009	0.42
	IQ group x RC structure	0.004	0.009	0.64

Note: The dependent variable is log-transformed reading time. VSL= Visual statistical learning; RC=relative clauses; IQ= nonverbal IQ.
#p<.10 *p<.05. ***p<.001.

As shown in Table 7, the main effect of VSL capacity group in RTs for the

relativizer (*de*), head noun and the word following head noun region did not reach

significance. Moreover, the interaction between VSL capacity group and RC

structure was not significant for the above 3 regions. In contrast, the non-verbal IQ

group significantly predicted RTs for both Chinses SRCs and ORCs at the relativizer

(de), but was marginal at the word following head noun region, as evidenced by the

main effect of non-verbal IQ group. The negative coefficient indicates that

individuals with higher non-verbal IQ capacity processed the Chinese RC structures

faster overall in those regions. While at the head noun, a marginally significant

interaction between non-verbal IQ group and RC structure was found.

In sum, there was a significant association between non-verbal IQ and online

processing of Chinese RCs.Unlike previous studies that reported a significant

correlation between SL and sentence comprehension (Misyak & Christiansen, 2012;

Kidd & Arciuli, 2016), VSL did not predict RTs for Chinese RCs. The present

findings suggest that VSL may not be a robust predictor of L1 Chinese sentence

processing, given the significant linkage had been reported mainly in alphabetic

languages.

I next present the analyses of the participants' performance on the English RC

comprehension task, and the results of the individual differences analyses.

The role of VSL and language proficiency in L2 sentence processing.

I now turn to the second part of this study to investigate the relationship

between participants' VSL capacity and the English self-paced reading times, and

examine whether second language proficiency mediates the relationship between

VSL and L2 language processing. Participants ($n = 40$) with higher than 80%

accuracy to all comprehension questions in English self-paced reading task were

analysed.

Table 8 shows the descriptive statistics for the bilinguals' performance on VSL

task, vocabulary, self-rated English proficiency, working memory and nonverbal IQ

among forty participants in English sample, whereas the correlations between all

variables are shown in Table 9.

Table 8

Descriptive Statistics for Predictor Variables in English Sample

	M	SD	Range	Skewness	Kurtosis
VSL	0.7	0.15	(0.48,1)	0.36	2.29
Vocabulary	173	21.89	(119,203)	-0.76	2.84
Eng.Prof.Rating	25.7	3.84	(14,35)	-0.06	1.84
WM	45.9	27.02	(17,110)	1.06	3.14
Nonverbal IQ	37.73	5.72	(25,48)	-0.39	2.74

Note.VSL: visual statistical learning; WM: working memory; Eng.Prof.Rating :English
proficiency rating
VSL task: min score=0, max score=1
Vocabulary: min score=0, max score=228
English proficiency rating: min score=0, max score=35
WM task: min score= 0, max score=120.
Nonverbal IQ: min score= 0, max score= 48

Table 9

Simple Bivariate Pearson Correlations between Predictor Variables in English
Sample

	VSL	WM	Nonverbal IQ	Vocabulary
Nonverbal IQ	-.274	.083		
WM	.037			
Vocabulary	-.227	-.218	.180	
Eng.Pro.Rating	-.077	-.089	-.006	.308

Note. VSL= visual statistical learning; WM: working memory; Eng.Prof.Rating :English
proficiency rating

None of the predictor variables were significantly correlated. Group

performance on the VSL task was significant above chance (one-sample t test tested

against 0.5 chance performance, $t(39) = 8.15$, $p < .001$, Cohen's $d = 1.29$). The

performance of VSL task was not associated with nonverbal IQ and working

memory. This is consistent with the prior research in L2 acquisition (Frost et al.,

2013; Brooks et al., 2017). As mentioned earlier, the reliability of the WM scores is

in doubt since the particularly low performance of the participants on the WM task

(mean below 40%). Therefore, performance on the WM task was eliminated from

further analyses.

Studies of English RC processing typically analyse differences in reading times

between SRCs and ORCs at the main verb (e.g., Misyak, Christiansen, & Tomblin,

2010; King & Just, 1991). However, previous studies have also observed differences

within the RC (e.g. L2 English learners: Wang, Ma, Wang, Troyer, & Li, 2015; L1

native English speakers: McCauley & Christiansen, 2015). Therefore, RTs to regions

both within the RC and at the main verb were analyzed.

The reading times for the mentioned regions of English RCs were log-

transformed since the residuals were highly skewed. The data were modelled using

an identical process to that described in the analysis of the Chinese data. Five

predictor variables, VSL, vocabulary, English proficiency rating, non-verbal IQ and

number of character of the whole RC sentence, were entered as fixed effects. All

variables were converted to z-score to reduce the collinearity. The final models for

the mentioned regions of English RCs are shown in Table 10.

SL IN SENTENCE PROCESSING

Table 10

Main effects of VSL, Vocabulary, English Proficiency rating, non-verbal IQ and no. of characters by region of interest in English sample

Region		Coefficient	SE	p value
Clause-internal region	Intercept	6.03	0.10	<.001***
Word 3	RC structure	0.01	0.02	0.50
	VSL	-0.03	0.06	0.60
	Vocab	-0.08	0.06	0.22
	Nonverbal IQ	-0.06	0.06	0.33
	Eng.Prof.R	-0.05	0.06	0.44
	VSL x RC structure	0.003	0.01	0.77
	Vocab x RC structure	-0.02	0.01	0.12
	Nonverbal IQ x RC structure	-0.001	0.01	0.90
	Eng.Pro.R x RC structure	-0.01	0.01	0.58
	No.of.characters	0.04	0.01	**<.001***
Clause-internal region	Intercept	6.12	0.08	<.001***
Word 4	RC structure	0.05	0.02	**.003**
	VSL	-0.02	0.06	0.79
	Vocab	-0.06	0.06	0.35
	Nonverbal IQ	-0.05	0.06	0.40
	Eng.Prof.R	-0.03	0.06	0.63
	VSL x RC structure	-0.01	0.01	0.31
	Vocab x RC structure	0.02	0.01	0.11
	Nonverbal IQ x RC structure	0.02	0.01	0.15
	Eng.Pro.R x RC structure	0.02	0.01	0.17
	No.of.characters	0.04	0.01	**<.001***
Main Verb	Intercept	6.19	0.07	<.001***
	RC structure	0.02	0.01	0.25
	VSL	-0.05	0.05	0.36
	Vocab	-0.05	0.05	0.34
	Nonverbal IQ	-0.05	0.05	0.33
	Eng.Prof.R	-0.06	0.05	0.26
	VSL x RC structure	0.005	0.01	0.62

SL IN SENTENCE PROCESSING

Region		Coefficient	SE	p value
Main	Vocab x RC structure	0.002	0.01	0.81
Verb	Nonverbal IQ x RC structure	0.01	0.01	0.25
	Eng.Pro.R x RC structure	0.001	0.01	0.89
	No.of.characters	0.03	0.01	**<.001***

Note: The dependent variable is log-transformed reading time. RC=relative clauses.

Eng.Prof.Rating :English proficiency rating

**p<.01. ***p<.001.

None of the individual difference predictor variables significantly predicted the

reading time in the main verb region and the the relative clause-internal regions.

Only the number of characters significantly predicted the reading time at word 3,

word 4 and the main verb. However, a significant main effect of RC structure was

observed at word 4, in which a positive sign on the estimated coefficient indicated a

subject preference. That is, the participants processed SRCs more quickly in this

region of the sentence. In addition, no significant interaction was found between L2

English RC structure and the predictor variables in the main verb region and the

relative clause-internal regions respectively.

The current findings differ, however, from the past L2 studies that have found

both SL and vocabulary to be predictive of L2 proficiency (Brooks et al., 2017; Frost

et al., 2013; Jeon & Yamashita, 2014). Following the analysis in the previous section

that used a dichotomized VSL variable, we conducted a series of analyses that

investigated the relationship between L2 English RC processing and VSL,

Vocabulary, and English language proficiency, where each predictor variable was

dichotomized into high and low abilities.

Participants were divided into high or low groups for each variable based on a median split. To further investigate the interaction between the four variable groups and English RC structure, the log-transformed reading times for the region of interest were analysed as the residuals were strongly skewed before log transformation. The data were analysed in a linear mixed effect model with RC structure and each of the variable group as fixed effects (code as +1 for high capacity group and -1 for low capacity group), and random intercepts for items and participants. Table 11 displays the statistical analyses for the main effect of the four variable groups and English RC structure, along with their interaction at main verb.

SL IN SENTENCE PROCESSING

Table 11

Main effects of VSL group, Vocab group and English Proficiency rating group and their interaction with English RC structure

Group	Main Verb	Coefficient	SE	p value
VSL group	Intercept	6.19	0.07	<.001***
	RC structure	0.02	0.01	0.25
	VSL group	0.02	0.10	0.82
	VSL group x RC structure	0.02	0.02	0.33
Vocab group	Intercept	6.21	0.08	<.001***
	RC structure	0.02	0.01	0.25
	Vocab group	-0.08	0.11	0.47
	Vocab group x RC structure	0.005	0.02	0.82
Eng.Pro.R group	Intercept	6.22	0.08	<.001***
	RC structure	0.02	0.01	0.25
	Eng.Pro.R group	-0.12	0.10	0.25
	Eng.Pro.R group x RC structure	-0.02	0.02	0.49
Nonverbal IQ group	Intercept	6.20	0.08	<.001***
	RC structure	0.02	0.01	0.25
	Non-verbal IQ group	-0.04	0.13	0.72
	Non-verbal IQ group x RC structure	0.01	0.03	0.59

Note: The dependent variable is log-transformed reading time at Main verb

Eng.Prof.Rating :English proficiency rating

. ***p<.001.

As can be seen in Table 11, there were no differences in how any of the groups processed the English RC structures, as indicated by the non-significant interactions.

Summary

Taken together, the results of the present study are inconsistent with the hypotheses. Firstly, the current measure of VSL did not predict L1 Chinese RCs processing, but non-verbal intelligence was correlated with the processing of the two RC structures at head noun region. Secondly, VSL capacity and L2 proficiency (as

measured by self-rated L2 English skills and English receptive vocabulary

knowledge) failed to predict reading times in L2 English RCs processing.

