Language Acquisition

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hlac20

Establishing upper bounds in English monolingual and Heritage Spanish-English bilingual language development

Kristen Syrett, Jennifer Austin & Liliana Sanchez

To cite this article: Kristen Syrett, Jennifer Austin & Liliana Sanchez (2021) Establishing upper bounds in English monolingual and Heritage Spanish-English bilingual language development, Language Acquisition, 28:1, 39-64, DOI: 10.1080/10489223.2020.1803328

To link to this article: https://doi.org/10.1080/10489223.2020.1803328

Published online: 22 Aug 2020.

LANGUAGE ACQUISITION
2021, VOL. 28, NO. 1, 39–64
https://doi.org/10.1080/10489223.2020.1803328

Routledge
Taylor & Francis Group

Establishing upper bounds in English monolingual and Heritage Spanish-English bilingual language development

Kristen Syrett[a], Jennifer Austin[b], and Liliana Sanchez[c]

[a]Rutgers, The State University of New Jersey, New Brunswick; [b]Rutgers, The State University of New Jersey, Newark; [c]University of Illinois at Chicago

ABSTRACT

Quantificational elements such as *some* pose a challenge to young language learners, given their vague meaning and ability to take on an upper-bounded interpretation (relative to *all*) in certain contexts. The challenge is enhanced when a child is acquiring multiple languages that do not share a one-to-one mapping between their lexical entries with *some*. Such is the case with *some* in English and *unos* and *algunos* in Spanish. Indeed, Heritage Spanish-English bilinguals have been documented as diverging from monolingual children and adults in their interpretation of *algunos*, which is said to lexically encode this upper-bounded meaning, although early Heritage bilinguals do not demonstrate this knowledge robustly. In this article, we ask how pervasive this challenge is by (a) investigating whether the same pattern holds in English, where there are not two words for *some*, and (b) comparing the pragmatic process for some to other linguistic items that either invoke another pragmatic process (particularized conversational implicature) or a semantic upper bound. Our results strongly suggest that the extended process of fine-tuning of quantificational lexical entries within and across languages precedes a pragmatic comparison of alternatives, but at the same time, Heritage bilinguals demonstrate pragmatic awareness beyond generalized conversational implicatures.

ARTICLE HISTORY
Received 25 January 2019
Accepted 20 May 2020

1. Introduction and background

It is well known that languages differ in their inventory of universal and existential quantifiers as well as how these terms are interpreted (Bach et al. 1995; Barwise & Cooper 1981; Keenan & Paperno 2012). Even within a language, there is variability in how certain quantifiers are interpreted in a given context. For example, English has an existential quantifier, or indefinite, *some*, which can be assigned an upper-bounded "some but not all" meaning in contexts that license a pragmatic "generalized conversational implicature" (Grice 1989; Horn 1989). Thus, a speaker who utters (1) could wish to simply express that they have read some number of articles on the syllabus, or that they have read only *some but not all* of the articles on the syllabus (that is, there are some that they have read, and others they have not read).

(1) I read some (of the) articles on the syllabus.

The reasoning is as follows: If a speaker chooses to deliver an utterance with *some* instead of the stronger universal quantifier (e.g., *all*), then their use of the weaker lexical item on the scale—which is entailed by the stronger scalar alternative—invites the inference that the stronger term does not hold (Horn 1989, 2004). It has been experimentally demonstrated that this upper-bounded meaning of *some* is made especially salient (or even obligatory) in particular linguistic environments (i.e., the partitive *x of the y* construction (Degen & Tanenhaus 2015) or the H* pitch-accented *SOME* (Miller

CONTACT Kristen Syrett kristen.syrett@rutgers.edu Center for Cognitive Science, The State University of New Jersey., Piscataway, NJ 08854.

et al. 2005). That this meaning is a pragmatic implicature and is not semantically entailed is highlighted by the fact that the "some but not all" meaning is defeasible; one can cancel it by following (1) with something like, *In fact, I read them all.*

