Pragmatic abilities in bilinguals
The case of scalar implicatures

Ludivine Dupuy,¹ Penka Stateva,² Sara Andreetta,²
Anne Cheylus,¹ Viviane Déprez,¹ Jean-Baptiste van der Henst,¹
Jacques Jayez,¹ Arthur Stepanov² and Anne Reboul¹
¹Institute for Cognitive Sciences-Marc Jeannerod (ISC-MJ UMR 5304,
CNRS – University Lyon 1), France / ²Center for Cognitive Science of
Language, University of Nova Gorica, Slovenia

The experimental literature on the pragmatic abilities of bilinguals is rather sparse. The only study investigating adult second language (L2) learners (Slabakova, 2010) found an increase of pragmatic responses in that population relative to monolinguals. The results of studies on early bilingual children are unclear, some finding a significant increase in pragmatic responses in early bilingual children (preschoolers) relative to monolinguals (Siegal et al., 2007), while another (Antoniou and Katsos, 2017), testing school children, does not. We tested adult French L2 learners of English and Spanish (in their two languages) as well as French monolingual controls in Experiment 1 and Italian-Slovenian early bilingual children (in both languages) and Slovenian monolingual controls in Experiment 2. Our results were similar to those of Antoniou and Katsos (2017) in early bilingual children, but different from those of Siegal et al. (2007). We found no pragmatic bias in adult L2 learners relative to adult monolinguals.

Keywords: scalar implicature, quantifiers, bilingualism, early bilingualism, L2 learners

1. Introduction

Until recently, few studies (Siegal et al., 2007, 2009, 2010, Antoniou and Katsos, 2017, Slabakova, 2010) have investigated the general pragmatic abilities (implicature processing) of bilinguals of different types such as early bilingual children and adult second language (L2) learners. Overall (though not unanimously), these studies argued that bilingual speakers display enhanced pragmatic performance
relative to monolingual ones, notably in deriving implicatures. Siegal et al., 2007 found a pragmatic bias in early bilingual preschoolers relative to their monolingual peers, while Antoniou and Katsos, 2017 found no pragmatic bias in school-aged bilectal/bilingual children. Given the unclear results of these studies, the issue of whether bilingual speakers (either L2 learners or early bilingual children) derive more pragmatic inferences than monolingual ones remains open. In this regard, the following questions arise:

– Do bilingual speakers have a pragmatic bias leading them to a greater rate of pragmatic interpretations relative to monolingual ones?
– If they do, are there differences between early bilingual children and adult L2 learners in terms of pragmatic bias in comparison to their monolingual peers?
– In keeping with the neo-Gricean view according to which pragmatic interpretations are cognitively less costly than logical interpretations, can L2 learners be expected to prefer pragmatic interpretations, given that they are less proficient in their L2 than in their L1 (see Slabakova, 2010)?

The present paper addresses these questions, focusing more specifically on scalar implicatures.

1.1 Grice’s theory of implicature

The notion of implicature was introduced by Grice (1989) to account for the discrepancy between the logical interpretation of logical words (e.g., and, or, if…then, quantifiers) and their interpretation in discourse. To take the case of some, (1) will be true both if Paul has eaten all the cookies (logical interpretation) and if he has eaten only a portion of them (pragmatic interpretation). In discourse, however, (1) will usually be interpreted as meaning that he has eaten some but not all of them.

(1) Paul has eaten some cookies.

According to Grice, the lexical meaning of some is the logical meaning (at least some or some and maybe all), but the alternative interpretation (only some or some but not all) in (1) arises in discourse as a scalar implicature, computed on the basis of the maxim of Quantity. The latter enjoins the speaker to say as much as she can, but no more than is strictly necessary. In the case of (1), the speaker had a choice between the more informative all and the less informative some. Given that she chose to use some rather than all, the hearer can infer that Paul did not eat all the cookies, or that the speaker does not know whether he ate all of them.

Grice’s pioneering work was followed by two different and opposing approaches, proposed by the neo-Griceans (see, e.g., Horn, 2004, Levinson, 2000) and the post-Griceans (see e.g., Sperber and Wilson, 1995, Noveck and Sperber,
Bilingualism and scalar implicatures

(2007). The neo-Griceans endorsed the notion of a specific class of implicatures, based on lexical triggers (such as some in (1)), of which scalar implicatures are the central example. Horn (2004) proposed that scalar implicatures are based on lexical scales, characterized by the fact that the stronger items in a scale (e.g., all) entail the weaker items (e.g., some), while the weaker items implicate the negation of the stronger items (e.g., some ~⇒ not all). Levinson (2000) went further and proposed that some is interpreted by default as some but not all. The lexical (or logical) meaning (i.e., some and maybe all) is only accessed if the pragmatic interpretation is contextually contradicted either through the situation or through the discourse.

By contrast, post-Griceans deny that scales play any role in the pragmatic interpretation of utterances such as (1). Rather, the interpretation process is a process of pragmatic enrichment (a process of strengthening), leading from the logical interpretation of the term (e.g., some and maybe all) to the pragmatic and more informative interpretation (e.g., some but not all).

The two accounts make quite opposite predictions relative to the derivation of scalar implicatures. According to the neo-Gricean account, the pragmatic interpretation is less costly than the logical interpretation. According to the post-Gricean account, the logical interpretation is less costly than the pragmatic interpretation. Consequently, the predictions relative to the developmental curve for pragmatic interpretations in children and to the reaction times for pragmatic interpretations in adults are different.

1.2 Experimental approaches to scalar implicatures in monolinguals

The experimental protocols used to test scalar implicature processing mostly exploited the fact that utterances with weak scalar terms give rise to two distinct interpretations. Sentences like (1) license both interpretations (2) and (3):

(2) Paul ate some and maybe all cookies. (Logical interpretation)

(3) Paul ate some but not all cookies. (Pragmatic interpretation)

Most studies used essentially similar paradigms. Participants were presented with an utterance that can be considered under-informative relative either to a context or to encyclopedic knowledge. That is, target sentences infelicitously contained a weaker term in a scale when a stronger term would have been more informative. Participants were then asked to judge these utterances, in a sentence evaluation task choosing between true/false, agree/disagree, well said/not well said. While both answers are correct, the participant favors either the logical (true/agree/well said) or the pragmatic interpretation (false/disagree/not well said) of the utterance depending on her answer choice. This type of sentence evaluation task formed the
staple of experimental work on scalar implicatures, with two main lines of investigation targeting the cost of pragmatic interpretations:

– developmental studies comparing pragmatic interpretations in children of different age-groups;
– reaction time studies in adults.

