行政院國家科學委員會專題研究計畫 期末報告

中文及英文量詞『一些』與『全部』的推論歷程：閱讀時間及事件相關電位研究

計 畫 類 別 ：個別型
計 畫 編 號 ：NSC 101-2410-H-002-210-
執 行 期 間 ：101 年 08 月 01 日至 102 年 07 月 31 日
執 行 單 位 ：國立臺灣大學外國語文學系暨研究所

計 畫 主 持 人 ：林祐瑜

公 開 資 訊 ：本計畫涉及專利或其他智慧財產權，2 年後可公開查詢

中 華 民 國 102 年 10 月 30 日

中 文 摘 要 ： 量詞處理歷程指的是在例如像『一些，全部』的量詞表中，當說話者用『一些』這個詞在句中時，他們心理拒絕使用另一個詞『全部』來表達。在語用學的領域中，目前有兩種理論可茲說明：『預設推論理論』及『相關理論』。這兩種理論對於人們是如何處理量詞推論做出了完全不同的預測。截至目前為止，量詞處理歷程大部分仍是依靠這兩個語用學理論或是本能想法來做闡述，鮮少有研究用實驗法或其他線上實驗方式來看這個議題。基於實驗法早已被許多學科，如心理語言學，所採用，且其他非線上實驗也產生了許多相抵觸的實驗結果，本文擬提出一個兩年期的線上實驗計畫來研究這個議題。實驗分成：(1)以中文為母與者是如何處理『一些』及『全部』推論； (2)學英文的中文母語人士是如何處理英文『一些』和『全部』的推論。本實驗研究成果顯示中文母語使用者及外語學習者皆將『一些』理解成『一些但非全部』。由反應時間來看，當受試者理解成『一些但非全部』時，反應時間皆顯著快於『一些但有可能全部』。這個實驗結果支持了 Default 推論過程，也解決了理論上的爭端。

中文關鍵詞： 中文量詞處理

英 文 摘 要 ： The study presented here reports Mandarin second language learners' processing of scalar inferences. Scalar inferences represents the condition when a speaker uses a weaker expression such as some in a pragmatic scale like <some, all>, s/he has the intention to reject the more stronger use of the other word like all in the utterance. Different theories, e.g., Default Inference and Relevance theory, have made different predictions regarding how people process scalar inferences and experimental approaches would help to identify which one is correct. Currently, most of the studies on scalar inferences looked at native speakers' performance. In other words, very few studies have been trying to look at this issue among second language learners. The current study thus serves as the pioneer study to use online experimental methods to investigate language learners' performance in interpreting scalar inferences. Results showed that even though Mandarin second language learners of English were slower in making Truth value judgment task and

reaction times, their processing of scalar inferences were very much like that of first language speakers. They showed faster reaction times when interpreting some as some but not all. The results found support for Default Inference account. In other words, second language learners' interpretation of some as some but not all seems to arise by default. In addition, the current study also found that working memory capacity plays a critical role during scalar inference processing.

英文關鍵詞： Mandarin scalar inferences, psycholinguistics

Time Course and Working Memory Capacity in Second Language Learners' Processing of Scalar Inferences

Abstract

The study presented here reports Mandarin second language learners' processing of scalar inferences. Scalar inferences represents the condition when a speaker uses a weaker expression such as *some* in a pragmatic scale like <some, all>, s/he has the intention to reject the more stronger use of the other word like *all* in the utterance. Different theories, e.g., Default Inference and Relevance theory, have made different predictions regarding how people process scalar inferences and experimental approaches would help to identify which one is correct. Currently, most of the studies on scalar inferences looked at native speakers' performance. In other words, very few studies have been trying to look at this issue among second language learners. The current study thus serves as the pioneer study to use online experimental methods to investigate language learners' performance in interpreting scalar inferences. Results showed that even though Mandarin second language learners of English were slower in making Truth value judgment task and reaction times, their processing of scalar inferences were very much like that of first language speakers. They showed faster reaction times when interpreting *some* as *some but not all*. The results found support for Default Inference account. In other words, second language learners' interpretation of *some* as *some but not all* seems to arise by default. In addition, the current study also found that working memory capacity plays a critical role during scalar inference processing.

Introduction

Language comprehension was executed in a rapid manner. Within millisecond, a lot of information has been exchanged among speakers, who need to decode incoming strings of words, construct syntactic representations, get the meaning and finally have to respond to their listeners. Sometimes, if the meaning of utterances is not literal, listeners have to read between the lines so as to get the inferences hidden in the conversation. Althouh it may sound like a daunting task at first, first language speaker usually can finish these tasks within milliseconds. However, for second language learners, any single task mentioned above is by itself very difficult, let alone second language learners need to process all of them at the same time, just like first language speakers. It is therefore no wonder that second language learning has always been considered a difficult task, especially when learners have reached adulthood. Many adult second language speakers, although equipped with superior linguistic knowledge and proficient second language skills, still display non-nativelike proficiencies from native speakers. For example, studies showed that English article system still poses a strong challenge even for advanced English learners (Master, 2002). In addition, many different theories have been proposed to account for the fact that second language learners generally fail to attain a native like proficiency level (Bley-Vroman, 2009; Clahsen & Felser, 2006; Schwartz, 1998; White, 1988).

To examine the validity of different theories and to understand the cognitive systems of second language learners, many different techniques and online methods have been used to investigate second language learners' online processing such as reaction times, event-

related potential (EPRs), fMRI and even PET (Chee, Caplan, et al., 1999; Chee, Tan, & Thiel, 1999; Dehaene et al., 1997; Klein, Milner, Zatorre, E. Meyer, & Evans, 1995; Klein, Milner, Zatorre, Zhao, & Nikelski, 1999; Osterhout & McLaughlin, 2000). Many different theoretical questions can be investigated between first language and second language comprehension. For example, one of the major questions is whether the neural substrates are the same between first and second languages. Despite the discrepancies found among different studies (Chee, Caplan, et al., 1999; Chee, Tan, et al., 1999; Dehaene et al., 1997; Klein et al., 1995, 1999), Perani et al. (1998) also found that proficiency level and age in a second language play an important role in second language processing.

So far, most of the studies investigating second language learning have focused on syntax, phonology, psycholinguistics, sociolinguistics, pedagogy and so on. Few have looked at second language learning in pragmatics, let alone even fewer studies have tried to use online methods to address pragmatic issues. Based on this, the current paper tried to bridge this gap by addressing scalar inferences in second language learners. To be more specific, we investigated Mandarin learners' processes of English scalar inferences. Before we delve into our research design, we will first briefly introduce scalar inferences and some first language studies on them.

Theoretical Background

During a conversation, more often than not, a listener has to go beyond the literal linguistic expressions, or read between the lines, to get the implicit meaning, or

implicatures, from the speaker. How the interlocutors derive these inferences was first characterized by a philosopher of language, Paul Grice (1975), who systematically studied the meanings between interlocutors and laid the foundations of the modern study of pragmatics. He proposed Cooperative Principle and four maxims to specify how people follow these rules to make conversation efficient.

> ***Cooperative Principle***: Contribute what is required by the accepted purpose of the conversation.
>
> > **Maxim of Quality**: Make your contribution true; so do not convey what you believe false or unjustified.
> >
> > **Maxim of Quantity**: Be as informative as required.
> >
> > **Maxim of Relation**: Be relevant.
> >
> > **Maxim of Manner**: Be perspicuous; so avoid obscurity and ambiguity, and strive for brevity and order.

A typical implicature can be found in the following short conversation:

> (1)
>
> A: Did you do your laundry?
>
> B: I was busy

In this conversation, it is quite obvious that B didn't address A's question directly. On the contrary, B replied with "I was busy" to imply that he did not do his laundry. According

to Grice's Cooperative Principle, if B is being cooperative, his contribution, or implicature, would be helpful in contributing to understand this conversation. Grice has characterized the implicature like the above example as *conversational implicature* since it occurs out of specific context, instead of arising from the literal meaning of the words in the conversation. Conversational implicatuers thus can be separated from conventional implicature, which is hypothesized to be the standard or customary implicature that can be inferred from the utterances, as illustrated by the following classic example:

(2)

(a) He is an Englishman; he is, therefore, brave.

(b) His being an Englishman implies that he is brave.

