

The computation and suspension of presuppositions by L1-Mandarin Chinese L2-English speakers

Second Language Research

1–27

© The Author(s) 2021

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/0267658321993873

journals.sagepub.com/home/slr**Shuo Feng** 

Peking University, Beijing, China

Abstract

The Interface Hypothesis proposes that second language (L2) learners, even at highly proficient levels, often fail to integrate information at the external interfaces where grammar interacts with other cognitive systems. While much early L2 work has focused on the syntax–discourse interface or scalar implicatures at the semantics–pragmatics interface, the present article adds to this line of research by exploring another understudied phenomenon at the semantics–pragmatics interface, namely, presuppositions. Furthermore, this study explores both inference computation and suspension via a covered-box picture-selection task. Specifically, this study investigates the interpretation of a presupposition trigger *stop* and *stop* under negation. The results from 38 native English speakers and 41 first language (L1) Mandarin Chinese learners of English indicated similar response patterns between native and L2 groups in computing presuppositions but not in suspending presuppositions. That is, L2 learners were less likely to suspend presuppositions than native speakers. This study contributes to a more precise understanding of L2 acquisition at the external interface level, as well as computation and suspension of pragmatic inferences.

Keywords

Chinese, pragmatic inferences, presuppositions, semantics–pragmatics interface, suspension

1 Introduction

Recent trends in second language (L2) acquisition have led to a proliferation of studies that examine persistent non-native-like performance by L2 speakers, even at the highly proficient level. The issue of interfaces has received considerable attention in that studies have shown that semantics or syntax alone is less problematic for L2 learners than their interfaces with other components of the language (Sorace, 2005; Sorace and Filiaci, 2006).

Corresponding author:

Shuo Feng, Institute of Linguistics and Applied Linguistics, School of Foreign Languages, Peking University, 5 Yiheyuan Road, Beijing, 100871, China

Email: shuo.feng@pku.edu.cn

Highlighting the difficulties of integration and mapping between different levels of systems, the Interface Hypothesis (Sorace, 2011; Sorace and Filiaci, 2006) was proposed to address L2 acquisition at different interfaces. According to this hypothesis, linguistic properties at internal interfaces (involving mappings of different modules of the grammar) are possibly acquired by L2 learners at a native-like level, whereas properties at external interfaces (i.e. grammar interactions with other cognitive systems) are argued to be more difficult to acquire and compute, even at the phase of ultimate attainment.

Regarding the learnability issue of external interfaces, there is mounting evidence that supports the vulnerability of external interfaces at the end state of L2 acquisition (Tsimplici and Sorace, 2006; Valenzuela, 2006). Other studies challenging the Interface Hypothesis have suggested that linguistic properties at external interfaces, especially at the syntax–discourse interface, are acquirable by L2 learners to a native-like level, despite possible delays (Destruel and Donaldson, 2017; Ivanov, 2012; Rothman, 2009). Since the first publication investigating L2 acquisition of scalar implicatures at a new interface, namely, the semantics–pragmatics interface (Slabakova, 2010), several attempts have been made to suggest that computing scalar implicatures is not a problem for L2 speakers (Feng and Cho, 2019; Lieberman, 2009; Miller, et al., 2016; Slabakova, 2010; Snape and Hosoi, 2018). In fact, L2 speakers are more likely to generate scalar inferences than native speakers. Slabakova (2010) indicated that the high rate of computing scalar implicatures by L2 speakers may be due to difficulties and challenges L2 speakers face in suspending inferences. However, the issue of suspending pragmatic inferences by L2 speakers has scarcely been tested. Therefore, by employing a covered-box paradigm, this article examines not only the computation of pragmatic inferences but also, more importantly, the suspension or cancellation of pragmatic inferences by L2 speakers.

In addition to addressing L2 inference suspension, this article has another objective. In contrast with scalar implicatures, presuppositions, also at the semantics–pragmatics interface, have received much less attention in the L2 literature. This study aims to investigate L2 acquisition of presuppositions by Mandarin Chinese learners of English. While both are at the same semantics–pragmatics interface, scalar implicatures and presuppositions are different: scalar implicatures arise from the hearer’s reasoning about what is said and what is not said by the speaker (Grice, 1975; Horn, 1972); presuppositions are conventionally encoded in the lexical entries of the expression and mutually accepted by the interlocutors (Beaver and Geurts, 2011; Heim, 1982; Karttunen, 1974; Schwarz, 2015; Stalnaker, 1974). Psycholinguistic studies have shown that differences between scalar implicatures and presuppositions also exist in behavioral and processing measures (e.g. reading times), as well as between English-speaking adults and children (Bill et al., 2016; Chemla and Bott, 2013; Huang and Snedeker, 2009; Noveck, 2001; Romoli and Schwarz, 2015, among others).

This study aims to contribute to the growing area of L2 research on the acquisition of pragmatic inferences by exploring an understudied domain in L2 acquisition, i.e. presuppositions. In addition to acquiring the lexical expression of the presupposition trigger investigated in the present study, L2 speakers with mature cognition and pragmatic abilities have no need to acquire something brand-new or absent in their first language (L1). Given that the semantics and pragmatics of the target presupposition trigger are taken to be universal (Grice, 1989; Simons, 2006; von Stechow and Matthewson, 2008¹), the learning task

for L2 speakers is to map the already available interpretation that is mapped similarly in L1 to the new L2 lexical item. Therefore, any L1~L2 discrepancy in processing presuppositions provides an exciting opportunity to advance our knowledge of how L2 speakers process and integrate different readings at the semantics–pragmatics interface.

II Presuppositions at the semantics–pragmatics interface

I *The inferences from presuppositions*

Presuppositions are defined as common background assumptions or assumed preconditions of an utterance. They do not convey any new information but rather background information that interlocutors take for granted (Stalnaker, 1973, 1974). They are different from the assertive content, which is the main point of an utterance. For example, the assumption or presupposition of (1) is *John used to go to school* while the assertion is that *John was not going to school after a certain point of time*. Presuppositions are usually derived by lexical items or linguistic constructions, which are called presupposition triggers.

- (1) John stopped going to school.

Presuppositions attract linguists' attention, first of all, because presuppositions are ubiquitous. That is, lexical items and constructions are widely accepted to be presupposition triggers. For instance, lexical triggers are definite descriptions (e.g. *the*), factive predicates (e.g. *know*, *regret*), aspectual predicates (e.g. *stop*), iteratives (e.g. *again*, *return*) and implicative predicates (e.g. *manage*). Constructional or structural triggers are temporal clauses, cleft sentences and counterfactual conditionals. This article focuses on the inference of presupposition stemming from the lexical trigger *stop*, a change-of-state predicate, which indicates that the activity expressed in the verb phrase had been going on prior to a certain time in the context.²

There are three key characteristics of presuppositions (Stalnaker, 1973, 1974). First, presuppositions do not convey any new information but rather background information that interlocutors take for granted. Second, unlike assertions, presuppositions are not affected by the embedded environment and can still be 'projected' to the entire content of the utterance (Chierchia and McConnell-Ginet, 1990). Consider the following sentences (2a–c):

- (2) a. John didn't stop going to school.
 b. Did John stop going to school?
 c. If John stopped going to school, then he should be happy.

The presupposition *John used to go to school* does not change regardless of the fact that the embedded situation is under negation (2a), in a question (2b) or in a conditional clause (2c). The presence of the presupposition across all linguistic operators is referred to as the presupposition projection. Among those linguistic operators, the question of how presuppositions interact with negation is extremely interesting to researchers in

that presuppositions can be absent under negation in some conditions. The third characteristic of presuppositions is defeasibility, i.e. the absence of presuppositional inferences in certain environments, at least at the global level.

Presuppositions can be suspended or cancelled when the embedded environment is under negation. Let us consider the following example (where ‘~’ indicates a presupposition).

- (3) a. The King of France is wise.
 b. The King of France is not wise.
 c. The King of France is not wise; (in fact), there is no King of France!
 d. ~ There is a King of France.

(Birner, 2013: 158)

The utterance (3a) has the presupposition (3d), and when (3a) is negated as in (3b), the presupposition remains the same. The question of defeasibility comes from (3c), in which the presupposition is cancelled in the second clause. However, we should note the marked nature of (3c): this sentence occurs in a very specific context. In Horn’s (1985) analysis, an utterance such as (3c) is often referred to as ‘metalinguistic negation’. One of the most important features of metalinguistic negation is that it is a kind of non-truth-functional negation, in contrast to truth-functional, descriptive and ordinary negation. For instance,

- (4) a. Xiaoming was not born in Peking, he was born in Shanghai.
 b. Xiaoming was not born in Peking, he was born in Beijing.