Convergent results of developmental studies showed that there is a developmental trajectory with younger children producing fewer pragmatic interpretations than older children (see Braine and Rumain, 1981, Noveck, 2001, Gualmini et al., 2001, Papafragou and Musolino, 2003, Guasti et al., 2005, Pouscoulous et al., 2007, though for divergent results, see Feeney et al., 2004, Papafragou and Tantalou, 2004, Barner et al., 2011, Katsos and Bishop, 2011, Foppolo et al., 2012, Skordos and Papafragou, 2014).

The results of adults’ reaction time studies have been relatively straightforward, showing through sentence evaluation tasks\(^1\) that pragmatic interpretations take significantly more time than logical interpretations (Bott & Noveck, 2004, Huang & Snedeker, 2009, Bott & al., 2012, but see Feeney et al., 2004).\(^2\) Thus, most experimental results on the production of pragmatic interpretation for scalars in monolingual speakers seem more compatible with the post-Gricean than with the neo-Gricean account.

1.3 Experimental work on pragmatic implicatures in multilingual speakers

As noted above, there is relatively little work on pragmatic implicatures in bi- and multilingual speakers. In this respect, one should distinguish between studies targeting L2 learners and studies targeting early bilingual children.

1.3.1 Scalar implicatures in L2 learners

The main study on scalar implicatures in adult L2 learners has been conducted by Slabakova (2010). On the assumption that L2 learners have diminished processing capacities in their L2, Slabakova derived the following two predictions:

\(^1\) Some studies used categorical sentences (e.g., Some elephants are mammals), while others used contextually infelicitous sentences.

\(^2\) However, some experimental results are less straightforward: for instance, Grodner et al. (2010) showed that given enough contextual support, pragmatic interpretations for underinformative some sentences do not take longer than the interpretation for control items with all or none, in a real world visual paradigm using eyetracking.
1. If neo-Griceans are right, L2 adult learners should derive more pragmatic interpretations for SI than adult native speakers; 
2. If post-Griceans are right, L2 adult learners should derive fewer pragmatic interpretations for SI than adult native speakers.

Slabakova tested four groups of participants: native English speakers tested in English in the USA; native Korean speakers tested in Korean in Korea; advanced Korean L2 learners of English, tested in English in the US; intermediate Korean L2 learners of English, tested in English in the US.

She reported the results of two experiments. The first one, using a sentence evaluation paradigm, adapted Noveck’s (2001) experimental material, consisting of contextless categorical sentences (e.g. Some elephants have trunks). The second, using experimental material borrowed from Feeney et al. (2004), was a paradigm where the target sentence was presented in the context of a short story involving the speaker intentionally misleading the hearer, using both pictures and sentences:

(3) Charlotte finds three pieces of candy on the kitchen table – Charlotte eats the first candy – Charlotte eats the second candy – Charlotte eats the third candy – Charlotte’s mum says “Charlotte, what have you been doing with the candies?” – Target sentence: Charlotte says “I have eaten some of/all the candies”.

The task was again a sentence-evaluation task. Though there were more pragmatic interpretations overall in the second than in the first experiment, in both experiments the L2 learners (with similar results in the advanced and intermediate groups) gave significantly more pragmatic interpretations than did the native speakers (the results of the two groups of native speakers were similar in both experiments). We hypothesize that the higher rate of pragmatic interpretations for all groups in the second experiment is due to the fact that the context implies that the speaker has a deceptive intention when she produces the underinformative sentence. This makes the pragmatic interpretation more relevant than the logical interpretation, as it is clearly the interpretation intended by the speaker. According to Slabakova, Korean L2 learners of English gave more pragmatic answers in all conditions (with or without context) because they had a higher cognitive burden as they were tested in English, hence went for the easiest solution (consistently with the neo-Gricean default account). Note that this interpretation relied on the untested assumption that these Korean L2 learners of English would have given significantly fewer pragmatic answers if they had been tested in their L1 (Korean) than they did in English. Because the L2 learners were not tested in both the L1 and the L2, we are unable to evaluate whether the pragmatic bias that Slabakova (2010) found in the L2 would occur only in the L2. Furthermore, her conclusion does not give weight to all the empirical evidence against Levinson’s (2000) lexical
default position, gathered from studies using the same paradigm (namely, sentence evaluation tasks). Indeed, previous experiments seem to contradict the idea, central to Levinson’s account that the pragmatic interpretation is less costly than the semantic interpretation (see Section Experimental approaches to scalar implicatures in monolingual speakers, above).

1.3.2 Implicatures in early bilingual children

The few studies dedicated to the development of pragmatic abilities of early bilingual children (Siegal et al., 2007, Siegal et al., 2009, Siegal et al., 2010, Antoniou and Katsos, 2017) examined whether bilingualism triggers a potential advantage in general pragmatic abilities (i.e. sensitivity to pragmatic deviation and comprehension of different kinds of implicatures related to the Gricean maxims of Quantity, Quality, Relevance and Politeness) and affects positively executive control functions (e.g., Bialystok, 1999, 2010, 2011, Bialystok et al., 2009). Both studies of Siegal and colleagues present data from children between 3 and 6 years of age and report that all children demonstrate a very high level of adherence to the Cooperative Principle but the bilingual participants are nevertheless advantaged in the pragmatics tasks. A potential explanation considered in these two studies is the enhanced executive control abilities in bilingual speakers due to the inhibition of the non-target language(s). However, this hypothesis was not uncontroversially supported. On the other hand, scalar implicature processing was given specific attention in only some of the research, and the corresponding literature leaves one with two diverging claims:

1. There is evidence that early bilingual children significantly outperform monolingual children in the processing of scalar implicatures. Since the results on the pragmatics test and the language proficiency test were negatively correlated, the bilingual advantage was suggested to have a non-linguistic character (Siegal et al., 2007).