In example (2), when (2a) is expressed, it entails (2b). This implication does not arise out of a specific context and is purely based on the linguistic or literal meaning of the words. Grice described this kind of implicature as conventional implicature. In addition to context that distinguishes conversational implicatures from conventional implicatures, another two features also characterize conversational implicatures: *cancelable* and *reinforceable.* For example, what B said in (1) can be interpreted or inferred differently under a total different context. In other words, the implicature that *B didn't do the laundry* would be cancelled without the context of A's utterance. These two features do not hold for conventional implicatures since whenever (2a) is expressed, it will always entails (2b).

Scalar inferences

Another form of conversational implicature concerns scalar implicature (Horn, 1972; Levinson, 1983). It is the uses of words like "all" or "some" in conversation. For example, when a speaker says, *"I ate some of the cake"*, s/he implicates "I did not eat all of the cake." *Some* in this way is regarded as less informative term than *all*. The main idea behind scalar inference is that the use of lower ranking alternative, such as *some* in this case, by the speaker prompts the hearer to realize the speaker's reluctance to use the higher ranking one such as all. To put it in another way, people's choice of using *some* in the conversation may be used to indicate that the higher ranking one is not true or warranted. <some, all> is not the only set of scalar in scalar inferences, other possibilities arise by the use of numerals, modals, adverbs, lexical contrast and partial contextual orderings. Here are some examples of them.

(3) Numerals:

John: Do you have three pens?

Mary: I have two.

Implicature: Mary does not have three pens.

(4) Lexical contrast:

John: Did you succeed in getting in the team?

Mary: I tried.

Implicature: Mary did not succeed in getting into the team.

(5) Partial Contextual Ordering

John: Did you go to work today?

Mary: I went shopping.

Implicature: Mary did not go to work.

Default Inferences

The derivation of implicature in the original Gricena's theory has been found to be complex and onerous. It therefore does not seem to suit the requirement that conversations happen in a fast manner. Within second, interlocutors need to go through many different linguistic stages such as encoding and decoding phonological representations, constructing syntactic representation, and deriving the (underlying) meaning of the sentences. If the implicature derivation is too costly, it would not benefit human cognitive development or mental processing. Levinson, however, pointed out that scalar inferences should belong to generalized conversational implicatures (Levinson, 1983). In this way, they would be derived automatically without considering the context and will thus speed up the processes. For Levinson, since the generation of scalar inferences is by default, it can only be cancelled when the context requires it.

Relevance Theory

Relevance theory has different accounts on scalar inferecens (Sperber & Wilson, 2004). It postulates that when hearing an utterance, a hearer either narrows down or broadens the speaker's linguistic expressions. Consider the example of *some* in the above cases, interpreting *some* as *some and possibly all* is an example of broadening the scope of *some*

while interpreting *some* as *some but not all* is an example of narrowing down its meaning. Unlike Gricean's or Neo-Gricean's accounts, relevance theory posits that it is quite onerous for a hearer to derive *some* either as *some but not all* or *some and possibly all* during a conversation. In the case of *some*, a hearer needs to consider which meaning of *some* matches the real scenario so that they can derive precisely. For example, if a hearer does not find a scenario that matches the use of *some and possibly all*, he would realize a narrowing use of *some* must be applied instead. Moreover, hearers in real world either do not need to know the precise situation or it is unlikely for them to know the real situation to make precise application. Under this circumstance, relevance theory states that the narrowing down or broadening of linguistic expression does not occur automatically or by default. Instead, inferences such as *some but not all* or *some and possibly all* occur only when there is a need, e.g., a context, for the hearer to do so. When these inferences occur, it would be cognitively demanding for the listeners. In this way, the Relevance theory actually predicts just the opposite from the Default theory in the scalar inferences. In the example of "some", the Relevance theory assumes that the listener would tend to go with the logical interpretation. If there is a need for the listener to go further, s/he would adopt the "some but not all" pragmatic interpretation.

Psychological reality and experimental issues

Despite so many efforts working on unraveling conversational implicatures, surprisingly few studies have been trying to look at the psychological reality or mental derivations of conversation implicatures or scalar inferences. However, we need to point out that the cognitive resources during processing scalar inferences can be easily done in

experimental approaches. Up to date, the field on pragmatics relies heavily on pragmatic theorists' knowledge in theoretical linguistics and intuitions. Gricean-, Neo-Gricean or Relevance theories have been trying to offer their accounts of pragmatics but neither of them has been very explicit about the derivational process in human cognitive system. At this juncture, experimental approaches are very helpful in several ways. First of all, it helps to tease apart competing theories. Second, it sheds light on how human cognitive system work through implicatures. Experiments using reaction times usually ask participants to interpret the sentence in either a strict logical or pragmatic way without considering what the subjects might use in the real world. Their rationale is that if the Default model is correct, faster reaction time should be observed when an automatic inference is prompted. However, if the inference involves cancellations, it would slow people down, resulting in longer reaction time. Opposite pattern would be observed if the Relevance theory is correct.

Previous experiments in first language

Experiments on scalar inferences have not drawn researchers' attention until very recently. As one of the earliest researchers using experimentation to study pragmatics, Noveck (2001) noted that Smith (1980) found that 4-7 year-olds, although having mastered syntax in many ways, including the use of quantifiers, considered *some* as compatible with *all* when they had to respond to questions such as "Do some birds have wings?". Similarly, Braine and Rumain (1981) and Paris (1973) also found 7-9 year-olds tended to interpret "A or B" as compatible with "A and B". These earlier experiments all

show that children tend to interpret weaker scalar terms as strong ones and thus are considered initially more logical than pragmatic in nature.

In a classic experiment, Noveck (2001) gave 7- to 11-year old French-speaking children and adults a series of statements containing scalar inferences of *some* such as “Some giraffes have long necks” and asked them to judge the acceptance of the stimuli sentences. Noveck found that French children and adults behaved differently in their acceptance of the sentences containing *some*. While adults tended to reject the sentences with the use of *some* as *some and possibly all* (41%), children overwhelmingly accepted them (89% for 7-year-olds and 85% for 11-year-olds). Based on Noveck’s results, children are said to be more logical than pragmatic than adults. Similar findings have been found by other researchers as well.

Papafragou and Musolino (2003) questioned that the methodology that Noveck used would mask the true scalar inferences of children. They claimed the results that were found in Noveck’s study might be an artifact of the specific task that Noveck used in the study. In their second experiment, Papafragou and Musolino therefore tried to enhance 30 5-year-old Greek children’s awareness of the experimental goal and found that there was an increase in the acceptance of the pragmatic interpretation of *some* in children’s responses. Nevertheless, when children’s awareness was not enhanced, they still interpret *some* more as *some and possibly all*. What is noted is that children’s rejection rate of pragmatic infelicity increased when children’s attention was directed to pragmatic infelicity.

As the pioneer researchers to use online techniques to examine competing theories as Default Inference and the Relevance theory, Bott and Noveck (2004) measured subjects' reaction time and correct rate when subjects were asked to respond to underinformative stimuli such as "Some elephants are mammals". Logically, this stimuli sentence is considered correct since *some* can be interpreted as *some and possibly all*. Pragmatically, this stimuli sentence is considered wrong if people interpret *some* as *some but not all*. The rationale of Bott and Noveck's experiments was that if Default Inference was correct, people would spend less time when they were asked to judge the stimuli sentence based on the criterion of interpreting *some* as *some but not all*. Nevertheless, if the Relevance theory is correct in predicting scalar inferences, people would spend less time when they are required. In addition to the targeted underinformative sentences, there were also lots of other types of sentences serving as controls such as "Some mammals are elephants", "Some elephants are insects", "All elephants are mammals", "All mammals are elephants" and "All elephants are insects". Bott and Noveck found that subjects took longer time to react when they were required to judge the underinformative sentence pragmatically and were also more prone to making mistakes. Given that subjects spend less time using logical strategies, the results actually were more in line with the Relevance Theory. In the fourth experiment, Bott and Noveck gave subjects either 900 milliseconds or 3000 milliseconds to judge the target and the controls. They found that under time pressure, subjects tended to respond True to underinformative sentences, meaning that it was less likely for them to derive inferences.