Although past individual differences studies have demonstrated the correlation

between VSL and linguistic performance (Arciuli & Simpson, 2012; Kidd & Arciuli,

2016), there is a growing awareness that SL as a capacity must be at least partially,

or even substantially, modality -specific (e.g., Conway & Christiansen, 2005, 2009;

Seigelman & Frost, 2015). That is, even is performance on SL tasks is driven by

domain-general computational principles, there must be modality-specificity because

the input for the tasks is undoubtedly limited by constraints particular to individual

sensory systems. As such, it could be that auditory SL is a better candidate to explain

individual differences in syntactic processing. Therefore, in Study 2 I attempted to

further explore the role SL in syntactic processing by using *both* an auditory and a

visual SL task.

Chapter 4

Study 2: The Role of Visual and Auditory SL in
L1 Chinese Sentence Processing

Study 1 yielded few significant results. While it is certainly possible that none

of the individual differences variables I measured are related to syntactic processing

in L1 Mandarin and L2 English speakers, this is highly unlikely given past research.

An alternative explanation concerns participant attention: a significant number of the

participants ($n = 5$ for Mandarin, $n = 20$ for English) did not answer greater than

80% of the comprehension questions in the self-paced reading task correctly.

Additionally, the participants' scores on the Listening Span task were very low.

Given the ambiguity in the results, I decided to conduct an additional study with a

different set of participants. In this study, I concentrated on the role of SL in L1

processing of Mandarin SRCs and ORCs. As in Study 1, I used a visual SL (VSL)

task, but in addition measured auditory SL (ASL) using a newly developed task

(Isbilen, McCauley, Kidd, & Christiansen, 2017).

SL across modalities and language

Although a growing number of studies have reported that individual differences

in SL capacity predicted linguistic abilities across auditory and visual modalities

(e.g., Conway, Bauernschmidt, Huang, & Pisoni, 2010; Kidd & Arciuli, 2016; Frost

SL IN SENTENCE PROCESSING

et al., 2013), SL capacity has been found to be constrained by modality (Conway &

Christiansen, 2005, 2009). Conway and Christiansen (2005) used non-linguistic

sequential input to compare the SL performance across (i) visual, (ii) auditory and

(iii) haptic modalities. The result showed that the auditory modality exhibited better

statistical learning when compared with visual and touch modalities, based on

overall learning of sequences. The authors further investigated learners' sensitivity to

the beginnings or final portion of the sequences in each modality, and found that

learners in auditory modality were more sensitive to the final portion of input

sequences when compared with visual and haptic stimuli, controlling for perceptual

and training effects. The authors concluded that the existence of modality constraints

affect SL across the senses. This raises something of a conundrum: past studies have

shown cross-modal correlations between SL and language (e.g., VSL predicting

auditory language, Conway et al., 2010; Kidd, 2012; Kidd & Arciuli, 2016), but it

appears that SL is at least partially modality-specific. This is one potential

explanation for the lack of association found in Study 1: if SL is partially modality-

specific then the effect may be difficult to capture (especially with less than optimal

participant attention). For this reason, I measured ASL in addition to VSL in Study 2.

Thus, the purposes of Study 2 were to test whether ASL capacity may also play a

role in on-line sentence processing, and whether VSL capacity also predicts the

performance, even when other cognitive variables are considered. To make data

collection more manageable within the constraints of a MPhil thesis I only tested L1

Mandarin speakers on-line processing of Mandarin RCs.

Method

Participants

Fifty-two (34 females, 18 males; age: *M* = 21.69 years, *range*: 17-34) native

Mandarin-speaking undergraduate and postgraduate students were recruited via the

Psychology Research Participation Scheme at the Australian National University. All

participants provided their informed consent and participated for course credit or

received $20 as compensation for their time. The participants were all native

speakers of Mandarin, with no reported history of uncorrected visual, auditory or

neurological impairments. The data of participants with less than 80% accuracy on

comprehension questions were removed from the analysis. Based on the inclusion

criteria, 41participants were selected to be valid in Study 2 (13 males, 28 females;

age: *M* = 21.4 years, *SD* = 3.2, *range* = 17-34).

Materials and Procedures

The same VSL task, non-verbal IQ task, and verbal WM tasks from Study 1

were used in Study 2. An ASL task was also used to measure participants' auditory

SL capacity. Study 2 also used different items for the Mandarin self-paced reading

task. Specifically, I used the materials that, unlike in the materials in Study 1,

controlled for the syntactic category in post-head noun region. Some studies have

found complexity effects in this region (Vasishth et al., 2013), but it could not be

analysed in Study 1 because the materials were not tightly controlled in this region.

The use of the new materials therefore provided more controlled linguistic stimuli,

enabling additional analyses to be conducted in the post-head noun region. Each task

is discussed in turn.

Chinese Sentence Comprehension Task

As in Study 1, the Chinese stimulus set had 160 test sentences, comprised of 64

Relative Clause (RC) sentences (32 SRCs and 32 ORCs) and 96 filler sentences in

total across eight lists.

Modified Mandarin Self-paced Reading Task

Thirty-two Chinese single-embedded RC sentence pairs were used, with each

pair containing a SRC and ORC version of the sentence.

Twenty-four pairs of Chinese single embedded RCs were taken from Vasishth et

al.'s (2013) Experiment 2, and the remaining 8 pairs were adapted from Chen, Ning,

Bi and Dunlap (2008). All noun phrases were animate, since past studies have shown

that animacy moderates the complexity associated with ORCs in Mandarin (e.g., Wu,

Kaiser, & Andersen, 2012). The following are examples for the two types of Chinese

RCs.

(22a) Subject-extracted RC

检举　　厂商　　的　　市民　　逼疯了　　官员　　还　　得意洋洋。

[Jiǎnjǔ　chǎngshāng de]　shìmín　bīfēngle　guānyuán　hái　déyìyángyáng.

[___*i* prosecuted the manufacturer that] **citizens***i*　drove mad　the officials　still triumphant

V1　　　　N1　　　de　　Head N　　Main V　　NP3　　Adv.　　Adj

(22b) Object-extracted RC

店员　　　打昏　　的　　歹徒　　见到了　　记者　　并　　立即　　报警。

[Diànyuán　dǎ hūn　de]　dǎitú　jiàndàole　jìzhě　bìng　lìjí　bàojǐng

[shop assistant knocked out ___*i* that]　**ruffian***i*　saw the reporter and immediately called the police

N1　　　　　V1　　　de　　Head N　　Main V NP　conj.　Adv.　VP3

In sentence (22a), the head noun *citizens* occupies the subject position within the

RC, as denoted by the underscore gap. In contrast, in sentence (22b), the head noun

ruffian occupies the object position. The test sentences varied in length, from 15 to

20 characters (mean length = 17 characters), while the length of the RCs was

composed of five to six characters. All the target sentences are listed in the

Appendix.

As in Study 1, the Chinese RCs were interspersed within 96 filler sentences,

which contained three different types of structures: (i) Chinese pivotal sentences, (ii)

passive sentences ('BEI' structure with an agent), and (iii) declarative sentence.

There were 32 tokens of each filler type. All target structures and fillers were

presented in simplified characters. After reading the complete sentence on a

computer screen, participants were required to answer a Yes/No comprehension

question. To enhance participants' engagement for the self-paced reading task,

a comprehension question was asked for every RC test sentence. As such, sixty-four

yes-no comprehension questions in total were generated for Chinese RC sentences,

in which half of the comprehension questions asked about the main clause and the

other half asked about the relative clause. For each category of filler sentences, eight

comprehension questions were generated.

Procedure

The same self-paced moving window method from Study 1 was used (Just,

Carpenter & Woolley, 1982) to present the test sentences via E-prime (Schneider,

Eschman & Zuccolotto, 2002). Eight experimental lists were created for the Chinese

self-paced reading tasks as in Study 1. Each trial incorporated six practice sentences

to allow participants to become familiar with the self-paced presentation format of

the experiment. Every four sentences, on average, contained either ORC or SRC

sentence and one of each type of filler sentence. Thus, participants were individually

tested on 32 RCs which consisted of 16 SRCs and 16 ORCs, coupled with 96 fillers

in total. In each list, yes-no comprehension questions were approximately asked after

every fourth sentence to ensure that participants understood the test sentences. Each

participant received a total of 56 comprehension questions, in which consisted of 32

for the target structures and 24 for the filler sentences.

The procedure of each trial was identical to Study 1. Reaction times in

milliseconds for every word segment and the number of correct answers to the

comprehension questions were recorded.

ASL Task

Participants' auditory SL capacity was measured using the novel ASL task

developed by Isbilen, McCauley, Kidd and Christiansen (2017). The task was chosen

because it has high test-retest reliability (rs > .7).

The input stimuli consisted of 18 syllables (*ba, bi, bu, da, di,du,ga,*

gi,gu,la,li,lu,pa,pi,pu,ta,ti, tu), which were used to create six trisyllabic nonce words

(*latugi, piduba,dipapu,tagalu,bubida,guliti*). The six nonce words were then used to

create a continuous auditory input sequence that consisted of 72 randomized blocks

of the six words (i.e., each word appeared 72 times, with each word appearing, on

average, once every six words). The MBROLA speech synthesizing software (Dutoit

et al., 1996) was used to generate the input sequence. Each syllable was

approximately 200ms long, with a 75 milliseconds(ms) mute pause between

syllables. The input stream followed a statistical distribution, where words had high

transitional probabilities within syllables (approximately 1.0), but low transitional

probabilities at word boundaries (TP = 0.17). This followed classic studies of ASL

(e.g., Saffran, Aslin, & Newport, 1996). SL was measured in two manners: (i) using

a two-alternative forced choice (2AFC) task, and (ii) a novel chunking task, called

the *Statistically Induced Chunking* Recall (SICR).

For the 2AFC component of the task, six extra foil nonwords were created

(*tabudu, ligibi, paditi ,bapula, lugada, pigutu*), with different transitional

probabilities (TPs) from the target input non-words (TPs = 0). The SICR component

of the task required participants to recall either (i) two words from the training

sequence, which were presented as a six-syllable chunk (e.g., *latugipiduba*), or (ii)

randomly assembled sequences of six syllables (all syllables coming from the

training set). There were 12 experimental items and their corresponding foil items in

the SICR stimuli, which gave a total of 24 six-syllables items. The logic of the SICR

paradigm is that chunking is likely to play a vital role in on-line processing, since a

listener must hold sequences of language in verbatim memory to analyze its form

(Christiansen & Chater, 2016).