(Huang, 2011: 55)

In (4a), it is a case of a standard and truth-functional negation, and what has been negated is the semantic content of the first clause regarding the truth value: the city where Xiaoming was born was Shanghai, not Peking. The negation in (4b) is metalinguistic in that it does not negate the truth-value fact about which city Xiaoming was born or not born in. Rather, it is objected to some property of the depiction (i.e. Peking is the Wade spelling of Beijing) that is in the scope of negation. In fact, speakers can negate any aspect of a sentence, including spelling in (4b) or presupposition in (3c). Note that metalinguistic negation occurs only when the presupposing sentence is negated. For example, the presupposition of (3d) is impossible to be cancelled in (5) in which the presupposing utterance *The King of France is wise* is affirmative.

- (5) * The King of France is wise; (in fact) there is no King of France!

One way of analysing this theoretically (Heim, 1983) is that presuppositions can contribute to an utterance at different levels, i.e. globally and locally. For (3b), when the presupposition (3d) is applied globally to the entire sentence, the inference is present, and the reading is that there is a King of France and he is not wise. In the case of (3c), the presupposition cannot be interpreted globally due to the inconsistency in the second clause. To reconcile the problem, one can accommodate the presupposition by interpreting it locally

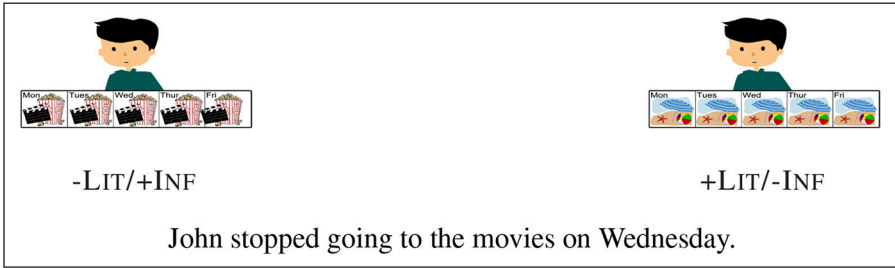


Figure 1. Visible pictures for a stimulus *John stopped going to the movies on Wednesday.*
Source. Adapted from Bill et al., 2018.

within the scope of the negation and, as a result, the existence of the King is part of the negation; therefore, the presupposition is cancelled. Note that the suspension via local accommodation occurs only when in force, i.e. triggered by explicit contextual information that nullifies the presupposition.

2 L1 processing of presuppositions

A large portion of experimental investigations on meaning have focused on exploring different aspects of meaning. In particular, scalar implicatures have drawn researchers' attention regarding both native speakers' and L2 speakers' comprehension. In recent years, researchers have examined native speakers' processing of presuppositions (Bill et al., 2016; Bill, Romoli and Schwarz, 2018; Chemla and Bott, 2013; Romoli and Schwarz, 2015; Schwarz, 2014). One question of particular interest is the availability of presupposition inference in processing. By employing a visual world eye-tracking paradigm, Schwarz (2014) looked at the timing of fixation on a display as the auditory stimulus containing *again* or *stop* (in affirmative sentences) unfolded. Participants could only distinguish between the target picture and the other pictures with the presupposed inference. Despite the fact that *again* and *stop* have been argued to belong to different types of presupposition triggers, the presupposed inference of both triggers was immediately available to native speakers in online processing. However, a follow-up experiment testing *stop* in negated sentences suggested a processing delay of the negated *stop*.

Bill et al. (2018) and Romoli and Schwarz (2015) are worth discussing here in that by using the covered-box paradigm, they explored *stop* in affirmative and negated contexts, as well as the issue of inference suspension. The covered-box paradigm is inherently a picture-matching task: given a linguistic stimulus, participants choose one of the pictures (either a visible picture or a black covered box) that matches the stimulus (more explanation about this method is in Section IV). Regarding *stop* in affirmative contexts, Bill et al. (2018) discovered that over 97% of selections were covered pictures when the visible pictures in the two conditions were similar to those in Figure 1: (1) (-Lit/+Inf) is inconsistent with the truth-conditional literal meaning but true to the presupposition and (2) (+Lit/-Inf) is false to the presupposition but consistent with the truth-conditional meaning. Moreover, selecting the covered box in the +Inf condition was more rapid than in the -Inf condition, which indicated that



Figure 2. Visible pictures for a stimulus *John didn't stop going to the movies on Wednesday*.
Source. Adapted from Bill et al., 2018.

rejections based on the presupposed inference of *stop* were slower than rejections based on the truth-conditional literal meaning.

The visible pictures targeting a negated *stop* in Bill et al. (2018) are shown in Figure 2. The visible pictures either contained the inference that *John went to the movies before Wednesday* (in the +Lit/+Inf picture) or lacked the inference but were consistent with the truth-conditional literal meaning that *John went to the movies after Wednesday* (in the +Lit/-Inf picture). They unsurprisingly discovered that when *stop* was under negation, pictures that were consistent with both the presupposition and the truth-conditional literal meaning were chosen at the ceiling. However, when the visible pictures were only compatible with the literal meaning, not to the presupposition, visible pictures were selected at 62% of the time, which was significantly lower than the +Lit/+Inf condition. They also found that selecting visible pictures in the +Lit/+Inf condition was faster than selecting visible pictures in the +Lit/-Inf condition. In other words, generating the inference of presupposition (global reading) was faster than suspending the inference (local reading), which involves local accommodation. Similar results were reported in Romoli and Schwarz (2015), who found that the absence of presuppositions resulted in higher processing costs. This is in line with Chemla and Bott (2013), who provided the first experimental finding to substantiate the claim that local interpretations are derived instead of being part of the basic meaning; compared with global readings, local readings are generally dispreferred.

Extending the investigation of the availability of local readings to English-speaking children, Bill et al. (2016) tested another presupposition trigger, i.e. *win* under negation between English-speaking adults vs. children.³ The visible picture in Figure 3 shows the no-inference reading in which the presupposed inference is suspended, corresponding to (6c). This is because the visible picture where the bear is baking at home displays no presupposition and selecting the visible picture indicates a suspension of presuppositions, i.e. *the bear did not win the race. . .because the bear did not even participate*. Selecting the covered box indicated the generation of the presupposition inference (6b); therefore, the reading was that *the bear participated but did not win the race*.

- (6) a. The bear didn't win the race.
 b. Inference: ~ The bear participated in the race.
 c. No-inference: The bear didn't participate in the race.



Figure 3. Examples of the visible picture and the stimulus in the no-inference condition.

There were three groups of English-speaking participants: adults, 4–5-year-olds, and 7-year-olds. The results revealed that adults' rate of selecting the visible picture was approximately 65%, indicating that they were more likely to suspend the presupposition and derive the local reading. However, 4–5-year-olds rejected the visible picture and thus selected the covered box near 100%, whereas 7-year-olds were more likely to choose the visible picture than 4–5-year-olds, at approximately 40% of the time. Both groups of children were less likely than adults to locally accommodate the interpretation and derive the local reading. They favored the covered box, calling for the pragmatically more felicitous global interpretation. The behavioral discrepancy in deriving inferences between adults and children was argued to be children's insensitivity to contextual cues (a more detailed discussion can be found in Section VII).

III L2 acquisition at the semantics–pragmatics interface

As mentioned in Section I, no previous L2 research has investigated how second language speakers interpret presuppositions. Thus, in this section, I will review L2 studies on scalar implicatures that are situated at the same interface. More importantly, scalar implicatures can also be explicitly cancelled, deriving a logical reading. For example, on the scale *<never, sometimes, always>*, *sometimes* implies *not always* and vice versa. However, the scalar inference can be cancelled in Speaker B's utterance in (7–8). By employing a covered-box experiment, Feng and Cho (2019) examined L1-Chinese L2-English speakers' computation and suspension of the scalar expressions *sometimes* and *not always*. They found that L2 speakers computed and suspended the inference of *sometimes* at a native-like level; however, the inference of *not always* posed difficulties to L2 speakers, especially the suspension of *not always* which involved calculation and processing alternative meanings.

- (7) A: Bob was very sick last week. But he sometimes went to school last week.
 B: Yes, in fact, he always went to school last week. (sometimes and possibly always)
- (8) A: Bob was very sick last week. So, he didn't always go to school last week.
 B: Yes, in fact, he never went to school last week. (not always and possibly never)
 (Feng and Cho, 2019: 3)

In fact, L2 speakers' ability to compute scalar inferences has been confirmed in several L2 studies with different language pairings. For instance, Slabakova (2010) investigated how L1-Korean L2-English learners process the quantifiers *some* and *all*. For a universal statement such as (9), participants will reject such sentences if they rely on a pragmatic *some but not all* calculation on *some*. Acceptance of these sentences indicates a logical meaning *some and possibly all* that involves inference suspension.

- (9) Some elephants have trunks.

