L2 tolerance of pragmatic violations of informativeness

Evidence from ad hoc implicatures and contrastive inference

Shuo Feng
Peking University

This study sets out to investigate second language (L2) speakers' derivation of pragmatic inferences and tolerance of violations of informativeness in two types of inferences, i.e., ad hoc implicatures and contrastive inference. The results of a graded judgment task revealed that pragmatic tolerance is inference-specific: L2 speakers were overly tolerant of underinformative statements in ad hoc implicatures than in contrastive inference. In addition, L2 speakers were found to be more relaxed with overinformativeness than underinformativeness in contrastive inference. The fact that L2 speakers tend to be redundant (overinformative) than ambiguous (underinformative) is further discussed with the Pragmatic Principles Violation Hypothesis (Lozano, 2016). This study hopes to contribute to a more find-grained understanding of L2 speakers' abilities of deriving pragmatic inferences.

Keywords: pragmatic tolerance, ad hoc implicatures, contrastive inference, underinformative, overinformative

1. Introduction

One salient and fascinating aspect of human language is the ability to express and comprehend meanings that go beyond strict semantics or the literal meaning. One classic example is (1).

(1) a. I ate some of the cookies.
   b. ~ I ate some but not all the cookies.

The hearer of (1a) may infer that the speaker ate only some of the cookies, not all of them, as in (1b). Otherwise, *all* would be used since it is the maximally informative choice. Inference like (1b), known as *scalar implicature*, is a type

of implicatures wherein a string of words is ordered on a scale with respect to the informational strength, such as *some...most...all* (Horn, 1972; Grice, 1989). According to Grice (1975, 1989), scalar implicatures are also generalized (conversational) implicatures that are usually part of a lexical scale. In contrast, another kind of implicatures, *ad hoc implicatures*, is particularized (conversational) implicatures which arise from real-world context. For example,

(2)   The bag with an apple is pretty.

Imagine two bags are present in the context, one with an apple and the other one with an apple and a banana. (2) implies that the pretty bag is the one with an apple only but not with both an apple and a banana. The context sets up a contrast between the utterance (2) and a stronger alternative *Both bags are pretty* which arises from this particular context. While theories differ in terms of the computation of generalized and particularized implicatures (e.g., Levinson, 2000; Sperber & Wilson, 1986), one salient difference between the two types of implicatures is that the former require knowledge of lexical alternatives on a Horn scale (e.g., *some, all*) and the latter are derived ad-hoc from a specific context.

Grice's Maxim of Quantity (Grice, 1975) under the *Cooperative Principle*, which is especially relevant to the computation of pragmatic implicatures, encourages interlocutors to provide information as much as required by the purpose of the conversation. If the speaker in (1) follows the Quantity Maxim and says (1a) to truly mean (1b), the use of the relatively weaker scalar item *some*, not the stronger *all*, suggests that the speaker is not ready to give a stronger statement that *I ate all of the cookies*. Assuming the speaker intends to mean that *I ate all of the cookies*, (1a) is logically true but pragmatically infelicitous since it is underinformative and violates the Maxim of Quantity. This reasoning process is analogous to inferring ad hoc implicatures: in the same context described above hearers would infer (2) as the pretty bag is only with an apple. In a case that the pretty bag is the one with an apple and a banana, (2) is underinformative and it is optimally informative if the speakers utters (3).

(3)   The bag with an apple and a banana is pretty.

It should be noted that while both (1a) and (2) can be underinformative with certain contextual information provided and the reasoning procedures of computing scalar implicatures and ad hoc implicatures are similar, alternatives are derived in distinctively different ways: alternatives in scalar implicatures are linguistically pre-determined on a scale whereas alternatives in ad hoc implicatures are determined by a particular situation in a context. It is worth noting that other theories are proposed to explain scalar implicatures. For example, Chierchia and colleagues' work (Chierchia, 2006; Chierchia, Fox & Spector, 2012) argued that

an implicit grammatical operator is responsible for scalar implicatures and implicatures are enriched in a recursive, compositional manner of computing meaning with semantic and pragmatic calculation. Carston (1998) proposed that some scalar inference might contribute to what is said rather than to what is implicated.

In addition to ad hoc implicatures, this paper also looks at another type of pragmatic inference (i.e., contrastive inference) expressed in referential expressions (REs). Reference refers to the fact that speakers use words and expressions to uniquely identify objects and entities in the world. For every entity, there is a wide range of expressions available to speakers, e.g., definite noun phrases (e.g., *the mug*), pronouns (e.g., *it*) and specific and explicit modified noun phrases (e.g., *my favorite black mug*). For instance, a big mug in the context is known to a speaker and a hearer and the speaker informs the hearer sentence (4). If a small mug is also available in the context, the referential expression "the mug/it" becomes ambiguous since the hearer would ask "which one". The more appropriate and informative referential expression would be sentence (5).

(4)   My mom bought the mug/it for me.

(5)   My mom bought the big mug for me.

Expressing reference is also a cooperative behavior and follows the Grice's Maxim of Quantity. Speakers are cooperative by offering the appropriate amount of information that is no more or less than required by the context and purpose of the exchange. A referring expression is optimally informative if the hearer is able to uniquely identify the object in the context (e.g., *the big mug* with two mugs differ in size) and underinformative if the reference is ambiguous (e.g., *the mug* in the same situation). A third condition is overinformative/overspecified in a context that only one big mug exists and the speaker says *the big mug*. Generally speaking, upon hearing *the big mug*, interlocutors tend to recognize that there are other mugs around in the reference set and the non-big ones are not considered. This recognition is a pragmatic inference, namely contrastive inference, in that it goes beyond what is linguistically encoded in the sentence.

The current study aims to investigate L2 speakers' ability of drawing pragmatic inferences and tolerance of violations of informativeness in the two types of inferences, i.e., ad hoc implicatures and contrastive inference. One advantage with a focus on violations of informativeness in these two types of inferences is that it provides an opportunity to bring this investigation with the most microscopically studied pragmatic inference in the literature, i.e., scalar implicatures. Recent L2 studies on scalar implicatures have discovered that L2 speakers are able to compute the pragmatically enriched meaning by rejecting underinformative sentences (Miller, Giancaspro, Rothman & Slabakova, 2016; Slabakova, 2010;

Snape & Hosoi, 2018). The large volume of L2 literature on REs has shown that while L2 speakers are overinformative by preferring more prolix forms (NPs and overt pronouns) than required in topic-continuity contexts in both production and comprehension, they are not underinformative since ambiguity leads to a communicative breakdown (e.g., Lozano, 2009, 2018; Bel & García-Alcaraz, 2015; Cunnings, Fotiadou & Tsimpli, 2017; Contemori & Dussias, 2016; Clements & Domínguez, 2017; Jegerski, VanPatten & Keating, 2011; Ryan, 2015). This behavior is captured by pragmatic proposals such as the Pragmatic Principles Violation Hypothesis (PPVH; Lozano, 2016). However, commonly studied REs (such as overt/null pronominals) in these L2 research are greatly influenced by L2 speakers' first language in terms of the null-subject parameter and referential mechanisms. In the current paper, simple and complex NPs (differing in adjective modification) with contrastive inference, as well as ad hoc implicatures, are mainly determined by contextual information, showing little cross-linguistic influence. This allows a more fine-grained investigation of the mechanisms that affect L2 speakers' violation of pragmatic principles. Furthermore, while attempts have been made to look at implicatures other than scalar implicatures (see for example Antoniou & Katsos, 2017; Antoniou, Veenstra, Kissine & Katsos, 2019), it remains unclear to what extent can the findings of L2 acquisition of implicatures and PPVH be applied to other inferences. Therefore, by juxtaposing ad hoc implicatures with contrastive inference, this paper aims to contribute to the growing area of pragmatic tolerance of over/under-informativeness in L2 pragmatics by providing empirical evidence to encompass different pragmatic inferences (i.e., ad hoc implicatures and contrastive inference) that can be accounted for by a single theory such as the PPVH.

2.   Native speakers' processing of ad hoc implicatures and contrastive inference[1]

A large body of psycholinguistic literature has demonstrated that adults are adept at deriving scalar implicatures (Bott, Bailey & Grodner, 2012; Breheny, Ferguson & Katsos, 2013; Breheny, Katsos & Williams, 2006; Huang & Snedeker, 2009 and among many others). However, young children up to 9 years old have difficulties in computing scalar implicatures relying on *some* and other quantifiers (e.g., Noveck, 2001; Papafragou & Musolino, 2003; Huang & Snedeker, 2009; Feeney, Scrafton, Duckworth & Handley, 2004; Su, 2013; Guasti, Chierchia, Crain,

---

1.   This section focuses more on native-speaking children's derivation of ad hoc implicatures and contrastive inference.

Foppolo, Gualmini & Meroni, 2005). Although children's acceptance of underinformative sentences in binary judgment tasks used in the previous studies seemed to suggest that they lack mature pragmatic competence, Katsos and Bishop (2011) argued that children were able to detect the infelicitous situation where a more informative statement could be made by using a ternary judgment task (as summarized in the Pragmatic Tolerance Principle). What prevented children from penalizing underinformative statements might be their inability to generate the set of relevant alternatives for particular scales (Barner, Brooks & Bale, 2011). Studies that made alternatives associated with scale words explicitly available and relevant significantly improved children's performance in computing scalar implicatures (Skordos & Papagfragou, 2016; Tieu, Romoli, Zhou & Crain, 2016). Similarly, the manipulation of providing sufficient contextual cues brought out children's adult-like performance on deriving scalar implicatures (Guasti et al., 2005). The influence of task manipulation by offering contextual information has even been certified in studies testing native-speaking adults. According to Degen and Tanenhaus (2011, 2015), an inference supported by multiple cues of information in the linguistic and discourse context (e.g., naturalness of *some*, presence of other alternatives and question-under-discussion) would be processed rapidly and easily.