Other languages beyond English have multiple lexical items that are assigned such an existential quantifier meaning comparable to *some*. These items and the scalar implicature associated with them have been investigated experimentally in children and adults. For example, within the Romance languages, French has both *certains* and *quelques* (Bott & Noveck 2004; Marty & Chemla 2013; Pouscoulous et al. 2007), Italian has both *alcuni (de)* and *qualche* (Foppolo, Guasti & Chierchia 2012), and Spanish has *unos* and *algunos*. In each case, the two competing quantifiers are proposed to have different meanings reflected in their lexical semantics. The latter item is claimed to encode (at some level or another)—and therefore signal to a listener—an upper-bounded meaning. That this may be an implicature and not a necessary part of the truth conditional meaning is illustrated with the Spanish example in (2), a version of the English example in (1), which shows that the upper-bounded implicature can be canceled.

(2) *Leí algunos artículos sobre el programa.* (*De hecho*, [*los leí*] *todos.*)

Read.1S.PST some articles on the syllabus (In fact (them read) all)
'I read some articles on the syllabus. In fact, I read them all.'

The contrast between *unos* and *algunos* as well as the lexical meaning of *algunos* in particular have received attention in the theoretical literature in recent years (Fabrégas 2010; Gutiérrez-Rexach 2001; Martí 2007, 2008, 2009). And in fact, previous research has shown that by age 4–5, Spanish monolingual children may not only appreciate the "some but not all" meaning of *algunos* (Miller et al. 2005; Vargas-Tokuda, Gutiérrez-Rexach & Grinstead 2008), and do so unequivocally in the partitive form (Katsos et al. 2011), but they also recognize a distinction between *unos* and *algunos* in that the latter (but not the former) encodes an upper bound (Vargas-Tokuda, Gutiérrez-Rexach & Grinstead 2008). At the same time, however, Syrett et al. (2017a, 2017b) have shown that even Spanish monolingual children may fail to calculate implicatures at a robust rate under certain circumstances and methodological conditions. This finding is consistent with a range of research on scalar implicatures in both children and adults demonstrating variability of performance induced by methodological choices (Foppolo, Guasti & Chierchia 2012; Katsos & Bishop 2011; Pouscoulous et al. 2007; Skordos & Papafragou 2016).

Given the presence of multiple lexical items within a language that encode an existential quantificational meaning, but which differ with respect to their implicature triggering, the following question arises: What happens when a child is acquiring multiple languages, where at least one language has two words for "some" and where none of the lexical entries overlaps entirely with another? In their investigations of Spanish, Syrett et al. (2017a, 2017b) have shown that in contrast to their monolingual peers, Heritage Spanish-English bilingual preschoolers do not appear to readily calculate the "some but not all" implicature with *algunos*. However, in a task that highlights the conversational goals of interlocutors and makes relevant an upper-bounded meaning in a pragmatic context, Heritage Spanish-English bilingual children *do* assign an upper-bounded meaning to *algunos* but still do not readily distinguish its meaning from *unos* with respect to upper bounds. Thus, while monolingual Spanish children may differentiate between *unos* and *algunos* and recognize the upper-bounded meaning highlighted by the latter, children acquiring another language alongside Spanish (where this other language, English, is the socially dominant language) exhibit a more protracted process in identifying lexical meaning.

It is perhaps not surprising that Heritage bilingual children should display a unique pattern when assigning the intended meaning to *algunos* and *unos* for three main reasons. First, lexical development is often protracted in bilingual children compared to monolinguals (Bialystok et al. 2010; Oller, Pearson & Cobo-Lewis 2007), in part because bilingual children must reconcile multiple competing lexical entries within and across two languages, which both overlap and diverge in meaning (Yan & Nicoladis 2009). In this case, a child acquiring *some, unos*, and *algunos* is faced with lexical features that in some cases neatly

overlap (i.e., the existential quantification component), and at the same time do not (i.e., how context specifically alters the meaning, whether and how an upper bound is signaled, and so on). To capture the inherent challenges involved in this process, one might appeal to structural alignment across representations (Gentner & Markman 1997; Markman & Gentner 1993): While these lexical entries share some common features, it is their distinctive features that must be sorted out—and this naturally takes time because it requires accumulating evidence across speakers and contexts. That preschoolers appeal to structural alignment in analogy and comparison in language learning (Gentner & Namy 1999, 2006; Namy & Gentner 2002) makes this possibility plausible for both languages being acquired.