Procedure

The ASL began with a familiarization phase, accompanied by a cover task in

which participants had to detect repeated syllables (as in Arciuli & Simpson, 2012).

The purpose of the cover task was to prevent participants from strategically

remembering syllable sequences, which aims to ensure that the linguistic knowledge

that they learn is implicit. Repeated syllables could occur in any of the three

positions in a target word (e.g., *latugi* → *lalatugi , latutugi, latugigi*). Participants

were required to press the space bar as soon as they heard a repeated syllable. Each

of the three variants of the target input nonwords occurred 4 times. The

familiarization stream lasted approximately 11 minutes.

Following training, participants' knowledge of the non-words was assessed via

(i) a series of 2AFC trials, and (ii) the SICR task. In each of the 2AFC trials,

participants listened to one trisyllabic target word heard in training and one foil,

which consisted of three syllables in the training set, but which never occurred

adjacently (i.e., the TPs = 0). The order of presentation of the target and foil word

was counterbalanced. Participants were required to identify which of the two 3-

syllable nonword that had previously been presented in the familiarization stream.

There were 36 2AFC trials in total.

In the SICR task, participants were asked to repeat the syllables in an audio

recording, which consisted of 24 six-syllable items. After listening to each item,

participants were required to repeat the entire string of syllables following 500 msc

delay (as marked by a high-pitched tone). Participants' syllable-by-syllable

production was recorded. In scoring the SICR component, participants received 1

mark for each accurately recalled statistically chunked syllable and random syllable

respectively, whereas no mark was given for the incorrect responses. Participants'

scores then were computed to generate a proportion correct score respectively (i.e.,

number of correctly recalled syllables divided by total number of syllables).

WM, Non-Verbal IQ, VSL

The same verbal WM task, non-verbal IQ test and VSL task from Study 1 were

used. The procedures of the above tasks were identical to Study 1.

Results

Table 12 shows the descriptive statistics for VSL task, ASL (2AFC), ASL

chunked items, Non-verbal IQ and WM performance among 41 participants. The

simple bivariate correlations between the measures are shown in Table 13.

Table 12

Descriptive Statistics for Predictor Variables in Chinese syntactic processing

	M	SD	Range	Skewness	Kurtosis
VSL	0.67	0.16	(0.36,0.95)	0.02	-0.97
ASL 2AFC	0.57	0.10	(0.33,0.83)	0.36	0.83
ASL chunked items	0.5	0.12	(0.19,0.71)	-0.41	-0.24
WM	92.76	11.56	(74,118)	0.06	-0.93
Nonverbal IQ	38.63	5.55	(21, 47)	-1.29	2.57

Note: VSL: visual statistical learning; WM: working memory
ASL (2AFC): audio statistical learning (two-alternative forced choice task);
VSL task: min score= 0, max score=1.
ASL (2AFC): min score=0, max score=1
ASL chunked item: min score =0, max score=1
WM task: min score= 0, max score=120.
Nonverbal IQ: min score= 0, max score= 4

SL IN SENTENCE PROCESSING

Table 13

Simple bivariate Pearson Correlation Between Predictor Variables in Chinese syntactic processing

	ASL chunked items	WM	IQ	VSL
WM	-.237			
IQ	.204	-.015		
VSL	.208	-.217	-.179	
ASL(2AFC)	.002	-.088	.079	-.091

Note: VSL: visual statistical learning; WM: working memory

ASL (2AFC): audio statistical learning (two-alternative forced choice task);

*p<.05 **p<.01 (two-tailed test)

Group performance on the VSL task was significantly above chance (one-sample t test tested against 0.5 chance performance, t (40) = 6.85, p < .001, *Cohen's d* = 1.07). The ASL (2FAC) task performance was significantly greater than chance, [t (40) = 4.28, p < .001, *Cohen's d* = 0.67]. The overall mean SICR accuracy rate of correctly recalling the statistically chunked items and random items was 50% (*M* = .50, *SD* = .11). While the accuracy differences between the statistically chunked items (*M* = .50, *SD* = .12) and the random items (*M* = 0.49, *SD* = .10) did not reach significance (t (40) = 0.67, p > 0.1). The correlation between 2AFC and SICR measures (i.e. chunked items) in the ASL task was not significant. These data are consistent with Isbilen et al. (2017, Experiment 1), suggesting that the two mechanisms tap into different components of SL. None of the predictor variables were significantly correlated.

Figure 4 demonstrates the mean raw reading time by sentence region of Chinese relative clause types. The mean reading times that beyond 2000ms were excluded

from the analysis, as per Study 1.
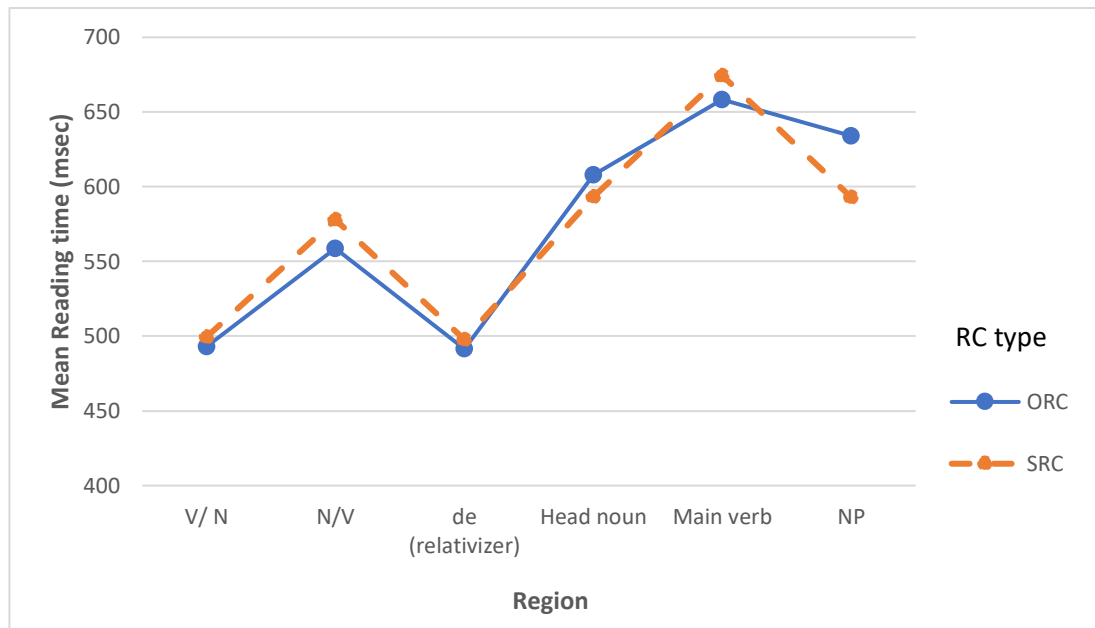
Figure 4. Mean raw reading times of each region of interest in Chinese RCs.

I first analysed the group performance on the test structures. The pattern of

reading times in the first five words in Figure 4 was comparable to those observed in

the past Chinese RC processing studies, where greatest mean reading times were

observed at the head noun and the word following head noun regions (e.g., Gibson &

Wu, 2013; Lin & Bever, 2006; Vasishth, Chen, Li & Guo, 2013). Reading times in

the following regions were examined: the third word (*de*), the fourth word (head

noun), the fifth word (main verb), and the sixth word (Noun phrase). RTs were log

transformed as the residuals were strongly skewed. The data were analysed in a

linear mixed effect model with RC structure as fixed effect. The procedure to

determine the random effects structure here was identical to Study 1. Thus, the

SL IN SENTENCE PROCESSING

random-effect structure included a by-participants random intercept, by-items

random intercept and by-participants random slopes for RC structure. The statistical

analyses for the mentioned regions are shown in Table 14.

Table 14

Main effect of RC structure by region of interest

Region		Coefficient	SE	P value
de (relativizer)	Intercept	6.29	0.20	<.001***
	RC structure	-0.0085	0.01	0.57
Head noun	Intercept	6.55	0.25	<.001***
	RC structure	0.02	0.02	0.41
Main verb	Intercept	6.34	0.27	<.001***
	RC structure	-0.0017	0.02	0.93
Noun phrase	Intercept	6.42	0.20	<.001***
(NP)	RC structure	0.02	0.01	0.09[#]

Note: A negative sign on the estimated coefficient indicates an object preference. [#]$p < .10$

Table 14 shows that there were no significant differences in the processing of

SRCs and ORCs across the regions of interest. The only result that approached

significance was the analysis of the NP region following the main verb ($p = .09$).

Interestingly, these null effects were observed despite the use of materials that have

shown previous differences across structures (Vasishth et al., 2013). However, it

should be acknowledged that others have also found differing results in Mandarin

using the same materials, which attests to the elusive nature of the structural effect in

Chinese, a point I take up in the General Discussion.

Regardless of the fact that no overall effect was found for structure in the group

analyses, this does not preclude the possibility of differences in RC processing being

attributable to individual differences. In fact, the null effect may obscure the

presence of individual differences. Therefore, I investigated the relationship between

participants' performance on SL tasks and L1 Chinese SRCs and ORCs reading

times (RTs). Five variables: VSL, ASL (2AFC), ASL SICR performance, non-verbal

IQ and working memory were examined using the linear mixed-effects modeling

with crossed random effects (Baayen, Davidson, & Bates, 2008) in R package *lme4*

(Bates, Maechler & Bolker, 2012).

All predictor variables were converted to z-score to reduce the collinearity and

make the coefficients more interpretable. SRCs were coded as -1 and ORCs as +1.

Residuals of linear mixed effects models were inspected to avoid large deviations

from the normality. Random slopes and intercepts were included to explain the

variance of random effects for participants and items. Models with the maximal

random effects structures were fit and then compared to simplified models using AIC

(Barr et al., 2013). The model with the smallest AIC value (i.e., the best fitting

model) was selected. Each of the five predictors entered as main effect and the

interaction between each predictor and RC structure were analyzed.