The body of literature on young children's computation of scalar implicatures demonstrated that their difficulties with deriving scalar inference might come from immature pragmatic competence, tolerance towards pragmatic violations and limited lexical knowledge to the full set of alternatives and manipulation in test design (e.g., decrease the linguistic complexity and provide more contextual cues) can bootstrap their performance (Degen & Tanenhaus, 2011, 2015; Guasti et al. 2005; Papafragou & Musolino, 2003). Ad hoc implicatures, not lexically encoded on a scale, do not require a lexical retrieval of the scale as in the case of scalar implicatures and are purely dependent on context. Therefore, the more straightforward and context-driven way of computing alternatives in ad hoc implicatures are easier for young children to compute, as evidenced in the literature (Foppolo, Mazzaggio, Panzeri & Surian, 2020; Horowitz, Schneider & Frank, 2018; Wilson & Katsos, 2021). For example, Foppolo et al. (2020), aiming to find out which inference is easier to derive for native children, discovered that deriving scalar, not ad hoc, implicatures was demanding for children before pre-school age and this difficulty disappeared with development. Children as young as three and a half years old are already adult-like in interpreting ad hoc implicatures (Stiller, Goodman & Frank, 2015; Yoon & Frank, 2019).

Similar to the considerable amount of literature on the acquisition of implicatures, native speakers' production and comprehension of REs has been a central topic in many subfields of linguistics (e.g., Sonnenschein, 1982; Maes, Arts &

Noordman, 2004; Kronmuller, Morisseau & Noveck, 2014; Sedivy, Tanenhaus, Chambers, Carlson, 1999; Davies & Katsos, 2010). Numerous production studies have shown that adult speakers often provide redundant or unnecessary information (Belke, 2006; Engelhardt, Bailey & Ferreira, 2006; Sonnenschein, 1982; Pechman, 1989). Such overdescription, for example, is a referential expression with a modifier (*the big mug* or *the mug on the table*) but appears in a context that does not have more than two objects of the same type (two or more mugs). Regarding how overdescription affects comprehension, the results has been mixed since previous studies concluded that extra information either facilitates (Arts, 2004; Davies & Katsos, 2010) or hinders comprehension (Engelhardt et al., 2006; Sedivy, 2007). Sonnenschein (1982) examined how redundant and contrastive messages influence young children's comprehension of REs. Redundant instructions have two discriminating features of a referent which is more than minimally required to be informative, whereas contrastive instructions have the minimum necessary information (one discriminating feature) of the referent to be informative. Sonnenschein discovered that redundant information, not contrastive information, facilitated older children's (9 years old) comprehension rather than younger children (5 years old). Unlike children who benefit from redundant information, in a classic study, Sedivy et al. (1999) showed that on hearing "the tall glass" adult native speakers identified the referent more rapidly in a contrastive condition (e.g., one tall glass and one short glass) than a singleton tall glass condition in which the sentence is overinformative. A more recent study recorded event-related potentials (Engelhardt, Demiral & Ferreira, 2011) further certified the detrimental effect of redundant information in comprehension.

Veenstra, Hollebrandse and Katsos (2017) and Davies and Katsos (2010) who explored children's derivation of ad hoc implicatures and contrastive inference respectively are worth mentioning since the test conditions in these two studies are insightful for the current experiment. Figure 1 shows test conditions in the binary and ternary judgment tasks in Veenstra et al. (2017). The results indicated that children who accepted underinformative sentences in the binary judgment task did not reward the character with the largest strawberry in the ternary judgment task, suggesting that these children were sensitive to informativeness but also tolerant of violations of informativeness.

In Davies & Katsos (2010), participants were presented with an array of four objects under the instruction "Pass me the X". In the overinformative condition, participants saw four objects with only one star accompanied by "Pass me the small star" where the prenominal adjective *small* is redundant since the star is uniquely identifiable. The underinformative condition has two stars but with the instruction "Pass me the star" and the optimal message should be "Pass me the small/big star". Similar to the results in Veenstra et al. (2017) and Katsos

In the basket, there is a…	
shoe	True (1 object)
ball	False (1 object)
shoe and a ball	True (2 objects)
shoe and a hat	False (2 objects)
ball	Under-informative
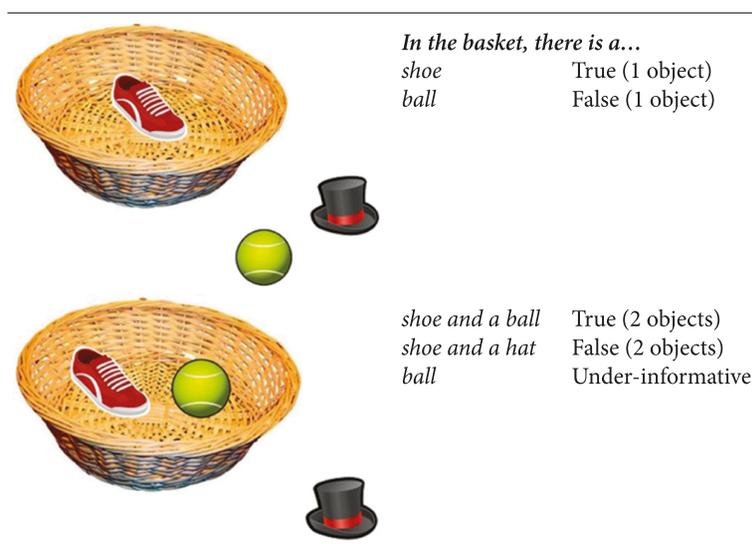
Figure 1. Conditions in the binary and ternary judgment tasks (adopted from Veenstra et al., 2017)

and Bishop (2011), in the ternary task, children rejected overinformative utterances more frequently than logically false statements but less frequently than the optimal counterparts. Confirming the Pragmatic Tolerance Hypothesis again, children were sensitive to informativeness but also tolerant of overinformative utterances in a binary response task.

3.   L2 acquisition of implicatures and referential expressions

L2 speakers are different from native-speaking children in that they have mature mental representations of the kind of language-universal semantic and pragmatic forms and referential function with a fully developed cognitive system during the course of their acquisition of the native language. The universality of pragmatic principles provides an interesting testing ground for L2 acquisition of linguistic properties at the semantics-pragmatics interface (e.g., scalar implicatures). Recently, several studies investigating the learnability issue at the semantics-pragmatics interface have been carried out on the interpretation of scalar implicatures (Miller et al., 2016; Slabakova, 2010; Snape & Hosoi, 2018). L2 speakers are found to be successful at deriving scalar implicatures in an L2 despite of different language pairings, proficiency levels and complexity of lexical items. The very first study that explored L2 acquisition of scalar implicatures was Slabakova

(2010) who looked at how L1-Korean learners interpret two scalar quantifiers, *some* and *all*, in L2-English. Given that Korean quantifier *etten* is roughly equivalent to English *some*, Korean learners should not encounter much difficulties in interpreting *some* since the learning task for them is to map the already available reading in L1 to the new lexical item in L2. In the first experiment, participants were presented with a logically true but pragmatically infelicitous universal statement "Some elephants have trunks" and were asked to respond either "Agree" or "Disagree". The results indicated that being able to compute the pragmatic inference of *some*, Korean learners responded "Disagree" more frequently than native English speakers. Similar results were replicated in the second experiment which, unlike the first experiment, participants judged underinformative sentences in a story context.

One might ask whether Korean learners' successful generation of scalar implicatures might come from positive transfer from their L1. Recent L2 studies demonstrate that this is not the case (Miller et al., 2016; Snape & Hosoi, 2018). For instance, Japanese quantifier *ikutsuka* 'some' has a non-partitive reading and is more likely to be interpreted as *some and possibly all*. This cross-linguistic difference might lead to negative transfer for Japanese learners acquiring the English *some* and *some of*, resulting in high acceptance of the logical reading of the partitive *some of*. Thus, Japanese learners were faced with more difficulties than Korean learners. Snape and Hosoi (2018) tested intermediate and advanced Japanese learners of English. Participants were asked to respond "Yes" or "No" to a question "Are some of the strawberries in the red circle?" when presented with 14 out of 14 strawberries in a red circle. A rejection to the question suggested a successful computation of the inference *some but not all*. The results showed that L1 influence was absent in Japanese learners' computation of scalar implicatures and they even rejected the question more frequently than native English speakers (although this L1~L2 difference was not significantly different), regardless of their English proficiency.