Second, children who are exposed to a language spoken at home, but who experience decreased exposure to this minority-heritage language and increased exposure to the socially dominant language of the community as they enter the school system, often experience variability in their linguistic representations (Flores & Barbosa 2014; Montrul & Sánchez-Walker 2013). Differences in the quantity and quality of input in the heritage language, as well as differences in activation for production and comprehension, have been identified as determining factors for differential outcomes among Heritage speakers when compared to monolingual speakers of the heritage language (Montrul 2018; Putnam & Sánchez 2013; Rothman 2007). In their corpus analysis of Heritage Spanish-English bilingual children's narrative proficiency, Goldin, Syrett & Sánchez (2020) have demonstrated production of perspective-taking deictic verbs such as *venir* 'come' and *ir* 'go' that aligns more with monolingual speakers of the more flexible and dominant language, English, relative to speakers of the more restricted language, Spanish, in that the Heritage bilingual children are significantly more likely to allow for a nonspeaker perspective to be adopted.

Third and finally, the word meanings of *unos* and *algunos* themselves are dependent upon variable, pragmatic features of the context, which may be challenging. Integrating such pragmatic information can be a difficult task for bilingual children. There is evidence that cross-linguistic influence is most likely to occur at the syntax-pragmatics interface (Hulk & Müller 2000; Müller & Hulk 2001; Serratrice 2007; Serratrice, Sorace & Paoli 2004), especially where two languages overlap with respect to a particular property, thereby making it difficult for the bilingual child to navigate between language-universal interpretive strategies and language-specific rules (Müller & Hulk 2001).[1] What's more, among the most vulnerable areas of a heritage language to succumb to the more dominant language are the lexicon and aspects of syntax that rely on the context (see Montrul 2004). While the phenomenon of interest here is not at the syntax-pragmatics interface, it involves an aspect of lexical meaning that is inherently pragmatic and dependent upon the presence and relevance of lexical alternatives. We might therefore predict a protracted process of sorting out these lexical meanings in Heritage speakers.

The perceived difficulty Heritage Spanish-English bilingual children show when calculating upper-bound scalar implicatures with *algunos* gives rise to two main questions. First, how general is this pattern across the languages they are learning and across lexical items? Is the pattern observed in previous work with Spanish-English bilinguals restricted to Spanish, where there are two competing existential quantifiers that mean *some*, or does it also extend to the other language they are learning (English)? Second, is the observed pattern restricted to the existential quantifier *some* (in either language) and to entailment-based scalar implicatures specifically, or is it more broadly observed with other lexical items and contexts that invite quantity implicatures?

In asking the second question about the source of the response pattern in bilinguals, we can begin by looking at the steps involved in the calculation of a scalar implicature with *some* (whatever its form in the target language). Barner & Bachrach (2010) outline the steps for the upper-bound implicature associated with *some* as follows. First, compute the basic meaning of the sentence with the lexical item. Second, generate the set of alternatives (similar sentences containing scalemates of *some*, including stronger

[1]While research by Siegal and colleagues (Siegal, Iozzi & Surian 2009; Siegal, Matsuo & Pond 2007) has suggested that performance by bilinguals with pragmatic implicatures may exceed that of monolinguals, reasons for questioning the strength of this claim are outlined in Syrett et al. (2017a). Briefly, they concern the choice of languages and their competing existential quantifiers, performance compared to chance level, and task design.

alternatives). Third, restrict the alternatives to those with stronger expressions that entail the weaker ones (e.g., *all*). Finally, negate these stronger equivalents to generate the alternative. Foppolo, Guasti, & Chierchia (2012) outline a slightly different set of steps, which nevertheless require generating the basic meaning, activating alternatives, but then selecting the most informative interpretation of the original sentence, given Gricean maxims. It has been established that children's lack of knowledge of the relevant alternatives impacts their ability to compute scalar implicatures (Barner, Brooks & Bale 2011). However, simple knowledge of these alternatives may not be sufficient; children must also recognize that other scalar terms and the sentences in which they appear are *relevant* alternatives, given the context or nature of the task (Skordos & Papafragou 2016). Thus, simply activating an alternative (like *all*) is not enough: Children must appreciate that this alternative is *relevant* to the conversational goals. Thus the bilingual child is not only faced with this process as is the monolingual child, but with the added challenge of sorting out additional lexical meanings and the increased cognitive load of generating the relevant scalar alternatives in a particular language and context. One might predict that by moving from a generalized conversational implicature that relies on comparisons of overlapping lexical meanings arranged in an entailment-based scale in a given context to a particularized conversational implicature that removes this lexical-semantic component that bilingual children might exhibit a higher rate of success.

