# Processing pragmatic inferences in L2 French speakers

Emilie Destruel

University of Iowa

Abstract

A large amount of literature exists on how native speakers derive and process pragmatic inferences, yet few studies have examined the issue in second language learners, despite a controversial debate of second language (L2) ultimate attainment of phenomena situated at external interfaces. This study contributes to the debate on the integration of external interfaces in highly proficient end-state adult L2 grammars. In an effort to bridge the empirical gap in the past literature, this paper reports on two sentence-picture verification tasks designed to test the processing cost of the exhaustive inference associated with cleft sentences in the L2 French of English learners with different proficiency levels. Truth-value judgments and reaction times were recorded in contexts that violated or supported the exhaustive inference. Overall, results show that L2 learners diverge from the natives in their online processing of the exhaustive inference only and that proficiency plays an important role in predicting their behavior. Nevertheless, what post-hoc observations of the data reveal that L2 length of exposure to native input might be an even better predictor, since only those L2 speakers who have had more exposure to native French input compute the exhaustive inference in clefts as fast as French natives.


Keywords:  French, exhaustivity, processing, inferences, clefts, proficiency

## 1. Introduction

Since Grice (1975), it is largely accepted that the overall meaning of sentences amounts to more than what is lexically encoded by the words that constitute them. Beyond the strict logical, semantic meaning, listeners must (correctly) infer pragmatic information about the speaker's intended meaning, and this to arrive at successful communication. One commonly cited example is the inference conveyed by the expression *some*, for instance in the sentence *Some students attended the conference*. While this sentence is compatible with the logical interpretation that *some and possibly all students attended the conference*, listeners will typically infer that, out of set of students, *some but not all students did*. This type of inference is known as a scalar implicature and is derived because, when a speaker chooses to use a weaker, less informative form like *some* rather than a stronger alternative like *all*, she is taken to implicate that the proposition conveyed by the stronger form does not hold (Grice 1989).

Besides lexical expressions like *some*, certain grammatical structures also convey pragmatic inferences. In this paper, I study one well-known example, i.e., cleft sentences. These sentences express a single proposition via a bi-clausal structure but can be 'unclefted' into a mono-clausal counterpart that is semantically equivalent (see 1c), even though the two sentences may not always necessarily be interchangeable in all discourse contexts (Lambrecht 1994). Pragmatically, clefts convey an exhaustive inference: In addition to communicating that *Camille is drinking a martini*, the English it-cleft in (1a) and the French counterpart in (1b) express that *no one else is drinking a martini* (Horn 1981; Lambrecht 1994; Clech-Darbon et al. 1999; Büring & Kriz 2013).

(1)     a. It's Camille who is drinking a martini.          English *it*-cleft

        b. C'est Camille qui boit un martini.          French *c'est*-cleft

      c. Camille is drinking a martini.                       Canonical sentence

The presence and status of this inference is often diagnosed by a suspension test—a hallmark feature of pragmatic inferences is that they can be suspended, for instance by adding supplemental information. Therefore, while (1a) and (1b) can be felicitously continued by *… and so is Louis*, an exclusive sentence such as *Only Camille is drinking a martini* cannot, because it entails exhaustivity (i.e., exhaustivity is lexically encoded rather than pragmatically inferred).

To date, the literature on pragmatic inferencing primarily stems from psycholinguistics and first language (L1) acquisition. In the former field, one debated issue concerns the time-course of implicatures, that is at what point during online comprehension does an implicature arise compared to semantic information. Contradicting results are found that either support the Default hypothesis (Levinson 2000) under which implicatures are computed immediately and effortlessly but their cancelation is costly, or the Literal-First hypothesis (Huang & Snedeker 2009) that assumes the semantic, logical interpretation of sentences is computed rapidly and that all other inferences require extra time and resources. In the L1 acquisition literature, a number of studies have tested children's ability to understand implicatures, and found that, without considering the context or pragmatic relevance, children are not sensitive to these inferences. In fact, they treat weaker under-informative scalar terms logically (Chierchia et al. 2001; Noveck 2001; Musolino and Lidz 2002).

Comparatively, much less is known on inferencing behaviors in a second language (L2). Issues regarding how L2 learners process implicatures, how their performance differ from that of native speakers, and what factors influence and affect cognitive computation, remain not very well understood. Even though there is a steadily growing number of studies on the topic (Lieberman 2009; Slabakova 2010; Lin 2016; Miller et al. 2016; Antoniou & Katsos 2017; Snape & Hosoi

2018), most have focused on inferences associated with scalar expressions, principally *some*, and very few have examined online processing patterns. Yet, this research avenue is particularly relevant to theories of L2 acquisition. Indeed, there is a consensus that L2 acquisition is not uniform, but that certain domains within the linguistic system seem to pose problems. To that end, the Interface Hypothesis posited by Sorace (2011) argues that L2 learners may have particular difficulties with phenomena that lie at the interface of two linguistic modules, predicting that integrating syntactic knowledge with discourse requirements will give rise to insurmountable issues, even at highest levels of attainment. Therefore, pragmatic inferencing may be problematic compared to mastery of grammar (Bardovi-Harlig & Dörnyei 1998)—the learning task involves going beyond L2 knowledge to decode the L2 input and the literal meaning to infer the intended meaning (Taguchi, Li & Liu 2013).

Given this background, the goal of this paper is to expand on prior L2 literature by addressing the following research questions: (i) can L2 learners be native-like in the domain of exhaustive inference calculation and processing in French?, and (ii) What is the role of proficiency and exposure in their online behavior? The experimental design uses a picture-sentence verification task that compares the truth-value judgments and reaction times for cleft, exclusive and canonical sentences in contexts that either support or violate exhaustivity. I rely on data from two control groups (English and French native speakers) as a baseline. Accordingly, this paper makes an empirical contribution by providing new data, not only in development, but also in the end-state of learners with extremely high proficiency. The paper also makes a theoretical contribution to the discussion on the L2 acquisition of phenomenon that lie at external interfaces.

The next section provides relevant background information, including a summary of the L2 literature on pragmatic inferences and a discussion of the differences between the English and

the French cleft. Then, I present the experimental task, and the results. I conclude the paper by offering a general discussion of the findings and relating them to the current theoretical literature on the topic.

## 2. Literature review

### 2.1 The Interface Hypothesis

Grounded in linguistic theory (Chomsky, 1995), the Interface Hypothesis (IH) developed by Sorace and Filiaci (2006) and Sorace (2011) assumes that language involves modules and interfaces between these modules, and proposes that structures that lie at given interfaces – external interfaces that connect linguistic modules with non-linguistic cognitive components – may give rise to enduring issues in L2 acquisition that remain insurmountable at even the highest levels of attainment. More directly pertinent to the current study, the keystone of the hypothesis is that in (certain types of) bilingualism, "the integration of syntactic and pragmatic conditions remains less than optimally efficient and [this] gives rise to optionality", owing to cognitive resource limitations, resource misallocation or "attentional resources used to inhibit the language not in use" (Sorace, 2011). Moreover, the IH carves out a limited role for the influence of the native (L1) language, proposing instead that difficulties at external interfaces shall remain regardless of the languages involved. In short, the IH predicts that difficulties may persist and residual optionality may linger at the end-state grammar of L2 learners, regardless of L1/L2 pairings, due to processing difficulties in external interfaces.