2. There is no robust evidence that suggests a multilingual or multilectal advantage in the comprehension of scalar implicatures (Antoniou and Katsos, 2017).

Based on previous claims that bilingual speakers are more sensitive to the communicative context (Genesee et al., 1995), Siegal at al. (2007) hypothesized that bilingual children should perform better on scalar implicature tasks as compared to their monolingual peers. This hypothesis was tested within the classical Gricean framework that views scalar implicatures as arising due to the Maxim of Quantity. Siegal et al. (2007) tested sensitivity to scalar implicatures with 21 monolingual English speaking children, 23 monolingual Japanese speaking children, and 20 bilingual children acquiring English and Japanese, aged 4 to 6. The experimental
material adapted from Papafragou and Musolino (2003) was presented in the respective native languages of the control monolingual children, and only in Japanese to the bilingual children. Participants were asked to evaluate the appropriateness of the description that a puppet gave of an event. For instance, the experimenter told a short story about a bear who was very good at putting hoops on a pole and pointed to a picture of the bear and hoops all of which were already on the pole – then the puppet described that event by producing an under-informative sentence: *The bear put some of the hoops on the pole.* The results, correlated with measures of participants’ linguistic proficiency, point to a strong multilingual advantage. All children were administered a BPVT (British Picture Vocabulary Test) or a JPVT (Japanese Picture Vocabulary Test). The bilingual population was reported to be significantly behind their monolingual peers in linguistic proficiency. Based on this result, Siegal and colleagues concluded that the source of the pragmatic bias is not linguistic.

Siegal et al.’s (2007) results which traced a bilingual pragmatic bias related to scalar implicatures were not replicated in Antoniou and Katsos (2017). That study included a large battery of pragmatic tests which similarly to Siegal et al. (2009, 2010) probed into knowledge of the Gricean maxims of Quantity, Quality, Relevance and Manner and metaphor understanding. The Maxim of Quantity was targeted also by assigning two tasks on scalar implicature comprehension: an action-based task and a binary sentence evaluation task. In the first task children had to modify a display (involving boxes and animals) in such a way that it matched the statement (e.g., *There are dolphins in some boxes*). In the second one, they had to decide if a picture of five cards can be truthfully described by a sentence which contained a scalar implicature trigger in the target condition (e.g., *There are stars on some of the cards*). Three groups of participants between 6 and 9 years of age were compared on these tasks: (i) monolingual speakers of Modern Standard Greek, (ii) bilectals speaking Modern Standard Greek and Cypriot Greek, and multilingually growing children who spoke Cypriot Greek, Standard Modern Greek and English (or other languages in addition). There was no statistical difference between the three groups in their performance on the pragmatic tasks (and no difference between the scalar task results and those of the other implicature tasks). Language proficiency (tested through the Word Finding Expressive Vocabulary Test and the Peabody Picture receptive Vocabulary Test) was lower in the multilingual and bilectal groups than in the monolingual group, despite their identical pragmatic performance, suggesting that these groups may use different pragmatic strategies.
2. The current experiments

Our main goal was to test in a more regimented way the existence of a pragmatic bias in bilingual speakers relative to monolingual speakers. We targeted populations like those who participated in the experiments described above (see section 1.3.1 *Scalars in L2 learners* and section 1.3.2 *Implicatures in early bilingual speakers* above), i.e., adult L2 learners and early bilingual children (aged 10–10.11, i.e., in the same range but more homogeneous in age than the subjects in the studies by Antoniou and Katsos as mentioned above).

We wanted to use the same material with children and adults to ensure a close comparison. To this end, we created a sentence evaluation task in which the target sentence was presented relative to a context (a short story presented in pairs of image-sentence), with sets of five objects, where *some* would be felicitous when only two objects (i.e., less than half, see Stateva et al. in preparation for the rationale) in the set were affected by the action described in the target sentence. We eliminated all traces of deceptive behavior in the scenarios on the hypothesis that it introduces an artifactual bias toward more pragmatic answers. Finally, in the target sentences, we settled for the construction *some xs* rather than the partitive *some of the xs*, to avoid the bias of the encoded linguistic meaning in the partitive (see Degen and Tanenhaus, 2011), as we wanted to test the pragmatic abilities of bilingual speakers, not their ability to grasp conventional meaning. Participants were asked to answer a metalinguistic question that seemed more likely to be interpreted as targeting felicity rather than truth. As the target sentence was presented by a puppet named Pipo, the question was *Is Pipo right?*

We tested the following groups of participants: early bilingual Slovenian/Italian children; adult French L2 learners of either English or Spanish; a control group of monolingual Slovenian children; a control group of monolingual French adults.

Given previous experimental results (see section *Experimental work on pragmatic implicatures in multilingual speakers*), we tested two predictions, on the hypothesis that bilingual speakers (both adult L2 learners and early bilingual children) have a pragmatic bias that leads them to give more pragmatic answers than monolinguals do:

1. Early bilingual Italian/Slovenian children will produce significantly more pragmatic interpretations than monolingual children;
2. Adult French L2 learners of English or Spanish will produce significantly more pragmatic answers than adult monolingual French speakers.

---

3. The expression *well said* is not always easy to translate in languages other than English.
On the additional hypothesis that Slabakova’s (2010) explanation is correct – adult L2 learners favor pragmatic interpretations in their L2 because these are easier to access –, we also tested a third prediction:

3. Adult French L2 learners of English or Spanish will produce significantly more pragmatic interpretations in their L2 than in their L1.

We conducted two experiments; one with adult French L2 learners of English or Spanish and adult monolingual French speakers as controls; one with early bilingual Slovenian/Italian children and monolingual Slovenian children as a control group.

This study was carried out in accordance with the recommendations of the Comité de Protection des Personnes Sud Est II, who gave it its agreement (IRB number: 11263). Participation of participants in Italy was sanctioned by Comitato Etico del Dipartimento di Scienze Neurologiche e del Movimento dell’Università degli Studi di Verona. All participants gave written informed consent in accordance with the Declaration of Helsinki.