(Slabakova, 2010: 2452)

Native English adults, native Korean adults and L1-Korean L2-English speakers participated in the experiments. The results suggested that Korean learners of English not only successfully computed scalar inference in their L2 but also computed the pragmatic reading of *some* – i.e. rejected sentences such as (9) – more frequently than native English speakers and native Korean adults. Moreover, even when the target language has a more complicated scalar system (e.g. two scalar items are mapped to *some*), L2 speakers are still able to interpret scalar inferences at a native-like level. Unlike Korean, which has only one lexical item roughly equal to the English scalar term *some*, Spanish has two: *algunos* and *unos*. *Unos* traditionally encodes the logical reading *some and possibly all*, whereas *algunos* usually triggers the pragmatic reading *some but not all* if no explicit contexts require a cancellation. Miller et al. (2016) reported that English speakers were able to obtain a native-like judgment on the two Spanish scalar terms irrespective of the fact that English has a different scalar implicature system. In addition to investigating L2 learners' acquisition of scalar items, Snape and Hosoi (2018) asked a question that had not been addressed in previous studies: does L2 proficiency level play a role in the acquisition of scalar implicatures? By looking at intermediate and advanced L1-Japanese L2-English learners, Snape and Hosoi discovered that the two proficiency groups did not significantly differ in their responses. Similar to the Korean participants in Slabakova's (2010) study, the two Japanese groups both interpreted more pragmatically than the native English speakers.

In light of these L2 studies, it seems that generating pragmatic inferences is not difficult for L2 speakers. A pilot study of the current research also suggested that the pragmatic (or global) reading of *stop* and *didn't stop* was preferred by Chinese learners of English. Therefore, an L1~L2 discrepancy was predicted to occur in suspending the inference. The difficulty of suspending inferences was already indicated in Slabakova's explanation of Korean learners' behavior of being overpragmatic. She explained that native English speakers were able to conjure up situations to make example (9) sound felicitous (for instance, some elephants' trunks got cut due to accidents), whereas Korean learners, due to limited processing capacity, were less able to produce alternative contexts to accept *some*

as *some and possibly all*. Snape and Hosoi (2018) also suggested that L2 speakers tried to avoid using more processing resources in computing the more effortful logical answer. However, their test design did not allow them to scrutinize L2 behavior related to suspending an inference. By using a covered-box paradigm, the present study manipulates visible pictures (displaying the inference or not) and categorizes participants' behavior of rejecting or accepting a particular picture as computing or suspending the inference.

IV The current study

1 Chinese presuppositions

This study focuses on *stop* which is a lexical presupposition trigger suggesting a change of state. The advantage of using a lexical verb rather than other triggers, such as the definite articles *the* and cleft sentences is that crosslinguistic differences exist in the latter group, and potential L1 influence may not be facilitative for a successful acquisition of these items (von Stechow and Matthewson, 2008). The general mechanism that gives rise to the presupposition trigger *stop* is not language-specific only to English or Germanic languages, although the work on presuppositions has been mostly conducted in these two languages (Bill et al., 2018; Romoli and Schwarz, 2015; Schwarz, 2015; Tiemann, 2014; Tiemann et al., 2011).

Overall, presupposition triggers are inferred in similar manners in Chinese and English (He, 1988; Lan, 1999). Chinese has change-of-state verbs as presupposition triggers: *tingzhi* 'stop', *kaishi* 'start', *fangqi* 'give up', *jixu* 'continue', and *jieshu* 'finish' (Bao, 2005; He, 1988; Lan, 1999; Lei, 2013). They all presuppose that the subject enters a state different from its original one. For instance, sentence (10a) contains the presupposition trigger *tingzhi* 'stop' and requires the presupposition that *John used to beat his child* as in (10c). Sentence (10b) is (10a) under negation, and the same presupposition (10c) is still applied.

- (10) a. *Yuehan tingzhi da tade haizi.*
 John stop beat his child.
 'John stopped beating his child.'
- b. *Yuehan meiyou⁴ tingzhi da tade haizi.*
 John no stop beat his child.
 'John didn't stop beating his child.'
- c. *~ Yuehan guoqu yixiang da haizi.*
 John past always beat child.
 'John used to beat his child.'

(He, 1988: 106)

2 Research questions

This study investigates L2 speakers' interpretation of the presupposition trigger *stop* and *stop* under negation via a covered-box paradigm. The following research questions will be addressed:

1. Do native and L2 speakers differ in generating the presupposition of *stop* in affirmative and negated sentences? In other words, does it present a challenge to Chinese learners to compute presuppositions when the presupposing sentence is under negation?
2. Do native and L2 speakers differ in suspending the presupposition of *stop* in negated sentences?

For Chinese speakers in the current study, in addition to acquiring the lexical expression *stop* which has a roughly corresponding expression in Chinese, they do not face the task of learning anything new due to the universal semantics and pragmatics that are associated with the presupposition trigger. Therefore, the learning task for the acquisition of presuppositions by Chinese speakers is to transfer the arguably universal mechanism of interpreting presuppositions from their L1-Chinese. That is, they only need to map the already existing interpretation on *tingzhi* 'stop' in the L1 to the new lexical item *stop* in English. It seems that L2 speakers have little difficulty transferring universal semantics and pragmatics from L1 to L2 in computing inferences considering the successful computation of scalar implicatures (Miller et al., 2016; Slabakova, 2010; Snape and Hosoi, 2018). L2 speakers in the current study are predicted to have native-like computation of *stop*. However, L1~L2 discrepancies might occur in suspension since L2 speakers' being overly pragmatic in interpreting scalar implicatures may be attributed to difficulties in suspending the inference (Slabakova, 2010).

V Methodology

The method adopted in this experiment was the covered-box paradigm developed by Huang et al. (2013), which has been successfully used in exploring implicatures (Huang et al., 2013) and presuppositions (Romoli and Schwarz, 2015; Schwarz, 2014; Zehr et al., 2016). The covered-box paradigm includes one picture hidden under a black, covered box (Figure 4, right). Participants in this experiment were instructed that they should select the visible picture if the visible picture matches the sentence stimuli. If the visible picture does not match the stimuli, the match must be under the black box, and participants should choose the covered box. To investigate the interpretation of pragmatic inferences, especially suspension, the covered-box paradigm was chosen for an important reason. The rationale is that no-inference readings are extremely rare in daily communication, whereas inference readings are largely preferred by conversation interlocutors since they follow pragmatic principles. However, it is logically possible to derive no-inference readings through inference suspension. In a traditional picture-selection task, if the inference and the no-inference readings are shown to participants at the same time, presumably the participants would mostly favor the inference reading. Therefore, it is difficult to examine the likelihood of computing the no-inference reading via inference suspension. With the covered-box paradigm, the no-inference interpretation can be explicitly displayed through a visible picture, and participants are forced to consider whether the shown picture corresponds to the stimulus. A rejection of the visible picture (i.e. choosing the covered box) clearly indicates that this no-inference or nondominant interpretation is not available to the participants. The same rationale also applies to the

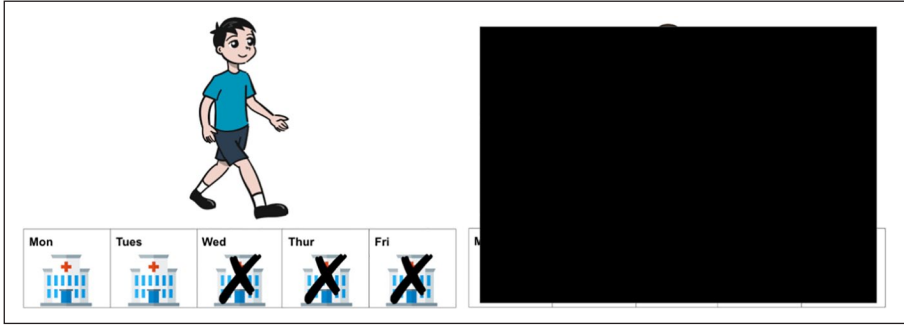


Figure 4. A test trial of the stimulus *Thomas stopped going to the hospital on Wednesday*.

inference or dominant interpretation. The visible picture in Figure 4 displays the dominant interpretation, which is true to the presupposition (*Thomas went to the hospital before Wednesday*) and true to the literal meaning (*Thomas did not continue to go to the hospital after Wednesday*).

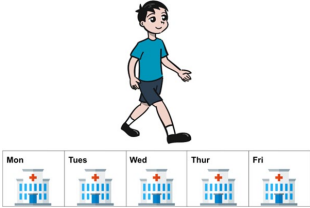
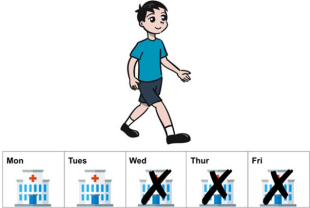
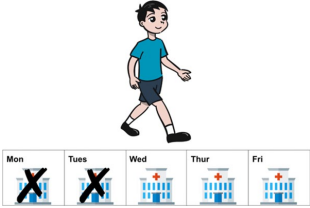
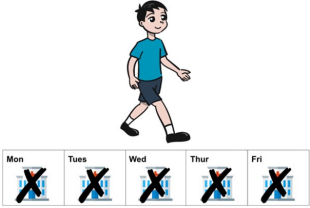
1 Test design

In this experiment, two factors were manipulated in a 2×2 design: Sentence type and Visible picture. The Sentence type factor has two levels, namely, negated and affirmative. The Visible picture factor has two levels, depending on whether the visible picture shows the presupposition inference (inference) or does not display the inference (no-inference). These two factors were crossed to create four conditions: (1) negated sentence with a visible picture depicting the inference in (11b), (2) negated sentence with a visible picture depicting a no-inference reading, as in (11c), (3) affirmative sentence with a visible picture depicting the inference in (12b), and (4) affirmative sentence with a visible picture depicting a no-inference reading, as in (12c). Half of the sentence stimuli used in the experiment were affirmative, and the other half were negated. Similarly, half of the visible pictures displayed the inference meaning, and the other half were in the no-inference condition.