Within the past decade, contradictory evidence has been found that either supported or opposed the IH. In her study on the interpretation of topicalization and clitic left dislocation, Valenzuela (2005) found evidence that upheld the IH, showing that near-native L2 learners

mastered the formal syntactic properties of these structures and displayed fossilization at the syntax-discourse interface. Similarly, Belletti, Bennati, and Sorace (2007) investigated how NSs of English with near-native L2 Italian mapped the pragmatic notion of focus and the preverbal versus postverbal placement of subject pronouns. These L2ers deviated from native norms in a way that suggested continued reliance on an infelicitious strategy, in this case transferred from their L1 English. Per contra, Slabakova, Kempchinsky, and Rothman (2012), who investigated the contrast between focus fronting and clitic left dislocations in the L2 Spanish of intermediate, advanced, and near-native speakers, found nativelike performance as early as the intermediate level. Further evidence challenging the IH emerges in the work of Donaldson (2011, 2012), who investigated various non-canonical sentences such as clefts and clitic right dislocations in L2 learners of French, and reported nativelike judgments, preferences and use of these structures in authentic spontaneous discourse and experimental tasks. The present study will contribute to this debate by providing empirical data on another phenomenon at the syntax-pragmatics interface; the L2 processing of pragmatic inferences in French clefts.

## 2.2 Prior literature on pragmatic inferences in L2

Until recently, very few studies have examined the L2 acquisition of pragmatic inferences, especially studies that rely on experimental techniques and examine online comprehension. In two of the earlier studies, Bouton (1992, 1994) compared native speakers' and non-native speakers' interpretation of implicatures in an attempt to investigate whether it was necessary to teach implicature to L2 learners. His findings were that L2 learners of English comprehended the implicature at a lower rate than native speakers, even though this rate improved over time.

Using a truth-value judgment task, a more recent study by Dekydtspotter and Hathorn (2005) investigated whether English learners of French could correctly observe the implied interpretation in the case of sentences like (2) that can surface in French either with a continuous (a) or a discontinuous quantifier (2b).

(2)     Something remarkable was observed by each of the researchers.

      a. Quelque chose de remarquable a été observé par chacun des chercheurs.

      b. Quelque chose a été observé de remarquable par chacun des chercheurs.

In (2a), two interpretations are available; either that *all the researchers observed a different remarkable object* or that *all the researchers observed the same object that is remarkable.* In the case of a discontinuous quantifier sentence like (2b), only the latter is possible. The authors found that language learners with high proficiency displayed native-like interpretation while low and intermediate proficiency learners did not show such a tendency.

Slabakova (2010) is the first study to directly test the L2 acquisition of implicatures by investigating how adult L1-Korean L2-English learners process the scalar expression *some* in their L1 and L2. Replicating the design in Noveck's (2001) study, participants had to say whether they agreed or disagreed with the statement made in sentences that included *some* (3) or *all* (4). The critical experimental items were sentences that involved under-informative *some* and that are logically true but pragmatically infelicitous (3b), compared to sentences that are pragmatically felicitous as in (3a).

(3)     a. Some books have colored pictures.

b. Some elephants have trunks.


(4)      All elephants have trunks.


If participants reject (3b), it provides clear evidence that they are able to derive the scalar implicature and compute the pragmatic interpretation of *some* that *some but not all elephants have trunks*. On the other hand, if participants accept sentences like (3b) as agreeable, then it indicates that they have suspended the implicature and generated the logical, semantic interpretation of *some* that *some and possibly all elephants have trunks*. The results showed that, even though Korean learners of English successfully acquired scalar implicatures in their L2, they were less logical than both NSs baseline, i.e., they more often rejected pragmatically infelicitous sentences like (3b). This was the case regardless of their proficiency level. I briefly note that similar results were found in a recent study by Snape & Hosoi (2018) who found that Japanese L2 learners of English were more pragmatic than English NSs regardless of their proficiency level; i.e., they had no problem deriving the scalar implicature and avoided further processing cost by not considering alternative contexts for the logical answers. In the same vein, Miller et al. (2016) who tested the interpretation of scalar implicatures in the L2 Spanish of native English advanced learners, found that both groups performed on par in an acceptability judgment task, and this despite the complexity of the distribution of the two indefinites determiners in Spanish.

One explanation Slabakova (2010) proposes for her findings is that that L2 learners may not have the cognitive resources to conjure up alternative contexts that would make under-informative sentences plausible. For instance, if one can think of a situation where some elephants' trunks got cut due to poaching, then the sentence in (3b) becomes felicitous. But calculating this

alternative context would require further processing effort from the L2 learners, which Slabakova argues they lack. Another possible explanation is that L2 learners may lack the cognitive resources to suspend pragmatic inferences that are derived automatically. Indeed, since implicature suspension arguably requires more cognitive efforts, it might be more difficult to do in an L2 under the assumption that less cognitive resources are available for L2 processing than L1 (Green 1986). Slabakova argues that her findings support the Default hypothesis under which the derivation of implicatures is effortless and automatic, whereas their cancelation comes with a processing cost.

The study by Lin (2016) provides further evidence for this explanation. In a series of real-time psycholinguistic tasks, the author investigated the processing mechanisms in the acquisition of *some* for Chinese learners of English. In the first experiment, participants were asked to judge whether follow-up target sentences, of the form in (5b), were true or false given a context, as in (5a). In the second experiment, Lin gave participants less time to respond, therefore testing whether inferential patterns would be modulated by a cognitively more demanding task (i.e., time pressure).

(5)     a. Mary had many sisters. Some of the sisters were married.

        b. Some and possibly all of the sisters were married.

Results showed that, when given unlimited time to respond (experiment 1), L2 learners were fast at computing the implicature *some but not all* but slower to reject it (i.e., to compute the logical interpretation). When pressed for time, they were more likely to reject the suspension of the implicature, further revealing that canceling an implicature requires more cognitive capacity—in line with Slabakova's idea. When L2 learners' processing capacity was experimentally constrained, they preferred the less demanding reading by computing the implicature.

In sum, this overview of the existing literature on L2 acquisition of the implicatures associated with the scalar term *some* in various languages suggests that generating the implicature does not pose difficulties to L2 learners, but what seems challenging is suspending it.

Finally, Destruel and Donaldson (2017) examined the exhaustive inference in declarative *c'est*-clefts in L2 French, using an offline forced-choice task in which participants had to select a continuation for three sentence types (canonicals, clefts or exclusive sentences) which they deemed appropriate. They found uncovering entirely nativelike performance among the most advanced learners, with cleft sentences being continued by and large with a moderate continuation, such as "yes, and…" or "yes, but…," suggesting that exhaustivity in clefts was interpreted as less categorical or less robust than with exclusives.

The experiment presented in section 3 provides further evidence supporting this tendency. It expands on prior work by testing the exhaustive inference associated with French cleft sentences; an inference that has been rarely studied to date in the L2 literature. In addition, it addresses the need for further research in L2 acquisition on the computation of pragmatic inferences by examining online data.