Discrepancies exist between Bott and Noveck's results and the results of the previous studies, such as Noveck (2001) and Papafragou and Musolino (2003). Recall that Noveck and other researchers found children using logical inferences more while adults used pragmatic inferences more. If these researchers' results were true, Bott and Noveck's results would be considered surprising. Nevertheless, care must be taken when interpreting Bott and Noveck's results. First of all, subjects in the study were forced to respond to underinformative sentences either logically or pragmatically. If either of the logical or pragmatic interpretations is the default one, how can one make sure that the subjects were really using the non-default one to make responses? Besides, it is also possible that the reaction times were spent on sorting out which instruction to use in the experiment or on fight to resist using the intuitive inferences. Second, in the second experiment where yes and no responses were equally set up as target responses, subjects' reaction times greatly increased to more than 4 seconds, as compared to less than 2 seconds in the first experiment. Spending more than 4 seconds is considered peculiar in sentence comprehension, given that thinking involves rapid processing. If longer than usual processing time was observed, it makes the results look even more suspicious. Third, in the fourth experiment where subjects were either given 900 or 3000 milliseconds, true responses to simple controls such as "Some mammals are elephants" were only 0.79 in both short and long lag. Given that subjects' responses were nearly perfect in the first two experiments when answering "Some mammals are elephants", it is quite strange that the proportion correct dropped so much in the fourth experiment. Last, the stimuli used in the experiments required subjects to apply real world knowledge so as to answer the scalar inferences. For example, subjects need to know the relationship

between elephants and mammals and the fact that elephants actually subsumes under mammals category. Prior to the experiment, the researchers did not make sure that all the subjects were familiarized with these categorization concept. To sum up, considering the real world knowledge might contribute to the long reaction times found in the experiments since subjects may be trying to find out exceptions.

Given that previous studies either did not look at adults' online processing of scalar inferences or were not satisfactory in methodologies, this paper aimed to improve on these part and bridge the theoretical gap by examining the time course of scalar inferences as in scales like <all, some> and the function of working memory capacity with truth value judgment task and reaction times. Detailed description of the three experiments will be presented in the following sections.

Previous studies in second languge

Bouton (1992) used an offline method to examine 30 second language speakers' interpretation of implicature and found that even though his subjects made improvements over time, their interpretation performance was still significantly different than that of native speakers. Several drawbacks occur in Bouton's study. First, the this study was a continuation of an even earlier 1986's study and given that the author decided to use the same group of subjects, it turned out that not many subjects wished to take the same test again. Second, Bouton did not explained the association among four different tests used in his study, which makes us concern about the validity of his study. Recently, Dekydtspotter and Hathorn (2005) Truth value judgment task to investigate French

continuous and discontinuous interpretation of quantifiers among second language learners of French and found that language learners with high language proficiency displayed native-like implicature processes while low-intermediate language learners did not show such tendency. Slabakova (2010) examined the interpretation of some in either felicitous context or the infelicitous context. In felicitous context, one of the test stimuli was like “Some books have color pictures” while in infelicitous context, the test stimuli could be “Some elephants have trunks”. The results showed that Korean learners of English used more pragmatic interpretation in second language processing than first language comprehension.

Given that the existing few studies on second language scalar inferences still rely heavily on truth value judgment tasks or other offline methods, it is of great interest to apply online technique to further examine this issue. The current study used both online reaction time data and truth value judgment task to investigate the following issues: First, which theory, Default inference or Relevance theory, can be used to account for second language learners’ performance in scalar interpretation? Second, would second language learners’ performance change under time pressure? Third, does working memory capacity play a role in processing scalar inferences? We would address these questions in the following three experiments.

Experiment 1

Given that most of the second language studies on scalar inferences did not directly address the issue of Default inference and the Relevance theory, the current experiment

will draw on the findings from the first language research. The purpose of the first experiment was therefore to reexamine Default Inference and Relevance Theory. Instead of forcing subjects to judge stimuli sentences either logically or pragmatically or to apply real world knowledge online, subjects in the first experiments were introduced a context which depicted a condition that could support both a logical or a pragmatic interpretations. Note that all the stimuli presented on the computer screens were in English. For example, a context presented on the computer screen might say, “John has many dictionaries. Some of the dictionaries are used.” Again, two possible interpretations can be read from the second sentence. First, it is possible to interpret *some* as either *some but not all* or *some and possibly all*. Second, no real world knowledge is needed to evaluate the stimuli. After reading these two sentences, subjects were then presented with a question sentence on the screen: “Some and possibly all of the dictionaries are used” or “Some but not all of the dictionaries are used.” A total of 40 set of target sentences were used in the first experiments with 20 of them using “some and possibly all” and 20 using “some but not all”. These target stimuli were then randomized with 40 control sentences so that subjects did not detect the manipulation of the experiment. Controls were designed not to include scalar inferences so that they could serve as the basis in the experiment. A typical control sentences would read, “Jennifer has four borrowed books. They were from the library.” The question sentence says, “Does Jennifer have four borrowed books from the library.”

After reading the question sentence, subjects need to respond by pressing either “True” or “False” button. Two kinds of responses were recorded: subjects’ reaction times and their

responses to the question sentence. Reaction times were measured as the time that subjects took to respond to the question as either true or false.

Predictions

The design of the experiment allowed us to examine Default inference account and Relevance Theory among Mandarin learners of English. According to Default Inference account, less time would be spent on interpreting *some* as *some but not all* since it is the default interpretation. Even though the theorists make predictions based on the first language, second language speakers, if using Default inference in the cognitive system, should spend less time in *some but not all* conditions as well. Relevance Theory, on the contrary, predicts that less time would be observed in interpreting *some* as *some and possibly all* and we would expect second language learners to show the same pattern if Relevance theory is correct. If the results of the first experiment showed less reaction times on *some but not all*, Default Inference account is supported. If, however, the result showed less reaction times on *some and possibly all*, Relevance Theory is supported.

Procedures

Thirty native Mandarin speakers were recruited in the first experiment. They were all college or grad school students, with age ranging from 19 to 30. Since the subjects in the current study were not native speakers of English, care was taken to ensure that all the words and the simple structures were familiar to them. Even though most of the vocabulary belongs to beginners' level, before participating in the experiment, subjects

needed to go through all the vocabulary used in the study and make sure that they understood all of them.

After coming that, subjects were told that they would be processing sentences in English online and their responses will be recorded. Before the actual experiment began, they were presented with instruction on how to do the experiment. During the 15 practice trials, subjects were asked to answer each question as quickly and accurately as possible. On the computer screen, subjects would read stimuli like "John has many dictionaries. Some of the dictionaries are used." This first screen would last for 2500 ms. After that the screen would be replaced by a question either as "Some but not all of John's dictionaries are used" or "Some and possibly all of John's dictionaries are used". Subjects were prompted to make a True/False judgment by pressing right/left shift keys.

Statistical Analysis and Results

Given that the responses in the True/False judgment in the experiment are not continuous, traditional ANOVA analysis should not be applied in this context. Instead, categorical data analysis should be performed. First, simple chi-square analysis was done for initial analysis. The results showed that both Chi-square and Likelihood Ration Chi-square has reached significance ($\chi^2(2)=1436.23$, $p<.00$; Likelihood Ratio (2)= 1413.68, $p<.00$). Furthermore, due to the binary features of the response variables, binary logistic regression was applied for further analysis. SAS output revealed that the convergence criterion has been satisfied. There is a main effect of the answer "yes" among three types of stimuli: *Some but not all*, *some and possibly all* and *control* sentences (Wald

$\chi^2(2)=839.47$, p<.00). In addition, subjects' acceptance of *some but not all* was significantly more than their acceptance of *some and probably all* (Wald$\chi^2(1)=65.22$, p<.00). Similarly, subjects' acceptance of the control sentences was reliably higher than their acceptance of *some and possibly all* (Wald $\chi^2(1)=330.4$, p<.00), as indicated in Figure 1.