In contrast to the few studies that have explored how L2 speakers derive implicatures, there has been a great number of studies on REs by investigating English definite and indefinite articles (e.g., Crosthwaite, 2014; Ekiert, 2010), overt pronominal subjects (e.g., Cunnings et al., 2017; Mitkovska & Bužarovska, 2018) as well as topic/focus information structure (e.g., Yuan, 1995; Donaldson, 2012). In particular, the choice of REs in subject position (e.g., overt pronominals, null pronominals, NPs), within the domain of anaphora resolution (AR, how an anaphoric expression relates to the reference to the antecedent in the discourse), has been a fruitful avenue of studies using experiments and corpora (e.g., Bel & García-Alcaraz, 2015; Cunnings et al., 2017; Contemori & Dussias, 2016; Clements & Domínguez, 2017; Judy, 2015; Lozano, 2009, 2016; Rothman, 2009; Quesada

& Lazono, 2020). This choice of REs in subject position is greatly influenced by the type of language. Null-subject languages such as Spanish and Greek allow null pronominal subjects, while nonnull subject languages such as English and German require overt pronominal subjects. For the null-subject languages, the alternation between null and overt pronoun in subject position is not free but mediated by information-status factors such as discursive constraints (topic-shift or topic-continuity, contrastive focus).

Previous L2 studies (e.g., Cunnings et al., 2017; Contemori & Dussias, 2016; Clements & Domínguez, 2017; Jegerski et al., 2011) are largely dominated by the investigation of the position of antecedent strategy (PAS). That is, in a background with two antecedents present (e.g., *Marry* and *Jenni*), how an anaphor (e.g., an overt pronoun *she*) biases toward one of the antecedents. It has been found that L2 speakers generally have problems in topic-continuity contexts but not in topic-shift contexts (Contemori & Dussias, 2016; Lozano, 2009, 2016; Clements & Domínguez, 2017) and more importantly, this difficulty has been reported even when cross-linguistic influence is absent, i.e., L1 and L2 converge on the null-subject parameter and AR mechanism (Lozano, 2018; Judy, 2015; Bel & García-Alcaraz, 2015). In particular, in topic-continuity contexts, L2 speakers over-accepted/produced redundant overt subjects that were not required. This non-native like performance of producing redundant REs in topic-continuity contexts was accounted for by the Interface Hypothesis (Sorace, 2011) in that simultaneous integration of syntactic (grammatical knowledge of overt/null pronominals) and discourse (information structure constraints) information is taxing for L2 speakers. Overt pronouns, as the default strategy, posed processing burden minimally to L2 speakers. In other words, such overinformativeness is preferred since it can ease the processing load of keeping track of the saliency of referents in the discourse. However, what is less clear is the nature of overinformativeness in L2 pragmatics when the complexity of syntactic and discourse information is reduced. Thus, the current study sets out to investigate the mechanisms that underpin L2 speakers' overinformativeness by employing two modifications in design: use simple/complex NPs with adjective modification which is more straightforward than language-specific property (null vs. nonnull subjects) and provide contextual information in a picture which is easier to identify potential referents than keeping track of long and complicated textual information.

Another important finding in the literature of L2 acquisition of REs is that while being overinformative, L2 speakers were not ambiguous, supported by L2 speakers' performance in contrastive focus contexts. A contrastive focus scenario is when two potential antecedents differing in gender are introduced in the context, the subject referring to one of the two antecedents has to be an overt pronoun. For example, if the sentence "Although Marry and John worked very

hard for the final exam, he/she/null subject received higher grade" is in Spanish, the subject in the second clause has to be an overt pronoun to clarify disambiguation. A few L2 studies that investigated pronominal subjects in contrastive focus (e.g., Rothman, 2009; Lozano, 2018; Judy, 2015) have reported L2 speakers' native-like performance of providing overt pronouns. L2 speakers' behavior of "better to be redundant than ambiguous" is captured and predicted by PPVH (Lozano, 2016). PPVH categorizes violations of pragmatic principles into two groups depending on which principle is being broken: violations of the Quantity/ Informativeness principle (e.g., being overinformative of using overt pronouns in a topic-continuity context where the antecedent is clear) are mild since the consequence is redundancy without any communicative breakdown; strong violations are violations of the Manner principle (e.g., being ambiguous of the antecedent in a topic-shift context) and are rather infrequent since the ambiguity cannot be resolved, leading to a communicative breakdown. These different reactions to pragmatic violations are not unique to L2 speakers but observed in native adult and child speakers (e.g., Alonso-Ovalle, Fernandez-Solera, Frazier, Clifton. 2002; Jegerski et al., 2011). PPVH, originated in production (corpus) data, was certified in Lozano (2018) who tested L2 speakers' comprehension of null and overt pronouns in topic-continuity, contrastive focus and emphatic contexts in Spanish. The results showed that L2 speakers preferred overt pronouns not only in contrastive-focus contexts to avoid ambiguity but also in topic-continuity contexts which were redundant.

It should be noted that the core rationale in PPVH of using some pragmatic principles as benchmarks is analogous to the understanding of ad hoc implicatures and contrastive inference mentioned in the introduction. For example, similar to the abovementioned contrastive focus scenario, the contrastive inference context in the current study also has two potential antecedents (two objects belong to the same type, e.g., two socks) that differ in one feature, such as size, and the use of a more complex NP (e.g., *the long sock*) is preferred over a simple NP (e.g., *the sock*) to avoid ambiguity. The use of a simple NP leads to a communicative breakdown since the referent of the NP cannot be resolved (hearers would wonder "which sock"), violating the Manner principle. However, if there is only one (long) sock in the context, the complex NP *the long sock*, although redundant by violating the Quantity/Informativeness principle, is clear in its referent. Another important observation from this example is that REs in the current study are simple and complex NPs with contrastive inference (instead of null/overt pronominals or repeated Ns/NPs as what are often investigated in the literature) and this different type of REs does not alter the basic idea of PPVH. Therefore, another goal of the current research is to explore if PPVH is a general pragmatic strategy for L2 speakers that can encompass different pragmatic situations and lin-

guistic phenomena (i.e., ad hoc implicatures and contrastive inference with simple and complex NPs) other than topic continuity/shift and anaphoric REs that it is originally based on.

4.   The present study

4.1   Research questions

The present study aims to answer the following three research questions:

1. Are L1-Mandarin Chinese L2-English speakers sensitive to and pragmatically tolerant of underinformativeness and overinformativeness in interpreting ad hoc implicatures and contrastive inference?
2. Is L2 speakers' level of pragmatic tolerance different between ad hoc implicatures and contrastive inference?
3. Does L2 speakers' pragmatic tolerance pattern differ from the native controls?

Based on previous L2 studies on scalar implicatures (Miller et al., 2016; Slabakova, 2010; Snape & Hosoi, 2018) who found that L2 speakers were able to derive pragmatic enriched meanings, L2 speakers should be able to derive the two types of pragmatic inference in the current study. Furthermore, according to PPVH (Lozano, 2016) which claims that pragmatic violations that lead to communicative breakdown are worse than the ones resulting in mild problems such as redundancy, L2 speakers are predicted to: (a) rate underinformative sentences lower than overinformative sentences in REs; (b) rate underinformative sentences in REs lower than the ones in ad hoc implicatures in that underinformativeness in reference creates a confusion (participants would think "which one do you refer to"), whereas in ad hoc implicatures it only leads to an incomplete or ambiguous description of the context.

4.2   Test design and materials

This study used a sentence judgement task inspired by Veenstra et al. (2017) and Davies and Katsos (2010) for testing L2 speakers' interpretation of ad hoc implicatures and REs. Participants were told that a fictional character *Jack* is either describing objects in a basket or giving an instruction on the objects in a basket. Presented with an array of everyday objects, participants were asked to rate the naturalness of the description or instruction on a 7-point likert scale (1 = unnatural, 7 = natural).

In the design for ad hoc implicatures, all the test items were presented with a carrier description sentence "In the basket, there is a/an [noun phrase]". Descriptions and pictures were divided into five conditions, see Table 1.

Table 1. Conditions and experimental displays for ad hoc implicatures

Conditions	Experimental displays
Underinformative	In the basket, there is an apple.
Optimal 1	In the basket, there is a potato.
Optimal 2	In the basket, there is a book and a coke.
False 1	In the basket, there is a hat.
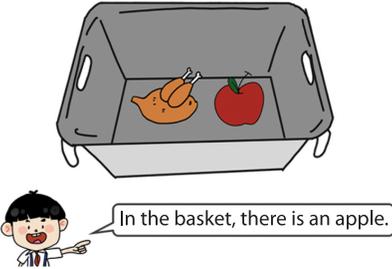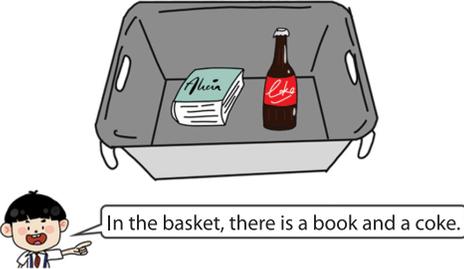
Table 1. *(continued)*

| Conditions	Experimental displays
False 2	
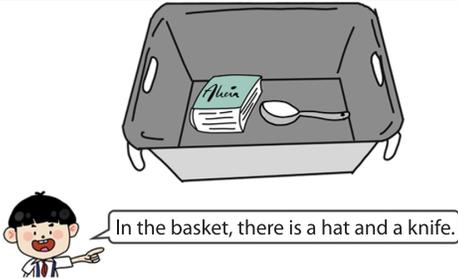
The critical condition is the underinformative condition where the character Jack only mentions one object in the description with two objects existing in the basket. The other four conditions served more like filler conditions in that the four types of fillers are either completely correct or wrong, whereas the target items are ambiguously felicitous depending on different readings. In the two optimal conditions, Jack's description matches with what is in the basket. The description in the two false conditions is completely false since it mentions objects that do not exist in the basket.[2] There were 6 items in the critical condition and 16 items in the filler conditions (4 items * 4 conditions).