3. Experiment 1a

Experiment 1a aimed at 1) replicating (or not) Slabakova’s (2010) results, avoiding the methodological problems that her study faced, as well as 2) assessing the impact of late second language acquisition on the mechanism of computating SIs in order to determine whether or not learning a foreign language has an impact on pragmatic abilities.

3.1 Method

3.1.1 Participants

The participants were 90 students, all native speakers of French (age range 19–26, \(M = 21.9\), twenty five were men), who were divided into three language groups: 30 French monolinguals \((M = 23; SD = 2.0)\), 30 French learners of English \((M = 21.2; SD = 0.59)\) and 30 French learners of Spanish \((M = 21.4; SD = 1.3)\).

The French monolingual speakers were recruited in Lyon and Bordeaux, France, and were categorized as monolingual based on a questionnaire regarding (1) their ability to understand and speak a L2, (2) their language environment and (3) their current exposure to formal L2 teaching. Those who qualified for this study had a poor level of foreign language knowledge and had not been exposed to any foreign language for at least six months, be it through formal classroom instruction or informal natural exposure.
The L2 learners were recruited in the Grande Ecole of Management in Strasbourg, France. The participants had all started to learn their L2 at school between 11 and 13 years of age and were still currently exposed to their L2 both in classes and for communication purposes. Their proficiency level in their L2 (English or Spanish) was established based on their level in the Common European Framework of Reference for Languages. They were upper-intermediate learners (B2 level), that is, proficient enough to interact normally with native speakers.

3.1.2 Design and materials

The eight experimental items (infelicitous some) stories consisted of 7 image-sentence pairs, followed by a question screen, as shown in Figure 1:

![Example of a storyboard for the Infelicitous Some condition](image_url)

Figure 1. Example of a storyboard for the Infelicitous Some condition. This figure illustrates the experimental material that was used in the target condition.

Control items were four stories for false all, four stories for true all and four stories for felicitous some. The control items for false all followed the same type of scenario...
but the target sentence used *all* instead of *some*. The felicitous *some* control items followed a slightly different scenario as shown below:

![Storyboard for the felicitous *some* condition](image)

**Figure 2.** Example of a storyboard for the felicitous *some* condition. This figure illustrates the experimental material that was used in one of the control conditions.

The bilingual participants were tested in both their L1 (French) and L2 (English or Spanish). They were tested on all the test items with half of them presented in their first or dominant language and half in their second or non-dominant language. We thus chose a design with four test items in each control condition (i.e. true *all*, false *all*, felicitous *some*) and eight test items in the target condition (infelicitous *some*). Finally, we decided to use written material.

### 3.1.3 Procedure

The experiment lasted approximately 15 minutes. It was self-paced and administered on a computer screen. Following a set of instructions, participants immediately began the experiment. The monolingual participants saw all 20 randomized
items (8 target infelicitous some, 4 felicitous some, 4 false all, 4 true all) in their native language. The groups of L2 learners saw two blocks of items, one in their L1 (French) and the other in their L2 (English or Spanish). The presentation order was counterbalanced. Each block was composed of 2 tokens of each set of control items (true all, false all, felicitous some) and 4 tokens of the set of target experimental items (infelicitous some). Each item was presented equally in both languages.

3.2 Results and discussion

3.2.1 Data treatment and exclusion criteria
We calculated the mean number of correct responses to the control sentences per participant. The accuracy rate was overall high and showed little inter-group variation (see Table 1). Four participants were removed from the study as they made too many mistakes in the control conditions (rate of error ≥ 25%).

<table>
<thead>
<tr>
<th>Groups</th>
<th>Language</th>
<th>True all</th>
<th>False all</th>
<th>Felicitous some</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2 English (n = 28)</td>
<td>English</td>
<td>96.7 (18.1)</td>
<td>93.3 (25.2)</td>
<td>88.3 (32.4)</td>
</tr>
<tr>
<td>L2 English (n = 28)</td>
<td>French</td>
<td>91.7 (27.9)</td>
<td>96.7 (18.1)</td>
<td>95 (22)</td>
</tr>
<tr>
<td>L2 Spanish (n = 29)</td>
<td>French</td>
<td>98.3 (12.9)</td>
<td>93.3 (25.2)</td>
<td>93.3 (25.2)</td>
</tr>
<tr>
<td>L2 Spanish (n = 29)</td>
<td>Spanish</td>
<td>98.3 (12.9)</td>
<td>100 (0)</td>
<td>83.3 (37.6)</td>
</tr>
<tr>
<td>Monolinguals (n = 29)</td>
<td>French</td>
<td>100 (0)</td>
<td>95.8 (20)</td>
<td>98.3 (12.9)</td>
</tr>
</tbody>
</table>

Response Analysis. Figure 3 presents the percentage of pragmatic responses (i.e., “false”) for under-informative some in the three groups of participants.

The analysis of the results for the French monolingual group revealed a rather low rate of pragmatic answers for French monolinguals (M = 0.5) in keeping with previous research (see Slabakova, 2010 and the general discussion below).

We analyzed the results by means of a mixed logistic regression, a technique appropriate for proportional binomial or multinomial non-ordinal responses (see Hilbe, 2009, for a detailed presentation). We fitted mixed logistic regression models using the glmer function of the R package lme4 (Bates et al., 2015). Pairwise contrasts were estimated with the R package multcomp (Hothorn et al., 2008).

The linear mixed model for logistic regression used Group (monolinguals or L2 learners) and Language (French or L2) as fixed effects and random intercepts on participants; random intercept and slopes on items (stories); and random intercept on the experimental variation (monolinguals in French; Spanish learners in French; Spanish learners in Spanish; English learners in French; English learners
Bilingualism and scalar implicatures

in English). It revealed a main effect of Group (monolinguals vs. L2 learners: $p < .05$) but no difference between the two languages used by L2 learners ($p = .96$).

In keeping with Slabakova’s (2010) results, the group of L2 learners made significantly more pragmatic interpretations than the control group of French monolingual adults. Additionally, given that Slabakova tested Korean native speakers while we tested French monolingual speakers, and that the L2 tested were English and Spanish, the choice of language itself seems to make no difference. The increase of pragmatic answers for L2 learners thus seems to be a solid cross-linguistic phenomenon.