In the present study, ASL(2AFC), Non-verbal IQ and working memory did not

significantly contribute to the models for either RC structure, and no interactions

between the above variables and RC structure were found. In contrast, performance

on the VSL and the SICR component of ASL task did contribute to the models. The

final models for each region of interest are shown in Table 15. Note that WM,

nonverbal IQ, ASL 2AFC, and SICR foils repetition did not contribute to any of the

models at any point in the sentence; therefore, only those variables that significantly

or marginally significantly contribute to explaining variance in RTs were kept in the

model (i.e., VSL and ASL chunked item recall).

SL IN SENTENCE PROCESSING

Table 15

Main effects of VSL, ASL, RC structure and their interaction by region of interest.

Region		Coefficient	SE	p value
Relativizer (*de*)	Intercept	6.29	0.20	<.001***
	RC structure	-0.01	0.02	0.57
	VSL	**-0.07**	**0.04**	**.06+**
	VSL x RC structure	-0.01	0.01	0.21
	ASL chunked items	0.01	0.04	0.81
	ASL chunked items x RC structure	-0.01	0.01	0.59
Head noun	Intercept	6.55	0.25	<.001***
	RC structure	0.02	0.02	0.41
	VSL	**-0.10**	**0.06**	**.08+**
	VSL x RC structure	-0.01	0.01	0.39
	ASL chunked items	-0.03	0.06	0.66
	ASL chunked items x RC structure	**-0.03**	**0.01**	**.04***
Main verb	Intercept	6.34	0.27	<.001***
	RC structure	-0.002	0.02	0.91
	VSL	**-0.07**	**0.04**	**.08+**
	VSL x RC structure	-0.0009	0.01	0.94
	ASL chunked items	0.0002	0.04	0.10
	ASL chunked items x RC structure	-0.004	0.01	0.75
Noun phrase (NP)	Intercept	6.43	0.20	<.001***
	RC structure	0.02	0.01	**0.09+**
	VSL	-0.05	0.04	0.20
	VSL x RC structure	0.01	0.01	0.30
	ASL chunked items	0.02	0.04	0.64
	ASL chunked items x RC structure	0.004	0.01	0.78

Note: The dependent variable is log-transformed reading time; VSL= visual statistical learning, ASL=auditory statistical learning, RC=relative clauses.

+p <.10. *p<.05. ***p<.001.

As can be seen in Table 15, VSL was marginally significant in predicting RTs for both Chinese SRCs and ORCs at the relativizer (*de*), head noun and main verb. The negative coefficient shows that higher VSL capacity is related to faster reading time for the above three regions. The interaction between VSL and RC structure was not significant for the relativizer; head noun and main verb.

At the head noun region, no significant main effect of ASL chunked items effect was observed in RTs, but a significant interaction between ASL chunked items and RC structure was found.

To further investigate the interaction between ASL chunked items and Chinese RC structure that shown in Table 9, participants were divided into good chunking ability group and poor chunking ability group using a median split. The good chunkers group consisted of 21 participants and the poor chunkers group consisted of 20 participants. Mean reading times at head noun across Chinese SRCs and ORCs for individual measured between good and poor chunking ability participants were shown in Figure 5.
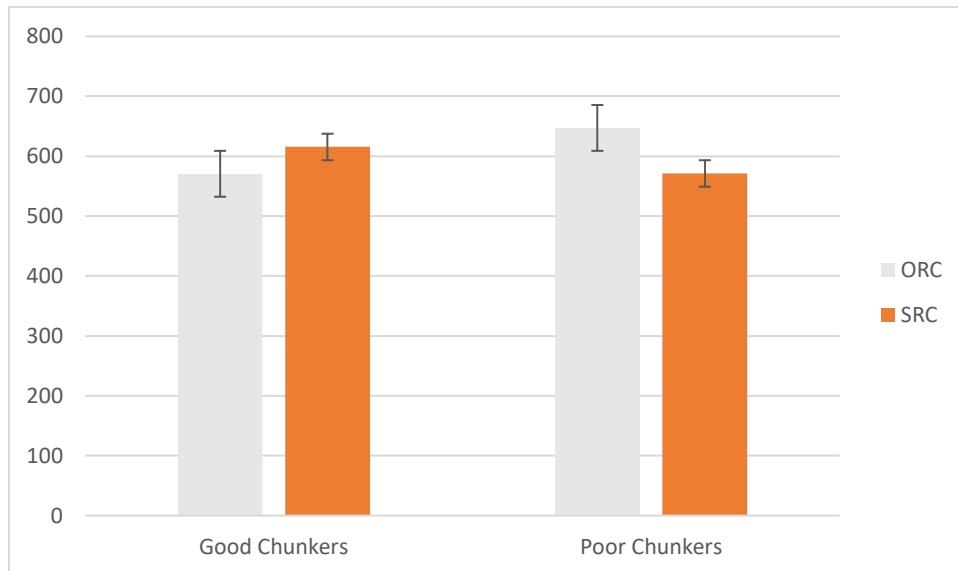
Figure 5. Mean reading times at head noun across Chinese SRCs and ORCs for good chunkers and poor chunkers

The difference in reading time for the head noun between SRCs and ORCs was greater for poor chunkers than for good chunkers, as indicated by the significant interaction in linear mixed-effects model. This finding is in line with other studies, which have shown that poor chunk sensitivity individuals (McCauley & Christiansen, 2015), poor statistical learners (Misyak, Christiansen, & Tomblin, 2010), and less experienced readers (Wells, Christiansen, Race, Acheson, & MacDonald, 2009) reported greater differences between SRCs and ORCs processing when compare with high-achieving readers.

To further explore the effect of RC structure in poor chunkers and good chunkers, the data at head noun was divided into good chunkers and poor chunkers groups, and were analyzed using a linear mixed-effect model. Table 16 demonstrates the statistical analyses for the main effect of RC structure by chunkers group.

Table 16

Main effect of RC structure by chunkers group

Chunkers group	Head noun region	Coefficient	SE	p value
Good	Intercept	6.65	0.31	<.001***
	RC structure	-0.03	0.02	0.23
Poor	Intercept	6.58	0.29	<.001***
	RC structure	**0.06**	**0.02**	**.007****

Note: The dependent variable is log-transformed reading time at head noun as the residuals were strongly skewed before log transformation. The data were fitted in a linear mixed effect model with RC structure as fixed effect; for random effects, including by participant random slopes for RC structure, random intercepts for items and participants. A negative sign on the estimated coefficient indicates an object preference. RC=relative clauses.
*p< .05. **p<.01. ***p<.001.

Table 16 shows a significant main effect of RC structure in the poor chunkers group. The positive coefficient indicates that participants with poor chunking ability exhibited a significant subject-relative preference at head noun (coefficient: 0.06, *SE* = 0.02, *p* < .01), which implicated a significantly faster mean RT of SRCs at the head noun in poor chunkers group compared to ORCs (571.1ms, *SE* = 21.48ms for SRCs, and 647.12ms, *SE* = 25.27ms for ORCs).

Notably, there was no significant RC structure effect in good chunking ability group, which indicated that good chunkers show similarities in processing the two RC structures at head noun region. Compared to poor chunkers, good chunkers encountered fewer difficulty in parsing ORCs, as reflected by lower RTs at the head noun. This result is in line with McCauley and Christiansen (2015), who found a numerical object-preference in English-speaking participants with good chunking ability.

Summary

Taken together, the results of Study 2 support the hypothesis that individual

differences in SL capacity are related to on-line sentence processing. There were two

results of interest. Firstly, VSL was marginally related to processing in general, with

better scores associated with shorter processing times overall. Secondly, SICR ability

was associated with individual differences in the processing of SRCs and ORCs:

individuals with low chunking ability showed a SRC advantage, whereas those with

high chunking ability showed no difference between the two structures.

The implications of the above results would be detailed in Chapter 5: *General*

Discussion.

Chapter 5

General Discussion

The current study investigated whether there are associations between SL and

on-line sentence processing in bilingual speakers who speak Chinese as their first

language (L1) and English as their second language (L2), and sought to examine

whether second language proficiency mediates the relationship between VSL and L2

language processing. To this end, two studies were conducted. In Study 1, it was

hypothesized that VSL ability would be independently associated with L1 (Chinese)

and L2 (English) RC processing. Specifically, it was predicted that Chinese-English

bilinguals who had higher VSL scores would show better performance in both L1

and L2 self-paced reading tasks, as reflected by faster reading times, even when

other linguistic and cognitive variables were considered. It was also hypothesized

that L2 proficiency would mediate the relationship between VSL and L2 processing.

However, the results in Study 1 revealed that only nonverbal intelligence predicted

L1 Chinese RCs processing, with VSL capacity and L2 proficiency not predicting L2

English RCs processing.

As discussed earlier, one possibility in explaining the lack of association

between VSL and syntactic processing found in Study 1 is that SL may be partially

modality-specific. Therefore, an ASL task was employed in addition to VSL task in

Study 2, which only investigate the relationship between SL and L1 on-line

processing.

In contrast to Study 1, in Study 2 I observed that individual differences in SL

ability were associated with L1 on-line Chinese sentence processing. The results

were twofold. First, VSL was marginally correlated with Mandarin RC processing in

general, with higher VSL scores associated with shorter processing time as a whole.

Second, ASL measured via a novel chunking task showed that SICR ability

successfully predicted individual differences in processing times for L1 Mandarin

SRCs and ORCs. Notably, individuals with low chunking ability exhibited a SRC

advantage, whereas those with high chunking ability exhibited no difference in

processing the two structures. In this General Discussion I consider these results

detail.

Null effect of VSL and language proficiency in L2 English sentence processing

Unlike previous studies that have revealed SL to be correlated with L2

linguistic performance as measured by L2 exam score (Kaufman et al., 2010), L2

literacy acquisition (Frost, Siegelman, Narkiss, & Afek, 2013) and L2 morphology

acquisition (Brooks & Kempe, 2013; Brooks, Kwoka, & Kempe, 2017; Granena,

2013; McDonough & Trofimovich, 2016), the current measure of VSL did not

predict L2 English RC processing. One possible explanation for the lack of

significant correlations between VSL and L2 English sentence processing is that SL

is partially modality-specific (Conway & Christiansen, 2009; Siegelman et al.,

2017). Notably, Conway and Christiansen (2005) found that learners in auditory

modality exhibited better statistical learning of the final portion of input sequences

when compared with visual and haptic modalities. Therefore, SL in different

modalities might result in differential attention to different points in sequences. If so,

further research should investigate how these differences relate to language.