- | | | | |
|------|----|---------------|---|
| (11) | a. | Negation: | Thomas didn't stop going to the hospital on Wednesday. ⁵ |
| | b. | Inference: | ~ Thomas went to the hospital before Wednesday. |
| | c. | No-inference: | Thomas didn't go to the hospital before Wednesday. |
| (12) | a. | Affirmative: | Thomas stopped going to the hospital on Wednesday. |
| | b. | Inference: | ~ Thomas went to the hospital before Wednesday. |
| | c. | No-inference: | Thomas didn't go to the hospital before Wednesday. |

To convert examples (11) and (12) into visual stimuli for use in the covered-box paradigm, the 5-day calendar-strip design was adapted, which has been commonly used to investigate the availability of presupposition interpretations (Bacovcin et al., 2018; Bill et al., 2015; Romoli and Schwarz, 2015; Schwarz, 2014). In the experiment, the calendar

Table 1. Examples of visible pictures and test sentences in target conditions using the verb *go*.













Sentence type	Negated	Affirmative
Visible picture	<i>Thomas didn't stop going to the hospital on Wednesday.</i>	<i>Thomas stopped going to the hospital on Wednesday.</i>
Inference		
No-inference		

strip contains icons of various activities and locations from Monday to Friday. A continuous appearance of an activity or a location means that this action has been repeated every day. A 'X' sign has been adopted to represent the meaning that the action did not happen on that day, making the event of *not going to the hospital* more salient, as shown in Figure 4.⁶ Table 1 displays four sample visible pictures for the four target conditions using the verb *go*. The two inference pictures were consistent with an interpretation that derived the presupposition because it showed that Thomas went to the hospital on Monday and Tuesday. The crucial manipulation of the two no-inference pictures was the 'X' sign on Monday and Tuesday, blocking an inference reading.

It should be noted that the presupposition trigger in this experiment, i.e. *stop*, is a lexical item, and generation and suspension mechanisms of the presupposition inference do not differ on the verb after *stop*. To achieve a more comprehensive understanding of L2 acquisition on presuppositions, the experimental stimuli also included three verbs other than *go* in the *stop + verb-ing* and *didn't stop verb-ing* constructions, e.g. *cook*, *play*, and *drink*. There are two reasons that these three verbs were chosen: (1) similar to *go*, they are common action verbs; (2) they are easily paired with simple objects that can be shown as icons in the calendar-strip design (e.g. *cook curry*, *play basketball*, *drink tea*). Similar to *stop + go*, the target conditions of the three verbs were also created through a 2 × 2 design. The same two factors, namely, Sentence type and Visible picture, were crossed to create four target conditions of *cook*, *play*, and *drink* (for the four conditions of the three verbs, see Appendix 1).

Each verb was paired with four different objects. For instance, the verb *go* was used with *museum*, *hospital*, *movies*, and *school*. The verb *drink* was paired with *wine*, *coke*, *beer*, and *orange juice*. Using the four objects, four sentences were created for each verb

Table 2. Examples of visible pictures and test sentences of presupposition control items.

Negated					Affirmative				
<i>Stanley didn't stop going to the gas station on Wednesday.</i>					<i>Camille stopped drinking coffee on Wednesday.</i>				
									
									

for each target condition. Therefore, for each verb, 16 target sentences were created (4 objects * 4 target conditions = 16 items). After all of the stimuli had been finalized, four counterbalanced presentation lists were created using a Latin square block design.

There were eight control items: four of them involved presuppositions, and the other four involved scalar implicatures. Among them, two items used negated sentences, and two used affirmative sentences. The visible pictures of controls allowed the inference interpretation of *stop*, but they were false to the literal meaning, as in Table 2.

Three types of fillers that were not related to research questions were included in this experiment. For each type, half of the visible pictures matched the stimuli, and the other half did not, calling for the selection of the covered box. The first type was simple negated and affirmative sentences without *stop*, but with the four verbs, i.e. *go*, *cook*, *play*, and *drink*. The second type of filler was created using *again*, such as *Naomi played baseball again on Wednesday during the week*. The last type used *twice* in the sentence stimuli. For instance, *Josh cooked pasta twice last week*.

To summarize the test design, four presentation lists were created for presuppositions. Each list contained 16 items (4 verbs * 4 objects in the token set). Each list was randomized twice, leading to a total of 8 presentation lists that were ready to be implemented in *E-prime*. Together with controls and fillers, each participant completed a total of 16 target, 8 control and 52 filler trials.

2 Participants and procedures

There were two groups of participants in the experiment: L1-English native speakers (n = 38) and L1-Mandarin Chinese L2-English speakers (n = 41). They were students at a midwestern university in the United States.

All participants finished two main tasks: a covered-box task and a proficiency test. The participants were also asked to provide minimal demographic information about gender, age and years of studying English. The proficiency test was based on the Common European Framework of Reference for Languages (CEFR) containing 40 items with a

Table 3. Participants' background information and proficiency scores.

	Age at testing (years)		Years studying English		Proficiency score (max = 40)	
	M (SD)	Range	M (SD)	Range	M (SD)	Range
Native English (n = 38)	20.4 (4.1)	18–32	n/a	n/a	38.8 (1.0)	37–40
High intermediate to advanced Chinese (n = 41)	23.7 (5.8)	18–50	13.7 (4.3)	5–20	34.9 (2.4)	30–40

maximum score of 40. A summary of the participants' information is shown in Table 3. The proficiency test and the categorization of advanced and intermediate learners were adopted from Cho (2017), where learners with scores above 34 were considered to be advanced learners and those who scored between 26 and 33 belonged to the intermediate level.

Regarding the experimental procedures, first, the participants were given a written consent form to sign. Second, the participants completed an icon recognition task that was used to ensure that they correctly understood the icons. Then, the participants completed six practice trials using the covered-box paradigm to familiarize themselves with the task. Then, they started the experimental trials. Last, the participants finished a proficiency test and were asked to provide demographic information. All participants completed the covered-box task on a computer where E-prime was used to display stimuli and collect data. The choice of pictures was achieved by clicking on the selected picture via a mouse. For both groups, it took on average 10 minutes to finish the covered-box task, and the overall experiment lasted approximately 30–40 minutes.

3 Data analysis

For the purpose of the analysis, the percentage of covered or visible pictures selected and response times (RTs) were the two dependent variables in the study. Responses were coded with regard to whether the visible or the covered picture was selected. RTs were calculated as the time taken to select a picture.

The data were trimmed in three steps. First, control items in the experiment were designed to check if participants understood the task and were paying attention to the items. The correct response to all control items was the covered box. If participants selected the visible picture for two or more controls, the participants' data were removed. This did not result in removing any data since there was only one participant from each language group who had selected the visible picture only once. Then, the data were further trimmed at ± 3 standard deviations (SDs) or more from the mean subject RTs (Jegerski, 2014). Last, if participants clicked on the region that was outside of the picture or accidentally clicked on the sentence, the data from that trial were removed. The trimming at the last two steps (extreme data points and meaningless data points) resulted in the loss of 1.7% of trials for L1-Chinese L2-English learners and 1.8% of trials for English speakers.

The picture selection percentages were analysed using the generalized logistic mixed-effects regression model. To correct the skewed distribution of the RT data, RTs were log transformed and analysed using the linear mixed-effects regression model. The maximal

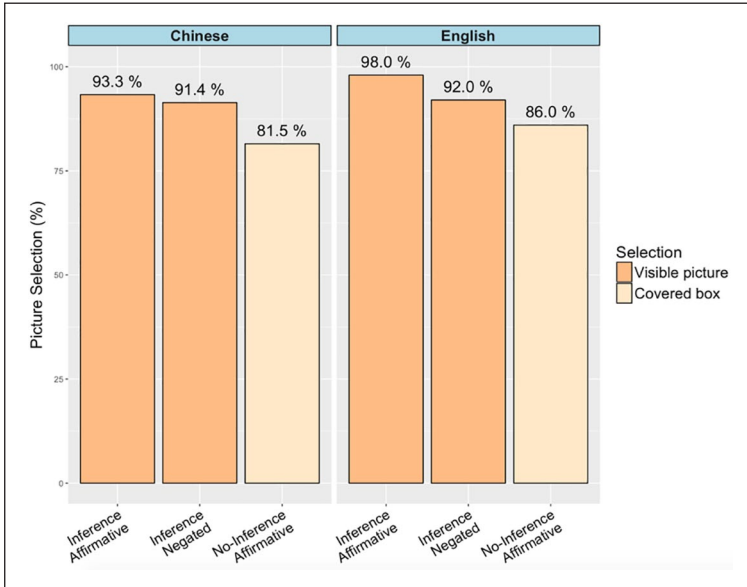


Figure 5. Picture selection percentages in the two *inference* and *no-inference affirmative* conditions (represent inference computation).

random effect structure that would converge was used, as recommended by Barr et al. (2013), and all fixed effects were centered prior to the analysis. Specific models are illustrated in the next section.