Test materials for REs were created by a 2×2 design with contrast (in the visual display) and modification (in the instruction sentence) as the independent variables (see Table 2).

The presence of a contrast in the visual display (e.g., a long sock vs. a short sock) and the presence of an adjective modification (e.g., *Pass me the long sock*) are the optimal and informative condition in that only by the adjective modification can the referred object be identified. The absence of a contrast with only one object appeared in the display does not require additional information of the referent and therefore, the instruction without an adjective is informative (e.g., *Pass me the sock* presented with only one sock). Otherwise, an instruction with an adjective seems to be unnecessary and redundant in such condition, result-

---

2. The False 2 condition in the current study differs from False (2 objects) in Veenstra et al. (2017). In their study, the description in the False 2 condition is partially correct in that it mentions one object in the basket and the other one is a new object (e.g., "In the basket, there is a shoe and a hat" with a shoe and a ball in the basket). Although the coordinating description is wrong, the results showed that participants were confused, offering medium-sized strawberries with longer reaction times in such condition. To remove such confusion, neither of the objects in the description of the False 2 condition in the current experiment appear in the basket, resulting in a completely wrong description.

Table 2. Conditions and experimental displays for REs

		Contrast	
		+	−
Modification	+	Optimal-2	Overinformative
		Pass me the long sock.	_Pass me the long sock._
	−	Underinformative	Optimal-1
		Pass me the sock.	_Pass me the sock._

ing in an overinformative instruction (e.g., _Pass me the long sock_ presented with only one sock). The critical condition is the underinformative condition where the referent is ambiguous in the instruction, i.e., _Pass me the sock_ presented with two objects with a contrasting feature (e.g., a long sock and a short sock). There were 12 items for the under-/over-informative conditions (6 items * 2 conditions) and 8 for the two optimal conditions (4 items * 2 conditions). Additionally, four false items were included where the referent in the instruction does not exist in the basket.

Taken together, there were 6 targets (in the underinformative condition) in ad hoc implicatures and REs respectively and 34 fillers (16 + 18). Since the current experiment were part of a large experimental project, there were other 24 test items for another experiment. In sum, the participants finished a total of 70 items for approximately 10–15 minutes.

4.3   Participants and procedures

The experiment was administered online by using Credamo (www.credamo .com), a reliable Chinese data-collection platform which is similar to Qualtrics Online Sample. Native speakers were recruited on the online crowdsourcing web-site *Prolific* (which is similar to Amazon Mechanical Turk) and they were directed to the Credamo website where the experiment was hosted. After signing a consent form, both native and L2 speakers were presented with an example item for ad hoc implicatures and REs respectively, followed by four practice trials to further familiarize themselves with the task. Participants received monetary compensation for their participation.

Forty-nine Chinese speakers of English (33 female) and twenty-one native speakers of English (11 female) participated in this study (Table 3). Regarding L2 speakers' English proficiency (see Appendix 1), participants were asked to report English proficiency tests they have taken and the scores.[3] All of the participants have taken the College English Test (CET), a large-scale national English proficiency exam in China with two levels, i.e., CET-4 and CET-6. CET-4, as a mandatory English exam all Chinese undergraduate students, aims for "a relatively high level of competence in reading, and intermediate level of competence in listening, writing and speaking" (National College English Syllabus for Non-English Majors, 1999, p.1). CET-6, available to students who obtained the CET-4 certificate, is more difficult. Fifteen participants reported CET-4 scores, four participants reported either TOEFL or IETLS and the rest of them reported CET-6 scores. Therefore, the L2 participants were roughly considered as intermediate to advanced learners.

Table 3. Participants' background

	Age at testing (years)		Years studying English	
	M (SD)	**Range**	**M (SD)**	**Range**
Native speakers ($n = 21$)	31.2 (6.7)	21–46	n/a	n/a
L1-Chinese L2-English speakers ($n = 49$)	21.4 (2.2)	18–28	13.1 (2.9)	6–19

3. Since the current paper did not aim to explore L2 development particularly, L2 speakers' proficiency was used to make sure that they were able to understand the experiment and finish the task. However, since some previous studies have shown that as proficiency increases, L2 speakers become more nativelike in anaphor resolution (e.g., Bel, Sagarra, Comínguez & García-Alcaraz, 2016), future research with a consistent proficiency test should be conducted to further explore L2 development and ultimate attainment. I thank an anonymous reviewer for pointing this out.

4.4   Data analysis

Data analysis was conducted in *R* version 3.5.1 (R Core Team, 2018). The dependent variable was ratings on a 7-point likert scale and were analyzed using ordinal regression models rather than linear regression models (Liddell & Kruschke, 2018). All the models were cumulative link mixed models using *clmm()* function from the *ordinal* package (Christensen, 2019) in *R*. One advantage of using ordinal modeling rather than linear modeling is that ordinal regressions do not assume that participants treat the ratings as equally ordered. Another advantage has to do with the fact that raw ratings, instead of z-scored ratings, are calculated in the model. Z-scored ratings in linear models are beneficial in keeping the measure without interruption and ruling out between-participant variation. However, the use of *clmm* eliminates the necessity for continuity and random effects in mixed models already account for between-subject variation.

Unless otherwise mentioned, following the common practice as suggested by Barr, Levy, Scheepers and Tily (2013), cumulative link mixed effect models with maximal random effects structure were fitted by using the *clmm*() function from the *ordinal* package in *R*. These models included random intercepts and slopes for participants and items.

5.   Results

The ratings of the four conditions in REs by the two participant groups are illustrated in Figure 2. The results indicated that both native and L2 speakers were able to give the lowest scores in the false condition and highest scores in the two optimal conditions, with overinformative and underinformative in the middle. Overinformative conditions were rated higher than underinformative conditions. High ratings in the two optimal conditions did not seem to be different.

A cumulative link mixed effect model was fitted, with the 7-point Likert scale ratings (1 = very unnatural, 7 = very natural) as the dependent variable. The independent variables were modification (2 levels: presence or absence of an adjective modification), contrast (2 levels: presence or absence of a contrast in visual displays) and language (2 levels: English and Chinese).

The results in Table 4 indicated significant effects of contrast and modification, as well as a significant interaction effect between the two factors. It suggests that utterances with an adjective modification were rated higher than utterances with no modification; utterances with the appearance of a contrast in the display were rated higher than utterances without the contrast. Further pairwise planned comparisons indicated that underinformative and overinforma-
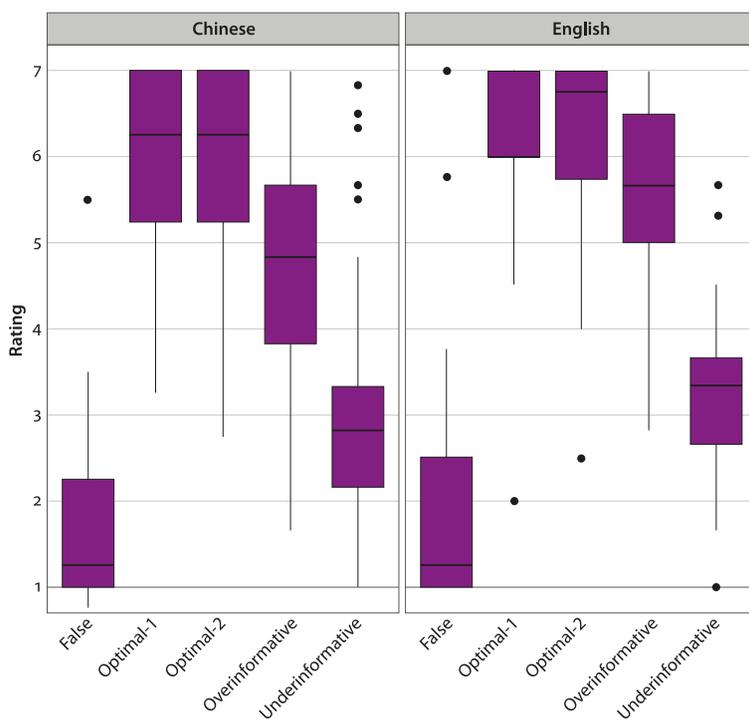
Figure 2. Mean ratings for all conditions in REs

Table 4. Output of the cumulative link mixed model for REs (reference level for contrast: -contrast; reference level for modification: -modification; reference level for language: English)

Effect	Estimate β	Std. Err.	z value	p value
Contrast (-contrast vs. +contrast)	−5.305	0.709	−7.485	< 7.15e-14 ***
Modification (-modification vs. +modification)	−2.132	0.597	−3.568	< 0.0001 ***
Language (English vs. Chinese)	−1.198	0.572	−2.095	0.036 *
Contrast : Modification	7.149	1.001	7.096	1.29e-12 ***
Contrast : Language	0.768	0.778	0.987	0.323
Modification : Language	−0.016	0.551	−0.029	0.977
Contrast : Modification : Language	−0.299	1.113	−0.269	0.788

Note.