Alternatively, the null effect of VSL in L2 English sentence processing may be

due to the individual differences in attention on the L2 self-paced reading task, as a

significant number of the participants ($n = 20$ for English) did not answer greater

than 80% of the comprehension questions in the self-paced reading task correctly.

Additionally, another factor that could possibly account for the low comprehension

scores is the allocation of the comprehension questions in the L2 self-paced reading

task. Only a subset of the test sentences was paired with a corresponding

comprehension question. Thus, it is possible that participants may naturally pay less

attention if they are aware that their understanding is not always tested. Along these

lines, Roberts (2012) has suggested that employing an additional grammaticality

judgment task maintains participants' attention, and could perhaps be used in future

studies. Using grammaticality judgment may also serve as a stronger test of syntactic

processing. For instance, Indefrey (2006) reported that stronger Broca's area

activation in L2 sentence processing when late L2 bilinguals were required to make

additional grammatical judgment on the experimental sentences. Overall, the

suggestion here is that the current L2 self-paced reading task may not fully tax

participants' syntactic knowledge in addition to not engaging participants as much as

it could have. Accordingly, if SL is partially modality-specific, together with the low

participants' engagement in the L2 reading task, the effect of VSL in L2 sentence

processing might be difficult to capture.

Although several empirical studies have revealed that SL predicted the learning

of particular L2 grammatical structure in terms of accuracy measure (e.g., Brooks &

Kempe, 2013; Brooks, Kwoka, & Kempe, 2017; Granena, 2013; McDonough &

Trofimovich, 2016), there has been no published evidence directly showing that SL

is implicated in the online L2 sentence processing. Thus, the present findings may

suggest that differences in detecting statistical distributions in the visual domain are

not related to the initial stages of L2 syntactic processing. One possibility is that L2

learners may rely more on semantic cues than syntactic information when processing

L2 sentences (Clahsen & Felser, 2006; Ullman, 2001). Conway et al. (2010) argued

that the SL-language relationship they observed in L1 speakers was due to the

common encoding of predictability in sequences required across their two tasks.

Specifically, they suggested that superior SL abilities lead to more robust

representations of word order probabilities in spoken language. Similarly, it is

possible that L2 learners in the present study were less sensitive to the word order

probabilities in their L2.

Relations between SL and L1 Chinese sentence processing

The hypothesis that SL would be associated with L1 syntactic processing was

supported in Study 2. That is, individuals' visual and ASL ability was independently

associated with L1 on-line Chinese RC sentences processing. The current data are in

line with the previous individual differences studies showing a correlation between

SL and L1 sentence processing (e.g., McCauley & Christiansen, 2015; Misyak,

Christiansen, & Tomblin, 2010), and extend those data by demonstrating the SL-

language link in a non-alphabetic language (i.e., Chinese). Several implications can

be made from the results of Study 2.

Firstly, VSL was marginally significant in predicting reading times for both

Chinese SRCs and ORCs at the relativizer (de), head noun and main verb, with

higher VSL scores associated with shorter processing times overall. Given that VSL

ability did not interact with the RC structure, the link between the higher ability to

detect transitional probabilities of a stream of visual shapes and the faster reading

time performance in general may be due to individual variation in perceptual

processing. The success on the current VSL task was based on participants' ability to

encode the visual shapes and retrieve the transitional probabilities within the shapes

from the continuous input stream. As such, SL performance is indexed by the

efficiency of processing the visual stimuli for encoding them into internal

representations and learning their statistical distributions (e.g., Frost et al., 2015;

Bogaerts, Siegelman, & Frost, 2016).

The efficiency of decoding visual stimuli is also the determinant of general

perceptual fluency in reading. LaBerge and Samuel (1974) proposed a model of

automaticity in reading in which visual stimuli are transformed into meanings via a

sequence of processing stages. At the first stage of word decoding, graphemic

information is analyzed by feature detectors, which in turn convert into visual codes.

The learning of visual codes in reading incorporates the perception of letters,

spelling patterns, words and word groups (i.e., equivalent to the perception of visual-

orthographic structure, morphemes, characters and words in Mandarin Chinese).

Through repeated exposure to print, readers could be able to decode and recognize

words automatically as a linguistic unit for comprehension with increasing activation

rate of visual codes, which in turn lower the processing times in reading. Therefore,

the results of Study 2 suggest that VSL skill may tap into general fluency in L1

Chinese sentences reading.

Secondly, ASL as assessed via a novel chunking task (Isbilen et al., 2017)

significantly predicted individual differences in processing times for L1 Chinese

SRCs and ORCs at the head noun, as indicated by the significant interaction between

SICR ability and RC structure. The SICR task used in the current study tested

participants' chunking ability. To deal with the limited auditory memory capacity

(e.g., Miller, 1956; Ericsson, Chase, & Faloon, 1980; Saults & Cowan, 2007) and the

short exposure of input stimuli during processing, participants must rapidly chunk

the co-occurring syllables in the incoming sequence into groups of items/words, so

that they could facilitate the recall of chunked items/words at a later stage of the

SICR task. In much the same way as on-line language processing, language learners

are required to rapidly chunk linguistic units and pass these representations to higher

levels of representation to overcome the Now-or-Never Bottleneck (Christiansen &

Chater, 2016). Consistent with this approach, the results of Study 2 showed that the

low-level chunking of syllable sequences predicted on-line processing at higher

levels (i.e. L1 Mandarin RC sentences processing). This finding provides evidence to

support the past studies showing that chunking ability shapes on-line sentence

processing (e.g., McCauley & Christiansen, 2015; McCauley, Isbilen, &

Christiansen, 2017). Specifically, the current result demonstrated that individuals

with low chunking ability exhibited a SRC advantage, whereas those with high

chunking ability exhibited no difference in processing the two structures.

The present findings thus provide insight for understanding the source of

subject-object asymmetry in Chinese RC processing. From the group-level

performance perspective, past studies focusing on the structure effect in Chinese

single embedded RCs have reported mixed results across two dimensions: (i) the

direction of the processing preference, and (ii) the location in the sentence where

participants experience processing difficulty. Specifically, an ORC preference in

Mandarin has been found in self-paced reading studies (e.g., Chen, Ning, Bi, &

Dunlap, 2008; Gibson & Wu, 2013; He, Xu, & Ji, 2017), whereas others have found

a SRC advantage (e.g., Lin & Bever, 2006; Vasishth et al., 2013). With respect to the

critical regions, for instance, a significant SRC preference was found at the

relativizer (de) and the head noun in single-embedded RCs (Lin & Bever, 2006), but

Vasishth et al. (2013) have shown an SRC preference at the word following the head

noun region instead, even though they used the exact same sentence materials.

Similarly, ORC preferences have been found across different regions.

One explanation for the mixed results is that there are significant individual

differences in Chinese RCs processing that are not observed in other languages,

which may contribute to the fact that there have been a lot of mixed results.

Therefore, the current finding of a significant effect of RC structure in poor chunkers

for the head noun is noteworthy, as the lower chunking ability is linked to slower

reading times for the head noun in ORCs than in SRCs. Compared to poor chunkers,

individuals with high chunking ability exhibited faster reading times at the head

noun in ORC. As ORCs are less frequent in Mandarin than SRCs (e.g., Wu, 2009;

Wu, Kaiser, & Andersen, 2011), the current data corroborate Kidd and Arciuli's

(2016) interpretation on the SL-grammar link to structural frequency. They

demonstrated that higher SL ability is associated with more robust knowledge of

infrequent ORCs among English-speaking children.

SL and modality

The fact that no correlation was found between the current visual and ASL

tasks, in conjunction with the relative weak association between VSL and L1

Chinese RC sentences processing when compared to ASL, suggests that SL is

partially modality-specific, where SL capacity cannot be predicted by a single set of

modality-independent computations. This view is consistent with previous studies

supporting the notion that a multi-faceted system of mechanisms underlie

performance on SL tasks (e.g., Conway & Christiansen, 2005, 2006; Siegelman &

Frost, 2015; Siegelman, Bogaerts, Christiansen, & Frost, 2017).

The suggestion that SL is modality-independent appears at odds with the past

studies that have found cross-modal associations between SL and language. In some

sense the data from Study 2 may provide some insight, since both VSL and ASL

were differentially and independently associated with online processing of Mandarin.

It could be that different components of SL in different modalities may contribute to

language. In a recent study, Bogaerts et al. (2016) demonstrated that adults' VSL

performance is determined by the interaction between the efficiency in encoding the

visual shapes (modality-specific) and extracting their transitional probabilities

(modality-general) from the continuous input stream. The suggestion here is that

both modality-specific and the domain-general computational mechanisms constrain

SL in an interactive way. Specifically, they argued that specific distributional

properties of shapes may facilitate their encoding efficacy in a bidirectional manner,

in which these processes conjointly determine participants' actual underlying SL

capabilities. Thus, they hypothesize two important mechanisms in guiding SL:

encoding efficiency and distributional learning. Taking this into account, it may be

that encoding efficiency, in the form of domain-general concepts such as *processing

speed*, explain cross-modal SL-language relationships, whereas more domain-

relevant concepts like distributional learning, which relies on language-specific

representations, operates solely within domains. Such an explanation is consistent

with the results of Study 2, since VSL was only related to processing in general, but

ASL was differentially related to structures that have different distributional

properties in Mandarin.

Limitations and future research

One main limitation of the current research lies in the correlational nature of

experimental design. For instance, although I have interpreted the relationship

between ASL and RC processing in Study 2 to be due to variability in chunking

ability, it is possible that good language ability explains the relationship. Future

research that incorporates longitudinal and/or training studies will be important if

this issue is to be settled.

Another limitation may arise from the nature in which SL is measured in the

tasks that I used. As is typical for many of these tasks, learning is measured

following familiarization using mostly recognition (although this was not the case

for SICR). There are several problems with assessing learning in this way. Firstly,

measuring learning following familiarization does not tap into the *dynamics* of the

learning process. Other tasks, such as serial reaction time tasks, have the capacity to

index learning throughout familiarization, and performance on them has been linked

to language (e.g., Kaufman et al., 2010; Kidd, 2012). Thus, a more dynamic index of

SL may better capture the underlying mechanism that SL and language may share.