In this study, we begin to answer these questions and probe the process of implicature calculation in bilinguals relative to monolinguals by exploring whether and how children who are beginning to experience a decrease in input in the home language (Spanish) and an increase in input in the socially dominant language (English) (Heritage Spanish-English bilingual speakers) compare with English-speaking monolingual children when integrating lexical and pragmatic information. There is no *a priori* reason to think that bilingual children should pattern any differently from monolingual children in the last part of the implicature calculation process (negating alternatives). However, they may diverge from monolinguals in their knowledge of the first two steps: identifying the target lexical meaning and generating the stronger lexical alternatives, especially since bilingual children have been found to have lower levels of receptive and expressive vocabulary (Bialystok et al. 2010), despite similar levels of exposure to input in both languages (Thordardottir 2011). While this article is designed as an investigation into Heritage bilingual children's ability to assign an upper bound, it is at the same time an attempt to demonstrate that by investigating bilinguals in this respect, we may gather further support for the discrete components of the implicature calculation process and identify a possible commonality in the pragmatic knowledge of children of the same age across two linguistic populations.

The goals of this article are thus twofold. First, we seek to determine how Heritage Spanish-English bilinguals compare with English monolinguals in their computation of quantity-based implicatures with *some*, thereby extending our knowledge of their ability to calculate upper-bounded implicatures with *some* beyond Spanish. Second, we seek to determine how general this (in)ability to assign an upper bound is—whether it is restricted to scalar implicatures with *some* or extends to other lexical items, which may, like *some*, invite generalized conversational implicatures associated with lexical meaning, or instead involve particularized conversational implicatures based on the context. In this article, we present two experiments designed to answer these questions.

Previewing our findings, we show that while Heritage Spanish-English bilingual children pattern fairly similarly with monolingual English-speaking children in their rate of implicature calculation with *some*—and that this ability changes predictably with the availability of alternatives and the conversational goals highlighted in the context of the experimental paradigm—they diverge from adult performance in a way that monolingual children do not. At the same time, they *can* assign an upper bound when the implicature does not depend on lexical meaning. Thus, the challenge for bilinguals lies in the first two steps of the process: computing the basic meaning of the sentence featuring the target quantificational lexical item and generating the set of relevant interpretational alternatives. These findings are consistent with recent evidence on the role of the developing lexicon (Foppolo, Guasti & Chierchia 2012) and knowledge and accessibility of lexical alternatives (Barner, Brooks & Bale 2011; Skordos & Papafragou 2016) in young children's ability to calculate scalar implicatures. Thus, once bilingual learners possess sufficient knowledge of alternatives and can flexibly

recalibrate how to calculate pragmatic meaning based on the context at hand, they should be able to calculate scalar implicatures (Miller et al. 2016; Slabakova 2010).

2. Methodological decisions

In order to accomplish these research goals, we adopted two specific and complementary experimental paradigms, implementing certain key design choices in each. We outline these tasks and manipulations in the following. Crucially, neither task called upon the children to retrieve real-world knowledge from memory or generic knowledge of object kinds to calculate the implicature (see Bergen & Grodner 2012; Bott & Noveck 2004; Guasti et al. 2005; Noveck 2001).