* *p* < 0.05   ** *p* < 0.01   *** *p* < 0.001.

tive utterances were rated significantly lower than their corresponding utterances in the Optimal-2 and Optimal-1 conditions respectively ($Z = -10.059$, $p < .0001$; $Z = 5.217$, $p < .0001$). The rating differences between overinformative and underinformative statements also reached statistically significance ($Z = 6.077$, $p < .0001$). However, the two optimal conditions were not significantly different from each other ($Z = 0.138$, $p = .999$). Although the effect of language is significant, the abovementioned response pattern did not differ between the two language groups since language does not interact with other factors.

Chinese and native speakers' reading patterns of ad hoc implicatures are presented in Figure 3. Generally speaking, native English speakers were more or less at ceiling for rejecting logically false utterances (in false-1 and false-2 conditions) and accepting optimal utterances (in optimal-1 and optimal-2 conditions). Moreover, their ratings of underinformative utterances were lower than optimal utterances but higher than false utterances, as were Chinese speakers' ratings. However, Chinese speakers showed greater variation in their ratings in false and optimal utterances.
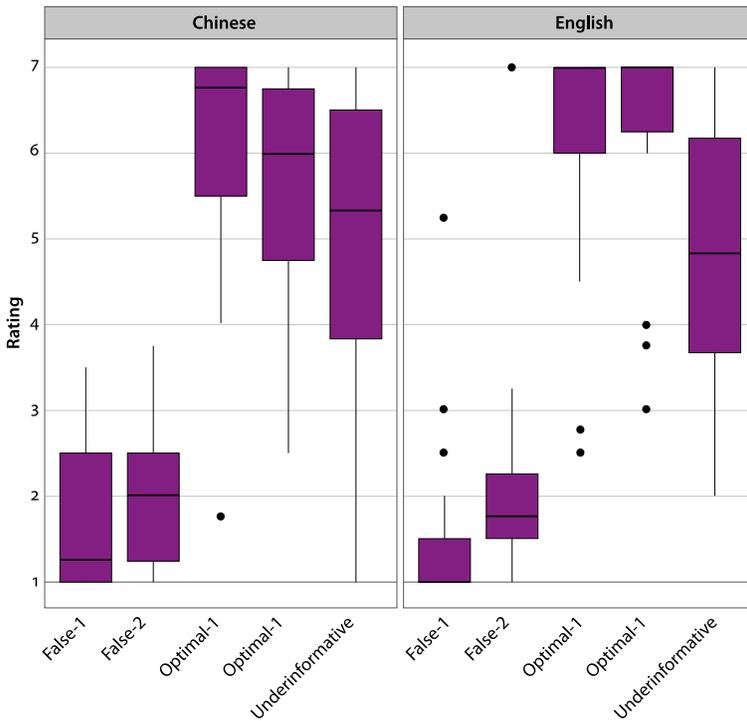

Figure 3. Mean ratings for all conditions in ad hoc implicatures

The data were fitted with a cumulative link mixed effect ordinal regression, i.e., ratings as the dependent variable, condition (5 levels: false-1, false-2, optimal-1, optimal-2, underinformative) and language (2 levels: English and Chinese) as fixed effects. For a complete description of the model, see Table 5. The ratings of optimal-1, as the reference level, were significantly different from those of false-1, false-2 and underinformative, but not optimal-2. The main effect of language is not significant when L2 speakers were compared to native speakers. The significant interaction suggested that if rating difference between Chinese learners and native speakers was compared in the optimal-1 and optimal-2 condition, the difference between groups was more pronounced in the optimal-2 condition, as is clear in Figure 3.

Table 5. Output of the cumulative link mixed model for ad hoc implicatures (reference level for condition: Optimal-1; reference level for language: English)

Effect	Estimate β	Std. Err.	z value	p value
Condition				
optimal-1 vs. false-1	−9.979	1.369	−7.291	< 3.08e-13 ***
optimal-1 vs. false-2	−8.060	1.170	−6.891	< 5.55e-12 ***
optimal-1 vs. optimal-2	0.060	0.746	0.080	.936
optimal-1 vs. underinformative	−3.537	0.807	−4.486	1.16e-05 ***
Language (English vs. Chinese)	−0.762	0.918	−0.831	.406
Condition : Language				
false-1 : Chinese	1.575	1.457	1.081	.406
false-2 : Chinese	0.693	1.230	0.564	.573
optimal-2 : Chinese	−1.422	0.623	−2.282	.025 *
underinformative : Chinese	1.069	0.767	1.393	.164

Note.

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

Lastly, the analysis looked at comparing participants' ratings between ad hoc implicatures and references between the two language groups (see Figure 4). In the statistical analysis, overinformative condition was excluded in that it is unique in the REs. A *clmm* model was fitted, including condition (3 levels: optimal, underinformative, false), inference type (2 levels: ad hoc implicatures and contrastive inference) and language (2 levels: English and Chinese) as fixed effects (see Table 6). The results showed that ratings in the underinformative condition (as the reference level) were significantly different from the other two conditions.

There was also a significant effect of inference type, but no effect of language. Since the two-way interaction between condition and inference type reached significance for one condition and marginally significant for another condition, further pairwise analysis was conducted. It showed that ratings between the two inference types were similar when they were compared in the false and optimal conditions (false: $Z=-0.431$, $p=0.998$; optimal: $Z=-1.535$, $p=0.642$), but not in the underinformative condition ($Z=-5.350$, $p<.0001$). In other words, underinformative utterances in ad hoc implicatures were rated significantly higher than the ones in REs for both language groups.
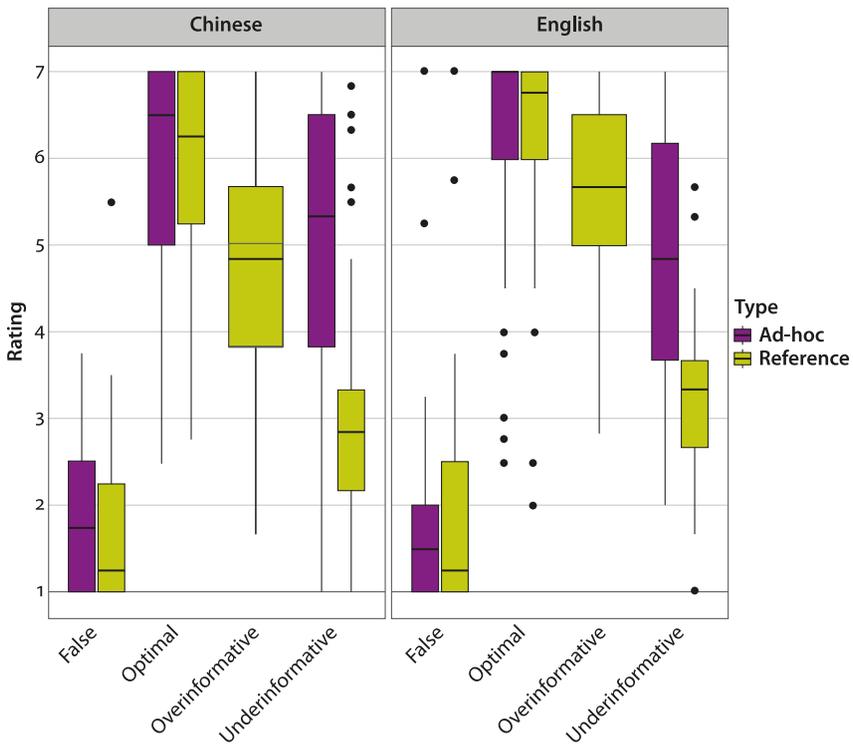
Figure 4. Mean ratings for ad hoc implicatures and REs by the two groups of participants

Table 6. Output of the cumulative link mixed model (reference level for condition: Underinformative; reference level for inference type: Contrastive inference; reference level for language: English)

Effect	Estimate β	Std. Err.	z value	p value
Condition				
underinformative vs. false	−2.771	0.857	−3.233	.001 **
underinformative vs. optimal	4.827	0.672	7.181	6.92e-13 ***
Inference type (contrastive inference vs. ad hoc implicatures)	2.022	0.577	3.504	.0004 ***
Language (English vs. Chinese)	−0.290	0.430	−0.673	.501
Condition : Inference type				
false : ad hoc implicatures	−2.279	0.722	−3.155	.002 ***
optimal : ad hoc implicatures	−1.271	0.667	−1.906	.057
Condition : Language				
false : Chinese	−0.413	0.883	−0.468	.640
optimal: Chinese	−0.685	0.661	−1.036	.300
Inference type : Language (ad hoc implicature: Chinese)	0.597	0.514	1.163	.245
Condition : Inference type: Language				
false : ad hoc implicatures: Chinese	0.326	0.536	0.609	.543
optimal : ad hoc implicatures: Chinese	−0.825	0.494	−1.668	.095

Note.

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

6.   Discussion

The present study aimed to answer the following three research questions:

1.   Are L1-Mandarin Chinese L2-English speakers sensitive to and pragmatically tolerant of underinformativeness and overinformativeness in interpreting ad hoc implicatures and contrastive inference?
2.   Is L2 speakers' level of pragmatic tolerance different between ad hoc implicatures and contrastive inference?
3.   Does L2 speakers' pragmatic tolerance pattern differ from the native controls?