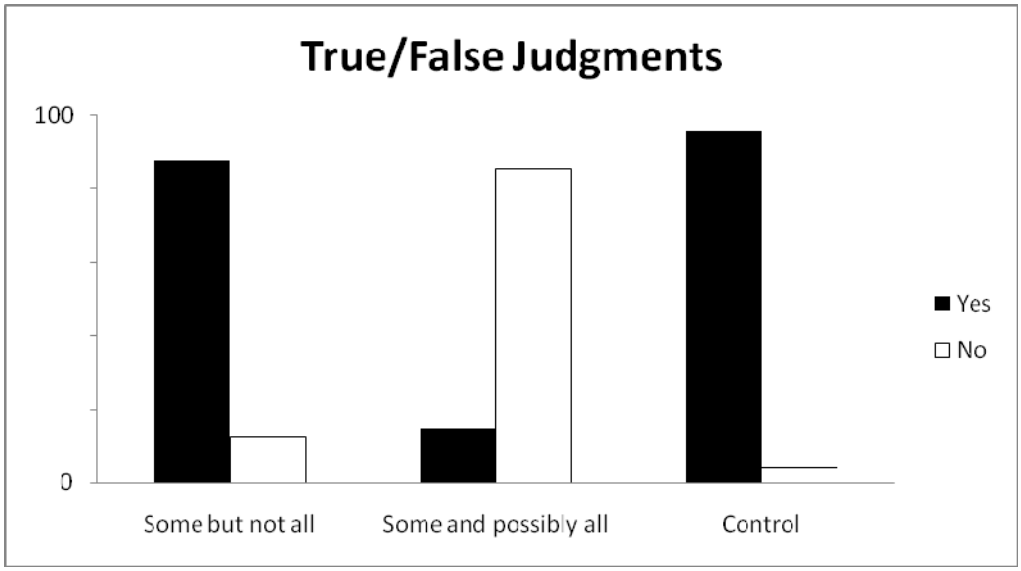

Figure 1: True/False judgments in Experiment 1

The results of the true/false judgments showed that Mandarin second language learners of English had a much higher acceptance rate of responding "yes" to *some but not all* than *some and possibly all*. In fact, their acceptance rate of *some but not all* was very similar to their rejection rate of *some and possibly all* conditions. When subjects were presented with *some and possibly all*, they tended to reject this interpretation 85% while accepting this interpretation only 15%. A further discussion of this acceptance rate will be given in the discussion section.

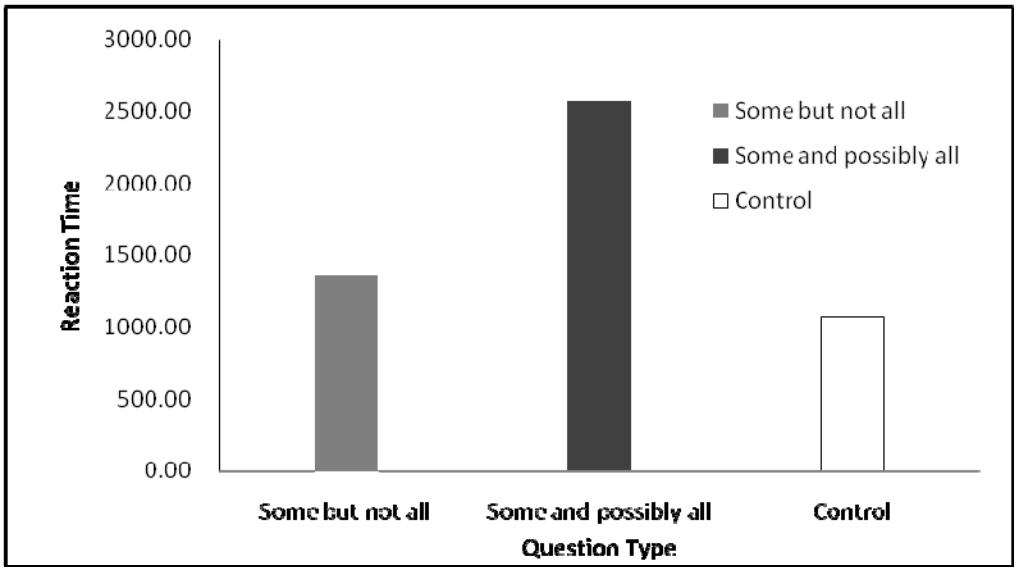

Figure 2: Reaction times in Experiment 1

Given that reaction time data belongs to Gamma distribution, it has to be log-transformed so that it won't violate the basic assumptions of ANOVA analysis. After reaction times were log-transformed, the results showed there is a significance difference in reaction times among three types of conditions, i.e., *some but not all*, *some and possibly all* and *controls* (Subject analysis: $F_1(2, 87)=166.7$, $p<.000$; Item analysis: $F_2(2, 87)=82.26$, $p<.000$). In addition, omnibus post-hoc comparisons revealed that there were significance differences in reaction times in each post-hoc comparison (all subject and item analyses: $p<.01$). Therefore, subjects were faster in accepting *some* as *some but not all* than *some and possibly all* (both subject and item analyses: $p<.000$). Figure 2 shows that subjects only spent around 1350 milliseconds when responding to *some but not all*. However, they spent almost twice as long (average RT: 2500 milliseconds) when responding to *some and possibly all*. Based on the rejection rate in Figure 1, the results indicates even though

second language learners did not accept to interpret *some* as *some and possibly all*, it took them some time to ponder over it. It is possible that they were trying to come up with a scenario that would fit *some and possibly all* interpretation.

When comparing the control sentences with the logical and pragmatic interpretations of some, the results showed that even though subjects were faster in responding to *some but not all*, the reaction times were still significantly longer than the reaction times of the control sentences (both subject and item analyses: $p<.000$). The average time that people spent on responding to control sentences were around one second while they spent 1350 milliseconds on *some but not all* condition.

Discussion

Contrary to the findings in Bott and Noveck (2004), the results of the first experiment showed that Mandarin second language learners spent significantly less time when interpreting *some* as *some but not all* when processing English scalar inferences. When the prompting question was, for example, "Some but not all of John's books were used", subjects took an average of 1350 ms to respond to the question. However, when the prompting question was "Some and possibly all of John's books were used", second language learner needed to spend 2500 ms on average to respond. Even though this finding comes from Mandarin second language learners of English, it can still be used to support the Default theory, rather than the Relevance theory. Note that the results of the first experiment not only supported the Default Inference, they were also in line with the previous experimental results, which found that adults used more pragmatic inferences

than children. Given that the second language subjects in the current study were also adults, the results thus did not contradict with the previous findings.

Regarding the reaction times that subjects took in responding to *some and possibly all*, it has been observed that subjects spent almost twice as much time as they did in responding to *some but not all*. Judging from their higher acceptance rate of *some and possibly all*, we hypothesized that subjects might be trying to come up with a scenario that can fit the situation, therefore lengthening the processing time.

Finally, do scalar inferences take people longer time to process? According to the results of experiment 1, the answer seems to be positive. When scalar inferences such as *some* are involved, people usually needed more time to consider, as reflected by the slightly and still significantly longer reaction times comparing control sentences and *some but not all* conditions. The results can be used to indicate that cognitive resources are required for implicature calculation.

Experiment 2

The results in the first experiment indicated that subjects prefer to interpret *some* as *some but not all*, as reflected by the reaction times and the True/False responses. However, since subjects were not confined to make their responses within a certain period of time, it is arguably true that they might be able to come up with scenarios that are compatible with either *some but not all* or *some and possibly all*. Since the purpose of the experiment

was *not* to let subjects brainstorm so that they can come up with a possible scenario to fit the interpretation, it is therefore necessary to know if subjects still make the same responses under the pressure of time. Experiment 2 thus set out to examine whether subjects interpret *some* as *some but not all* or *some and possibly all* under the pressure of time. It is hypothesized that when people are under pressure to respond within a short range of time, they would not be able to brainstorm to search for a possibility to fit the other interpretation.

The results of the first experiment showed that people tended to spend around 2500 milliseconds when they were presented with *some* as *some and possibly all* and around 1300 milliseconds when presented with *some* as *some but not all*. It seems that when presented with *some and possibly all*, subjects needed another 1200 milliseconds to evaluate the possibility of this interpretation of some. Experiment 2 thus gave subjects 1500 milliseconds to respond to the stimuli sentences. This was to make sure that subjects could still process the test stimuli but couldn't use real world knowledge or other techniques to come up with a possible scenario.

The prediction is that when presented with *some but not all* interpretation, subjects' reaction times and the true/false judgments in Experiment 2 should approximate those in Experiment 1. However, for *some and possibly all* condition, given that subjects did not have enough time to search for an interpretation to fit the unlikely situation, the acceptance rate of some and possibly all will drop.