## 2.3 French clefts: function, interpretation, processing and L2 acquisition

Both in English and French, it is widely assumed that clefts signal focus, meaning that the clefted element evokes alternatives that are relevant for the interpretation of the sentence (Rooth 1992; Krifka 2008). Thus, in (1b) *C'est Camille qui boit un martini*, the focus element is *Camille* and the congruent (explicit) question is of the form *Who is drinking a martini?* Several authors note, however, that French relies more heavily on clefting due to prosodic constraints on the sentence initial position (Hamlaoui 2007; Carter-Thomas 2009; Bourns 2014), making clefts the

default strategy to answer subject focus. Pragmatically, clefts are also known to associate with a specific inference, exhaustivity, whereby the clefted element is interpreted as if under the scope of an exclusive expression like 'only' (Horn 1981; Clech-Darbon et al. 1999; deCat 2007). Yet, recent studies provide data that suggest that there are differences in the strength and systematicity of the exhaustive inference in the two languages. In their study, Destruel and DeVeaugh-Geiss (2018) used a sentence-picture verification task in which English and French native speakers (NSs) saw a series of pictures that either violated or supported an exhaustive interpretation. Pictures violated exhaustivity if, out of four characters, there was more than one depicted as performing the action encoded by the predicate, and supported it if exactly one was doing so. After each picture, participants heard a sentence in one of three forms; cleft, exclusive or canonical. They were asked to judge whether the sentence they heard appropriately described the picture (i.e., 'true') or not ('false'). Their reaction times were recorded. Results show that French clefts are less strongly exhaustive than their English counterparts and are processed faster in contexts that do not support the inference. Specifically, in contexts that violated exhaustivity, even though English NSs commonly accepted clefts as appropriately describing the picture (59%), they do so to a significantly lesser extent than French NSs (76%). But accepting clefts comes at a cost in English: English NSs were found to be significantly slower to say that clefts appropriately described the picture when it violated exhaustivity compared to rejecting it. The authors argue this finding is in line with a Default view of implicatures under which these inferences are taken to arise automatically: In English, the exhaustive inference in clefts is quickly present after the trigger appears—accepting clefts is costly because of the need to go through (a potential second processing step of) the suspension of the unsupported inference. In contrast, processing times for

French NSs showed no significant difference, that is, clefts were accepted quickly in non-exhaustive contexts.

Finally, several studies have examined the L2 acquisition of French clefts, and more specifically their emergence in focus contexts. Yet, these have produced differing results. On the one hand, earlier work such as Trévise (1986), Barning (1997), Watorek (2004) and Bartning & Hammarberg (2007) all find that clefts are acquired quite late and that learners with as many as six years of classroom instruction still exhibit infelicity in the production of these forms. On the other hand, Donaldson (2012) and Reichle and Birdsong (2014) report no significant differences between L2 learners and NSs—the learners appropriately produced and understood clefts in contexts that licensed narrow focus. To date, Destruel and Donaldson (2017) is the only study that investigated the L2 acquisition of the pragmatic properties of French *c'est*-clefts. The authors found that the high and (to a lesser degree) the intermediate-proficiency groups exhibited interpretation patterns comparable to French NSs. This was taken as evidence against the Interface Hypothesis (Sorace 2011), suggesting that L2 learners can indeed achieve native-like proficiency for a phenomenon that lies at the syntax-pragmatics interface.

In sum, from an L2 perspective, this short comparison of English and French clefts indicates that there may be limitations for learners since a 1:1 mapping between form and interpretation is not quite straightforward—the input might be misleading.

### 3. The present study

The present experiment involves a single sentence-picture verification task for which two dependent variables were recorded, truth-value judgments (TVJ) and response times (RTs). It tested three types of sentences, *c'est*-clefts, sentences that lexically encode the exhaustivity (i.e.,

exclusive sentences) and those that do not (i.e., canonical SVO sentences). The design is directly taken from Destruel & DeVeaugh-Geiss (2018) – although it does not differ from this previous study methodologically, the present study expands on it by testing L2 learners of French. The specific research questions are as follow:

(6)     Can L2 learners be native-like in the domain of clefts' exhaustive inference calculation and processing in French?

(7)     What is the role of proficiency in L2 offline and online behavior, especially for *c'est*-clefts?

As discussed in section 2, despite both conveying exhaustivity, English and French clefts do so to a different degree. Native speakers differ in their underlying processing mechanisms; only English NSs show delays when having to suspend the exhaustive inference. Thus, the learning task for L2 learners in the present experiment involves transferring the supposedly universal inference computation mechanism from the L1. If Slabakova (2010) is correct and L2 learners do lack the cognitive resource to suspend pragmatic inferences, L2 learners should be slower at accepting clefts in non-exhaustive contexts. Moreover, if proficiency and exposure also play a role on inferential behavior, learners who are more proficient and have had more exposure to native French input should be more likely to pattern like French NSs, i.e., to be faster at suspending exhaustivity in clefts.

## 3.1 Participants

The two control groups include the data from English and French native speakers taken from Destruel and DeVeaugh-Geiss (2018). The English native group, ENSs for short, (n = 64; 40

females, age range 18-21 years old) consisted of undergraduate students enrolled in a first semester language class at a Midwestern university. The French native group, FNSs for short, (n = 64; 35 females, age range 18-45 years old) included 89% of undergraduate students and 9% of graduate students in a university in Southern France, as well as 2% of young professionals working as staff at the university. All native speakers were monolingual speakers, had little knowledge of the other language tested.

The L2 speaker group tested in this study includes a total of 67 L2 learners of French (35 females). The participants received either monetary compensation or extra credit towards their class grade. Participants were recruited based on the age of onset: All were L1 English native speakers who started learning French as an L2, after adolescence, either in high school or at the university level (mean age of acquisition = 16.3 years old, range = 15-20 years old). Furthermore, in order to tap into different proficiency levels, L2 learners were recruited from different sources. First, I recruited undergraduate college students, all majors or minors in French at a Midwestern university (n = 30, 19 females; age range 20-24, mean = 22.2 years). All had completed their entire education in the U.S and were enrolled in upper-division French courses at the time of testing. Only 8 had studied abroad, 5 of them having taken part in a summer program (2 months) and 3 having spent a full academic year in France (9 months). Second, I recruited graduate students in French literature MA and Ph.D. programs at a Midwestern university (n = 12, 8 females; age range 25-29, mean = 26 years). Of these students, 6 reported having spent time in France (average = 1 year). All had studied French for 6 to 10 years (average = 7). Finally, I recruited professionals residing in the United States who reported using French regularly in their profession (n = 10, 7 females; age range 42-57, mean = 52.4 years)—of which 9 had spent time in France (average = 3

years)—and expatriates who resided in France at the time of testing (n = 15, 5 females; age range 39-55, mean = 43.7 years; average time spent in France = 12 years).

## 3.2     Pretests

*3.2.1. Proficiency test*

To avoid conflating experience or institutional level with L2 proficiency (see Tracy-Ventura, McManus, Norris, & Ortega, 2014)*,* L2 speakers completed an independent proficiency test; the cloze-test found in Tremblay's (2011). This test consisted of a non-technical French newspaper article in which 45 out of 314 words were blanked. Thus, the maximum score possible was 45/45. The scoring protocol was borrowed from Tremblay (2011) and the cutoff points used were taken from Destruel & Donaldson (2017) since uniform cutoff points for the test have not yet been established and agreed upon in the literature (Tremblay, personal communication, 2015). Moreover, the L2 speakers tested in this study differed from those in Tremblay in at least two respects. First, the L2 speakers I tested had overall a higher proficiency level. For instance, the highest score in Tremblay's lowest group was 13/45 whereas the lowest score in this participants' pool was 16/45. Second, I was interested in a more fine-grained distinction of the speakers with high level of attainment—distinguishing between advanced and highly advanced (or near-native) proficiencies—given that several studies suggest that there are subtle linguistic properties that seem to be acquired by L2 learners only beyond advanced proficiency levels (Hopp 2009, Slabakova et al. 2012).

Given the scoring scheme adopted, three proficiency groups emerged: (i) a beginner group (n = 20) that included the L2 speakers with scores of 26 or lower; (ii) an intermediate group (n =

25) that included the L2 speakers with scores between 27 and 36; (iii) an advanced group (n = 22), with L2 speakers who scored 37 and above.

We shall note that the French native speakers did not take this proficiency test and were instead given a default 45/45 score. An anonymous reviewer correctly remarks that this is a potential shortcoming, which I will discuss in section 5.