However, Slabakova’s explanation cannot be sustained, given that the L2 learners in experiment 1a drew the same proportion of pragmatic interpretations in their L2 and in their L1. If her interpretation in terms of the choice of the less costly interpretation in the L2 were correct, one would expect L2 learners to draw significantly more pragmatic interpretations in their L2 than in their L1. But L2 learners in our study had identical performance in the two languages which goes against Slabakova’s conclusion. Rather, this suggests that L2 learning induces a pragmatic bias.

Following a suggestion by the editor, we analyzed individual results to see whether the pattern reported in the literature on SIs (see, e.g., Noveck 2001, Noveck and Posada 2003) – the tendency for participants to split into two categories ‘logical’ or ‘pragmatic’ in the presence of underinformative sentences – holds across
languages. We created a contingency table (Table 2 below) and then grouped together participants based on the majority of their responses in either the logical or pragmatic categories. Individuals were categorized as predominantly pragmatic if they gave pragmatic answers to six out of eight items (i.e., we used a 75% cut-off point). We ran a Chi-Square test. The test rejected the hypothesis of independence ($\chi^2 = 21.6, p < .001$), thus indicating that our L2 learners were either ‘logical’ or ‘pragmatic’ with respect to SI processing and that they kept the same strategy in both their languages. This is compatible with what has been observed in the previous literature (e.g., Noveck, 2001; Noveck and Posada 2003).

Table 2. Contingency table showing the individual relationship between performance in L1 and performance in L2 for Experiment 1a

<table>
<thead>
<tr>
<th>Percentage of pragmatic answers in L2 (English or Spanish)</th>
<th>0%</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>100%</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of pragmatic answers in L1 (French)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0%</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>25%</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>50%</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>75%</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>100%</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>7</td>
<td>20</td>
<td>30</td>
</tr>
<tr>
<td>Sum</td>
<td>12</td>
<td>2</td>
<td>6</td>
<td>10</td>
<td>27</td>
<td>57</td>
</tr>
</tbody>
</table>

In our experiment, we used a within-subject design for the L2 learners group (the participants saw half the stimuli in their L1 and the other half in their L2) and the instructions given made that explicit. Since the participants knew they would be tested in two languages, we can legitimately wonder whether L2 learning results in enhanced pragmatic abilities or if being tested in two languages makes participants more aware of pragmatic cues. Experiment 1b is aimed at teasing apart these possibilities.

4. Experiment 1b

Experiment 1b is a follow-up study that relies on a between-subjects design. It was carried out to establish whether the effect found in Experiment 1a reflects a genuine increase of pragmatic abilities in L2 learners or whether it results from an artifact of the within-subject design.
4.1 Method

4.1.1 Participants
Forty-six French second year L2 learners of English (age range 19–27, \( M = 21.1; \ SD = 2.2, \) 18 were men) enrolled in an English Studies Degree at the University of Lyon, France, were recruited for this experiment.

4.1.2 Design
As we used a between-subject design, half the participants (\( n = 23 \)) saw all 20 randomized items in their L1 and the remaining ones (\( n = 23 \)) saw all the stimuli in their L2 (English).

4.1.3 Materials and procedure
The stimuli and procedure were the same as those in Experiment 1a.

4.2 Results and discussion

4.2.1 Data treatment and exclusion criteria
Table 3 shows the mean number of correct responses to the control sentences per participant. No participants were removed from the study (rate of error < 25%).

Table 3. Accuracy rate and standard deviation for the three control conditions of Experiment 1b

<table>
<thead>
<tr>
<th>Groups</th>
<th>Language</th>
<th>True all (SD)</th>
<th>False all (SD)</th>
<th>Felicitous some (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2 English (( n = 23 ))</td>
<td>English</td>
<td>95.7 (20.5)</td>
<td>97.8 (14.7)</td>
<td>95.7 (20.5)</td>
</tr>
<tr>
<td>L2 English (( n = 23 ))</td>
<td>French</td>
<td>97.8 (14.7)</td>
<td>97.8 (14.7)</td>
<td>97.8 (14.7)</td>
</tr>
</tbody>
</table>

Response Analysis. The percentage of pragmatic responses for the L2 learners tested in Experiment 1b along with the percentage of pragmatic answers for the monolinguals and L2 learners obtained in Experiment 1a are displayed in Figure 4.

We conducted a Mann-Whitney test using Statistica. It revealed no effect of Language (L1 French versus L2 English) on the rate of pragmatic answers (\( U = 262.5, Z = 0.04, p = .965 \)) suggesting that L2 learners do not produce more pragmatic interpretations in their L2 than in their L1. This is in keeping with the results of Experiment 1a and is additional evidence against Slabakova’s conclusion. Additionally, when tested in only one of their languages, the rate of pragmatic answers of L2 learners is rather low (in French = 48%, in English = 41%) and is no longer statistically different from monolinguals either in French (\( U = 330, Z = 0.064, p = .949 \)) or in English (\( U = 332.5, Z = -0.018, p = .985 \)).
Experiment 1a showed that when tested in both languages L2 learners draw more implicatures than monolingual participants. These results could be interpreted as showing that learning a foreign language enhances pragmatic abilities. However, the results of Experiment 1b were quite different: the L2 learners behaved much like the monolingual speakers, giving a comparable amount of pragmatic answers. This suggests that learning a second language does not enhance pragmatic abilities but that having to switch between two languages makes one more aware of pragmatic cues.