In Experiment 1, we employed a Truth-Value Judgment Task (TVJT) (Crain & Thornton 1998). Although it has been argued that tasks that elicit a binary dependent measure (specifically, the TVJT) may potentially mask underlying pragmatic competence (Katsos & Bishop 2011; Veenstra, Hollebrandse & Katsos 2018), Skordos & Papafragou (2016) have shown that this is not necessarily so. When quantificational lexical alternatives are made accessible and relevant, behavioral responses in a binary forced-choice judgment task can still reveal ability to calculate implicatures through an increased rate of *yes*/*no* responses. Moreover, within the TVJT, we made sure to include both true/false *all* statements and felicitous/infelicitous *some* statements within each experimental session, making the stronger alternative salient and quantity relevant (see Foppolo, Guasti & Chierchia (2012); contra Papafragou & Musolino (2003)). In this way, we invited bilinguals to participate in a task in which we were relatively confident about the relative rate of "some not all" implicature calculation by monolingual children and adults as a baseline.

In Experiment 2, we employed a video-recorded communicative context task (following Syrett et al. 2017a), which is particularly effective in highlighting communication between a speaker and hearer, as well as the role of the context and speaker goals in boosting the "some not all" interpretation (see also, Guasti et al. 2005; Papafragou & Tantalou 2004). In this task, we intermingled utterances with *some* with other quantificational items and nonquantificational items inviting an "and nothing more" particularized conversational implicature (Breheny, Ferguson & Katsos 2013; Stiller, Goodman & Frank 2015), thereby allowing us to compare implicature calculation with *some* to the establishment of upper bounds in other linguistic contexts.

In addition, in Experiment 2, we also manipulated the sentential frame of the speaker utterances between subjects (see Kennedy & Syrett, under review). In one condition, the modal frame of the utterance (*You may take* . . .) semantically induces an *upper* bound of permissibility, so that the specified quantity should not be exceeded. In a second condition, the modal frame (*Please put* . . .) semantically specifies a *lower* bound in the request, so that any upper bound is pragmatically assigned. This difference across subjects allows us to explicitly compare the frequency an upper-bounded interpretation being assigned and associate it with either a semantic or pragmatic linguistic source. The combination of this modal frame with the variation in implicature types (generalized or particularized) examined in Experiment 2 allows us to more carefully probe and compare the ability of monolingual and bilingual children to assign upper-bounded interpretations.

3. Experiment 1: Truth value judgment task

3.1. Participants

The participants included the following three groups:

(i) 28 native English-speaking adults (1 male, 27 females; 3 additional nonnative speakers excluded); all adults in both experiments were university-level undergraduate students taking introductory Linguistics or Cognitive Science classes.

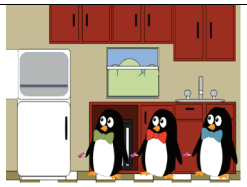
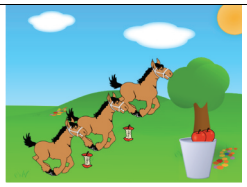
(ii) 15 English monolingual children (5 boys, 10 girls; 4;04 to 5;02, M: 4;10, SD: 3.2 months) (1 additional excluded for "yes" bias: responded "yes" to every single item)
(iii) 15 Heritage Spanish-English bilingual children (7 boys, 8 girls; 3;08 to 5;01, M: 4;05, SD: 3.9 months) (12 additional excluded: exhibited "yes" bias [n = 3]: responded "yes" to every single item); exhibited "no" bias (n = 1) (responded "no" to all but three items, including controls); responded to more than half of the controls incorrectly (n = 2); were noncommunicative with the experimenter and puppet, ostensibly as the result of a possible language barrier (n = 5); refused to interact with puppet in task (n = 1). The Heritage Spanish-English bilingual children in both experiments lived in New Brunswick, in Middlesex County, New Jersey, in a city where 49.3% of the population speaks Spanish at home. Their preschool served a majority of Latino children with a similar demographic background. The English monolingual children attended preschools in neighboring counties, with a predominantly English-speaking population. All studies were conducted on the preschool premises.

3.2. Materials and procedure

Experiment 1 employed a Truth-Value Judgment Task (Crain & McKee 1985; Crain & Thornton 1998). In this task, an experimenter narrates a series of stories to participants one at a time, following a script, while a puppet (played by a second experimenter) watches alongside the child. The premise of the task is that the puppet is learning and needs the child's help. As each story unfolds, both the child and puppet watch and listen carefully. At the end of each story, the puppet delivers a target response to which the child either responds affirmatively (and gives the puppet a nibble of a cupcake) or rejects (and gives the puppet a drink of milk). In this way, the child assesses the truth or falsity of the proposition expressed by the puppet's statement.