### 3.2.2. Biographical questionnaire

Prior to the main task described hereafter, all L2 speakers completed the Bilingual Language Profile (BLP) questionnaire (Gertken, Amengual & Birdsong 2014), provided to participants in their L1 (i.e., English). This questionnaire is a biographical information survey that provides a general linguistic profile through self-reports taking into account multiple dimensions: Age of acquisition of the L1 and L2, frequency and contexts of use, competence in different skills, and attitudes toward each language. Table 1 reports the results of this information crossed per proficiency group, as defined by Tremblay's cloze test.

<INSERT TABLE 1 ABOUT HERE>

## 3.3. Main task

### 3.3.1 Materials

This sentence-picture verification task (replicated from Destruel & DeVeaugh-Geiss, 2018), included two within-subject factors. Participants saw a picture in one of the following four conditions (Figure 1):

- 'Wrong' (Figure 1a): none of the four characters depicted has the property described by the predicate.

- '–Violation' (Figure 1b): only one character has the property asserted in the sentence, supporting the exhaustivity inference that 'no one other than x has property Z'.

- '+Violation' (Figure 1c/1d): at least one alternative character is also performing the described action. This is the condition of interest since it violates exhaustivity.

All pictures included four characters of roughly the same size, color, and shape (unless otherwise required by the descriptive adjective in the sentence). In order to avoid recognition effects, the location of the target character was counterbalanced across the four positions in the picture.


<INSERT FIGURE 1 ABOUT HERE>


The second factor was the form of the sentence participants heard, illustrated in (8): A cleft, exclusive or SVO sentence with prosodic focus (see Appendix A for all test items).


(8)  a. It's a [BLOND]$_F$ baby who is shaking a rattle.          Cleft condition

     *C'est un bébé [blond]$_F$ qui secoue un hochet.*

    b. Only a [BLOND]F baby is shaking a rattle. [SEP]          Exclusive condition

     *Seul un bébé [blond]$_F$ secoue un hochet.* [SEP]

    c. A [BLOND]$_F$ baby is shaking a rattle.          SVO condition

     *Un bébé [blond]$_F$ secoue un hochet.* [SEP]

All the sentences were designed including the same basic constituents, i.e., the indefinite article *un(e)*/'a', a descriptive adjective, a [+human] subject noun, a transitive verb and a [–animate] object. The clefted element was always the grammatical subject, and therefore the focus element was always a descriptive adjective that qualified this grammatical subject. The desired informational structure and intonational pattern were achieved by recording all items, by a French adult native speaker, prior to the experiment in a sound-proof booth. The experimenter (also a native French speaker) interacted with the native speaker by asking the congruent wh-question of the form 'Which x did Z?' that triggered each test item (e.g., 'Which baby is shaking a rattle?' for example 8 above).

We note a potential issue with the test items which relates to the form of the indefinite article. Indeed, in French, the indefinite article *un/une* has the same lexical form as the numeral 'one'. If results show that participants interpret the stimuli as strictly exhaustive, it will be difficult to decide whether they did so because of the meaning they attribute to the sentence structure itself, or because of the fact that they interpreted the sentence as meaning 'exactly one x' due to the indefinite. To alleviate this potential issue, the experimenter debriefed all L2 participants following the main task, asking them whether they had interpreted the indefinite as the numeral 'one'. No participants had.

For each of the three sentence forms, I created 40 lexicalizations, therefore 40 different scenes for the pictures (e.g., 'blond baby shaking rattle' was one). In addition to the 40 experimental items, I created 40 fillers (consisting of other non-canonical structures such as definite descriptions and passives, all with focus on the adjective), and randomized the total into eight experimental lists (due to the two types of +violation pictures) via a 2 x 3 Latin square design.

*3.3.2 Procedure*

All L2 participants completed the main task in a quiet room using the stimulus presentation software SuperLab Pro 4.5 (Cedrus Corporation, USA). Before the experiment began, they saw a training session that consisted of two practice trials. On every trial, participants looked at a single centrally-located picture displayed on the computer screen for 2000ms, after which they heard the stimuli in a set of headphones. The picture remained on the screen as the sentence finished playing. Participants were asked to judge as fast and accurately as possible whether the sentence they heard appropriately described the picture seen by pressing a 'true' or 'false' button (counterbalanced) on a USB Response Pad (RB-530). Truth-value judgments (TVJ) and response times (RTs) were recorded. Between each trial, a white screen appeared for 2500ms. Figure 2 illustrates how the procedure unfolded and what the RTs analyzed correspond to. Overall, the entire experiment lasted between 40 and 55 minutes.

<INSERT FIGURE 2 ABOUT HERE>

*3.3.3 Data preparation and Analysis*

RTs were analyzed from the offset of the audio stimulus (sentence) since sentence duration varied across sentence conditions, thus allowing to directly compare RTs. Furthermore, because participants were permitted to execute their response at any point after the sentence started playing, they could have in principle responded immediately after the onset of the adjective (e.g., 'blond') without considering the rest of the sentence; analyzing RTs from the offset of the sentence also allowed to exclude results from participants who may have been guessing or anticipating

heuristically what the end of the sentence might be. Any response made while the sentence was still playing was removed from the final analysis, which represented than 0.4% of the data analyzed. All RT data were log-transformed in order to improve normality and reduce the impact of skew due to, e.g., fatigue or sequential effects (Whelan, 2008). Although response times were measured for all picture conditions, I only present results for the '+violation' picture (the condition of interest) for two reasons: First, there are no particular predictions for the RTs of the other picture types, and second, it is difficult to make a meaningful comparison for the TVJ across picture types due to complete separation (0% acceptance rate) in the 'wrong' condition for clefts.

Given the recommendation in Bates et al. (2015), I report on results from generalized linear mixed-effects models used to analyze the binary TVJ data (estimates, standard errors, $z$-values, and $p$-values), which was coded as ('true':1, 'false': 0). Linear mixed-effects models were run for the continuous RT data (estimates, standard errors, and $t$-values, with any $t$-value above |1.96| considered statistically significant at $p <.05$).  Data was first analyzed by language group, based on whether participants were native speakers or L2 learners, and then I refined the analysis by looking at proficiency (treated as a continuous variable based on the results from the cloze test, with French native speakers assigned a max value of 45). Sum-coding was used for the two fixed-effects sentence form (clefts being coded as baseline level, -1) and language group (FNSs being coded as the baseline level, -1). All analyses were implemented using the *lme4* library in the R environment (GPL-2| GPL-3, v.3.3.3) (R Core 2017). In all cases, a maximal model was initially constructed, that included random intercepts and by-participant and by-item random slopes (the two random-effect factors), as well as all the relevant fixed effects and their interactions. If this model failed to converge, the random effect structure gradually simplified by first removing

interaction terms and the main effects. The results reported correspond to the model with the maximal RES that converged.

**3.4 Results**

3.4.1 Truth-value judgments data

Table 2 reports on the percentage of 'true' responses for each picture condition and each sentence form, for both native groups and L2 speakers (collapsed for proficiency). The low rates of acceptance for all sentences in the 'wrong' condition, along with the high rates of acceptance in the '–violation' condition indicate that all participants, regardless of group or proficiency, were engaged in the task.

<INSERT TABLE 2 ABOUT HERE>

Turning to the condition of interest ('+violation'), all groups display a similar trend: clefts are accepted as correctly describing non-exhaustive pictures, and this to a higher degree than exclusives but to a weaker degree than SVO sentences. In L1 French though, the descriptively higher number of 'true' TVJs (76%) in addition to the lack of statistical significance found when comparing clefts to SVO is compatible with *c'est*-clefts being merely weakly exhaustive, and having weaker exhaustive effects compared to their English counterparts (59%). L2 learners, despite remaining closer to their L1 peers than to the French norm (63% vs. 59% and 76% respectively of 'true' responses given for clefts in '+violation' pictures), are overall able to cancel the exhaustive inference when the context does not support it – they can provide logical answers, i.e., judging clefts to be compatible with the meaning that 'x and possibly someone else Zed'.