The answer to the first research question is yes. L2 speakers showed excellent semantic knowledge of the two types of inferences and sensitive tolerance of prag-

matic violations: they were at ceiling in judging sentences where the inference was optimal or logically false. They were also sensitive to intricate violations of overinformativeness in contrastive reference, being more tolerant of overinformativeness than underinformativeness. Regarding the second question, L2 speakers' level of pragmatic tolerance between the two types of inferences only differed in the underinformative conditions, i.e., they were more tolerant of violations of informativeness in ad hoc implicatures than in contrastive reference. Lastly, L2 and native speakers showed very similar response patterns and more importantly, both groups gave higher ratings of underinformative statements in ad hoc implicatures than in contrastive inference. Overall, findings in this study are consistent with previous L2 studies in demonstrating that L2 speakers were able to derive pragmatic inferences (Miller et al., 2016; Slabakova, 2010; Snape & Hosoi, 2018). Extending previous studies, the current study showed that L2 speakers were not only sensitive to pragmatic violations of underinformativeness, but also overinformativeness. L2 speakers, in the current experiment, were found to be more inclined to accept underinformative statements in deriving ad hoc implicatures than in contrastive reference. The overall scale of L2 speakers' pragmatic tolerance is (from the most tolerant to the least tolerant): ad hoc implicatures > scalar implicatures[4] > contrastive inference.

First, it is not surprising to see that context-driven ad hoc implicatures can bootstrap L2 performance since this type of implicatures is linguistically less complicated and often provided with sufficient contextual information. An example of underinformativeness in scalar implicatures is a context shows that John folded five out of five t-shirts but with a statement "John folded *some* t-shirts". Upon hearing *some*, participants simultaneously activate another scale mate on the same scale with *some*, i.e., *all*. Although the underinformative statement is logically true, the competition of using the alternative scale mate in "John folded *all* the t-shirts" is very strong, leading to a clear recognition that the statement with *some* is far from optimal. However, deriving ad hoc implicatures from "In the basket, there is an apple" is only based on the contextual alternative and the push for rejecting this sentence is much weaker than in the scalar implicature example. It might be the case that the dual process of retrieving alternatives in scalar implicatures, instead of being costly for L2 speakers since they already had mature cognitive abilities and knowledge of lexical scales, provided strong evidence that the underinformative statement is clearly less than optimal. But underinformative statements in ad hoc implicatures, solely depending on contextual alterna-

---

4. That scalar implicature is placed in the middle on the pragmatic tolerance scale is based on results reported in Feng (under review) who used a graded judgment task similar to this study to investigate how L2 speakers interpret scalar implicatures.

tives, were under less pressure to be rejected. Also, the sufficient contextual cues provided in the task of ad hoc implicatures facilitated L2 speakers in disambiguating different readings since the task effect that increased cues aid adults and children in deriving pragmatic inferences is already evident in the literature (e.g., Degen & Tanenhaus, 2011, 2015; Guasti et al., 2005), as discussed in Section 2.[5] Additionally, the tolerance scale confirms the expectation that contrastive inference received the lowest tolerance in that underinformativeness in contrastive inference is a stronger violation than the other two types of implicatures (PPVH; Lozano, 2016). For instance, on hearing "Pass me the sock" when there are two socks in the context, interlocutors would think "which sock are you referring to". The ambiguous referential expression *the sock* is underinformative in this context and leads to a communication breakdown.

Besides underinformativeness, another important focus of the current research is L2 speakers' pragmatic tolerance of overinformative statements in contrastive inference. According to PPVH (Lozano, 2016), overinformativeness is a mild violation of informativeness. For example, in a context where there is only one long sock, "Pass me the sock" is already sufficient enough for identifying the referent, whereas the overinformative statement "Pass me the long sock" carries extra information than is necessary for identification. However, unlike underinformative statements in contrastive reference which lead to a communication breakdown (referents cannot be identified), overinformative statements only has redundant information and the referent can be resolved. Adult native speakers are reported to often provide extra information in production (Belke, 2006; Engelhardt et al., 2006). Although overinformative statements might be regarded as felicitous as optimal statements since some studies suggested that extra information is beneficial for comprehension (e.g., Arts, 2004; Davies & Katsos, 2010), interlocutors would be more tolerant of overinformativeness (redundancy) than underinformativeness (ambiguity) in contrastive inference (Lozano, 2016). Confirming this prediction, L2 speakers in the current study rated overinformative statements significantly higher than underinformative statements (but still lower

---

5. Another potential influence on the high acceptance of underinformative utterances in ad hoc implicatures might come from Chinese counterparts of *there-be* sentences in English, i.e., *you-* and *shi-* sentences. *You* is solely existential, whereas *shi* has an exclusive reading (Huang, 1987; Paul, Lu & Lee, 2019). If the experimental display in the underinformative condition in ad hoc implicatures (see Table 1) is accompanied with a *you* sentence in Chinese, it is very likely that the sentence would be accepted since *you* purely indicates existence. Therefore, Chinese participants in the current experiment might not find the underinformative statement in ad hoc implicatures to be pragmatically infelicitous and severe enough for a complete rejection; therefore, they were more prone to accept it. I appreciate suggestions and comments from Xin Yan and Siyu Wang.

than optimal statements). Taken together, the results that L2 speakers are redundant (high tolerance of overinformativeness) but not ambiguous (lower tolerance of underinformativeness in contrastive inference than in ad hoc implicatures) are in line with PPVH, which was originally proposed to capture anaphor resolution of REs such as overt/null pronominals. This study demonstrates that PPVH can account for a wide range of pragmatic situations (topic-continuity/shift, ad hoc implicatures, contrastive inference) involving informativeness with different linguistic properties (overt/null pronominals, simple and complex NPs).

Lastly, L2 speakers' sensitivity to the subtle differences between the two types of violation of informativeness opens up a new venue for further studies. For instance, it still remains unclear how overinformativeness actually affects L2 comprehension. Is the redundant information "a waste" that has little impact on comprehension or "a scaffold/block" that accelerates/impedes L2 processing of contrastive inference? Future research is needed to better understand not just underinformativeness in implicature studies, but also how overinformativeness affects L2 processing of pragmatic inferences. Another direction for future research is to investigate L2 speakers' online processing of deriving pragmatic inference motivated by the dichotomy between "logical" and "pragmatic" responders in interpreting underinformative sentences reported in native speakers' literature (e.g., Bott & Noveck, 2004; Noveck & Posada, 2003). In addition, cognitive tests such as the Autism Spectrum Quotient (Baron-Cohen et al., 2001) can be implemented since it helps to address to what extent individual differences and personality-based factors can account for the variation in implicature derivation (Nieuwland, Ditman & Kuperberg, 2010; Yang, Minai & Fiorentino, 2018).[6]

7.   Conclusion

The goal of the current study was to investigate L2 speakers' pragmatic tolerance of violation of informativeness in two types of inferences, i.e., ad hoc implicatures and contrastive inference, comparing with scalar implicatures that are already given serious attention by L2 researchers. The underinformative and overinformative scenarios in the three types of inferences are operationalized as follows:

Underinformative in ad hoc implicatures – *In the basket, there is an apple* (in a context with an apple and a book).

---

6. I appreciate this suggestion from an anonymous reviewer. Some individual results from this experiment also call for more studies in this direction, e.g., Optimal-2 condition in ad hoc implicatures was rated below 5 by 17 (out of 49) L2 speakers but only 3 (out of 21) native speakers.

Underinformative in contrastive inference – *Pass me the sock* (in a context with a long sock, a short sock, a book and a hat).
Underinformative in scalar implicatures[7] – *John folded some of the t-shirts* (in a context that John folded all the five t-shirts).
Overinformative in contrastive inference – *Pass me the long sock* (in a context with a long sock, a pen, a book, a hat).

The results revealed that L2 speakers were sensitive to underinformativeness to avoid ambiguity since its violation leads to a communication breakdown, whereas they were relaxed with overinformativeness that simply leads to redundancy. The finding that L2 speakers are redundant but not ambiguous in interpreting ad hoc implicatures and contrastive inference supports PPVH and further demonstrates that PPVH can be extended to a much wider range of pragmatic inferences and linguistic properties than it is originally based on. This study has provided new empirical evidence with respect to L2 speakers' abilities of deriving pragmatic inferences in L2 pragmatics and hopes to offer some insights for future research in this direction.

References

Alonso-Ovalle, L., Fernandez-Solera, S., Frazier, L., & Clifton, C. (2002). Null vs. overt pronouns and the topic-focus articulation in Spanish. *Journal of Italian Linguistics*, *14*(2), 151–169.

---

7. This condition regarding scalar implicatures is not specifically tested in the current experiment. However, it is highly relevant to this study and has been referred to throughout the paper. Thus, I added it here for a clear comparison among these different types of inferences.

Antoniou, K., & Katsos, N. (2017). The effect of childhood multilingualism and bilectalism on implicature understand. *Applied Psycholinguistics*, *38*, 787–833. https://doi.org/10.1017/S014271641600045X

Antoniou, K., Veenstra, A., Kissine, M., & Katsos, N. (2019). How does childhood bilingualism and bi-dialectalism affect the interpretation and processing of pragmatic meanings? *Bilingualism: Language and Cognition*, *23*(1), 186–203. https://doi.org/10.1017/S1366728918001189

Arts, A. (2004). Overspecification in instructive texts. Doctoral dissertation, Tillburg University.