Secondly, the explicit nature of the recognition task, which requires participants to

explicitly reflect upon their choice, may not perfectly tap into knowledge acquired

via SL, which is typically considered to be *implicit* (Reber, 1993). Newer tasks are

attempting to capture SL both implicitly and as it happens (e.g., Seigelman,

Bogaerts, Krononfeld & Frost, in press). Future individual differences studies should

consider incorporating these tasks into their test battery.

Conclusion

The results of the current study demonstrate that individual difference in adults'

capacity for SL are associated with on-line L1 Chinese RC sentences processing,

whereas VSL capacity and L2 proficiency did not predict L2 English RC sentences

processing. Specifically, the associations found between SL capacity and on-line L1

Chinese RC sentences processing are suggestive of a complex set of underlying

mechanisms supporting language. Firstly, VSL was marginally related to L1 Chinese

sentence processing in general, with higher VSL scores associated with shorter

processing time overall. Secondly, the SICR ability that assessed via an ASL task

was associated with individual differences in processing of SRCs and ORCs.

Notably, individuals with low chunking ability exhibited a SRC preference, whereas

those with high chunking ability showed no difference between the two structures.

Thus, these findings provide a better understanding of how individual differences in

SL capacity contribute to the subject-object asymmetry in on-line Chinese RCs

processing. Accordingly, the current data suggest that a domain-general and

modality-specific SL capacities may jointly contribute to language.