Procedures

The procedure of experiment 2 was the same as experiment 1. Thirty native Mandarin college or grad school students were recruited in experiment 2. Again, they were asked to go through a vocabulary list to make sure that they knew all the words in the experiment. After that, subjects would be seated in front of a computer. They were then presented with instruction on how to do the experiment at the beginning of the session. Instead of allowing subjects to judge *some but not all* and *some and possibly all*, subjects were told that they only had 1500 milliseconds to respond to the question. Again, they were also asked to answer each question as quickly and accurately as possible. Fifteen trials of practices were given to subjects to familiarize them with the experiment.

Statistical Analysis and Results

Like experiment 1, a simple Chi-square analysis was done first. The results showed that both Chi-square and Likelihood Ration Chi-square has reached significance ($\chi^2(2)=1729.6644$, $p<.00$; Likelihood Ratio (2)= 1720.98, $p<.00$). Afterwards, binary logistic regression was applied for data analysis. SAS output revealed that the convergence criterion has also been satisfied. There is a main effect of the answer "yes" among three types of stimuli: Some but not all, some and possibly all and control sentences (Wald $\chi^2(2)=843.11$, $p<.00$). In addition, subjects' acceptance of *some but not all* was significantly more than their acceptance of *some and probably all* (Wald$\chi^2(1)=92.02$, $p<.00$). By the same token, subjects' acceptance of the control sentences was reliably higher than their acceptance of *some and possibly all* (Wald $\chi^2(1)=293.77$, $p<.00$), as indicated in Figure 3.

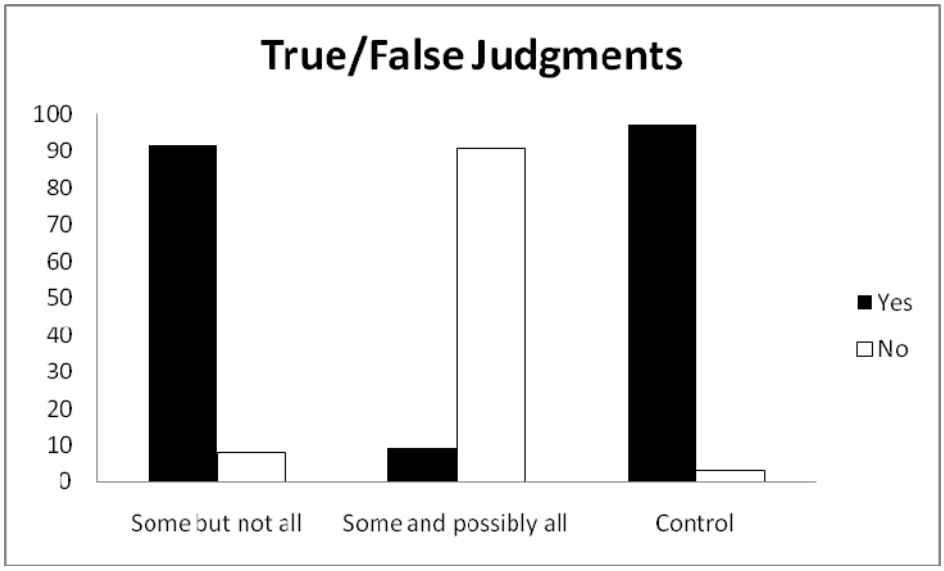

Figure 3: True/False Judgments in Experiment 2

The results of the true/false judgments in Experiment 2 showed that when second language learners were pressed for time, their performance on processing *some and possibly all* becomes even more distinct in that its rejection rate jumps to almost the same as the acceptance rate of *some but not all*.

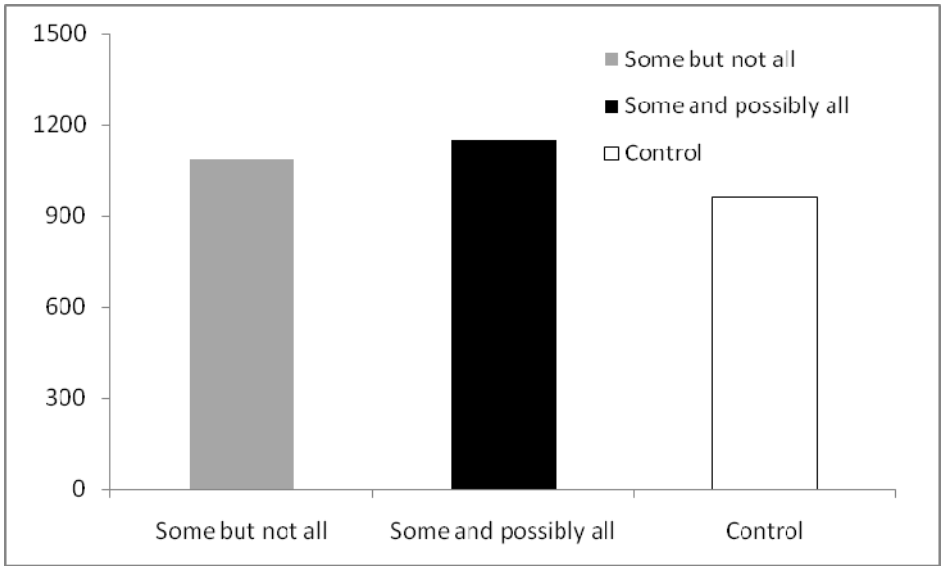

Figure 4: Reaction times in Experiment 2

If we compare Figure 2 and Figure 4, the first thing we will notice is that the reaction times for *some and possibly all* condition have significantly dropped. As a matter of fact, the reaction times for all the three types of sentences dropped considerably, probably due to the time pressure. Still, there was a significant differences in reaction time among the three types of interpretations (Subject analysis: $F_1(2, 87)=5.44$, $p<.01$; Item analysis: $F_2(2, 77)=8{,}211$, $p<.01$). Post-hoc analysis further revealed that there are reliable difference between *some but not all* and *control* group (Subject analysis: $p < .05$; Item analyses: $p <.01$) and also between *some and possibly all* and the *control* group (Subject analysis: $p < .00$; Item analyses: $p <.00$). Nevertheless, there was no reliable difference in reaction times between *some but not all* and *some and possibly all*.

Discussion

The second experiment set out to investigate whether second language learners' responses would change if they were pressed for time to respond. It has been hypothesized that during the experiment, subjects might be trying to come up with a scenario that can fit the alternative interpretation so as not to reject the stimuli question. Based on the reaction time results of the first experiment, the second experiment only gave subjects 1500 milliseconds to make their responses. Under the time pressure, subjects would not have enough time to generate a compatible scenario and would thus follow their intuitions more.

The reaction times in experiment 2 showed a sufficient drop of reaction time in subjects' responses when responding to *some and possibly all*. Statistics revealed that there is no difference in reaction times between *some but not all* and *some and possibly all*. This seems at first peculiar given that the results of the first experiment showed much lengthened time over *some and possibly all* conditions. However, this situation can be attributed to the fact that since these subjects were not native speakers, they were concerned about making responses in time.

The more interesting part lies in the true/false judgment where we could observe an increase in rejection rate of *some and possibly all* conditions, now almost as high as the acceptance rate of *some but not all* condition. The increasing rejection rate of *some and possibly all* and acceptance rate of *some but not all* indicate subjects' preference to interpret *some* remains to be *some but not all*. In addition, even though subjects were very much pressed in time to make responses, a reliable difference still appears when we

compared the control group with the other two groups. This can be used to indicate that scalar processing still requires more cognitive capacity than simple structures.

The results of experiment 2 further support the Default Inference account. Even though subjects were limited in time to make responses, their reaction times for *some and possibly all* were still significantly longer than the reaction times for *some but not all*. Interpreting *some* as *some and possibly all* thus seems to be more onerous than interpreting it as *some but not all*.

Experiment 3

The second experiment has shown that subjects' calculation of scalar inferences were influenced by the pressure of time. Since the time it takes to process language also correlates with cognitive ability, it thus poses another interesting question: whether the capacity of cognitive system plays a role in scalar inferences. To be more specific, we want to know if there is a difference in people's derivation of scalar inferences with regard to high working memory capacity versus low working memory capacity. Before delineating how to approach this research question, we will first briefly introduce working memory with regard to language processing.