In the current version, we created scenes by animating images on Powerpoint slides, which we presented on a computer. For children, the slides were administered on a 17″ Macbook Pro computer in a quiet setting in their preschool. For adults, the slides were administered on iMacs at individual consoles in a quiet laboratory setting; the narration was prerecorded by a female native speaker of English and synced with automated slide presentation. Children interacted with a puppet, while adults were given a response sheet on a clipboard to record their responses. Regardless of their answer, we occasionally asked children to tell the puppet *why* he was right or wrong (so that he could learn more), thereby ensuring that we did not unintentionally favor one response type over another or give children the impression that they had to provide additional information only when they rejected the puppet's statement (which could result in a *yes* bias stemming from children's wanting the puppet to be correct more often than not) (Lidz & Musolino 2002; Syrett & Lidz 2011). Before the experimental session proper, children participated in a training session (without any quantifiers) to become accustomed to interacting with and rewarding the puppet for his responses. The entire session took approximately 18–20 minutes.

Participants were presented with 12 trials in the experimental session: four control items, two *all* statements in a true scenario, two *all* in a false scenario, two *some* statements in a true and felicitous scenario, and two *some* in a true but infelicitous scenario. See Figure 1. Items were pseudorandomized, and these items were then presented in one of two counterbalanced orders randomly among participants. The full set of target responses are presented in Appendix A. The control items did not feature quantificational lexical items in the target statement. Rather, participants were asked to evaluate a statement in which the number of objects, the color of an object, or the object kind was mentioned. In all eight of the target quantificational items, the last of the three or four characters hesitated before either completing the activity or not, thereby encouraging the children to attend to the end of the story, satisfying the condition of plausible dissent, and increasing the contextual support for the quantificational sentences.

all in a true scenario	*all* in a false scenario
	
All of the penguins ate a popsicle.	All of the horses ate an apple.

some in a true, felicitous scenario	*some* in a true, infelicitous scenario
	
Some dogs played with a ball.	Some boys scored a goal.

Figure 1. Sample images and test sentences for quantificational target items in Experiment 1, with the key *some* item on the bottom right.

3.3. Results

The results, captured in terms of the dependent measure (mean acceptance of the target statement), are presented in Table 1. Because of the binary nature of the responses, and the importance of comparing the performance of the groups for each trial type, we performed pairwise comparisons (Mann-Whitney tests for between-group independent sample, and Wilcoxon signed-rank tests for within-group nonindependent samples, with an alpha level of.05). We used the same pairwise comparison approach for both experiments.

While the percentage of acceptance for the false controls (the leftmost column) may appear to be somewhat high for the child groups, an analysis of individual response rates reveals that for each group, the majority of children answered correctly to three or four of the four false controls, and for the remainder of the children in each group, they responded systematically to the test items (in particular, the true and false *all* targets) in a way that revealed an understanding of the task and a lack of a response bias and were also able to offer justifications. We therefore included their data. In the figures that follow, we zoom in on each of these five trial types to compare the three participant groups.

We turn first to the control items, captured in Figure 2. Each of the three groups of participants correctly rejected these items. In a similar vein, as Figures 3 and 4 show respectively, all three groups correctly accepted true items and rejected false items with the universal quantifier *all*. All three participant groups responded at or near ceiling for the true *all* test items. There was no significant difference between any pairing of groups (all p values from here on out, two-tailed) (adults vs. monolingual children: $U_A = 784$, $z = .5$, $p = .617$; adults v. bilingual children: $U_A = 784$, $z = .25$,

Table 1. Summary of mean participant acceptance of target statement by group for five trial types with standard deviation in parentheses below.