Barner, D., Brooks, N., & Bale, A. (2011). Accessing the unsaid: the role of scalar alternatives in children's pragmatic inference. *Cognition*, *118*(1), 84–93. https://doi.org/10.1016/j.cognition.2010.10.010

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism Development Disorder*, *31*, 5–17. https://doi.org/10.1023/A:1005653411471

Barr, D., Levy, R., Scheepers, C., & Tily, H. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bel, A., & Garcia-Alcaraz, E. (2015). Subject pronouns in the L2 Spanish of Moroccan Arabic speakers. In T. Judy and S. Perpinan (Eds.), *The acquisition of Spanish in understudied language pairings* (pp. 201–232). John Benjamins. https://doi.org/10.1075/ihll.3.08bel

Bel, A., Sagarra, N., Cominguez, J.P., & Garcia-Alcaraz, E. (2016). Transfer and proficiency effects in L2 processing of subject anaphora. *Lingua*, *184*, 134–159. https://doi.org/10.1016/j.lingua.2016.07.001

Belke, E. (2006). Visual determinants of preferred adjective order. *Visual Cognition*, *14*, 261–294. https://doi.org/10.1080/13506280500260484

Bott, L., Bailey, T.M., & Grodner, D. (2012). Distinguishing speed from accuracy in scalar implicatures. *Journal of Memory and Language*, *66*, 123–142. https://doi.org/10.1016/j.jml.2011.09.005

Bott, L. & Noveck, I.A. (2004). Some utterances are underinformative: The onset and time course o scalar inferences. *Journal of Memory and Language*, *51*, 437–457. https://doi.org/10.1016/j.jml.2004.05.006

Breheny, R., Ferguson, H.J., & Katsos, N. (2013). Investigating the timecourse of accessing conversational implicatures during incremental sentence interpretation. *Language and Cognitive Processes*, *28*(4), 443–467. https://doi.org/10.1080/01690965.2011.649040

Breheny, R., Katsos, N., & Williams, J.N. (2006). Are generalised scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences. *Cognition*, *100*(3), 434–463. https://doi.org/10.1016/j.cognition.2005.07.003

Carston, R. (1998). Informativeness, relevance and scalar implicature. In R. Carston and S. Uchida (Eds.), *Relevance theory: Applications and implications* (pp. 179–236). John Benjamins. https://doi.org/10.1075/pbns.37.11car

Chierchia, G. (2006). Broaden your views: Implicatures of domain widening and the "logicality" of language. *Linguistic Inquiry*, *37*, 535–90. https://doi.org/10.1162/ling.2006.37.4.535

Chierchia, G., Fox, D., & Spector, B. (2012). Scalar implicature as a grammatical phenomenon. In P. Portner, C. Maienborn, & K. von Heusinger (Eds.), *Semantics: An international handbook of natural language meaning* (pp. 2297–2331). Berlin: Walter de Gruyter.

Christensen, R. H. B. (2019). Ordinal: Regression Models for Ordinal Data. *R package version 2019*, 12–10.

Clements, M., & Domínguez, L. (2017). Reexamining the acquisition of null subject pronouns in a second language: Focus on referential and pragmatic constraints. *Linguistic Approaches to Bilingualism*, *7*(1), 33–62. https://doi.org/10.1075/lab.14012.cle

Contemori, C., & Dussias, P. E. (2016). Referential choice in a second language: Evidence for a listener-oriented approach. *Language, Cognition and Neuroscience*, *31*(10), 1257–1272. https://doi.org/10.1080/23273798.2016.1220604

Crosthwaite, P. (2014). Definite discourse-new reference in L1 and L2: A study of bridging in Mandarin, Korean, and English. *Language Learning*, *64*(3), 456–492. https://doi.org/10.1111/lang.12062

Cunnings, I., Fotiadou, G., & Tsimpli, I. M. (2017). Anaphora resolution and reanalysis during L2 sentence processing: Evidence from the visual world paradigm. *Studies in Second Language Acquisition*, *39*, 621–652. https://doi.org/10.1017/S0272263116000292

Davies, C., & Katsos, N. (2010). Over-informative children: Production/comprehension asymmetry or tolerance to pragmatic violations? *Lingua*, *120*, 1956–1972. https://doi.org/10.1016/j.lingua.2010.02.005

Degen, J., & Tanenhaus, M. K. (2011). Making inference: The case of scalar implicature processing. In L. A. Carlson, C. Holscher & T. F. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*, 3299–3304.

Degen, J., & Tanenhaus, M. K. (2015). Processing scalar implicature: A constraint-based approach. *Cognitive science*, *39*(4), 667–710. https://doi.org/10.1111/cogs.12171

Donaldson, B. (2012). Syntax and discourse in near-native French: clefts and focus. *Language Learning 62*, 902–930. https://doi.org/10.1111/j.1467-9922.2012.00701.x

Ekiert, M. (2010). Linguistic effects on thinking for writing: The case of articles in L2 English. In Z. Han & T. Cadierno (Eds.), *Linguistic relativity in SLA: Thinking for speaking* (pp. 125–153). Multilingual Matters. https://doi.org/10.21832/9781847692788-008

Engelhardt, P. E., Bailey, K. G. D., & Ferreira, F. (2006). Do speakers and listeners observe the Gricean maxim of quantity? *Journal of Memory and Language*, *54*, 554–573. https://doi.org/10.1016/j.jml.2005.12.009

Engelhardt, P. E., Demiral, S. B., & Ferreira, F. (2011). Over-specified referring expressions impair comprehension: An ERP study. *Brain and Cognition*, *77*, 304–314. https://doi.org/10.1016/j.bandc.2011.07.004

Feeney, A., Scrafton, S., Duckworth, A., & Handley, S. J. (2004). The story of some: Everyday pragmatic inference by children and adults. *Canadian Journal of Experimental Psychology*, *58*(2), 121–132. https://doi.org/10.1037/h0085792

Feng, S. (under review). L2 speakers' tolerance of pragmatic violations in interpreting scalar implicatures.

Foppolo, F., Mazzaggio, G., Panzeri, F., & Surian, L. (2020). Scalar and ad hoc pragmatic inferences in children: guess which one is easier. *Journal of Child Language*, *48*(2), 350–372. https://doi.org/10.1017/S030500092000032X

Grice, P. (1975). Logic and conversation. In: Davidson, D. and Harman, G. (Eds.), *The logic of grammar* (pp. 64–75). Encino, CA: Dickenson.

Grice, P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.

Guasti, M. T., Chierchia, G., Crain, S., Foppolo, F., Gualmini, A., & Meroni, L. (2005). Why children and adults sometimes (but not always) compute implicatures. *Language and Cognitive Processes*, *20*, 667–696. https://doi.org/10.1080/01690960444000250

Horn, L. (1972). On the semantic properties of logical operators in English. [Unpublished doctoral dissertation]. Bloomington, USA.

Horowitz, A. C., Schneider, R. M., & Frank, M. C. (2018). The trouble with quantifiers: Exploring children's deficits in scalar implicature. *Children Development*, *89*(6), e572–e593. https://doi.org/10.1111/cdev.13014

Huang, C. -T James. (1987). Existential sentences in Chinese and (in)definiteness. In E. Reuland (Ed.), *Between syntax and semantics* (pp. 63–85). Taylor & Francis.

Huang, Y. T., & Snedeker, J. (2009). Online interpretation of scalar quantifiers: Insight into the semantics–pragmatics interface. *Cognitive Psychology*, *58*, 376–415. https://doi.org/10.1016/j.cogpsych.2008.09.001

Jegerski, J., Van Patten, B., & Keating, G. D. (2011). Cross-linguistic variation and the acquisition of pronominal reference in L2 Spanish. *Second Language Research*, *27*(4), 481–507. https://doi.org/10.1177/0267658311406033

Judy, T. (2015). Knowledge and processing of subject-related discourse properties in L2 near-native speakers of Spanish, L1 Farsi. In T. Judy & S. Perpinan (Eds.), *The acquisition of Spanish in understudied language pairings* (pp. 169–199). John Benjamins. https://doi.org/10.1075/ihll.3.07jud

Katsos, N., & Bishop, D. V. M. (2011). Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition*, *120*(1), 67–81. https://doi.org/10.1016/j.cognition.2011.02.015

Kronmuller, E., Morisseau, T., & Noveck, I. (2014). Show me the pragmatic contribution: A developmental investigation of contrastive inference. *Journal of Child Language*, *41*, 985–1014. https://doi.org/10.1017/S0305000913000263

Levinson, S. (2000). *Presumptive Meanings*. MIT Press. https://doi.org/10.7551/mitpress/5526.001.0001

Liddell, T. M., & Kruschke, J. K. (2018). Analyzing ordinal data with metric models: What could possibly go wrong? *Journal of Experimental Social Psychology*, *79*, 328–348. https://doi.org/10.1016/j.jesp.2018.08.009

Lozano, C. (2009). Selective deficits at the syntax-discourse interface: Evidence from the CEDEL2 corpus. In Y.-I. Leung, N. Snape & M. Sharwood-Smith (Eds.), *Representational deficits in second language acquisition* (pp. 127–166). John Benjamins. https://doi.org/10.1075/lald.47.09loz

Lozano, C. (2016). Pragmatic principles in anaphora resolution at the syntax-discourse interface: Advanced English learners of Spanish in the CEDEL2 corpus. In M. Alonso-Ramos (Ed.), *Studies in corpus linguistics* (Vol. *78*, pp. 235–265). John Benjamins.