Appendix A

Mandarin Experimental Sentences of Study 1

1. 妈妈抱着的儿子开心地笑了。
   抱着儿子的妈妈开心地笑了。

2. 兽医治疗的母猫刚生小孩。
   治疗母猫的兽医刚生小孩。

3. 主任批评的教授终于退休了。
   批评教授的主任终于退休了。

4. 家长拜访的老师为人谦虚。
   拜访老师的家长为人谦虚。

5. 牧师赞美的教友笑嘻嘻地走过来。
   赞美教友的牧师笑嘻嘻地走过来。

6. 老板领养的孤儿很有爱心。
   领养孤儿的老板很有爱心。

7. 外公照顾的孙子昨天摔倒了。
   照顾孙子的外公昨天摔倒了。

8. 老师惩罚的学生将参加公听会。
   惩罚学生的老师将参加公听会。

9. 将军率领的士兵接受表扬。
   率领士兵的将军接受表扬。

10. 记者採访的部长突然辞职。
    採访部长的记者突然辞职。

11. 厂长雇用的秘书人缘很好。
    雇用秘书的厂长人缘很好。

12. 女孩崇拜的歌星很有才华。
    崇拜歌星的女孩很有才华。

13. 黑道威胁的律师躲藏在国外。
    威胁律师的黑道躲藏在国外。

14. 警察讨厌的黑客盗用了市民的银行帐户。
    讨厌黑客的警察盗用了市民的银行帐户。

15. 画家欣赏的建筑工人只吃蔬菜和海鲜。
    欣赏画家的建筑工人只吃蔬菜和海鲜。

16. 园丁鄙视的木匠拥有良好的声誉。
    鄙视园丁的木匠拥有良好的声誉。

17. 艺术家崇拜的运动员享用了一顿美味的晚餐。
    崇拜艺术家的运动员享用了一顿美味的晚餐。

18. 助教质疑的学生很不高兴所以四处投诉。
    质疑助教的学生很不高兴所以四处投诉。

19. 老板信任的工程师工作很认真效率又高。
    信任老板的工程师工作很认真效率又高。

20. 教授认识的作家很有名著作也很多。
    认识教授的作家很有名著作也很多。

21. 店员不喜欢的经理站在门口招揽生意。
    不喜欢店员的经理站在门口招揽生意。

22. 老太太遇见的女孩长得很漂亮。
    遇见老太太的女孩长得很漂亮。

23. 歌手羡慕的演员想往其它方面发展　。
    羡慕歌手的演员想往其它方面发展　。

24. 男孩拜访的邻居家境贫寒。
拜访男孩的邻居家境贫寒。

25. 私家侦探跟踪的警探想知道事情的真相。
跟踪私家侦探的警探想知道事情的真相。

26. 小丑模仿的喜剧演员很受群众欢迎 。
模仿小丑的喜剧演员很受群众欢迎 。

27. 明星爱上的诗人充满不切实际的幻想 。
爱上明星的诗人充满不切实际的幻想 。

28. 寡妇嘲笑的老处女很想交男朋友 。
嘲笑寡妇的老处女很想交男朋友 。

29. 流氓威胁的逃犯害怕警察抓捕整天提心吊胆。
威胁流氓的逃犯害怕警察抓捕整天提心吊胆。

30. 富豪邀请的官员心怀不轨但是善于隐藏。
邀请富豪的官员心怀不轨但是善于隐藏。

31. 居民协助的军官跟随间谍到了海边。
协助居民的军官跟随间谍到了海边。

32. 司机抱怨的乘客总是大声喧哗很令人受不了。
抱怨司机的乘客总是大声喧哗很令人受不了。

SL IN SENTENCE PROCESSING

English Experimental Sentences of Study 1

1. The banker that the waiter praised (praised the waiter) climbed the mountain just outside of town.

2. The lawyer that the reporter phoned (phoned the reporter) cooked the pork chops in their own juices.

3. The salesman that the fireman liked (liked the fireman) dominated the conversation about the horse race.

4. The waiter that the chef avoided (avoided the chef) drove the sports car home from work that evening.

5. The policemen that the teacher disliked (disliked the teacher) cut out the article with the dull scissors.

6. The judge that the cameraman ignored (ignored the cameraman) watched the report about the escaped fugitive.

7. The robber that the accountant insulted (insulted the accountant) read the newspaper article about the fire.

8. The minister that the comedian admired (admired the comedian) answered the telephone in the fancy restaurant.

9. The tenant that the landlord despised (despised the landlord) called the newspaper to complain.

10. The professor that the student criticised (criticised the student) blushed and looked away.

11. The editor that the author irritated (irritated the author) play tennis on Saturday.

12. The pilot that the attendant complimented (complimented the attendant) asked for a date.

13. The businessman that the secretary married (married the secretary) invited many people to the party.

14. The waitress that the mechanic divorced (divorced the mechanic) won a lot of money in the lottery.

15. The poet that the dentist encountered (encountered the dentist) caught a nasty cold.

16. The acrobat that the juggler loved (loved the juggler) trained aggressively in the gym each day.

17. The warden that the prisoner taunted (taunted the prisoner) strained his ears to hear the conversation.

18. The foreman that the manager challenged (challenged the manager) replayed the conversation in his head afterwards.

19. The surgeon that the scientist envied (envied the scientist) denied all the allegations.

20. The navigator that the captain revered (revered the captain) found the map without any trouble.

21. The technician that the programmer embarrassed (embarrassed the programmer) believed the treatment was unfair.

22. The policeman that the arsonist feared (feared the arsonist) made an excuse for not attending the briefing.

23. The director that the producer resented (resented the producer) struggled to keep his mouth shut.

24. The inventor that the typist visited (visited the typist) arranged a lunchtime meeting.

25. The performer that the neurologist shamed (shamed the neurologist) whispered something to his friends.

26. The servant that the cook comforted (comforted the cook) searched the kitchen for the broom.

27. The actor that the artist adored (adored the artist) went home feeling good about his day.

28. The builder that the painter pushed (pushed the painter) walked out before the job was finished.

SL IN SENTENCE PROCESSING

29. The scientist that the historian consulted (consulted the historian) submitted the article to a major journal.

30. The farmer that the labourer rescued (rescued the labourer) resolved to be more careful in the future.

31. The investor that the analyst greeted (greeted the analyst) donated a substantial sum of money to charity.

32. The motorist that the pedestrian shot (shot the pedestrian) decided violence would not solve anything.

SL IN SENTENCE PROCESSING

Appendix B

Mandarin Experimental sentence of Study 2

1. 熟识富人的经理遇见了牧师所以心里很高兴。
   富人熟识的经理遇见了牧师所以心里很高兴。

2. 配合家属的刑警恨透了嫌犯并打算破釜沉舟。
   家属配合的刑警恨透了嫌犯并打算破釜沉舟。

3. 告发校长的学生很信任父母并决定支持他
   校长告发的学生很信任父母并决定支持他。

4. 奉承老板的男子看不起专家并且讨厌l他。
   老板奉承的男子看不起专家并且讨厌他。

5. 勾引院长的少女撞到了议员而感到羞愧。
   院长勾引的少女撞到了议员而感到羞愧。

6. 欣赏董事长的女秘书暗恋着主任l而且不为人知。
   董事长欣赏的女秘书暗恋着主任而且不为人知。

7. 责怪市长的居民问候着总理并安慰着他。
   市长责怪的居民问候着总理并安慰着他。

8. 带来巡警的摊贩怒骂着农民并打了他。
   巡警带来的摊贩怒骂着农民并打了他。

9. 打昏店员的歹徒见到了记者并立即报警。
   店员打昏的歹徒见到了记者并立即报警。

10. 敬佩教练的选手招呼着市长并为他l斟酒。
    教练l敬佩的选手招呼着市长并为他l斟酒。

11. 雇用员工的律师斥责了经理并起诉他。
    员工雇用的律师斥责了经理并起诉他。

12. 陪伴厂长的职员打伤了暴民还谩骂了厂长。
    厂长陪伴的职员打伤了暴民还谩骂了厂长。

13. 责怪家长的老师误导了学生并要求停课。
    家长责怪的老师误导了学生并要求停课。

14. 检举厂商的市民逼疯了官员还得意洋洋。
    厂商检举的市民逼疯了官员还得意洋洋。

15. 包庇商人的政客低估了部长l因而懊悔不已。
    商人包庇l的l政客低估了部长因而懊悔不已。

16. 巴结队长的老翁赶走了书记结果适得其反。
    队长巴结的老翁l赶走了书记结果适得其反。

17. 杀死台商的少年不认识医师l因而没有注意他。
    台商杀死的少年不认识医师因而没有注意他。

18. 照顾祖母的男子吵醒了队长因而感到愧疚。
    祖母照顾的男子吵醒了队长因而感到愧疚。

19. 救活游客的农民很尊敬老板还答谢了他。
    游客救活的农民很尊敬老板还答谢了他。

20. 联络媒体的画家很爱慕歌手还决定娶她。
    媒体联络的画家很爱慕歌手还决定娶她。

21. 陷害雇主的劳工拜访了贵宾还带了礼品。
    雇主陷害的劳工拜访了贵宾还带了礼品

22. 玩弄女子的商人看到了警探并逮捕了他。
    女子玩弄的商人看到了警探并逮捕了他。

23. 邀集工人的民众没见到市长非常失望。
    工人邀集的民众没见到市长非常失望。

24. 回避客户的小姐找到了律师并进行了询问。
    客户回避的小姐找到了律师并进行了询问。

25. 威胁流氓的小偷害怕众人采取行动。
    流氓威胁的小偷害怕众人采取行动。

26. 憎恨政府的罪犯协助间谍实施叛变计划。
    政府憎恨的罪犯协助间谍实施叛变计划。

27. 追求保安的女孩要求同事不要乱说话。
    保安追求的女孩要求同事不要乱说话。

28. 约见广告商的制片人提醒导演时间不多了。
    广告商约见的制片人提醒导演时间不多了。

29. 关心丈夫的妻子珍惜婆婆为家庭的付出。
    丈夫关心的妻子珍惜婆婆为家庭的付出。

30. 热爱听众的播音员反对导播的做法。
    听众热爱的播音员反对导播的做法。

31. 等待护士的医生欢迎病人多和自己交流。
    护士等待的医生欢迎病人多和自己交流。

32. 模仿小丑的主持人推荐团长给大家。
    小丑模仿的主持人推荐团长给大家。

References

Alderton, D. L., & Larson, G. E. (1990). Dimensionality of Raven's advanced progressive matrices items. *Educational and Psychological Measurement*, *50*(4), 887-900.

Arciuli, J., & Simpson, I. C. (2012). Statistical learning is related to reading ability in children and adults. *Cognitive science*, *36*(2), 286-304.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*, *59*(4), 390-412.

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in psychology*, *4*.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, *68*(3), 255-278.

Bates, D., Maechler, M., & Bolker, B. (2012). lme4: Linear mixed-effects models using S4 classes. R package version 0.999999-0 [Computer software]. *Retrieved from rproject. org website: http://CRAN. Rproject. org/package0lme4*.

Bever, T. G. (1970). Cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and development of language* (pp. 279–362). New York: Wiley.

Bogaerts, L., Siegelman, N., & Frost, R. (2016). Splitting the variance of statistical learning performance: A parametric investigation of exposure duration and transitional probabilities. *Psychonomic bulletin & review*, *23*(4), 1250-1256.

Bors, D. A., & Forrin, B. (1995). Age, speed of information processing, recall, and fluid intelligence. *Intelligence*, *20*(3), 229-248.

Brooks, P. J., & Kempe, V. (2013). Individual differences in adult foreign language learning: The mediating effect of metalinguistic awareness. *Memory & cognition*, *41*(2), 281-296.

Brooks, P. J., Kwoka, N., & Kempe, V. (2017). Distributional effects and individual differences in L2 morphology learning. *Language Learning*, *67*(1), 171-207.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and brain Sciences*, *22*(1), 77-94.

Carrow-Woolfolk, E. (1999). *CASL: Comprehensive assessment of spoken language*. American Guidance Services.

Chen, B., Ning, A., Bi, H., & Dunlap, S. (2008). Chinese subject-relative clauses are more difficult to process than the object-relative clauses. *Acta Psychologica*, *129*(1), 61-65.

Christiansen, M. H. (2001). Intrinsic constraints on the processing of recursive sentence structure. Manuscript in preparation.

Christiansen, M. H., & Chater, N. (1999). Toward a connectionist model of recursion in human linguistic performance. *Cognitive Science*, *23*(2), 157-205.

Christiansen, M. H., & Chater, N. (2016). The Now-or-Never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*, *39*.

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, *27*(1), 3-42.

Cohen, L., & Mehler, J. (1996). Click monitoring revisited: An on-line study of sentence comprehension. *Memory & cognition*, *24*(1), 94-102.

Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(1), 24.

Conway, C. M., & Christiansen, M. H. (2006). Statistical learning within and between modalities: Pitting abstract against stimulus-specific representations. *Psychological science*, *17*(10), 905-912.

Conway, C. M., & Christiansen, M. H. (2009). Seeing and hearing in space and time: Effects of modality and presentation rate on implicit statistical learning. *European Journal of Cognitive Psychology*, *21*(4), 561-580.

Conway, C. M., Bauernschmidt, A., Huang, S. S., & Pisoni, D. B. (2010). Implicit statistical learning in language processing: Word predictability is the key. *Cognition*, *114*(3), 356-371.

Daltrozzo, J., Emerson, S. N., Deocampo, J., Singh, S., Freggens, M., Branum-Martin, L., & Conway, C. M. (2017). Visual statistical learning is related to natural language ability in adults: An ERP study. *Brain and language*, *166*, 40-51.

Dunn, L., & Dunn, L. (2007). Peabody Picture Vocabulary Test (4th ed.). Minneapolis, MN: Pearson Assessments. (PPVT-4)

Dutoit, T., Pagel, V., Pierret, N., Bataille, F., & Van der Vrecken, O. (1996, October). The MBROLA project: Towards a set of high quality speech synthesizers free of use for noncommercial purposes. In *Spoken Language, 1996. ICSLP 96. Proceedings., Fourth International Conference on* (Vol. 3, pp. 1393-1396). IEEE.

Elman, J. (1991). Distributed representations, simple recurrent networks, and grammatical structure. Machine Learning, 7, 195–224.

Ericsson, K. A., Chase, W. G., & Faloon, S. (1980). Acquisition of a memory skill. *Science*, *208*(4448), 1181-1182.

Evans, J. L., Saffran, J. R., & Robe-Torres, K. (2009). Statistical learning in children

with specific language impairment. *Journal of Speech, Language, and Hearing*

Research, *52*(2), 321-335.

Farmer, T. A., Fine, A. B., Misyak, J. B., & Christiansen, M. H. (2017). Reading

span task performance, linguistic experience, and the processing of unexpected

syntactic events. *The Quarterly Journal of Experimental Psychology*, *70*(3),

413-433.

Frazier, L. (1987). Syntactic processing: evidence from Dutch. *Natural Language &*

Linguistic Theory, *5*(4), 519-559.

Friedman, N. P., & Miyake, A. (2005). Comparison of four scoring methods for the

reading span test. *Behavior research methods*, *37*(4), 581-590.

Frost, R. (2012). Towards a Universal Model of Reading. *The Behavioral and Brain*

Sciences, *35*(5), 263–279. http://doi.org/10.1017/S0140525X11001841

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain

generality versus modality specificity: The paradox of statistical

learning. *Trends in cognitive sciences*, *19*(3), 117-125.

Frost, R., Siegelman, N., Narkiss, A., & Afek, L. (2013). What predicts successful

literacy acquisition in a second language?. *Psychological science*, *24*(7), 1243-

1252.

Gathercole, S. E., & Pickering, S. (2001). Working memory test battery for children

(WMTB-C). *London, UK: The Psychological Corporation*.

Gennari, S. P., Mirković, J., & MacDonald, M. C. (2012). Animacy and competition

in relative clause production: A cross-linguistic investigation. *Cognitive*

psychology, *65*(2), 141-176.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic

dependencies. *Cognition*, *68*(1), 1-76.

Gibson, E. (2000). The dependency locality theory: A distance-based theory of

linguistic complexity. In Y. Miyashita, A. Marantz, & W. O'Neil (Eds.), *Image,*

language, brain (pp. 95–126). Cambridge, MA: MIT Press.

Gibson, E., & Wu, H. H. I. (2013). Processing Chinese relative clauses in

context. *Language and Cognitive Processes*, *28*(1-2), 125-155.

Gibson, E., Desmet, T., Grodner, D., et al. (2005). Reading relative clauses in

English. *Cognitive Linguistics*, 16(2), pp. 313-353.

Glicksohn, A., & Cohen, A. (2011). The role of Gestalt grouping principles in visual

statistical learning. *Attention, Perception, & Psychophysics*, *73*(3), 708-713.

Gordon, P. C., Hendrick, R., Johnson, M., & Lee, Y. (2006). Similarity-based

interference during language comprehension: Evidence from eye tracking

during reading. *Journal of Experimental Psychology: Learning, Memory, and*

Cognition, *32*(6), 1304.

Granena, G. (2013). Individual differences in sequence learning ability and second

language acquisition in early childhood and adulthood. *Language*

Learning, *63*(4), 665-703.

Grodner, D., Gibson, E., & Watson, D. (2005). The influence of contextual contrast

on syntactic processing: evidence for strong-interaction in sentence

comprehension. *Cognition*, *95*(3), 275-296.

Hammill, D. D., Brown, V. L., & Larsen, S. C. (8). Wiederholt. JL (1994). *Test of*

Adolescent and Adult Language.

Hay, J. F., Pelucchi, B., Estes, K. G., & Saffran, J. R. (2011). Linking sounds to

meanings: Infant statistical learning in a natural language. *Cognitive*

psychology, *63*(2), 93-106.

He, W., Xu, N., & Ji, R. (2017). Effects of Age and Location in Chinese Relative

Clauses Processing. *Journal of Psycholinguistic Research*, 1-20.

Hellman, A. B. (2011). Vocabulary size and depth of word knowledge in adult-onset

second language acquisition. *International Journal of Applied*

Linguistics, *21*(2), 162-182.

Hopp, H. (2006). Syntactic features and reanalysis in near-native processing. *Second*

Hsiao, F., & Gibson, E. (2003). Processing relative clauses in

Chinese. *Cognition*, *90*(1), 3-27.

Hutton, I., & Kidd, E. (2011). Structural priming in comprehension of relative clause

sentences In search of a frequency x regularity interaction. *The Acquisition of*

Relative Clauses: Processing, Typology and Function, *8*, 227.

Indefrey, P. (2006). A Meta-analysis of Hemodynamic Studies on First and Second

Language Processing: Which Suggested Differences Can We Trust and What

Do They Mean?. *Language learning*, *56*(s1), 279-304.