5. Experiment 2

This experiment served two goals: (i) given the unclear results of the previous studies (Siegal et al., 2007 and Antoniou et al., 2013, Antoniou and Katsos, 2017), we wanted to test for a potential bilingual pragmatic bias; (ii) we also asked whether balanced early bilingualism affects pragmatic abilities.
5.1 Method

5.1.1 Participants and inclusion criteria

We invited monolingual Slovenian children and bilingual Slovenian-Italian children aged 10 – 11 for the current study. The age range was chosen to be close to the upper range of age of the participants in Antoniou and Katsos (2017) since the scalar implicature comprehension task is arguably the most demanding out of the set of pragmatic tasks. In addition, we wanted the bilingual children to have a good command of both of their languages, given the delays in language acquisition in bilingual children (see Bialystok et al., 2010). Monolingual participants were recruited in three elementary schools in Nova Gorica (Slovenia) and its surroundings. Bilingual participants were recruited in three Italian language elementary schools in Piran, Koper and Izola (Slovenia), and in three Slovenian language elementary schools in Gorizia (Italy). These areas are populated by recognized minority communities which have access to child education in mother tongue, as well as in the official state language.

For the purposes of this study, a special effort was made to target balanced early bilingual children. We also aimed at maximum homogeneity in the population with respect to their competence in both languages. To this end, information concerning the children’s linguistic background was (anonymously) collected through questionnaires that were distributed to the children’s parents together with consent forms before the study. The questionnaire asked:

a. for monolingual children, the language(s) that the child was exposed to during the periods of (i) between 0–2 years; (ii) between 2–6 (years), and (iii) after 6 years of age;

b. for bilingual children, the same as (a) for each of the two languages. In addition, we asked parents to evaluate how much the child used each language in different environments (e.g. with parents, siblings, other relatives and at school/daycare), on a five-point Likert scale ranging from ‘only Slovenian’ to ‘only Italian’ during each of these time periods (cf. Fernández, 2003).

Children were selected for the experiment on the basis of information given in parents’ questionnaires regarding (i) children’s age at the time of experiment; (ii) consistency of exposure to only one language in the case of monolinguals; and (iii) balance in exposure to the two languages in the case of bilingual children. The experiment was administered to 33 monolingual children. Our inclusion criteria for bilingual participants required that (i) a child has started receiving input in one of the two languages at age 0 and in the second not later than the end of the second year of life; and (ii) both Slovenian and Italian must be used on a daily basis in school and at least to some extent at home or within the local community.
On the basis of the questionnaire, we excluded eight participants from the bilingual group, because their competence in both languages was seriously unbalanced. One participant was excluded because of a cognitive disorder. As a result, the experiment was administered to 40 bilingual children.

For the selected bilingual group, we assessed the participants’ vocabulary knowledge of one of their two languages through the Peabody Picture Vocabulary Test which capitalizes on matching concepts to words in an increasing order of complexity (PPVT, Dunn and Dunn, 1981). To make sure that every child was sufficiently competent in both Slovenian and Italian to easily understand the details of the experimental tasks, we always administered the language proficiency test in the less dominant language. Although the bilingual participants were generally balanced speakers of both Slovenian and Italian, we used personal questionnaire information about the daily use of the languages in the school and home environment to decide which language input has been greater/smaller for each child. For the present study, we used the Italian PPVT (Stella, Pizzoli and Tressoldi, 2000), which is standardized for children from 4 to 12 years of age. In the test, there are different starting points, based on chronological age. We used a starting point two years before the starting points recommended in the test because of the claim that the lexicon of bilingual children is smaller than the one of monolingual children (e.g., Oller et al., 2007, Bialystok, 2010).

Since a Slovenian version of PPVT does not exist, for the purposes of this study we used our own adaptation of PPVT to Slovenian. This adaptation was created on the basis of translations from the Italian version by two Slovenian native speakers. We cannot guarantee the correspondence of the frequency of words between the two languages, but we assumed that concepts were mostly equally acquired cross-linguistically at the same age. Data from children whose raw score did not correspond to a standard score within the age range were excluded.

5.1.2 Design and materials
Experiment 2 had a within-subjects design. We used the same set of materials as described in Experiment 1a. Bilingual participants were tested on all the test items with half of them presented in Slovenian and half in Italian.

5.1.3 Procedure
Participants were tested individually in their respective school environment by two experimenters, one a native speaker of Slovenian, the other a native speaker of Italian. Children participated with parent’s signed permission and were told that they could stop at any point if they did not want to continue with the experiment. They were also told that they could ask for a break, if needed. They all received a sticker as a reward for participation. The testing included two sessions
for bilingual children: main session lasting approximately 15 minutes, followed by a language assessment session (approximately 15 minutes). The material was the same as in Experiment 1. As in Experiment 1, children received no training for the main session. The presentation of the material followed that of Experiment 1a: Slovenian monolingual participants saw all 20 items in Slovenian, while bilingual participants were randomly assigned to two groups, each of which saw half of all 20 items in Italian, and the other half in Slovenian and vice versa (counterbalanced for language, see Method of Experiment 1). The choice of items per group and manner of presentation followed closely the procedure in Experiment 1a, with one difference: the context and the target sentence for each item were read to the child by the experimenter who spoke natively the respective language. The experimenters used neutral intonation without prompting the choice of a specific interpretation. Children could see the target sentence at the point of answering whether the puppet describing the story was right in what he said. This way, we hoped to compensate for potential memory and attention limitations with child participants.

5.2 Results and discussion

The collected data were pre-processed using the following exclusion criteria:

1. performance on the PPVT;
2. performance on the control conditions during the experiment (similarly to Experiment 1 we excluded participants with rate of error ≥ 25%).

Five children were excluded based on their performance on PPVT which cast doubt on their being balanced bilinguals. Furthermore, two participants from each group were excluded because of a higher than 25% rate of error on control conditions. Accuracy for control conditions is presented in Table 4.