Crucially, this also shows that L2 learners do not take clefts to semantically encode exhaustivity, yet they associate a stronger exhaustive inference with clefts than with plain, canonical structures. This is a welcome result as it aligns with the findings in Destruel & Donaldson (2017), further showing that L2 learners recognize clefts as being no more than pragmatically exhaustive.

Statistically, Table 3 contains model coefficients for the generalized linear mixed-effect model predicting the binary TVJ data from the two fixed-effects language group and sentence form, and their interaction (Mod1 = TVJ ~ SentenceForm*LangGroup). Results show no effect of sentence form when comparing clefts to SVO but a main effect when comparing clefts to exclusives. There was a main effect of language group both when comparing FNSs to L2 learners and FNSs to ENSs. Finally, the two significant interactions found suggest a reliable difference between L1 English and L1 French clefts. When looking at follow-up analyses for between-groups comparisons on specific sentence types, results further reveal no significant effect of language group between ENSs and FNSs, and between FNSs and L2ers, in the exclusive condition ($\beta = 1.22$, SE = 0.11, $z = 0.46$, $p < .5$ and $\beta = 1.38$, SE = 0.15, $z = 0.32$, $p < .5$, respectively), but significant differences in the cleft condition ($\beta = 1.29$, SE = 0.109, $z = 3.26$, $p < .001$ for ENSs vs. FNSS, and $\beta = 1.03$, SE = 0.14, $z = 4.03$, $p < .001$ for FNSs vs. L2ers).

<INSERT TABLE 3 ABOUT HERE>

Refining the investigation by proficiency, a descriptive analysis of the results in Table 4 points to a split between the beginner group on the one hand, and the intermediate and advanced group on the other hand, with the beginner group seemingly driving the difference in rate of cleft's acceptability in the '+ violation' condition. Indeed, the beginner L2 learners perform close to

chance when judging clefts in that pictorial context. In contrast, the other two groups align with the trend observed in French NSs – they take clefts to be acceptable at a higher rate (67% and 74% respectively) than the English NSs do (59%), i.e., they derive exhaustivity less often.


<INSERT TABLE 4 ABOUT HERE>


Statistically, I ran a generalized linear mixed-effect model only on the data for L2 learners and FNSs. The model tested the effect of sentence form (sum coded for the three levels) by proficiency (analyzed as a continuous variable) on the TVJ given. Results revealed a main effect of sentence form when comparing clefts to exclusives ($\beta = 1.74$, SE $= 0.17$, $z = 4.16$, $p < .001$) but not when comparing clefts to SVO ($\beta = 1.65$, SE $= 0.12$, $z = 1.53$, $p < .21$). There was also a main effect of proficiency ($\beta = 0.82$, SE $= 0.14$, $z = 3.91$, $p < .001$) and an interaction between the two predictors ($\beta = 0.51$, SE $= 0.11$, $z = 3.67$, $p < .001$) – proficiency was positively correlated with acceptance of clefts in non-exhaustive contexts.


*3.4.2 Response Times*

Recall that RTs were analyzed from the offset of the target sentence until the time when participants pressed the T/F button to indicate their judgment. Figure 3 illustrates the results aggregating true and false responses in the '+violation' condition for all language groups. When exhaustivity did not hold, ENSs and L2 learners showed a delay when making judgments for clefts in comparison to the other two sentence forms – this is not the case with FNSs.


<INSERT FIGURE 3 ABOUT HERE>

These observations are statistically corroborated: Table 5 reports on the results from the linear mixed-model regression that predicted log RTs from sentence form by language group. Clefts hinder processing only in ENSs and L2 speakers, suggesting a significantly different processing behavior from the French native norm in these two groups. There were two significant interactions between sentence form and language group when comparing clefts to exclusives but not to SVO. This is further suggested by between-group comparisons. Indeed, when running follow-up analyses on the data for clefts, there is a significant effect of language group between FNSs and ENSs ($\beta = 2.74$, SE $= 0.21$, $t = 5.27$, $p < .001$), and between FNSs and L2ers ($\beta = 2.17$, SE $= 0.15$, $t = 3.38$, $p < .001$), which is not the case in the exclusive condition ($\beta = 1.37$, SE $= 0.04$, $t = 0.47$, $p < .5$, and $\beta = 1.04$, SE $= 0.11$, $t = 0.73$, $p < .5$).

<INSERT TABLE 5 ABOUT HERE>

Does proficiency play a role, and does processing exhaustivity in clefts remain subject to vulnerability in L2 speakers, especially those at the highest level of attainment? Visual inspection of Figure 4 reveals two important things. First, exclusive sentences are processed similarly across the three L2 proficiency groups, suggesting that even at the lowest level, semantic exhaustivity does not pose a problem. Second, for clefts, RTs continuously decrease with proficiency, with the beginner group displaying a much longer delay compared to the other L2 groups, and even more so compared to the two native groups. In the highest group, RTs for clefts remain greater than those in French NSs, but faster than the English NSs. These observations unveil a developmental

sequence whereby L2 learners are gradually becoming faster at accepting clefts as appropriate in contexts that go against the exhaustive inference they convey.

<INSERT FIGURE 4 ABOUT HERE>

Statistical analysis of the data upholds these descriptive findings: A linear mixed-effect regression model predicting RTs from proficiency and sentence form on the data for L2 learners and FNSs in the '+violation' condition yielded a main effect of proficiency ($\beta = 0.08$, SE $= 0.05$, $t = -3.72$, $p < .001$), suggesting an overall decrease in delay positively associated with higher proficiency; a main effect of sentence form when comparing clefts to exclusives ($\beta = 0.13$, SE $= 0.17$, $t = 5.76$, $p < .001$) but not to SVO sentences ($\beta = 0.09$, SE $= 0.11$, $t = 1.61$, $p < .16$); and only one significant interaction ('SentFormExc:Proficiency', $\beta = 0.12$, SE $= 0.14$, $t = 2.47$, $p < .001$).

To gain a better understanding of the source of the delay associated with clefts when exhaustivity is violated, I refined the analysis by looking at whether RTs were affected by the truth-value judgment made. Figure 5 reveals that, unlike for FNSs, ENSs and L2 learners show a difference in true vs. false judgments; these participants took much longer to provide a 'true' judgment, suggesting that rejecting clefts is quick but what seems costly is the suspension of the pragmatic inference in contexts that do not support it. I entered TVJ and language group as fixed-effect predictors in a mixed-effects logistic regression model predicting RTs for clefts in '+violation' pictures. Results revealed a main effect of TVJ ($\beta = 0.05$, SE $= 0.14$, $t = 2.76$, $p < .001$)., and a main effect of language group when comparing ENSs ($\beta = 0.28$, SE $= 0.26$, $t = 4.32$, $p < .001$) and L2 learners ($\beta = 0.22$, SE $= 0.15$, $t = 3.83$, $p < .001$) to FNSs. Finally, the two

interactions between the predictors were significant (TVJ:LangGroupENS: $\beta = 0.23$, SE $= 0.12$, $t$ $= 5.11$, $p < .001$, and TVJ:LangGroupL2: $\beta = 0.38$, SE $= 0.21$, $t = 4.86$, $p < .001$).


<INSERT FIGURE 5 ABOUT HERE>


Here as well, proficiency seems to have an effect: Figure 6 indicates that TVJ is affected by proficiency, with most of the RT delay in L2 learners being attributable to lower proficiency learners. Nevertheless, even at the most advanced level, L2 learners still display a delay in accepting clefts (selecting 'true' responses) when the pictorial context violates exhaustivity.