VI Results

I Percentage of picture selection

To recapitulate the logic of the covered-box method and predictions, when the visible picture showed inference, the participants were expected to choose the visible picture. When the visible picture did not show inference, in the affirmative condition, participants were predicted to select the covered box for presupposition computation since the affirmative context does not trigger suspension, as discussed in 2.1. In the negated and no-inference condition, presupposition suspension was indicated by selecting the visible picture, whereas computation was suggested by selecting the covered box. First of all, I will discuss the percentage of picture selection that was associated with presupposition computation. Next, I will discuss the percentage of visible picture selection in the no-inference and negated condition, which suggested presupposition suspension.

When the visible picture suggested an inference reading, both language groups favored the visible picture more than 90% of the time, as shown in Figure 5, regardless of the sentence type. These results indicated that both groups successfully computed presupposition inferences in affirmative and negated sentences, confirming my prediction. Note that under metalinguistic negation, presuppositions cannot be suspended in

Table 4. Fixed effects estimates and standard errors for the generalized logistic mixed-effects model of visible picture selection in the two inference conditions.

	Estimate	Standard error	z value	p value
Intercept	3.235	0.334	9.674	0.000***
Sentence type	0.497	0.210	2.364	0.018*
Language group	-0.392	0.228	-1.715	0.086
Sentence type \times language	-0.281	0.197	-1.426	0.154

Notes. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

an affirmative environment (see (5) in 2.1). Thus, in the no-inference and affirmative condition, participants were predicted to select the covered box if they computed the presupposition since the suspended reading in the visible picture was impossible. As in Figure 5, the percentage of covered box selection in the no-inference affirmative condition was over 80% of the time.⁷ A generalized logistic mixed-effect model was fitted to statistically analyse visible picture selection in the two inference conditions, i.e. selection as the dependent variable, sentence type (2 levels: affirmative and negated) and language group (2 levels: native and L2) as fixed effects, and participants and items as random intercepts. As shown in Table 4, the results showed a significant main effect of sentence type ($\beta = 0.497$, $SE = 0.210$, $z = 2.364$, $p = 0.018$), indicating that the percentage of visible picture selection in the affirmative condition was significantly different from the visible picture selection in the negated condition. However, the percentage of visible pictures selected between English and Chinese speakers was not significantly different. Overall, the percentage of inference computation was similar between Chinese and English groups.

Regarding the percentage of visible pictures selected in the no-inference negated condition, as visualized in Figure 6, the native speakers selected the visible picture at 64% of the time, whereas the L2 speakers selected the visible picture 38.5% of the time. The results of a generalized logistic mixed-effects model (language group as the fixed effect and participant and item as the random intercepts) indicated a main effect of language group ($\beta = -1.354$, $SE = 0.616$, $z = -2.198$, $p = 0.028$). This result revealed that the English speakers significantly more frequently selected the visible picture than the Chinese speakers, suggesting that it was more likely for the English speakers to suspend the inference than the Chinese speakers (for possible reasons regarding L2 speakers' low rates of suspension, see Section VII).

2 Response times

The results are organized as follows: first, regarding inference computation, I report RTs for selecting visible pictures in the two inference (affirmative and negated) conditions; then, RTs for selecting visible pictures and covered box in the no-inference negated condition are presented for the discussion of suspension.

In the two inference conditions, selecting the visible picture was expected since salient inference reading should be preferred, as confirmed by the high percentage of selecting the visible picture (see Figure 5). As shown in Figure 7, mean RTs in these two

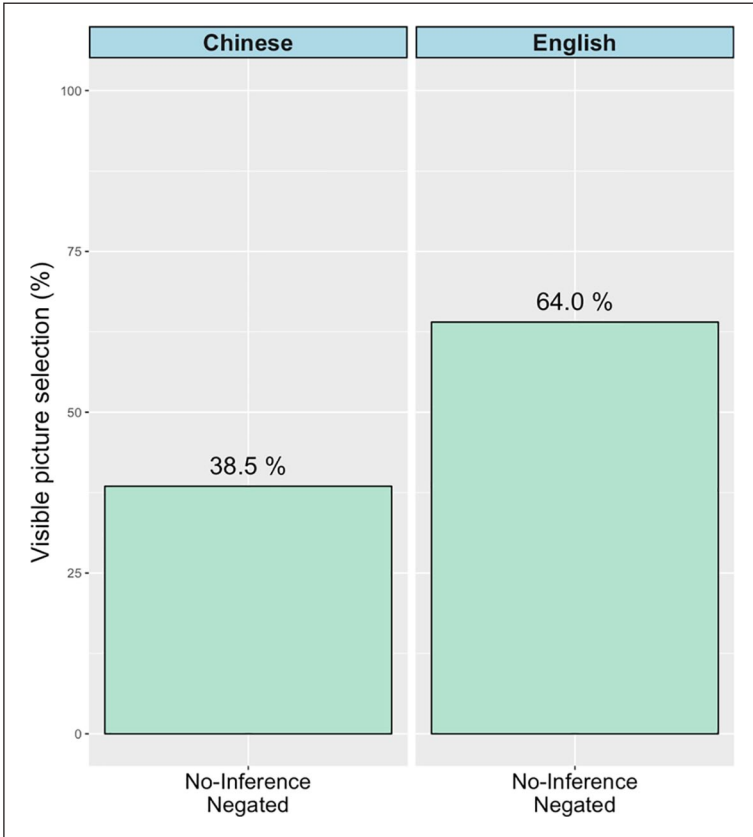


Figure 6. Visible picture selection percentages in the *no-inference negated* condition (represent inference suspension).

inference conditions depended on the sentence type, as both participant groups were faster in affirmative sentences than in negated sentences. I performed a mixed-effects linear regression with an interaction between sentence type and language, as well as random intercepts for participants and items. The main effects of sentence type and language and the interaction were all significant in Table 5. Post hoc analysis revealed that L2 speakers were significantly slower than native speakers in both affirmative and negated conditions (affirmative: $t = 5.907, p < 0.0001$; negated: $t = 3.567, p < 0.001$). More interestingly, regarding within-group RTs, native speakers were significantly faster in selecting the visible picture in affirmative sentences than in negated sentences ($t = -4.485, p < 0.0001$). It seems that the native speakers' rapid responses were reinforced in the affirmative context. There was a numerical effect in the same direction for the L2 speakers, but it was not significant ($t = -1.435, p = 0.152$). However, as pointed out by an anonymous reviewer, this result should be interpreted with caution as the L2 speakers' overall long RTs might have masked any intricate RT differences between the affirmative and negated conditions.

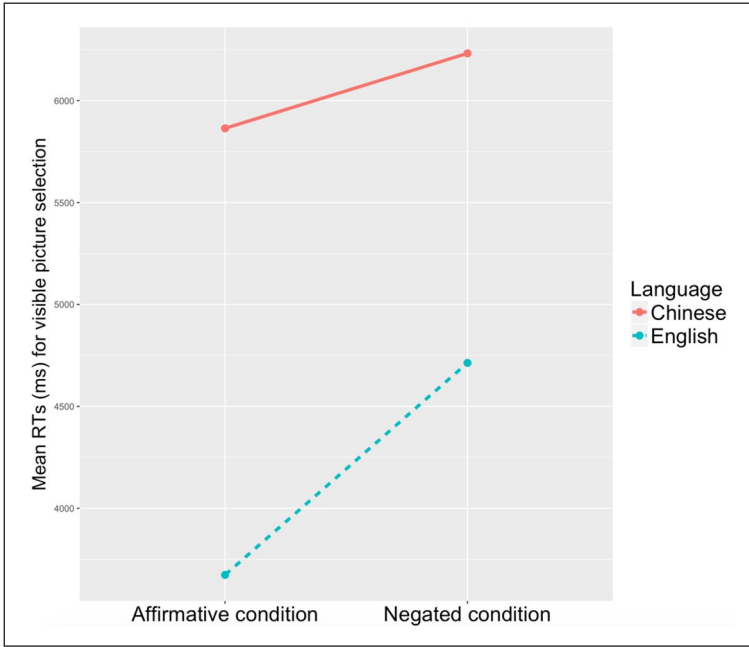


Figure 7. Mean response times (RTs) for selecting visible pictures in the two *inference* (affirmative and negated) conditions by native speakers (L1) and second-language (L2) speakers.

Table 5. Fixed effects estimates and standard errors for the mixed-effects linear model of RTs in the two inference conditions.