Table 4. Accuracy rate and standard deviation for the three control conditions of Experiment 2

<table>
<thead>
<tr>
<th>Groups</th>
<th>Language</th>
<th>True all</th>
<th>False all</th>
<th>Felicitous some</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bilinguals (n = 33)</td>
<td>Italian</td>
<td>100 (0)</td>
<td>100 (0)</td>
<td>97.1 (16.7)</td>
</tr>
<tr>
<td>Bilinguals (n = 33)</td>
<td>Slovenian</td>
<td>98.5 (11.9)</td>
<td>100 (0)</td>
<td>98.5 (11.9)</td>
</tr>
<tr>
<td>Monolinguals (n = 34)</td>
<td>Slovenian</td>
<td>99.2 (8.7)</td>
<td>99.2 (8.7)</td>
<td>84.8 (35.9)</td>
</tr>
</tbody>
</table>

For the subsequent analyses, we used the data from 31 monolingual participants (ages 10.1 – 10.11, M = 10.5; SD = 0.3) and 33 bilingual participants (ages 10.2–10.11, M = 10.5; SD = 0.3). The statistical analysis of the results was performed using the same procedures and statistical tools as those outlined in Experiment 1a.
Specifically, mixed logistic regression models taking a binary (yes, no) response as a dependent variable were fit to our data. First, we calculated the overall ratios of NO answers given by the Slovenian monolingual, as well as Slovenian/Italian bilingual children to the infelicitous *some* condition. These are the ratios of answers rejecting the felicitousness of the statement based only on the logical meaning of *some*, i.e. *some and possibly all* and ignoring the pragmatic contribution in the form of a scalar implicature. The answers for the bilingual group were broken down for each of their respective languages. The results are shown in Figure 5.

As Figure 5 indicates, the Slovenian monolingual participants produced a comparable amount of negative answers to the infelicitous *some* condition as the bilingual participants. Mixed logistic regression models taking Group (bilinguals, monolinguals) and Language (Slovenian, Italian) as fixed factors, random intercepts on Participants and random intercepts and slopes on Story revealed no significant main effect of Group ($p = .71$) or Language ($p = .23$). This suggests that the amount of scalar implicatures generated by children does not seem to be affected by bilingualism, nor does there seem to be a robust preference in generating more implicatures in any of the two languages within the bilingual group.

These results are in line with those of Antoniou and Katsos (2017) who tested 6 to 9 years old children and report a lack of significant advantage in the development of pragmatic abilities that can be linked to (early) bilingualism, in contrast with the results of Siegal et al. (2007). More specifically, we found no evidence that
early bilingual development affects the development of pragmatic abilities. It is also important to emphasize in this respect that the two groups in our study were carefully constructed so that each participant could come as close as possible to the ideally balanced early bilingual Italian-Slovenian speaker and a truly monolingual Slovenian speaker with as little as possible input from other languages outside of the school curriculum. In that sense, we can claim that the empirical basis for our conclusion is reasonably solid.

6. General discussion

Regarding adult L2 learners, the results of Experiment 1a using a within-subject design seem to support Slabakova’s (2010) results as L2 learners gave a higher rate of pragmatic answers than monolingual participants. Experiment 1b, using a between-subject design, shows that L2 learners do not have a higher rate of pragmatic answers than monolingual speakers, suggesting that the results of Experiment 1a were an experimental artifact. Again, the rate of pragmatic answers was similar in both languages. Thus, the results of both experiments also contradict Slabakova’s (2010) interpretation that deriving implicatures happens more often in L2 than L1, because it is easier than deriving the logical interpretation. Additionally, we did not find any pragmatic bias in early bilingual children in comparison with monolingual children. An analysis in terms of individual preferences toward pragmatic vs. logical answers showed no difference between monolingual and bilingual participants, and no difference in the bilingual group between the two languages. Thus, again, bilingual participants perform similarly to monolingual ones in terms of their pragmatic behavior, given that, just as monolingual participants, they tend to follow either a pragmatic or a logical path fairly consistently.

While previous experimental work on bilingual children (see Siegal et al., 2007, 2009, 2010, and Antoniou and Katsos, 2017) did not specifically address the neo-Gricean/post-Gricean debate, Slabakova (2010) did. Given that L2 learners produce similar rates of pragmatic answers in their L1 and in their L2, Slabakova’s explanation in terms of Levinson’s (2000) default account is not supported. It is not clear what prediction (if any) the neo- and the post-Gricean accounts would make regarding pragmatic performances in L2 learners or early bilinguals. It seems that the debate between neo- and post-Griceans is largely orthogonal to the issue of pragmatic competence in bilingualism.

Regarding the adult participants’ results in Experiment 1, the monolingual French speakers’ rate of pragmatic answers (46%) is in line with the results reported in previous research. To make an obvious comparison, in Slabakova’s first
monolinguals’ percentage of pragmatic answers was in the same range as ours (English speakers: 44.6%; Korean speakers: 38.8%). In Noveck’s (2001) third experiment, adult participants (French native speakers) gave a higher number of pragmatic answers (59%). In Pouscoulous et al.’s (2007) first experiment, again, adult participants (French native speakers) gave 47% of pragmatic answers.

On the other hand, in Experiment 1a, the within-subject conditions for L2 learners revealed a difference in the rate of pragmatic (i.e., negative) answers between the monolingual group and the two groups of L2 learners, which must be explained. There are several possibilities here. It is noteworthy, however, that the higher rate of pragmatic (negative) answers only exists in the within-subject experiment in L2 learners (Experiment 1a). In the between-subject experiment (Experiment 1b), there is no difference between the monolingual and L2 learner groups. This seems to preclude any explanation in terms of a general pragmatic bias in L2 learners, i.e., in terms of enhanced cognitive or pragmatic abilities. One would expect such a permanent advantage to manifest itself in both the within-and the between-subject experiments. It does not. This suggests that the better pragmatic performance of L2 learners in the within-subject experiment is just an experimental artifact: being warned that you will be tested in your two languages increases concentration and possibly metalinguistic awareness.

Let us now turn to the children’s results. The participants in our study are (at least) 10 year-olds but, while Noveck (2001) and Pouscoulous et al. (2007) found low rates of pragmatic answers in 9 year-olds, they used contextless categorical sentences. In experiments similar to ours (e.g., Feeney et al., 2004, Guasti et al., 2005) or those using a different type of responses (rewarding a character: Papafragou and Tantalou, 2004), the rate of pragmatic answers was similar to ours in younger children (4 to 8 years old, depending on the study). Also, the recent results of Foppolo et al. (2012), Katsos and Bishop (2011), and Skordos and Papafragou (2014), who all obtained comparable results with much younger children (5 and 6 years old), go in the same direction. Thus, the relatively high rate of pragmatic answers given by the child participants in Experiment 2 are not really unexpected.