<INSERT FIGURE 6 ABOUT HERE>


Statistically, I ran a mixed-effects logistic regression model predicting RTs for clefts in '+violation' pictures from the two fixed-effect predictors proficiency and TVJ and their interaction, and this on the L2 and FNSs data. Results revealed a main effect of TVJ ($\beta = 0.13$, SE $= 0.23$, $t = 3.28$, $p < .001$)., a main effect proficiency ($\beta = 0.11$, SE $= 0.22$, $t = 3.47$, $p < .001$) and a significant interaction between the two predictors ($\beta = 0.32$, SE $= 0.25$, $t = 2.83$, $p < .001$). This suggests that, as they get more proficient, L2 learners processing pattern for canceling exhaustivity (or put differently, accepting clefts in violation contexts) becomes more akin to that of FNSs, yet without completely matching the native norm.


## 5. Discussion

### 5.1 Summary of findings

Before summarizing the results, I ought to briefly discuss a potential limitation that may limit the strength of the inferential statistics and the conclusions drawn from them. As an anonymous reviewer notes, the scoring system for L1 proficiency (i.e., awarding them 45/45 points on the proficiency test as a default) is not quite appropriate – L1ers' scores could actually vary. Moreover, from a theoretical perspective, research in SLA is interested in how L2 proficiency influences nativelike performance. Thus, the inferential statistics on proficiency could have been impacted from this "default" score for L1ers, as roughly half the data (L1 data) are not 'real' data as such, and it is not certain that the significant effects and interactions in proficiency are truly L2 effects. A way to circumvent this worry would be to limit the proficiency analyses to the L2 data.

The results that emerge from the study on native English learners of French are the following. First, considering offline TVJ judgments, results taken overall suggest that, even though L2 learners judge clefts as being acceptable in contexts where exhaustivity is violated, they do so to a lower level than the French natives. But the study further reveals an important role played by proficiency in the ability to understand and derive implicatures, such that low proficiency learners seem mainly responsible for the discrepancy. Indeed, although this group of learners have acquired that clefts are not semantically exhaustive, and differ from exclusives in that respect, they still perform close to chance in contexts that violate exhaustivity. Results from the control condition further show that these 'beginners' have indeed acquired the words, syntactic computation and semantic meaning of the French *c'est*-cleft: they correctly reject clefts in 'wrong' contexts and accept them in contexts that support exhaustivity. Therefore, we can rule out the hypothesis that their poorer performance in '+violation' contexts comes from issues with understanding the cleft structure itself. Or put slightly differently, their lower acceptance rate is not due to a deficient understanding of the semantics of the target sentence due to a gap in their language knowledge.

Rather, I argue that they have no lexical or grammatical issues with the French *c'est*-cleft, but that the task remains to learn what pragmatic inference they can justifiably draw from this specific word order when it occurs in discourse. This finding parallels findings in past L2 studies, which have shown that L2 implicature interpretation skills are affected especially at early stages of language learning. In a pioneering series of studies, Bouton (1988, 1994) found that there was a significantly lower performance in implicature derivation for different types of implicatures in the non-native group. Similarly, Destruel & Donaldson (2017), who also examine the interpretation of the exhaustive inference in French clefts, find that low-proficiency learners differ from higher groups and the French natives. Interestingly though, the low-proficiency learners in their study attribute a weaker (or even absent) level of exhaustivity to clefts compared to the other groups – contrary to what is found in this study. One potential reason for this divergence might be the task used to test the inference. In their study, Destruel & Donaldson did not include contexts that violated exhaustivity but rather presented participants with three types of continuations to sentences, including clefts, one of which included conflicting information. In the present study, the condition of interest more directly violated exhaustivity, or at least right of the bat, when participants scrutinized the picture before hearing the target sentence. Thus, if exhaustivity arise as a default inference in English, accepting the cleft in such contexts meant that the inference had to be canceled to accommodate for the logical interpretation, which could demand more (processing) resources in the L2.

In the present study, the discrepancy in acceptance rate of clefts improved at the intermediate level and was completely remedied in the highest group – these learners converged with nativelike interpretation. Like many syntax–discourse phenomena (Bohnacker 2010; Donaldson 2011a, b; Sorace & Filiaci, 2006), a nativelike interpretation of exhaustivity involves

preferences rather than absolute constraints, subtleties that are unlikely to be formally taught in a classroom setting but that must be noticed and acquired uniquely from exposure to authentic language. In such input, the L2 learners' target for the *c'est*-cleft is a set of natives' preferences that shows roughly a 76% preference for a non-exhaustive interpretation. The fact that the highest proficiency learners acquired this (far from categorical) preference indicates not only that they are sensitive to natives' interpretive patterns but also that they are able to integrate discourse-pragmatic and syntactic information in a nativelike manner. The results from the two higher group are in line with findings in past literature as well. For instance, Taguchi (2008) reports that intermediate-level learners of Japanese have a better ability than those at the elementary level to understand implicatures. Taguchi (2011) also reported that more-proficient learners of English understand implicatures better and faster than lower-level learners. A similar effect was found by Roever (2005) for learners of English for both conventionalized and idiosyncractic types of implicatures. Thus, it seems clear that better language comprehension is positively correlated with the ability to modulate implicature comprehension.

The most novel results in the present study pertain to online computation of the exhaustive inference. Here, results overall show a significant delay for clefts in L2 learners when the pragmatic inference is not supported by the pictorial context, suggesting that clefts do hinder processing. But here as well, proficiency modulates these patterns, with beginner L2 learners being much slower at providing a (true) judgment, while we observe a speed up in the intermediate and the advanced groups. I contend that these results reflect optionality in the L2 grammar: Low-proficiency learners interpret the *c'est*-cleft without reference to discourse pragmatics, most likely because they are unable to cope with the cognitive demands of integrating syntactic and discourse-pragmatic information, as posited by the IH (Sorace, 2011, p. 14). Nevertheless, even at high level

of attainment, L2 learners still display a different processing pattern than FNSs, despite a positive learning curve. In the past literature, some studies have also examined the role of the learning context (institutional domestic setting vs. stay-abroad) on implicature comprehension. While in the latter case, learners have more opportunities to practice the language and get acquainted with the culture, the difference has not proved to give a definite edge in the ability to understand implicatures to learners who had studied abroad. Yet, a relevant factor is whether the length of residence was important. Taguchi (2008) did report a low to moderate effect of length of residence, and Taguchi (2011) specifically found a beneficial effect of this factor on the comprehension of idiosyncratic implicatures. In the present study, the length of residence of the high-proficiency learners was significantly longer than those in the two studies by Taguchi, and accordingly, an interesting post-hoc observation reveals that, when focusing on the data for the highly proficient learners, the length of exposure (measured by length of residence) to native French language seems to make a difference in processing patterns. Indeed, the learners who have resided in France for several years – being the most immersed in native French input – processed the exhaustive inference as fast as FNS (no apparent delay in RT), unlike those whose length of residence was shorter (and resided in the US at the time of testing). In short, the learners with the longest time of residence did prove to converge towards the native norm during online processing of clefts. So, exposure to native input might be the crux of differences in highly proficient learners, and a better predictor of successful acquisition than proficiency, supporting the gradual shift towards a better operationalization of exposure over measures of proficiency in the literature. Further work will need to be done to test for the exact effect of exposure by controlling for this factor more precisely.

**5.2 Implications for theories of L2 acquisition**

Motivated by the fact that very few studies have examined the processing patterns of pragmatic inferences in L2 acquisition, and that the IH argues there should be permanent residual difficulties at the interface between syntax and discourse, I set out to provide new empirical data on the learnability and processing patterns of an inference that is less commonly examined than scalar implicatures in the past literature, i.e., the exhaustive inference of clefts. I now consider the implications of my findings for L2 theory.