Participant group	Controls (false)	*All* (true)	*All* (false)	*Some* (true + felic, not-all)	*Some* (true + **infelic**, all)
adults	3.6%	100.0%	3.6%	92.9%	39.3%
	(.19)	(0.0)	(.19)	(.26)	(.49)
monolingual children	21.7%	93.3%	26.7%	86.7%	56.7%
	(.42)	(.25)	(.45)	(.35)	(.50)
bilingual children	39.0%	96.7%	30.0%	50.0%	76.7%
	(.49)	(.18)	(.47)	(.51)	(.43)

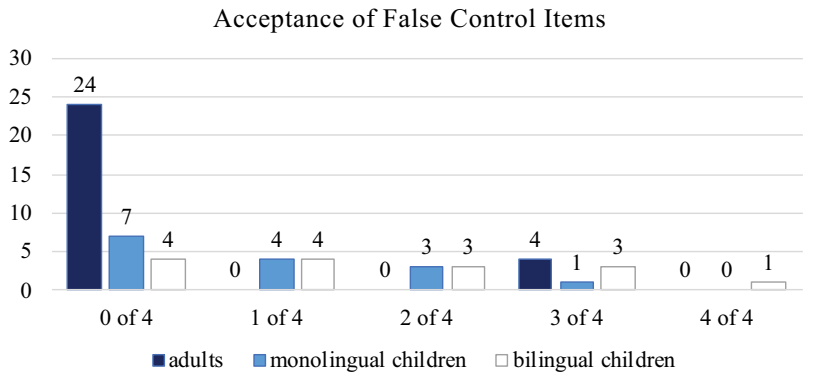


Figure 2. Individual acceptances of false control items.

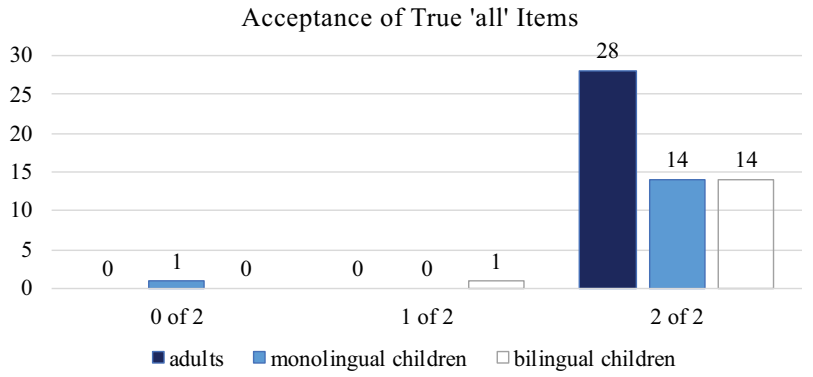


Figure 3. Individual acceptances of true *all* items.

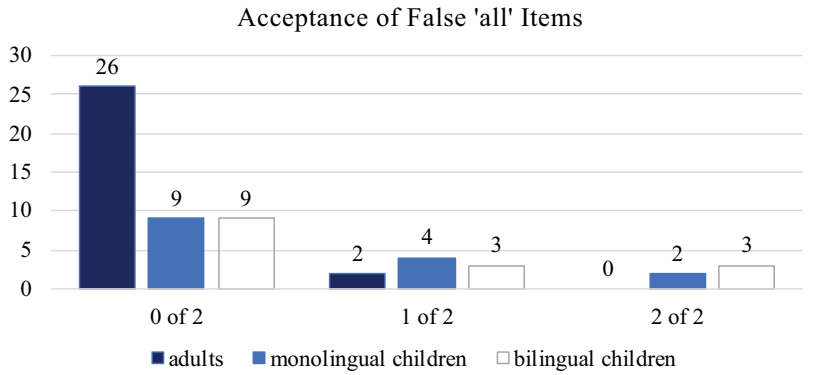


Figure 4. Individual acceptances of false *all* items.

p = .803; monolingual children v. bilingual children: U_A = 465, z = −.21, p = .834). Participants also responded in the predicted direction to the false *all* items. As with the control items, there was more variability with both groups of child participants than with adults. While there was no difference between child groups (U_A = 465, z = −.21, p = .834), and only a marginal difference between the adults and the monolingual children: U_A = 1034, z = −1.75, p = .08, the difference between adults and the bilingual children was significant (U_A = 1062, z = −2.01, p = .044).