Working Memory

Working memory (Baddeley, 1992) has long been a theoretical and central construct in cognitive psychology due to its methodological merits of measuring human working

memory capacity and its involvement in many complex cognitive behaviors, such as comprehension, reasoning and problem solving. With working memory, human beings are able to store and retrieve information, and keep important information active and accessible under various circumstances.

To see whether working memory capacity vary across people, Daneman and Carpenter (1980) developed a reading span task following Baddeley and Hitch's thinking that there are two major components in working memory: processing and storage. Subjects in this reading span task read aloud a series of different sets of sentences starting with two sentences while maintaining the last word in their working memory. When they finish the target set, subjects are required to recall the last words. If they succeed in recalling the words, they are presented with increasing longer sequence of sentences. Their working memory capacity is reflected by the number of items they can recall correctly in serial order. Daneman and Carpenter have shown that the reading span task correlates well with subjects' reading proficiency. Since Daneman and Carpenter's reading span task, a variety of working memory tasks has been developed to suit different purposes. For example, Waters, Caplan and Hildebrandt (G. Waters, Caplan, & Hildebrandt, 1987g) modified Daneman and Carpenter's reading span task by asking subjects to silently judge the acceptability of the sentence stimuli on the computer before writing down the last words. In addition, other kinds of span tasks have also been developed, including operation span task (Turner & Engle, 1989), counting span task (Case, Kurland, & Goldberg, 1982), and spatial span tasks (Shah & Miyake, 1996). Researchers also found

that these tasks also correlate highly with various cognitive abilities, such as language comprehension, learning, problem solving and reasoning.

Working memory and language comprehension

One of the major questions in psycholinguistic research concerns the nature of working memory capacity in sentence comprehension. During an incoming string of words, comprehenders need to get the lexical meaning of each word, build syntactic structures, and assign thematic roles. Unfortunately, not all these stages can be done simultaneously. Some of the initial items need to be held temporarily in the working memory to be matched with latter words. As indicated by a growing body of research (Caplan & G. Waters, 1998a; Just & Carpenter, 1992; Kolk, Chwilla, van Herten, & Oor, 2003; MacDonald & Christiansen, 2002; MacDonaldMarcel Adam & Carpenter, 1992), working memory resources help comprehenders process and store and integrate syntactic constituents in sentences like the following:

1. The athlete in the pub next to the school was hit by a stranger last night.
2. It was the athlete that the crowd in the stadium sneered last night.
3. The athlete that the stranger hit is currently in the police station.
4. The athlete that hit the stranger is currently in the police station.

In each of the sentences above, comprehenders need to temporarily store the noun phrase *the athlete* in the working memory in order to integrate it with the latter syntactic structure. In (1), comprehenders need to temporarily store *the athlete* and integrate it with

a passive structure. In (2), *the athlete* has to be held in the working memory long enough until it is integrated with *sneered*. Sentences (3) and (4) represent English relative clauses and even though these two sentences look alike, they require comprehenders to hold *the athlete* to different extent so that the thematic roles can be matched. In fact, many researchers (Caplan & G. Waters, 1998b; Ferreira, Anes, & Horine, 1996; Ferreira, Henderson, Anes, Weeks, & McFarlane, 1996; Frazier & Rayner, 1982) have indicated that complex syntactic structures are more difficult to process and people's processing time increases as a function of increasing processing load.

A number of studies have found that individual differences in working memory span constrains sentence comprehension (Friederici, Steinhauer, Mecklinger, & M. Meyer, 1998c; MacDonaldMarcel Adam & Carpenter, 1992; Mecklinger, Schriefers, Steinhauer, & Friederici, 1995; Osterhout & Holcomb, 1992d; G. S. Waters & Caplan, 1996e, 1996f). For example, when using ambiguous sentences as stimuli, MacDonald et al. (1992) claimed that subjects with high reading span were able to hold multiple representations that are compatible with ambiguous regions than subjects with low reading span. However, this result was not supported by some of the event-related potential findings. Mecklinger et al. (1995) found more positivity shift in garden-path sentences among subjects with higher reading span. This result suggests that high span readers might be more likely to be led down the garden path during ambiguous sentences and is incompatible with the previous suggestion by MacDonald et al (1992).

Procedures

The procedure of experiment 3 was the same as the procedure in experiment 2 except that every subject needed to complete a working memory test before participating in the scalar inference experiment. The procedure of working memory capacity followed Daneman and Carpenter's reading span task. Subjects were presented with a series of English sentences on a computer screen. The first set started with two sentences. After responding to the comprehension questions, subjects needed to repeat the very last word of each sentence in correct order. If they succeeded in repeating them, they would be presented with a set with three sentences. The trial would continue until they fail to repeat the last words in correct order. Following the criteria of previous experiments, we labeled subjects with reading span below 3 as Low and those with reading span above 3 as High.

Twenty native Mandarin students were recruited in experiment 3. Among these 20 subjects, 8 of them were labeled as High reading span group while 12 of them were labeled as Low reading span group. The rest of the procedure was like that in experiment 1.

Statistical analysis and results

Most of the procedures of analyzing the true/false judgment data and reaction time data were like the previous two experiments, except that in experiment 3, another between-group independent variable, i.e., reading span, was included in the analysis.

Again, simple chi-square analyses were performed for both high and low reading span subjects. In high reading span subjects, there was significant differences in responding

"yes" to stimuli sentences (χ^2 (2)=1284.99, p<.00; Likelihood Ratio χ^2 (2)=1194.24, p<.00). In low reading span readers, there was also significant differences in producing "yes" to stimuli sentences (χ^2 (2)=1300.02, p<.00; Likelihood Ratio χ^2 (2)=1286.37, p<.00).

Afterwards, logistic regression were performed with predictor variables set as sentence types (i.e., "Some but not all", "Some and possibly all" and "Control") and reading span (i.e., High vs Low reading spans). There was main effects for both sentence type (Waldχ^2 (3)=1572.39, p<.00) and reading span (Waldχ^2 (1)=29.28, p<.00), the latter of analyses indicated that high span readers produced significant more "yes" than low span readers. In multiple comparisons, subjects' acceptance of "Some but not all" was reliably higher than their acceptance of "Some and possibly all" (χ^2 (1)=67.35, p<.00). There was also a reliable higher acceptance rate of "Control" than "Some and possibly all" (χ^2 (2)=536.68, p<.00). Last, significant difference was found between "Some but not all" and "Control" (χ^2 (2)=106.8, p<.00). The graph is shown below.

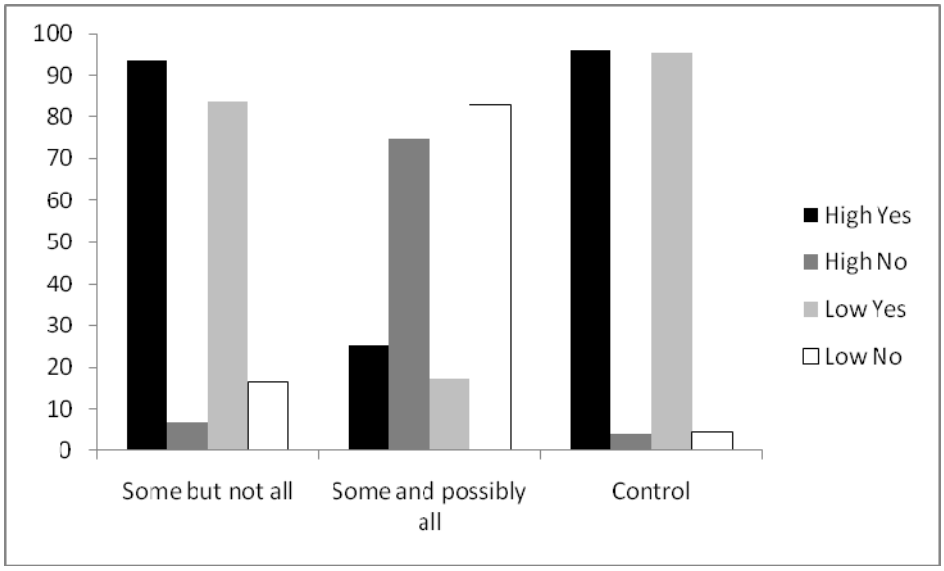

Figure 5: High vs Low working memory responses

All reaction time data were log-transformed before analysis. There were reliable reaction time differences among three stimuli sentences (Subject analysis: $F_1(2)=142.84$, $p<.00$; Item analysis: $F_2(2)=200.46$, $p<.00$). In post-hoc multiple comparisons, there were significant results in all three pairwise comparisons (all subject and item analyses: $p<.00$). In addition, there was also a main effect in reading span (Subject analysis: $F_1(2)=11.04$, $p<.01$; Item analysis: $F_2(2)=19.28$, $p<.00$), which indicates that high span readers were faster in making responses during the experiment.