	Estimate	Standard error	t value	p value
(Intercept)	8.392	0.042	201.225	0.000***
Sentence type	-0.080	0.022	-3.587	0.000***
Language	0.171	0.033	5.198	0.000***
Sentence type × language	0.042	0.015	2.817	0.005**

Notes. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Mean RTs in the no-inference negated condition for native and L2 speakers are shown in Figure 8, and an interaction between language group (native vs. L2 speakers) and selection type (visible picture vs. covered box) was found.⁸ As is apparent in Figure 8, the Chinese speakers were slower to select the visible pictures than the covered boxes (9150 ms vs. 8262 ms, respectively), while the reverse was true for the native speakers (covered box selection: 6950 ms vs. visible picture selection: 5743 ms). To statistically examine the results, a mixed-effects linear regression model was fitted with an interaction between selection and language, participants and items as random effects. As shown in Table 6, the interaction in Figure 8 was marginally significant ($\beta = 0.121$, $SE = 0.068$, $t = 1.762$, $p = 0.071$). Post hoc analyses indicated that accepting the visible picture was

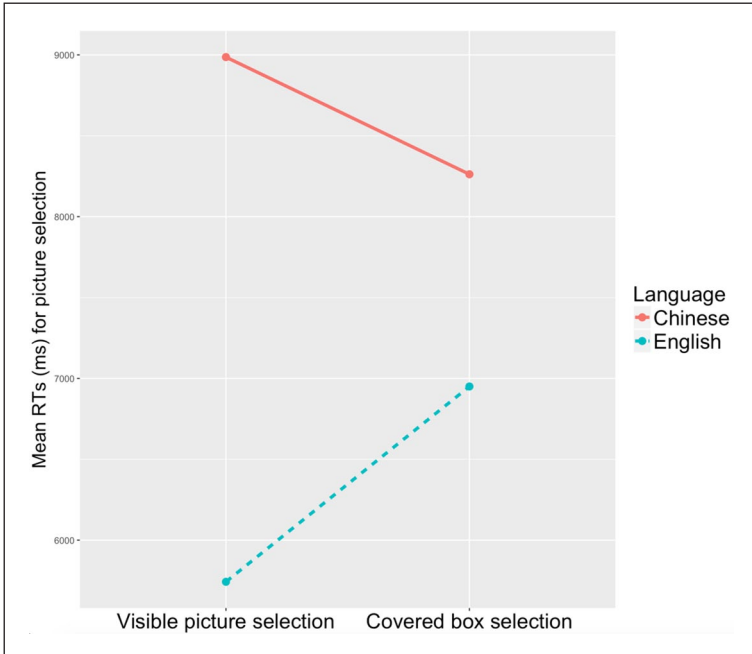


Figure 8. Mean response times (RTs) for selecting visible pictures and covered boxes in the *no-inference negated* conditions by native speakers (L1) and second-language (L2) speakers.

significantly faster than rejecting the visible picture (i.e. choosing the covered box) for the native speakers ($t = 2.03$, $p = 0.04$), while RT differences between visible picture selection and covered box selection were not significant for the L2 speakers ($t = -0.335$, $p = 0.74$). More importantly, the native speakers were only significantly faster than the L2 speakers in selecting the visible picture ($t = 4.148$, $p < 0.0001$), but not in selecting the covered box ($t = 1.694$, $p = 0.09$).

VII Discussion

I The generation of presuppositions

The aim of this study was to investigate Mandarin Chinese learners' acquisition of presuppositions at the semantics–pragmatics interface by looking at the presupposition trigger *stop* and *stop* under negation. I will first summarize the findings to answer the first research question: 'Do L2 and native speakers differ in generating the presupposition of *stop* in affirmative and negated sentences'? Put differently, is it more demanding to compute presuppositions for Chinese learners when they encounter negated presupposing sentences?

The results of the covered-box experiment indicated that the L2 speakers and native speakers generated a presupposition inference at a very similar rate in both inference (affirmative and negated) conditions. This suggested that when a shown picture was

Table 6. Fixed effects estimates and standard errors for the mixed-effects linear model of RTs in the no-inference negated condition.

	Estimate	Standard error	t value	p value
(Intercept)	8.817	0.065	135.997	0.000***
Selection	-0.090	0.069	-1.314	0.190
Language	0.091	0.054	1.694	0.092
Selection × language	0.121	0.068	1.762	0.071

Notes. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

compatible with an inference reading, the two participant groups preferred deriving the inference of presuppositions and did not differ in the rates of inference generation. Similarly, Bill et al. (2018) reported the preference of computation in processing *stop* in affirmative and negated contexts (as well as in two types of scalar implicatures) and further proposed ‘inference preference’, stating that ‘inference interpretations are preferred (for both SIs and Ps⁹)’ (p. 20). More importantly, the present study’s results indicated that the negated presupposing context did not present a greater challenge to the L2 speakers since the L2 speakers’ selection rate did not significantly differ from that of the native speakers. In fact, when scalar items are under negation, it is not demanding for L2 speakers to compute the inference either. For instance, the inference of *John didn’t always go to school last week* with an indirect scalar implicature *not always* is *John sometimes went to school last week*. In interpreting a scalar item in such a negated context, Chinese learners of English still overwhelmingly preferred generating the inference of *not always*, similar to native speakers’ computation rates (Feng and Cho, 2019).

2 The suspension of presuppositions

Regarding the presupposition suspension mentioned in the second research question – ‘Do native and L2 speakers differ in suspending the presupposition of *stop* in negated sentences?’ – the data revealed that when the inference was absent in the visible picture, the English speakers significantly more frequently selected the visible picture than the Chinese speakers, suggesting that the English speakers were more likely to suspend the inference than the Chinese speakers. Additionally, the native speakers were significantly faster than the L2 speakers in selecting the visible picture. This result indicated that the native speakers not only suspended the presupposition inference more frequently but also faster than the L2 speakers. However, the two participant groups’ RTs did not differ in selecting the covered box, indicating that it took a similar amount of time for the two groups to generate the inference in this condition.¹⁰

In summary, the answer to the first research question is that the L2 and native speakers did not differ in generating the inference of presupposition with the exception of longer RTs by the L2 speakers. However, the results regarding the second question demonstrated that the two participant groups did differ in the rate of and time to suspending the inference, i.e. the native speakers suspended the inference more frequently and faster than the L2 speakers.

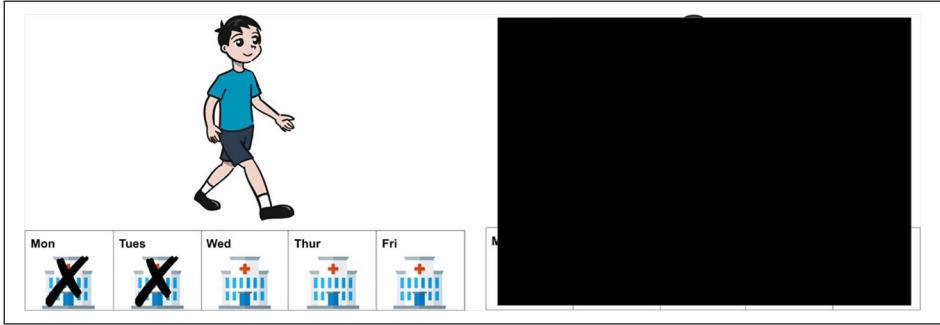


Figure 9. The visible picture and covered box for *Thomas didn't stop going to the hospital on Wednesday* in the no-inference negated condition.

In the following section, I will discuss some possible reasons for the L2 speakers' low percentage of suspending the inference of *stop* when facing a picture inconsistent with an inference reading. Selecting visible pictures that are not compatible with an inference reading requires participants to be sensitive to the contextual cues that trigger suspension and thus locally accommodate the presupposition, leading to higher processing cost than a global reading with the inference present (Bill et al., 2016; Chemla and Bott, 2013; Romoli and Schwarz, 2015). One possible situation that triggers suspending presupposition inference is contextual relevance (Romoli, 2014). That is, the local application of presuppositions has to be motivated by explicit contextual information. Thus, the ability to identify the no-inference cue in a visible picture is the key to suspending the inference. By using the same covered-box paradigm, Bill et al. (2016) found that adult native speakers of English were more likely to suspend the inference of presuppositions than child native speakers. In fact, the 7-year-old group had a suspension rate similar to that of the L2 speakers in the current experiment (i.e. approximately 40%). To account for the low rates by native-speaking children, Bill et al. explained that they were not sensitive enough to the contextual cues embedded in the visible picture which could have triggered suspension. In the current experiment, with a given sentence and a no-inference visible picture (as in Figure 9), the participants needed to consider the meaning of the sentence and compare with other possible alternative meanings. For instance, if L2 speakers were sensitive to the no-inference contextual cues in the visible picture in Figure 9 (i.e. Thomas did not go to the hospital prior to Wednesday), one of the alternative meanings that needed to be evaluated was the suspended-presupposition reading (Thomas did not stop going to the hospital on Wednesday. . . in fact, he did not go before Wednesday). Otherwise, participants were under no pressure to defeat the more accessible pragmatic meaning, and therefore, they would select the covered box. The L2 speakers' high rates of selecting the covered box may suggest that they were probably insensitive to the contextual relevance triggering suspension. It seems that suspending presuppositions is costly not only for native-speaking children but also for adult L2 speakers who have mature cognition and social communication skills.