Taken together, the findings in this experiment present a challenge for the IH: There is no evidence of optionality and indeterminacy at the most advanced levels of proficiency in offline judgments, and processing patterns approach the native norm as proficiency increases, and even more so when considering the amount of native input learners have been subject to. Issues only arise at the beginner and intermediate groups, with a clear developmental curve suggesting that these difficulties can be overcome at more advanced proficiency levels, especially when longer exposure to the L2 language occurs. Importantly, L2 learners in this study did not have any difficulty with the lexical items and the grammatical construction in test sentences, as indicated by their performance (speed and accuracy) on control items. Nevertheless, it is not clear from current research how exactly proficiency (broadly understood as proficiency in language knowledge and processing) affects implicature comprehension, but several authors posit that it is due to a difference in availability of cognitive resources; language processing in the L2 is costly and effortful, and thus, learners' resources are consumed by semantic processing, leaving no resources for further pragmatic processing (Cook & Liddicoat 2002; Slabakova 2010).

Another aspect of the IH is that difficulties at external interfaces obtain regardless of the languages involved because the variable application of interface rules constitutes a more general byproduct of bilingualism rather than a result of transfer of pragmatic features. Thus, even when

the L1 and L2 share strategies whereby positive transfer might be expected, the IH predicts that difficulties will occur whenever external modules are involved. In the case of implicatures, Slabakova (2010) argues that the calculation mechanism involved is purportedly universal, such that there should be no adverse L1 influence. So, in theory, from a transfer perspective, inference calculation should be 'easy'. Yet, this might not be exactly the case for the particular inference studied here. Indeed, as discussed in section 2.1, although English and French clefts are similar in that they convey exhaustivity, they vary in the inference's strength and systematicity. Destruel & Deveaugh-Geiss (2018) show that the inference is weaker and clefts are processed faster in non-exhaustive contexts in French. They argue that exhaustivity appears as the default interpretation in English, when no further context is provided, and that the processing cost in this language comes from the annulment of the unsupported inference, which is not the case in French. The results in this study seem consistent with a view that the L2 group is transferring some feature of their English grammar: the slower RTs and lower acceptance rate do point to L1 transfer whereby beginner and intermediate learners analyze clefts as exhaustive by default and cancel it in a subsequent step when unsupported. The analysis of the effect of TVJ on RTs in this study support the idea that the source of the processing limitation at these proficiency levels is arguably the inhibition of a pragmatic response in favor of a logical one, as proposed by Slabakova (2010). In other words, the costly process in L2 is the annulment of the exhaustive inference rather than its calculation, which in English arises by default. As language proficiency reinforces and exposure to native input increases, learners grasp how to 'un-associate' the inference to the non-canonical structure and more readily accept clefts in violation contexts.

This is in line with the results in Feeney et al. (2004), who, in their experiment 3, found that native adult speakers of English took significantly longer to provide logical responses to

infelicitous 'some' statements. The authors take this result to suggest that the subsequent inhibition of a response based on the pragmatic inference is accompanied by cognitive processing efforts. Lin (2016) finds similar results for Mandarin learners of English, who were about half as fast to interpret 'some but not all' rather than the logical response 'some and possibly all'. Taken together, the results in the present study and the studies just cited support the Default account posited by Levinson (2000) and that argues pragmatic inferences are fast and effortless, but their cancelation is costly.

## 6. Conclusion

So far, most prior studies on the acquisition of pragmatic inferences have mostly focused on learners' offline judgments or on children's interpretation of sentences containing scalar expressions. Very few have used online measures to examine the processing of inferences, and even fewer have looked at triggers other than scalar expressions. Moreover, considerable disagreement has arisen out of these offline studies, with certain findings supporting the Interface Hypothesis, while others did not.

In this study, I aimed to bridge this gap by using reaction times and true/false judgments to investigate the issue of ultimate attainment of phenomena at the syntax-discourse interface. The findings provided partial support for upper limit of L2 proficiency with respect to processing resources. Indeed, at very high levels of attainment, L2 speakers still have difficulties in the computation of the exhaustive inference associated with clefts, even though they have clearly acquired the syntax of the structure by a much lower level of proficiency. This is arguably due to processing limitations in the L2, and more specifically with limitations linked to the cost of canceling an inference that arises by default in L1 English when the context does not support it.

Nevertheless, these limitations seem to resolve with increasing amount of exposure to native input; L2 speakers who had lived in France for several years, had the highest self-reported daily use of French, and the highest self-reported French accent, did pattern with L1 French natives (i.e., they were as fast and accurate), suggesting that impairment at the syntax-discourse interface can be overcome, contra the predictions made by the IH. In sum, this work can inform future work on both information structure and pragmatic inferences in L2 French, and more generally on the learnability discrepancies in L2 acquisition.

## References

Antoniou, K. & Katsos, N. (2017). The effect of childhood multilingualism and bilectalism on implicature understanding. *Applied Psycholinguistics, 38*, 1–47.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1). http://doi.org/10.18637/jss.v067.i01

Bardovi-Harlig, K., & Dörnyei, Z. (1998) Do language learners recognize pragmatic violations? Pragmatic vs. grammatical awareness in instructed L2 learning. *TOESL Quarterly, 32*, 233–262.

Bartning, I. (1997). *C'est*—in native and non-native spoken French. *Studier i Modern Språkventenskap, 11*, 13–47.

Bartning, I., & Hammarberg, B. (2007). The functions of a high-frequency collocation in native and learner discourse: The case of French *c'est* and Swedish *det är*. *International Review of Applied Linguistics and Language Teaching, 45*, 1–43.

Belletti, A., Bennati, E., & Sorace, A. (2007). Theoretical and developmental issues in the syntax of subjects: Evidence from near-native Italian. *Natural Language & Linguistic Theory, 25*, 657–689.

Bohnacker, U. (2010). The clause-initial position in L2 Swedish declaratives: Word order variation and discourse pragmatics. *Nordic Journal of Linguistics, 33*, 105–143.

Bourns, S. K. (2014). Contrasting *c'est*-clefts and *it*-clefts in discourse. In S. K. Bourns & L. L. Myers (Eds.), *Perspectives on linguistic structure and context* (pp. 199–222). Amsterdam, The Netherlands: Benjamins.

Bouton, L. F. (1988). A cross-cultural study of ability to interpret implicatures in English. *World English, 7,* 183–196.

Bouton, L. F. (1992). The interpretation of implicature in English by NNS: Does it come automatically—without being explicitly taught? *Pragmatics and Language Learning, 3,* 53–65.

Bouton, L. F. (1994). Can NNS skill in interpreting implicature in American English be improved through explicit instruction? A pilot study. In L. Bouton & Y. Kachru (Eds.), *Pragmatics and Language learning*, vol. 5 (pp. 89–109). Urbana-Champaign: University of Illinois.

Büring, D. & Križ, M. (2013). It's that, and that's it! Exhaustivity and homogeneity presupposition in clefts (and definites). *Semantics and Pragmatics, 6*, 1–29.

Carter-Thomas, S. (2009). The French c'est-cleft: Functional and formal motivations. In D. Banks, S. Eason, & J. Ormrod (Eds.), *La linguistique systémique fontionnelle et la langue française* (pp. 127-156). Paris, France: L'Harmattan.

Chierchia, G., Crain, S., Guasti, M. T., Gualmini, A. & Meroni, L. (2001). The acquisition of disjunction: Evidence for a grammatical view of scalar implicatres. In A. H.-J. Do, L. Dominguez & A. Johansen (Eds.), *Proceedings of the 25th Boston University Conference on Language Development* (pp. 157–168). Somerville, MA: Cascadilla Press.

Chomsky, N. (1995).