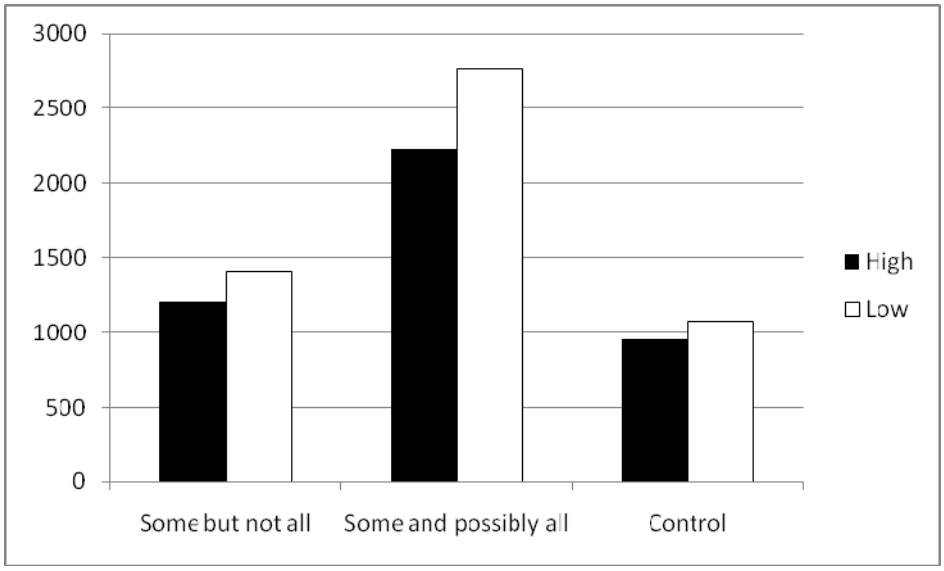

Figure 6: Reaction time in Experiment 3

It is also interesting to compare high versus low span readers' performance under each sentence type category. In "Some but not all", there was a marginally reliable differences in reaction times between high vs low readers ($F(1)=3.457$, $p=.06$). In "Some and possibly all", high span readers were much faster than low span readers ($F(1)=7.52$, $p<.01$). However, in "Control" group, the difference between the two groups was not significant ($F(1)=1.332$, $p=.253$).

Discussion

Experiment 3 aimed to examine whether cognitive capacity influence second language learners' processing of sentences containing scalar inferences. Previous studies have found that subjects with high reading span were faster in processing sentences than subjects with low reading span. In the current study, differences in reaction times and truth value judgments between high and low reading spans were also found.

In true/false judgments, high reading span subjects were more likely to accept *some but not all* interpretation than low span readers, although both of the two groups showed very high acceptance rate of this condition. Besides, high span readers were much faster than low span readers in accepting *some but not all* interpretations, as reflected by the significant differences in reaction time data.

What is more interesting and should be noted is that in *some and possibly all* conditions, high span second language learners were more likely than low span readers to accept this interpretation. The percentage for high span reader to accept some and possibly all was 25% while it was only 17% for low span readers. A legitimate account of such high acceptance rate in this condition is that high span readers had more cognitive capacities to come up with possible scenario to fit the *some and possibly all* condition. Both high and low reading span subjects showed lengthened reaction times in the *some and possibly all* condition. Despite so, high span readers still seemed to be a lot faster than low span readers. Finally, in control sentences, both high and low span readers were very fast in accepting the sentences and there was no significant differences in their responses.

Overall, experiment 3 found that working memory capacity did play a significant role in deriving scalar inferences among second language learners. When people have more capacity in working memory, they are faster in evaluating sentences and at the same time, they are more likely to come up with possibilities to fit the scenario.

Overall Discussion

This current study tried to investigate the scalar inferences among Mandarin second language learners of English by using experimental approaches that have rarely been used in studying pragmatic issues. Some of the studies working on scalar inferences have found that children use more logical interpretation when interpreting scalar expressions like “some” while adults tend to use more pragmatic interpretations. In the sentences like “Some of the students were absent”, adults tend to interpret it as “Some but not all of the students were absent” while for children, they tend to interpret it as “Some and possibly all students were absent”. Bott and Noveck, however, seemed to have found the opposite pattern. Despite some discoveries found in first language literature, surprising even fewer studies have been trying to look at this issue. Thus the current paper serves as one of the pioneer research to bridge the gap between first and second language scalar comprehension.

By using stimuli sentences that prevent subjects from using real world knowledge, the present study found that across three experiments, second language learners’ acceptance rate of *some but not all* was higher than their acceptance rate of *some and possibly all*. This evidence was also backed up by faster reaction times in the online measures. It indicates that subjects did not hesitate in interpreting *some* as *some but not all*. The overall reaction times for *some but not all* was about half as fast as that for *some and possibly all*. Note that such performance is very close to the performance of first language speakers. Despite the fact that second language learners used much more time

on interpreting the test stimuli, they still overwhelming chose *some but not all* interpretation as the default value for *some*.

Another interesting issue that can be observed in the current study is that scalar inferences did consume extra more time than simple structures. This can be found in the results that even though subjects' reaction times in *some but not all* were faster than *some and possibly all*, they were still slower than the reaction times of *control* stimuli. The difference between *control* stimuli and *some but not all* was that there was no scalar inferences involved in the *control* stimuli. This indicates that scalar inferences require extra efforts in cognitive abilities.

Theoretically, the faster reaction times and higher acceptance rate of some but not all could be used as evidence to support the Default Inferences, which hypothesized that scalar inferences are generated by default. Given that the reaction times in *some but not all* were faster than *some and possibly all* and were more similar to the reaction times for control stimuli, it can be used to indicate that the scalar inferences were derived by default, rather than arise during the requirement of the context, as indicated by the Relevance theory. In addition, the extra reaction times in *some and possibly all* conditions can be used to explain that when the default interpretation is not correct, people take time to cancel such interpretation.

Further evidence in support of Default account comes from the results of experiment 2, which limited subjects' responses to only one second. In this situation, subjects' rejection

rate of *some and possibly all* increased a lot than that in experiment 1. This indicates that when people do not have enough time to try to come up with a suitable scenario to fit *some and possibly all*, they tend to reject it. The acceptance rate of *some but not all* remained about the same.

Besides finding compelling evidence in support for the Default Inference account, this paper also adds into literature of experimental pragmatics the importance of working memory capacity. The results in experiment 3 showed that learners with high reading span are more likely to process scalar inferences faster and also more likely to generate possible scenario to fit *some and possibly all*, as indicated by its higher acceptance rate. However for subjects with low reading span, they were slower and possibly more prone to making pragmatic mistakes. Taken all together, the three experiments in this paper demonstrate that experimental procedures help tease apart the competing theories raised by pragmatic researchers. What's more, they also contribute to our knowledge of current pragmatic theories.

References

Baddeley, A. (1992). Working memory. *Science, 255*(5044), 556.

Bley-Vroman, R. (2009). The evolving context of the Fundamental Difference Hypothesis. *Studies in Second Language Acquisition, 31*(02), 175–198.

Bott, L., & Noveck, I. A. (2004). Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of memory and language*, *51*(3), 437–457.

Bouton, L. F. (1992). The Interpretation of Implicature in English by NNS: Does It Come Automatically–Without Being Explicitly Taught?. *Pragmatics and language learning*, *3*, 15.

Braine, M., & Rumain, B. (1981). Children's comprehension of "or": evidence for a sequence of competencies. *Journal of Experimental Child Psychology*, *31*, 46–70.

Caplan, D., & Waters, G. (1998a). Verbal working memory and sentence comprehension.

Caplan, D., & Waters, G. (1998b). Verbal working memory and sentence comprehension.

Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology*, *33*(3), 386–404.