In addition to the L2 speakers' insensitivity to contextual cues and the complexity of local accommodation, the difficulty of suspending the presupposition might also have

come from the fact that the L2 speakers were not certain about their language command in this specific situation involving inference suspension. Even though the L2 participants in the present study were high intermediate to advanced speakers with an average 13.7 years of studying English, with less exposure to English overall, they were still not certain and therefore reluctant in selecting the visible picture representing a less accessible and prominent reading. The pragmatic interpretation in the covered box, which was readily available to the L2 speakers, became a safe choice. In fact, even with the native speakers, they were least certain (indicated by giving the lowest confidence ratings) about sentences with inconsistent presuppositions (Bacovcin et al., 2018).¹¹ Furthermore, it should be noted that L2 speakers' 'being notably pragmatic' has been repeatedly reported in L2 research on scalar implicatures. For instance, Korean learners of English in Slabakova (2010) more frequently rejected the sentence *Some elephants have trunks* than native controls.¹² This behavior of 'being more pragmatic' was offered by an explanation that undoing pragmatic reading to derive logical reading is demanding for L2 speakers. More specifically, with limited processing resources, L2 speakers were less likely to conjure up a scenario to make this test sentence acceptable. However, even with the assistance of a visible no-inference reading at hand in the current experiment, the L2 speakers were still more inclined to reject the visible picture and opt for a pragmatic reading. This result suggested that processing the suspension of an inference is costly for L2 speakers.

Another related explanation is the nature of the presupposition trigger *stop*. As a change-of-state verb, *stop* has a representation of dynamic events that includes an initial state, a change and a final state. In terms of processing, by comparing change-of-state verbs with more static verbs denoting a stable state such as *love*, Gennari and Poeppel (2003) found that eventive verbs with more events to be activated required longer processing time and extra cognitive sources. In the current experiment, the meaning of *stop* was further complicated by introducing negation. When reading 'Thomas didn't stop going to the hospital on Wednesday', the L2 speakers also needed to generate three displaced temporal events. The first event (an initial state) was the presupposition: 'Thomas went to the hospital before Wednesday'. The second event (a change), was initially without the negation, i.e. 'Thomas stopped going on Wednesday'; then, after adding negation, it changed to 'Thomas didn't stop on Wednesday'. The third event (a final state) was 'Thomas kept going to the hospital from Monday to Friday'. The processing effect is reflected in the number of meanings participants have to process, as well as the contextual clues embedded in the environment. The greater the number of different readings and cues to be considered, the more processing effort is warranted. Regarding whether this suspension difficulty only applies to *stop* or not, further L2 research on suspending different types of presupposition triggers is needed.

Previous L2 studies on scalar implicatures (Miller et al., 2016; Slabakova, 2010; Snape and Hosoi, 2018) showed that deriving scalar implicatures is not a challenge to L2 speakers despite different language pairings and lexical items. This study's results of presuppositions that share the same external interface with scalar implicatures corroborate these earlier findings that L2 speakers succeeded in generating pragmatic inferences at this interface. The findings, alongside the abovementioned L2 research, seem to challenge the Interface Hypothesis (Sorace, 2011), which proposes that discrepancies between L2 speakers and native controls occur at the external interfaces, such as the

semantics–pragmatics interface. However, the picture is more complicated than the binary choice of either for or against the Interface Hypothesis. The present study goes one step further in examining L2 acquisition at the semantics–pragmatics interface by looking at both inference computation and suspension. The results demonstrated asymmetrical behavior between inference computation and suspension. That is, native-learner differences were detected in suspending the inference, not in computing the inference. This leads to an important methodological implication, which is the necessity of examining inference suspension in exploring language learners' pragmatic inference abilities.

VIII Conclusions

The main goal of the present study was to examine how L1-Mandarin Chinese L2-English learners interpreted presuppositions at the semantics–pragmatics interface via a covered-box experiment. The results suggested that while English native and L2 speakers shared similar response patterns in computing presuppositions, the two groups significantly differed in suspending presuppositions, i.e. L2 speakers were less likely to suspend presuppositions than native speakers. A number of possible explanations were discussed to account for L2 speakers' difficulty in suspending presuppositions. The asymmetrical responses behavior between inference computation and suspension by L2 speakers demonstrates the necessity of testing both in pragmatic inference research. This study hopefully sheds light on a more precise understanding of L2 acquisition at the external interface level.

Acknowledgements

I am truly grateful to anonymous reviewers for their invaluable comments. I would like to thank Dr. Jacee Cho for her insightful suggestions on the research and manuscripts. I also thank students and friends for participating in the experiment.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: I would like to acknowledge the support of The National Social Science Fund of China (# 20CYY002).

ORCID iD

Shuo Feng  <https://orcid.org/0000-0002-0837-2170>

Notes

1. For a detailed discussion of different approaches to the issue of universality of presupposition triggers, see von Stechow and Matthewson, 2008.
2. Note that some research, for example Sudo (2012), argues that for a presupposition trigger such as *stop*, the meaning of the change of state occurs twice in the lexical meaning, being part of the presupposition and the entailment. I thank an anonymous reviewer for pointing this out.

3. Bill et al. (2016) also compared reading patterns of scalar implicatures, but only the section on presuppositions is reported in this article.
4. Chinese has two negators, i.e. *meiyou* and *bu*. Note that to negate (10a) the negator *meiyou* is applied in (10b) rather than *bu* in that it has been argued that *bu* cannot occur with verbs denoting change of states or processes (Li and Thompson, 1981).
5. I acknowledge that the affirmative and negated test sentences have a habitual reading when 'on Wednesday' modifies 'going to the hospital' rather than 'stop' which is the target reading. Although data suggested that participants might still have computed the target reading (e.g. over 90% selection of the visible picture in the inference conditions), further test design should take this potential problem into consideration.
6. Compared with the design in Schwarz and his colleagues' study, there are two changes in the current experiment. The first change is the use of 'X' sign. In Schwarz and his colleagues' work, the use of different icons for places (e.g. school) to represent the meaning of not going to the target place (e.g. hospital) could be somewhat vague. However, the use of 'X' sign makes the meaning of 'not going to the target place' clear and salient. Second, since *going to the hospital* is an action (as are the other three action verbs used in the experiment), the static bust character icon in Schwarz and his colleagues' study, as in Figures 1 and 2, has been changed to a running pose to display the *going* event more dynamically.
7. The relatively lower percentage of deriving the inference (covered-box selection) in this condition is unexpected and it might be a problem with the design. Due to this issue, no further statistical analyses (including RTs) were performed in this condition.
8. I thank an anonymous reviewer for pointing out this interaction between the L2 and native speakers.
9. SIs stands for scalar implicatures and Ps stands for presuppositions.
10. One unexpected RT pattern was that it was easier for the native speakers to accept visible pictures (suggested by significantly shorter RTs in selecting visible pictures than covered boxes) whereas it was harder for the L2 speakers to accept visible pictures (although not significant). The native speakers' responses were surprising in that previous research (e.g. Bill et al., 2018) found the opposite RT pattern for native speakers (which was similar to the L2 speakers' response pattern in the current study). As pointed out by an anonymous reviewer, one possible explanation might be the different set-up for not going to the target location in the visible picture, i.e. a cross on the location icon in the current study, and an entirely different location icon in Bill et al. (2018). This question is left to future research.
11. L2 speakers' certainty and confidence level in interpreting implied meanings will need to be fleshed out in future research.
12. For a discussion of a possible artifact of the experimental design in Slabakova (2010), see Dupuy et al. (2019).

References

- Bacovcin HA, Zehr J, and Schwarz F (2018) To accommodate or to ignore?: The presuppositions of again and continue across contexts. *Glossa: A Journal of General Linguistics* 3: 16.
- Bao F (2005) Contrastive study of presupposition in English and Chinese. Unpublished MA thesis, Henan University, Henan, China.
- Barr D, Levy R, Scheepers C, and Tily H (2013) Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255–78.
- Beaver D and Geurts B (2011) Presuppositions. In: Maienborn C, von Stechow K, and Portner P (eds) *Semantics: An international handbook of natural language meaning: Volume 3*. Berlin: Mouton deGruyter.