We now come to the comparison of the respective advantages that bilingualism potentially offers to L2 learners or early bilingual children over their monolingual peers. As said above, the L2 learner results do not support a general pragmatic superiority of bilingual over monolingual speakers as the advantage found in the within-subject experiment (Experiment 1a) disappears in the between-subject

4. Slabakova’s (2010) second experiment, using the type of material illustrated in Example (3) in section 1.3.1 Scalar implicatures in L2 learners, involves the speaker trying to mislead her hearer by a less than informative utterance, which we consider a confounding factor (see section 1.3.1 Scalars in L2 learners).
Bilingualism and scalar implicatures

experiment (Experiment 1b). Regarding early bilingual children (Experiment 2), there is no statistical advantage in the results for the bilingual group over the monolingual group. Since the bilingual children were tested using a within-subject design, as in Experiment 1a, this may seem surprising. One might think that the tentative explanation proposed above for the L2 learners’ results in Experiment 1a, in terms of metalinguistic awareness, could apply here as well. However, the significance of the difference between monolingual participants and L2 learners in Experiment 1a is also due to the low rate of pragmatic answers in monolinguals (50%). In Experiment 2, on the one hand, the monolingual children have a higher rate of pragmatic answers (60%) than the monolingual adults in Experiment 1a. This suggests that the explanation for the lack of significance between the pragmatic performances of the two groups in Experiment 2 is due to the relatively high rate of pragmatic answers in the monolingual group. We have no specific explanation for this, apart from the hypothesis that the experimental material used seemed to be attractive to the children who commented favorably on the pictures. This may have made them more attentive leading to more pragmatic answers.

It is however noteworthy that L2 learners have a lower level of linguistic proficiency in their L2 than monolinguals. While we could not compare language proficiency of early bilingual and monolingual children, due to the absence of standardized linguistic test in Slovenian, it is a widespread finding that bilingual populations within the age range of the participants of Experiment 2 perform worse on proficiency tests (Bialystok et al., 2010, Antoniou and Katsos, 2017). Yet the pragmatic performances of L2 learners and of early bilingual participants are on a par with those of the two monolingual control groups, and this needs an explanation.

One possibility has to do with a link between pragmatic abilities and the Executive Control System. It has been argued that executive functions contribute to a bilingual advantage in behavioral tasks (Bialystok et al., 2012 and references therein; Costa and Sebastián-Gallés, 2014 and references therein). But Antoniou and Katsos (2017), who tested both for pragmatic abilities and for executive function, observed no such relation.

Another possibility, metalinguistic awareness, seems more plausible to account for the L2 bilingual pragmatic bias, but it has been linked more specifically to early bilingualism (see, e.g. Bialystok, 1988, 1993, Cromdal, 1999). While metalinguistic awareness possibly explains the results of the L2 learners in the within-subject design (Experiment 1a), it is hard to link it to L2 learning as such, given that it disappears in the between-subject design (Experiment 1b). It may however explain why bilinguals of both types do at least as well as monolinguals in the pragmatic tasks, despite the difficulties linked to a lower language proficiency.
A third possibility is that the pragmatic interpretation is not strongly linked to linguistic knowledge, making linguistic proficiency irrelevant to pragmatic tasks. This, however, needs further testing for which we now suggest some tentative directions. Work on embedded implicatures (see, e.g. Chemla and Spector, 2011) has shown that there are two strategies for interpreting scalar implicatures: the interpretation can be done either where the trigger occurs in the sentence (local strategy); or it can be done at the end of the sentence through a Gricean mechanism (global strategy). While these two strategies deliver the same interpretation on non-embedded scalar implicatures, they deliver different interpretations in embedded ones. On the assumption that only globally derived implicatures do not depend on grammatical knowledge, one can hypothesize that bilingual and monolingual speakers potentially prefer different strategies for scalar implicature processing. On this hypothesis, monolingual speakers would rely more on a linguistically based algorithm (hence adopting more local strategies, depending on lexically based alternatives), while bilingual speakers would favor more pragmatic processes (and hence adopt more global strategies, relying on Gricean reasoning). We leave this investigation for future research.

Notes about the contributors

Ludivine Dupuy and Penka Stateva are both first authors on this paper. Ludivine Dupuy and Sara Andreetta played an equal role in collecting the data, Ludivine Dupuy for Experiment 1, Sara Andreetta for Experiment 2.

Acknowledgements

We thank Patricia Tehami and Laura Suarez-Campos (two foreign language teachers at the Ecole de Management de Strasbourg) for their invaluable help in recruiting French L2 learners of English or Spanish. As well, we are thankful to the Ecole de Management de Strasbourg for offering us the use of a quiet office in which we could test the participants. We are grateful to Pia Živec and Petra Mišmaš for their contribution in preparing the material in Slovenian. We also thank Pia Živec for help in collecting the data in the early bilingual study. We gratefully acknowledge the help from the staff and children from the Slovenian and Italian schools and school centers in the Primorska region of Slovenia and the province of Friuli-Venezia-Giulia in Italy who collaborated and participated in our study.

The research leading to these results has received funding from the European Union’s Seventh Framework Program for research, technological development and demonstration under grant agreement no. 613465.
References

doi: 10.1017/S014271641600045X

doi: 10.1016/j.cognition.2010.10.010


doi: 10.1037/0012-1649.24.4.560


doi: 10.1111/1467-8624.00046


doi: 10.1037/a0025406

doi: 10.1177/152910060937084

doi: 10.1016/j.tics.2012.03.001

doi: 10.1017/S1366728999990423

doi: 10.1016/j.jml.2011.09.005

doi: 10.1016/j.jml.2004.05.006


doi: 10.1093/jos/ffq023

doi: 10.1038/nrn3709

doi: 10.1017/S0142716499000100


**Address for correspondence**

Anne Reboul
Institute for Cognitive Sciences-Marc Jeannerod
UMR 5304, 67 bd Pinel
69675 Bron cedex
France

reboul@isc.cnrs.fr

https://orcid.org/0000-0003-3976-5273

**Publication history**

Date received: 27 February 2017
Date accepted: 13 November 2017
Published online: 16 January 2018