Clech-Darbon, A., Rebuschi, G., & Rialland, A. (1999). Are there cleft sentences in French? In G. Rebuschi & L. Tuller (Eds.), *The grammar of focus* (pp. 83–118). Amsterdam, The Netherlands: Benjamins.

Cook, M. & Liddicoat, A. (2002). The development of comprehension in interlanguage pragamtics: The case of strategies in English. *Australian Review of Applied Linguistics, 25*, 19–39.

DeCat, C. (2007). *French dislocation: Interpretation, Syntax and Acquisition.* Oxford: Oxford University Press.

Declerck, R. (1988). Studies on copular sentences, clefts and pseudo-clefts. In E. S. Firchow et al. (Eds.), *Studies for Einar Haugen* (pp. 94–114). The Hague: Mouton.

Dekydtspotter, L. & Hathorn, J. (2005). *Quelque chose … de remarquable* in English-French acquisition: mandatory, informationally encapsulated computations in second language interpretation. *Second Language Research, 21*, 291–323.

Destruel, E. & DeVeaugh-Geiss, J. (2018). On the interpretation and processing of exhaustivity: Evidence of variation in English and French clefts. *Journal of Pragmatics, 138*, 1–16.

Destruel, E. & Donaldson, B. (2017). Second language acquisition of inferences: Evidence from the French *c'est*-cleft. *Applied Psycholinguistics, 38*, 703–732.

Donaldson, B. (2011a). Left dislocation in near-native French. *Studies in Second Language Acquisition, 33*, 399–432.

Donaldson, B. (2011b). Nativelike right-dislocations in near-native French. *Second Language Research, 27*, 361–390.

Donaldson, B. (2012). Syntax and discourse in near-native French: Clefts and focus. *Language Learning, 62*, 902–930.

Feeney, A., Scrafton, S., Duckworth, A., Handley, S. (2004). The story of *some*: everyday pragmatic inferenceby children and adults. *Canadian Journal of Experimental Psychology*, *58*, 121–132.

Gertken, L. M., Amengual, M., & Birdsong, D. (2014). Assessing language dominance with the Bilingual Language Profile. In P.  Leclercq, A. Edmonds, & H. Hilton (Eds.), *Measuring*

*L2 proficiency: perspectives perspectives from SLA* (pp. 208–225). Clevedon, UK: Multilingual Matters.

Green, A. (1986). A time sharing cross-sectional study of monolinguals and bilinguals at different levels of second language acquisition. *Brain Cognition, 5*, 477–497.

Grice, H. P. (1989) Studies in the Way of Words. Cambridge, Mass.: Harvard University Press.

Hamlaoui, F. (2007). French cleft sentences and the syntax–phonology interface. In M. Radišić (Ed.), *Actes du congrès annuel de l'Association canadienne de linguistique 2007* (pp. 1505–1518). Saskatchewan, Canada: Canadian Linguistic Association.

Horn, L. (1981). Exhaustiveness and the semantics of clefts. In V. Burke & J. Pustejovsky (Eds.), *Proceedings of the 11th Annual Meeting of the North East Linguistic Society* (pp. 125–142). Amherst, MA: Graduate Student Linguistic Association.

Huang, Y., & Snedeker, J. (2009). On-line interpretation of scalar quantifiers: Insight into the semantics-pragmatics interface. *Cognitive Psychology, 58*, 376–415.

Katz, S. L. (1997). *The syntactic and pragmatic properties of the c'est-cleft construction.* (Doctoral dissertation). University of Texas at Austin, Austin, TX. Retrieved from ProQuest Dissertations and Theses. (Accession Order No. AAT 9802920)

Kiss, K. E. (1998). Identificational focus versus information focus. *Language, 74*, 245–273.

Krifka, M. (2008). Basic notions of information structure. *Acta Linguistica Hungarica, 55*, 243–276.

Lambrecht, K. (1994). *Information structure and sentence form.* Cambridge, England: Cambridge University Press.

Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: MIT Press.

Lieberman, M. (2009). Necessary interpretation at the syntax/pragmatics interface: L2 acquisition of scalar implicatures. Paper presented at the Workshop on Mind Context Divide: Language Acquisition and Interfaces of Cognitive Linguistic Modules. University of Iowa.

Lin, Y. (2016). Processing of scalar inferences by Mandarin learners of English: An online measure. *PLoS One, 11*: e0145494.

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. In A. A. de la Fuente, E. Valenzuela, and C. Martinez Sanz, *Language Acquisition Beyond Parameters: Studies in honor of Juana M. Liceras. Studies in Bilingualism* (pp. 125–145). Amsterdam: John Benjamins.

Musolino, J. & Lidz, J. (2002). Children's command of quantification. *Cognition, 84*, 113–154.

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

Prince, E. F. (1978). A comparison of *wh*-clefts and *it*-clefts in discourse. *Language, 54,* 883–906.

Reichle, R. V., & Birdsong, D. (2014). Processing focus structure in L1 and L2 French: L2 proficiency effects on ERPs. *Studies in Second Language Acquisition, 36*, 535–564.

Roever, C. (2005). *Testing ESL Pragamtics*. Frankfurt: Peter Lang.

Rooth, M. E. (1992). A theory of focus interpretation. *Natural Language Semantics, 1*, 75–116.

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

Slabakova, R., Kempchinsky, P., & Rothman, J. (2012). Clitic-doubled left dislocation and focus fronting in L2 Spanish: A case of successful acquisition at the syntax–discourse interface. *Second Language Research, 28*, 319–343.

Snape, N. & Hosoi, H. (2018). Acquisition of scalar implicatures: Evidence from adult Japanese L2 learners of English. *Linguistic Approaches to Bilingualism, 8*, 163–192.

Sorace, A. (2011). Pinning down the concept of "interface" in bilingualism. *Linguistic Approaches to Bilingualism, 1*, 1–33.

Sorace, A., & Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian. *Second Language Research, 22*, 339–368.

Taguchi, N. (2008). Pragmatic comprehension in Japanese as a foreign language. *Modern Language Journal,* 92, 558–368.

Taguchi, N. (2011). Do proficiency and study-abroad experience affect speech act production? Analysis of appropriateness, accuracy and fluency. *International Review of Applied Linguistics in Language Teaching, 49*, 265—293.

Taguchi, N., Li, S. & Liu, Y. (2013). Comprehension of conversational implicatures in L2 Chinese. *Pragmatics and Cognition, 21*, 139—157.

Tracy-Ventura, N., McManus, K., Norris, J., & Ortega, L. (2014). "Repeat as much as you can": Elicited imitation as a measure of oral proficiency in L2. In Leclercq, P., Edmonds, A., and Hilton, H., (Eds.), *Measuring L2 Proficiency: Perspectives from SLA*. (pp. 143–166). Bristol, GB: Multilingual Matters.

Tremblay, A. (2011). Proficiency assessment standards in second language acquisition research. *Studies in Second Language Acquisition, 33*, 339–372.

Trévise, A. (1986). Is it transferable, topicalization? In E. Kellerman & M. Sharwood Smith
(Eds.), *Crosslinguistic influence in second language acquisition* (pp. 186–206). New
York, NY: Pergamon.

Valenzuela, E. (2006). L2 end state grammars and incomplete acquisition of Spanish CLLD
constructions. In R. Slabakova, S. Montrul, & P. Prévost (Eds.), *Inquiries in linguistic
development: In honor of Lydia White* (pp. 283–304). Amsterdam, The Netherlands:
Benjamins.

Watorek, M. (2004). Construction du discours par des apprenants de langues, enfants et adultes.
*Acquisition et Interaction en Langue Etrangère, 20*, 129–171.

Whelan, R. (2004). Effective analysis of reaction time data. *The Psychological Record, 58*, 475–482.