Isbilen, E. S., McCauley, S. M., Kidd, E., & Christiansen, M. H. (2017). Testing

Statistical Learning Implicitly: A Novel Chunk-based Measure of Statistical

Learning.

Jackson, C. (2008). Proficiency level and the interaction of lexical and

morphosyntactic information during L2 sentence processing. *Language*

Learning, *58*(4), 875-909.

Jeon, E. H., & Yamashita, J. (2014). L2 Reading comprehension and its correlates: A

meta-analysis. *Language Learning*, *64*(1), 160-212.

Jost, E., Conway, C. M., Purdy, J. D., Walk, A. M., & Hendricks, M. A. (2015).

Exploring the neurodevelopment of visual statistical learning using event-

related brain potentials. *Brain research*, *1597*, 95-107.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension:

individual differences in working memory. *Psychological review*, *99*(1), 122.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of experimental psychology: General*, *111*(2), 228.

Kaufman, S. B., DeYoung, C. G., Gray, J. R., Jiménez, L., Brown, J., & Mackintosh, N. (2010). Implicit learning as an ability. *Cognition*, *116*(3), 321-340.

Kidd, E. (2012). Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental psychology*, *48*(1), 171.

Kidd, E., & Arciuli, J. (2016). Individual differences in statistical learning predict children's comprehension of syntax. *Child development*, *87*(1), 184-193.

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.

Kim, R., Seitz, A., Feenstra, H., & Shams, L. (2009). Testing assumptions of statistical learning: is it long-term and implicit?. *Neuroscience letters*, *461*(2), 145-149.

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of memory and language*, *30*(5), 580-602.

Knowlton, B. J., & Squire, L. R. (1996). Artificial grammar learning depends on implicit acquisition of both abstract and exemplar-specific information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(1), 169.

Kwon, N., Gordon, P. C., Lee, Y., Kluender, R., & Polinsky, M. (2010). Cognitive and linguistic factors affecting subject/object asymmetry: An eye-tracking study of prenominal relative clauses in Korean. *Language*, *86*(3), 546-582.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive psychology*, *6*(2), 293-323.

Language Research, 22, 369–397.

Lewandowsky, S., Oberauer, K., Yang, L. X., & Ecker, U. K. (2010). A working

memory test battery for MATLAB. *Behavior Research Methods*, *42*(2), 571-

585.

Li, P., Zhang, F., Tsai, E., & Puls, B. (2014). Language history questionnaire (LHQ

2.0): A new dynamic web-based research tool. *Bilingualism: Language and

Cognition*, *17*(3), 673-680.

Lin, C. J. C., & Bever, T. G. (2006). Subject preference in the processing of relative

clauses in Chinese. In *Proceedings of the 25th west coast conference on formal

linguistics* (pp. 254-260). Somerville, MA: Cascadilla Proceedings Project.

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory:

comment on Just and Carpenter (1992) and Waters and Caplan (1996).

Mackey, A., Adams, R., Stafford, C., & Winke, P. (2010). Exploring the relationship

between modified output and working memory capacity. *Language

Learning*, *60*(3), 501-533.

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.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional

information can affect phonetic discrimination. *Cognition*, *82*(3), B101-B111.

McCauley, S. M., & Christiansen, M. H. (2015, July). Individual Differences in

Chunking Ability Predict On-line Sentence Processing. In *CogSci*.

McCauley, S. M., Isbilen, E. S., & Christiansen, M. H. (2017, July). Chunking

Ability Shapes Sentence Processing at Multiple Levels of Abstraction.

McDonough, K., & Trofimovich, P. (2016). The role of statistical learning and

working memory in L2 speakers' pattern learning. *The Modern Language

Journal*, *100*(2), 428-445.

Mecklinger, A., Schriefers, H., Steinhauer, K., & Friederici, A. D. (1995). Processing relative clauses varying on syntactic and semantic dimensions: An analysis with event-related potentials. *Memory & Cognition*, *23*(4), 477-494.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, *63*(2), 81.

Misyak, J. B., & Christiansen, M. H. (2012). Statistical learning and language: An individual differences study. *Language Learning*, *62*(1), 302-331.

Misyak, J. B., Christiansen, M. H., & Bruce Tomblin, J. (2010). Sequential expectations: The role of prediction-based learning in language. *Topics in Cognitive Science*, *2*(1), 138-153.

Misyak, J. B., Christiansen, M. H., & Tomblin, J. B. (2010). On-line individual differences in statistical learning predict language processing. *Frontiers in psychology*, *1*.

Miyamoto, E. T., & Nakamura, M. (2003, March). Subject/object asymmetries in the processing of relative clauses in Japanese. In *Proceedings of the 22nd West Coast conference on formal linguistics* (Vol. 22, pp. 342-355). Cascadilla Press Somerville, MA.

Onnis, L. (2012). The potential contribution of statistical learning to second language acquisition. *Statistical learning and language acquisition*, 203-235.

Pearlmutter, N. J., & MacDonald, M. C. (1995). Individual differences and probabilistic constraints in syntactic ambiguity resolution. *Journal of memory and language*, *34*(4), 521-542.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in cognitive sciences*, *10*(5), 233-238.

Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge
and academic reading performance: An assessment perspective. *Language
learning*, *52*(3), 513-536.

Qian, D. D., & Schedl, M. (2004). Evaluation of an in-depth vocabulary knowledge
measure for assessing reading performance. *Language Testing*, *21*(1), 28-52.

Qiao, X., Shen, L., & Forster, K. (2012). Relative clause processing in Mandarin:
Evidence from the maze task. *Language and Cognitive Processes*, *27*(4), 611-
630.

Rah, A., & Adone, D. (2010). Processing of the reduced relative clause versus main
verb ambiguity in L2 learners at different proficiency levels. *Studies in Second
Language Acquisition*, *32*(1), 79-109.

Raven, J. C., Raven, J. E., & Court, J. H. (1998). Progressive matrices. Oxford,
England: Oxford Psychologists Press.

Reber, A. S. (1993). *Implicit learning and tacit knowledge*. An essay on the cognitive
unconscious. New York: Oxford Psychology Series.

Roberts, L. (2012). Individual differences in second language sentence
processing. *Language Learning*, *62*(s2), 172-188.

Robinson, P. (2005). Cognitive abilities, chunk-strength, and frequency effects in
implicit artificial grammar and incidental L2 learning: Replications of Reber,
Walkenfeld, and Hernstadt (1991) and Knowlton and Squire (1996) and their
relevance for SLA. *Studies in Second Language Acquisition*, *27*(2), 235-268.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-
old infants. *Science*, 1926-1928.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of
distributional cues. *Journal of memory and language*, *35*(4), 606-621.

Saults, J. S., & Cowan, N. (2007). A central capacity limit to the simultaneous

storage of visual and auditory arrays in working memory. *Journal of*

Experimental Psychology: General, *136*(4), 663.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-prime, Version

1.1. *Pittsburgh, PA: Psychology Software Tools*.

Schvaneveldt, R. W., & Gomez, R. L. (1998). Attention and probabilistic sequence

learning. *Psychological Research*, *61*(3), 175-190.

Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability:

Theoretical perspectives and empirical evidence. *Journal of memory and*

language, *81*, 105-120.

Siegelman, N., Bogaerts, L., & Frost, R. (2017). Measuring individual differences in

statistical learning: Current pitfalls and possible solutions. *Behavior research*

methods, *49*(2), 418-432.

Siegelman, N., Bogaerts, L., Christiansen, M. H., & Frost, R. (2017). Towards a

theory of individual differences in statistical learning. *Phil. Trans. R. Soc.*

B, *372*(1711), 20160059.

Siegelman, N., Bogaerts, L., Kronenfeld, O., & Frost, R. (2017). Redefining

"Learning" in Statistical Learning: What Does an Online Measure Reveal About

the Assimilation of Visual Regularities?. *Cognitive Science*.

Speciale, G., Ellis, N. C., & Bywater, T. (2004). Phonological sequence learning and

short-term store capacity determine second language vocabulary

acquisition. *Applied psycholinguistics*, *25*(2), 293-321.

Sung, Y. T., Cha, J. H., Tu, J. Y., Wu, M. D., & Lin, W. C. (2015). Investigating the

processing of relative clauses in Mandarin Chinese: evidence from eye-

movement data. *Journal of psycholinguistic research*, *45*(5), 1089-1113.

Sung, Y. T., Tu, J. Y., Cha, J. H., & Wu, M. D. (2016). Processing preference toward
object-extracted relative clauses in mandarin chinese by l1 and l2 speakers: an
eye-tracking study. *Frontiers in psychology*, *7*.

Traxler, M. J., Morris, R. K., & Seely, R. E. (2002). Processing subject and object
relative clauses: Evidence from eye movements. *Journal of Memory and
Language*, *47*(1), 69-90.

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual
statistical learning. *Journal of Experimental Psychology: General*, *134*(4), 552.

Ueno, M., & Garnsey, S. M. (2008). An ERP study of the processing of subject and
object relative clauses in Japanese. *Language and Cognitive Processes*, *23*(5),
646-688.

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.

Vasishth, S., Chen, Z., Li, Q., & Guo, G. (2013). Processing Chinese relative clauses:
Evidence for the subject-relative advantage. *PloS one*, *8*(10), e77006.

Wang, H., Ma, L., Wang, Y., Troyer, M., & Li, Q. (2015). An ERP Study in English
Relative Clause Processing by Chinese-English Bilinguals. *Chinese Journal of
Applied Linguistics*, *38*(1), 3-35.

Wechsler, D. (1997). *WAIS-III: Wechsler adult intelligence scale*. Psychological
Corporation.

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.

Wu, F., Kaiser, E., & Andersen, E. (2010). Subject preference, head animacy and

    lexical cues: a corpus study of relative clauses in Chinese. In *Processing and*

    producing head-final structures (pp. 173-193). Springer Netherlands.

Wu, F., Kaiser, E., & Andersen, E. (2012). Animacy effects in Chinese relative clause

    processing. *Language and Cognitive Processes*, *27*(10), 1489-1524.

Wu, F.-Y. (2009). *Factors affecting relative clause processing in Mandarin.*

    (Doctoral dissertation). Retrieved from ProQuest Dissertations Publishing.

    (3368727)