- Bill C, Romoli J, and Schwarz F (2015) *Are some presuppositions scalar implicatures? Assessing evidence from reaction times*. Available at: https://www.academia.edu/17390559/Are_some_presuppositions_scalar_implicatures_Assessing_evidence_from_Reaction_Times (accessed February 2021).
- Bill C, Romoli J, and Schwarz F (2018) Processing presuppositions and implicatures: Similarities and differences. *Frontiers in Psychology* 3: 44.
- Bill C, Romoli J, Schwarz F, and Crain S (2016) Scalar implicatures versus presuppositions: The view from acquisition. *Topoi* 35: 57–71.
- Birner BJ (2013) *Introduction to pragmatics*. Oxford: Wiley-Blackwell.
- Chemla E and Bott L (2013) Processing presuppositions: Dynamic semantics vs. pragmatic enrichment. *Language and Cognitive Processes* 38: 241–60.
- Chierchia G and McConnell-Ginet S (1990) *Meaning and grammar*. Cambridge, MA: MIT Press
- Cho J (2017) The acquisition of different types of definite noun phrases in L2-English. *International Journal of Bilingualism* 21: 367–82.
- Destruel E and Donaldson B (2017) Second language acquisition of pragmatic inferences: Evidence from the French *c'est-cleft*. *Applied Psycholinguistics* 38: 703–32.
- Dupuy L, Stateva P, Andreatta S, et al. (2018) Pragmatic abilities in bilinguals: The case of scalar implicatures. *Linguistic Approaches to Bilingualism* 9: 314–40.
- Feng S and Cho J (2019) Asymmetries between direct and indirect scalar implicatures in second language acquisition. *Frontiers in Psychology* 10: 877.
- Gennari S and Poeppel D (2003) Processing correlates of lexical semantic complexity. *Cognition* 89: B27–41.
- Grice P (1975) Logic and conversation. In: Davidson D and Harman G (eds) *The logic of grammar*. Encino, CA: Dickenson, pp. 64–75.
- Grice HP (1989) *Studies in the way of words*. Cambridge, MA: Harvard University Press.
- He ZY (1988) *A survey of pragmatics*. Changsha: Hunan Education Press.
- Heim I (1982) The semantics of definite and indefinite noun phrases. Unpublished PhD thesis, University of Massachusetts, Amherst, MA, USA.
- Heim I (1983) On the projection problem for presuppositions. In: Flickinger DP (ed.) *Proceedings of WCCFL 2*. Stanford, CA: Stanford University, CSLI Publications, pp. 114–125.
- Horn L (1972) On the semantic properties of logical operators in English. Unpublished doctoral dissertation. Distributed by the Indiana University Linguistics Club, Bloomington, IN, USA.
- Huang Y (2011) Types of inference: Entailment, presupposition, and implicature. In: Bublitz W and Norrick NR (eds) *Foundations of Pragmatics*. Berlin: De Gruyter, pp. 397–421.
- Huang YT and Snedeker J (2009) Online interpretation of scalar quantifiers: Insight into the semantics–pragmatics interface. *Cognitive Psychology* 58: 376–415.
- Huang YT, Spelke E, and Snedeker J (2013) What exactly do numbers mean? *Language Learning and Development* 9: 105–29.
- Ivanov I (2012) L2 acquisition of Bulgarian clitic doubling: a test case for the Interface Hypothesis. *Second Language Research* 28: 345–68.
- Jegerski J (2014) Self-paced reading. In: Jegerski J and VanPatten B (eds) *Research methods in second language psycholinguistics*. New York: Routledge, pp. 20–49.
- Karttunen L (1974) Presupposition and linguistic context. *Theoretical Linguistics* 1: 181–94.
- Lan C (1999) Xiandai hanyuyusheyinfaxiangchutan [An exploration of the presupposition triggers in Modern Chinese]. *Waiyuxuekan [Foreign Language Research]* 3: 11–14.
- Lei K (2013) A look into the triggering of presuppositions in Chinese and English. *Theory and Practice in Language Studies* 3: 1988–95.
- Li CN and Thompson SA (1981) *Mandarin Chinese: A functional reference grammar*. Oakland, CA: University of California Press.

- Lieberman M (2009) Necessary interpretation at the syntax/pragmatics interface: L2 acquisition of scalar implicatures. Unpublished paper presented at the Workshop on Mind Context Divide: Language Acquisition and Interfaces of Cognitive Linguistic Modules. Iowa City, IA: University of Iowa.
- Miller D, Giancaspro D, Iverson M, Rothman J, and Slabakova R (2016) Not just *algunos*, but indeed *unos* L2ers can acquire scalar implicatures in L2 Spanish. *Language Acquisition Beyond Parameters* 51: 125–45.
- Noveck I (2001) When children are more logical than adults: Experimental investigations of scalar implicatures. *Cognition* 78L 165–88.
- Romoli J (2014) The presuppositions of soft triggers are obligatory scalar implicatures. *Journal of Semantics*, 32, 173–219.
- Romoli J and Schwarz F (2015) An experimental comparison between presupposition and indirect scalar implicatures. In: Schwarz F (ed.) *Experimental perspectives on presuppositions*. Berlin: Springer, pp. 215–40.
- Rothman J (2009) Pragmatic deficits with syntactic consequences? L2 pronominal subjects and the syntax–pragmatics interface. *Journal of Pragmatics* 41: 951–73.
- Schwarz F (2014) Presuppositions are fast, whether hard or soft: Evidence from the visual world. *Proceedings of SALT* 24: 1–22.
- Schwarz F (2015) Introduction: Presuppositions in context: Theoretical issues and experimental perspectives. In: Schwarz F (ed.) *Experimental perspectives on presuppositions*. Berlin: Springer, pp. 1–37.
- Simons M (2006) *Presuppositions without common ground*. Carnegie Mellon, Pittsburgh, PA. Available at: <https://www.cmu.edu/dietrich/philosophy/docs/simons/Presupposition%20without%20Common%20Ground.pdf> (accessed February 2021).
- Slabakova R (2010) Scalar implicatures in second language acquisition. *Lingua* 120: 2444–62.
- Snape N and Hosoi H (2018) Acquisition of scalar implicatures: Evidence from adult Japanese L2 learners of English. *Linguistic Approaches to Bilingualism* 8: 163–92.
- Sorace A (2005) Selective optionality in language development. In: Cornips L and Corrigan K (eds) *Syntax and variation*. Amsterdam: John Benjamins, pp. 55–80.
- Sorace A (2011) Pinning down the concept of ‘interface’ in bilingualism. *Linguistic Approach to Bilingualism* 1: 1–33.
- Sorace A and Filiaci F (2006) Anaphora resolution in near-native speakers of Italian. *Second Language Research* 22: 339–68.
- Stalnaker R (1973) Presuppositions. *Journal of Philosophical Logic* 2: 447–57.
- Stalnaker R (1974) Pragmatic presuppositions. In: Munitz M and Unger D (eds) *Semantics and philosophy*. New York: New York University Press, pp. 197–213.
- Sudo Y (2012) On the semantics of phi features on pronouns. Unpublished PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, USA.
- Tiemann S (2014) The processing of *wieder* (‘again’) and other presupposition triggers. Unpublished doctoral dissertation, Eberhard Karls Universität, Tübingen, Germany.
- Tiemann S, Schmid M, Bade N, et al. (2011) Psycholinguistic evidence for presuppositions: On-line and off-line data. *Proceedings of Sinn and Bedeutung* 15: 581–95.
- Tsimpli M and Sorace A (2006) Differentiating interfaces: L2 performance in syntax–semantics and syntax–discourse phenomena. *BUCLD Proceedings* 30: 653–64.
- Valenzuela E (2006) L2 end state grammars and incomplete acquisition of the Spanish CLLD constructions. In: Slabakova R, Montrul S, and Prévost P (eds) *Inquiries in linguistic development: In honor of Lydia White*. Amsterdam: Benjamins, pp. 283–304.
- von Fintel K and Matthewson L (2008) Universals in semantics. *The Linguistic Review* 25: 139–201.
- Zehr J, Bill C, Tieu L, Romoli J, and Schwarz F (2016) Presupposition projection from the scope of *none*: Universal, existential, or both? *Proceeding of SALT*, 26, 754.

Appendix I. Examples of four target conditions of cook, play, and drink

Stop + cooking

- (1) a. Negation: *Peggy didn't stop cooking burgers on Wednesday.*
b. Inference: ~ Peggy cooked burgers before Wednesday.
c. No-inference: Peggy didn't cook burgers before Wednesday.
- (2) a. Affirmative: *Peggy stopped cooking burgers on Wednesday.*
b. Inference: ~ Peggy cooked burgers before Wednesday.
c. No-inference: Peggy didn't cook burgers before Wednesday.

Stop + playing

- (3) a. Negation: *Greg didn't stop playing basketball on Wednesday.*
b. Inference: ~ Greg played basketball before Wednesday.
c. No-inference: Greg didn't play basketball before Wednesday.
- (4) a. Affirmative: *Greg stopped playing basketball on Wednesday.*
b. Inference: ~ Greg played basketball before Wednesday.
c. No-inference: Greg didn't play basketball before Wednesday.

Stop + drinking

- (5) a. Negation: *Emma didn't stop drinking orange juice on Wednesday.*
b. Inference: ~ Emma drank orange juice before Wednesday.
c. No-inference: Emma didn't drink orange juice before Wednesday.
- (6) a. Affirmative: *Emma stopped drinking orange juice on Wednesday.*
b. Inference: ~ Emma drank orange juice before Wednesday.
c. No-inference: Miranda didn't drink beer before Wednesday.