Lozano, C. (2018). The development of anaphora resolution at the syntax-discourse interface: Pronominal subjects in Greek learners of Spanish. *Journal of Psycholinguistic Research*, *47*(2), 411–430. https://doi.org/10.1007/s10936-017-9541-8

Maes, A., Arts, A., & Noordman, L. (2004). Reference management in instructive discourse. *Discourse Processes*, *37*, 117–144. https://doi.org/10.1207/s15326950dp3702_3

Miller, D., Giancaspro, D., Iverson, M., Rothman, J., & Slabakova, R. (2016). Not just *algunos*, but indeed *unos* L2ers can acquire scalar implicatures in L2 Spanish. *Language Acquisition Beyond Parameters*, *51*, 125–145. https://doi.org/10.1075/sibil.51.06mil

Mitkovska, L., & Bužarovska, E. (2018). Subject pronoun (non)realization in the English learner language of Macedonian speakers. *Second Language Research*, *34*, 463–485. https://doi.org/10.1177/0267658317747925

National College English Syllabus for Non-English Majors. (1999). Shanghai Foreign Language Education Press.

Nieuwland, M. S., Ditman, T., & Kuperberg, G. R. (2010). On the incrementality of pragmatic processing: An ERP investigation of informativeness and pragmatic abilities. *Journal of Memory and Language*, *63*, 324–346. https://doi.org/10.1016/j.jml.2010.06.005

Noveck, I. A. (2001). When children are more logical than adults: Experimental investigations of scalar implicatures. *Cognition*, *78*, 165–188. https://doi.org/10.1016/S0010-0277(00)00114-1

Noveck, I. A., & Posada, A. (2003). Characterizing the time course of an implicature: An evoked potentials study. *Brain and Language*, *85*(2), 203–210. https://doi.org/10.1016/S0093-934X(03)00053-1

Papafragou, A., & Musolino, J. (2003). Scalar implicatures: Experiments at the semantics–pragmatics interface. *Cognition*, *86*, 253–282. https://doi.org/10.1016/S0010-0277(02)00179-8

Paul, W., Lu, Y., & Lee, T. H. (2019). Existential and locative constructions in mandarin Chinese. *Linguistic Review*, *37*(2), 231–267. https://doi.org/10.1515/tlr-2019-2043

Pechman, T. (1989). Incremental speech production and referential overspecification. *Linguistics*, *27*, 89–110. https://doi.org/10.1515/ling.1989.27.1.89

Quesada, T., & Lozano, C. (2020). Which factors determine the choice of referential expressions in L2 English discourse? New evidence from the COREFL corpus. *Studies in Second Language Acquisition*, *42*(5), 959–986. https://doi.org/10.1017/S0272263120000224

R Core Team. (2018). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*.

Rothman, J. (2009). Pragmatic deficits with syntactic consequences?: L2 pronominal subjects and the syntax-pragmatics interface. *Journal of Pragmatics*, *41*, 951–973. https://doi.org/10.1016/j.pragma.2008.07.007

Ryan, J. (2015). Overexplicit referent tracking in L2 English: Strategy, avoidance, or myth? *Language Learning*, *65*(4), 824–859. https://doi.org/10.1111/lang.12139

Sedivy, J. C. (2007). Implicature during real time conversation: A view from language processing research. *Philosophy Compass*, *2*, 475–496. https://doi.org/10.1111/j.1747-9991.2007.00082.x

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, *70*, 19–47. https://doi.org/10.1016/S0010-0277(99)00025-6

Skordos, D., & Papafragou, A. (2016). Children's derivation of scalar implicatures: Alternatives and relevance. *Cognition*, *153*, 6–18. https://doi.org/10.1016/j.cognition.2016.04.006

Slabakova, R. (2010). Scalar implicatures in second language acquisition. *Lingua*, *120*, 2444–2462. https://doi.org/10.1016/j.lingua.2009.06.005

Snape, N., & Hosoi, H. (2018). Acquisition of scalar implicatures: Evidence from adult Japanese L2 learners of English. *Linguistic Approaches to Bilingualism*, *8*, 163–92. https://doi.org/10.1075/lab.18010.sna

Sonnenschein, S. (1982). The effects of redundant communications on listeners: When more is less. *Child Development*, *53*(3), 717–729. https://doi.org/10.2307/1129385

Sorace, A. (2011). Pinning down the concept of "interface" in bilingualism. *Linguistic Approaches to Bilingualism*, *1*(1), 1–35. https://doi.org/10.1075/lab.1.1.01sor

Sperber, D., & Wilson, D. (1986/1995). *Relevance: Communication and cognition (2nd ed.)* Harvard University Press.

Stiller, A., Goodman, N. D., & Frank, M. C. (2015). Ad-hoc implicature in preschool children. *Language Learning and Development*, *11*, 176–190. https://doi.org/10.1080/15475441.2014.927328

Su, Y. E. (2013). Scalar implicatures and downward entailment in child mandarin. *Journal of East Asian Linguistics*, *22*(2), 167–187. https://doi.org/10.1007/s10831-012-9101-z

Tieu, L., Romoli, J., Zhou, P., & Crain, S. (2016). Children's knowledge of free choice inferences and scalar implicatures. *Journal of Semantics*, *33*(2), 269–298. https://doi.org/10.1093/jos/ffv001

Veenstra, A., Hollebrandse, B., & Katsos, N. (2017). Why some children accept under-informative utterances. Lack of competence or pragmatic tolerance? *Pragmatic & Cognition*, *24*(2), 297–313. https://doi.org/10.1075/pc.00003.vee

Wilson, E., & Katsos, N. (2021). Pragmatic, linguistic and cognitive factors in young children's development of quantity, relevance and word learning inferences. *Journal of Child Language*, 1–28. https://doi.org/10.1017/S0305000921000453

Yang, X., Minai, U., & Fiorentino, R. (2018). Context-sensitivity and individual differences in the derivation of scalar implicature. *Frontiers in Psychology*, *9*, Article 1720. https://doi.org/10.3389/fpsyg.2018.01720

Yoon, E. J., & Frank, M. C. (2019). The role of salience in young children's processing of ad hoc implicatures. *Journal of Experimental Child Psychology*, *186*, 99–116. https://doi.org/10.1016/j.jecp.2019.04.008

Yuan, B. (1995). Acquisition of base-generated topics by English-speaking learners of Chinese. *Language Learning*, *45*(4), 567–603. https://doi.org/10.1111/j.1467-1770.1995.tb00455.x

Appendix

Table 1. L2 speakers' information

No.	Participants	Gender	Age	Years of studying English	Proficiency
1	Participant-01	Female	20	14	CET 6–374
2	Participant-02	Female	23	17	CET 6–400
3	Participant-03	Female	20	15	CET 4–614
4	Participant-04	Female	24	16	CET 6–600
5	Participant-05	Female	20	13	CET 6–449
6	Participant-06	Female	25	13	CET 6–523
7	Participant-07	Female	23	16	CET 6–560
8	Participant-08	Female	24	14	CET 6–427
9	Participant-09	Female	21	12	CET 4–560
10	Participant-10	Male	25	17	CET 6–451
11	Participant-11	Female	22	13	CET 6–436
12	Participant-12	Female	19	10	CET 6–558
13	Participant-13	Female	25	15	CET 6–450
14	Participant-14	Male	20	13	CET 4–440

No.	Participants	Gender	Age	Years of studying English	Proficiency
15	Participant-15	Female	21	9	CET 6–452
16	Participant-16	Female	23	14	CET 6–504
17	Participant-17	Female	19	13	CET 4–498
18	Participant-18	Female	19	14	CET 4–557
19	Participant-19	Male	23	16	CET 6–430
20	Participant-20	Female	20	17	TOEFL-101
21	Participant-21	Female	21	12	CET 4–370
22	Participant-22	Male	26	19	CET 6–450
23	Participant-23	Male	28	18	CET 6–449
24	Participant-24	Female	21	10	CET 4–536
25	Participant-25	Female	20	12	CET 6–460
26	Participant-26	Female	23	12	CET 6–448
27	Participant-27	Female	22	14	CET 6–450
28	Participant-28	Male	20	8	CET 6–525
29	Participant-29	Male	19	15	IELTS-8.0
30	Participant-30	Male	20	17	TOEFL-110
31	Participant-31	Male	20	12	CET 6–610
32	Participant-32	Male	18	10	CET 4–612
33	Participant-33	Female	19	9	CET 4–585
34	Participant-34	Male	19	16	CET 4–600
35	Participant-35	Female	18	12	IELTS-7.5
36	Participant-36	Female	19	12	CET 4–626
37	Participant-37	Female	18	15	CET 4–608
38	Participant-38	Male	22	7	CET 6–515
39	Participant-39	Female	23	10	CET 6–602
40	Participant-40	Female	21	6	CET 6–585
41	Participant-41	Male	21	15	CET 6–590
42	Participant-42	Female	21	9	CET 6–592
43	Participant-43	Female	21	10	CET 6–511
44	Participant-44	Male	21	14	CET 6–459
45	Participant-45	Male	21	14	CET 4–445
46	Participant-46	Male	22	14	CET 4–508
47	Participant-47	Female	22	13	CET 6–580
48	Participant-48	Female	23	16	CET 6–486
49	Participant-49	Female	22	10	CET 6–490

Address for correspondence

Shuo Feng
Yi He Yuan Road #5
Beijing 100871
China

shuo.feng@pku.edu.cn
https://orcid.org/0000-0002-0837-2170