Chee, M. W. L., Caplan, D., Soon, C. S., Sriram, N., Tan, E. W. L., Thiel, T., & Weekes, B. (1999). Processing of visually presented sentences in Mandarin and English studied with fMRI. *Neuron*, *23*(1), 127–137.

Chee, M. W. L., Tan, E. W. L., & Thiel, T. (1999). Mandarin and English single word processing studied with functional magnetic resonance imaging. *The Journal of Neuroscience*, *19*(8), 3050.

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, *27*(01), 3–42.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of verbal learning and verbal behavior*, *19*(4), 450–466.

Dehaene, S., Dupoux, E., Mehler, J., Cohen, L., Paulesu, E., Perani, D., Van de Moortele, P. F., et al. (1997). Anatomical variability in the cortical representation of first and second language. *Neuroreport*, *8*(17), 3809.

Dekydtspotter, L., & Hathorn, J. C. (2005). Quelque chose… de remarquable in English-French acquisition: mandatory, informationally encapsulated computations in second language interpretation. *Second Language Research*, *21*(4), 291.

Ferreira, F., Anes, M. D., & Horine, M. D. (1996). Exploring the use of prosody during language comprehension using the auditory moving window technique. *Journal of Psycholinguistic Research*, *25*(2), 273–290.

Ferreira, F., Henderson, J. M., Anes, M. D., Weeks, P. A., & McFarlane, D. K. (1996). Effects of lexical frequency and syntactic complexity in spoken-language comprehension: Evidence from the auditory moving-window technique. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(2), 324.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive psychology*, *14*(2), 178–210.

Friederici, A. D., Steinhauer, K., Mecklinger, A., & Meyer, M. (1998c). Working memory constraints on syntactic ambiguity resolution as revealed by electrical brain responses. *Biological Psychology*, *47*(3), 193–221.

Horn, L. R. (1972). On the semantic properties of logical operators in English.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological review*, *99*(1), 122.

Klein, D., Milner, B., Zatorre, R. J., Meyer, E., & Evans, A. C. (1995). The neural substrates underlying word generation: A bilingual functional-imaging study. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(7), 2899.

Klein, D., Milner, B., Zatorre, R. J., Zhao, V., & Nikelski, J. (1999). Cerebral organization in bilinguals: a PET study of Chinese-English verb generation. *NeuroReport*, *10*(13), 2841.

Kolk, H. H. J., Chwilla, D. J., van Herten, M., & Oor, P. J. W. (2003). Structure and limited capacity in verbal working memory: A study with event-related potentials. *Brain and Language*, *85*(1), 1–36.

Levinson, S. C. (1983). Pragmatics (Cambridge textbooks in linguistics).

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996).

MacDonaldMarcel Adam, M. C., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology*, *24*(1), 56–98.

Master, P. (2002). Information structure and English article pedagogy. *System*, *30*(3), 331–348.

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.

Noveck, I. A. (2001). When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition*, *78*(2), 165–188.

Osterhout, L., & Holcomb, P. J. (1992d). Event-related brain potentials elicited by syntactic anomaly. *Journal of memory and language*, *31*(6), 785–806.

Osterhout, L., & McLaughlin, J. (2000). What brain activity can tell us about second-language learning. *13th Annual CUNY Conference on Human Sentence Processing*.

Papafragou, A., & Musolino, J. (2003). Scalar implicatures: Experiments at the semantics-pragmatics interface. *Cognition*, *86*(3), 253–282.

Paris, S. G. (1973). Comprehension of language connectives and propositional logical relationships. *Journal of Experimental Child Psychology*, *16*(2), 278–291.

Perani, D., Paulesu, E., Galles, N. S., Dupoux, E., Dehaene, S., Bettinardi, V., Cappa, S. F., et al. (1998). The bilingual brain. Proficiency and age of acquisition of the second language. *Brain*, *121*(10), 1841.

Schwartz, B. D. (1998). The second language instinct. *Lingua*, *106*(1-4), 133–160.

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General*, *125*(1), 4.

Slabakova, R. (2010). Scalar implicatures in second language acquisition. *Lingua*, *120*(10), 2444–2462.

Smith, C. L. (1980). Quantifiers and question answering in young children. *Journal of Experimental Child Psychology*, *30*(2), 191–205.

Sperber, D., & Wilson, D. (2004). Relevance theory. *Handbook of Pragmatics. Oxford: Blackwell*, 607–632.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*(2), 127–154.

Waters, G. S., & Caplan, D. (1996e). The capacity theory of sentence comprehension: Critique of Just and Carpenter (1992).

Waters, G. S., & Caplan, D. (1996f). Processing resource capacity and the comprehension of garden path sentences. *Memory & cognition*, *24*(3), 342–355.

Waters, G., Caplan, D., & Hildebrandt, N. (1987g). Working memory and written sentence comprehension.

White, L. (1988). Island effects in second language acquisition. *Linguistic theory in second language acquisition*, 144–172.

國科會補助計畫衍生研發成果推廣資料表

日期:2013/10/30

國科會補助計畫	計畫名稱: 中文及英文量詞『一些』與『全部』的推論歷程：閱讀時間及事件相關電位研究
	計畫主持人: 林祐瑜
	計畫編號: 101-2410-H-002-210- 學門領域: 心理語言學

無研發成果推廣資料

101 年度專題研究計畫研究成果彙整表

計畫主持人：林祐瑜	計畫編號：101-2410-H-002-210-
計畫名稱：中文及英文量詞『一些』與『全部』的推論歷程：閱讀時間及事件相關電位研究	

成果項目			量化			單位	備註（質化說明：如數個計畫共同成果、成果列為該期刊之封面故事...等）
			實際已達成數（被接受或已發表）	預期總達成數(含實際已達成數)	本計畫實際貢獻百分比		
國內	論文著作	期刊論文	0	0	100%	篇	
		研究報告/技術報告	0	0	0%		
		研討會論文	0	0	0%		
		專書	0	0	0%		
	專利	申請中件數	0	0	0%	件	
		已獲得件數	0	0	0%		
	技術移轉	件數	0	0	0%	件	
		權利金	0	0	0%	千元	
	參與計畫人力（本國籍）	碩士生	0	0	0%	人次	
		博士生	0	0	0%		
		博士後研究員	0	0	0%		
		專任助理	0	0	0%		
國外	論文著作	期刊論文	0	2	100%	篇	
		研究報告/技術報告	0	0	0%		
		研討會論文	0	0	0%		
		專書	0	0	0%	章/本	
	專利	申請中件數	0	0	0%	件	
		已獲得件數	0	0	0%		
	技術移轉	件數	0	0	0%	件	
		權利金	0	0	0%	千元	
	參與計畫人力（外國籍）	碩士生	0	0	0%	人次	
		博士生	0	0	0%		
		博士後研究員	0	0	0%		
		專任助理	0	0	0%		

其他成果 (無法以量化表達之成果如辦理學術活動、獲得獎項、重要國際合作、研究成果國際影響力及其他協助產業技術發展之具體效益事項等，請以文字敘述填列。)	無

	成果項目	量化	名稱或內容性質簡述
科教處計畫加填項目	測驗工具(含質性與量性)	0	
	課程/模組	0	
	電腦及網路系統或工具	0	
	教材	0	
	舉辦之活動/競賽	0	
	研討會/工作坊	0	
	電子報、網站	0	
	計畫成果推廣之參與（閱聽）人數	0	

國科會補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估

■達成目標

□未達成目標（請說明，以 100 字為限）

□實驗失敗

□因故實驗中斷

□其他原因

說明：

2. 研究成果在學術期刊發表或申請專利等情形：

論文：■已發表 □未發表之文稿 □撰寫中 □無

專利：□已獲得 □申請中 ■無

技轉：□已技轉 □洽談中 ■無

其他：(以 100 字為限)

本研究因涉及中文母與使用者及外與使用者兩個現象之探討，故分成兩部分投稿。一部份已被期刊(Journal of psycholinguistic Research)接受，另一部份仍在撰寫中。

3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）(以 500 字為限)

Journal of Psycholinguistic Research 是心理語言學界相當重要的 SSCI 期刊之一。本國科會執行計畫之其一部份能被接受令人振奮。尤其是量詞研究至今尚無太多實驗方面的數據。本研究室少數語用學領域中採用實驗科學法平衡探討中文量詞的使用特性。此外，本文並用實驗結果解決理論中懸而未決的地方。