We now turn to the *some* trial types, beginning with the "some true, felicitous" trials, where not all of the entities had the relevant property, shown in Figure 5. Here, the monolingual children and adults both accepted the actions at a comparable rate ($U_A = 788$, $z = .47$, $p = .638$), in line with the semantics of the existential quantifier, but bilingual children accepted them at a significantly lower rate than the monolingual children ($U_A = 285$, $z = 2.43$, $p = .015$) and than adults ($U_A = 480$, $z = 3.26$, $p = .001$). When we took a closer look at their justifications for their responses, we saw that these children did not necessarily lack an understanding of the semantics of the quantifier. Rather, they resisted the application of it to smaller set sizes like two or three, for the reason that these are subitizable numerosities that allow for rapid, accurate judgments of quantity. Children thought these sets were not big enough to be "some" and were better described with these cardinality terms. For example, some children rejected the puppet's statement, offering explanations for their response by saying things like, *only two cows got prizes, because two of them did, only two cars*, etc. (Interestingly, these were the same responses other children gave for *accepting* the *some* items.) This pattern is, in fact, in line with previous findings (see, e.g., Katsos & Bishop 2011; Katsos et al. 2011). Other children also seemed to reject the *some* statement because they were attending to a salient outlier. Thus, these specific responses appear to be less informative about children's understanding of the precise quantificational nature of these lexical items but more informative about how these children are establishing a contrast among lexical alternatives (which is also relevant to the process of scalar implicature calculation).

Finally, we turn to the target *some* trials in which all of the entities had the relevant property and the target statement was therefore "true, but infelicitous," presented in Figure 6. These are the crucial trials for assessing scalar implicature calculation. While the monolingual and bilingual children did not

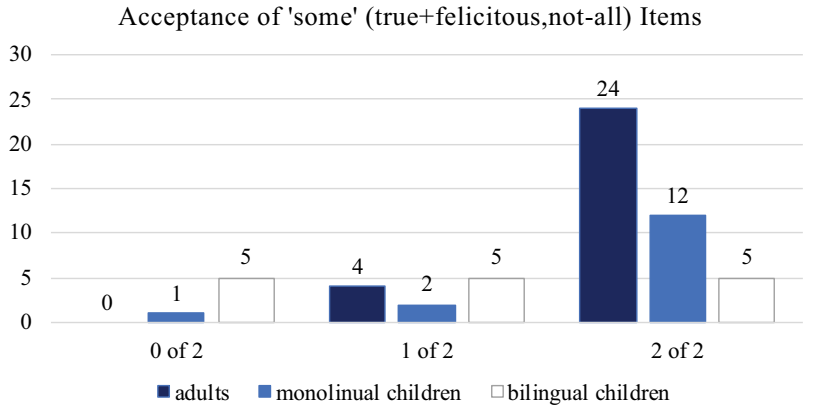


Figure 5. Individual acceptances of *some* (true + felicitous, not-all) items.

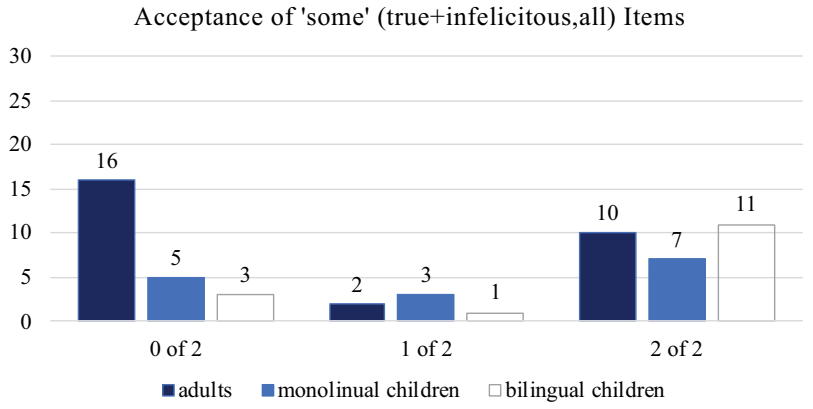


Figure 6. Individual acceptances of *some* (infelicitous, all) items.

Speaker Utterance	Addressee Action and Experimenter Question
You may take *some of the books*.	places all four books on the table, one by one
Please put *some of the books on the table.*	Experimenter: "Is what she did ok?"

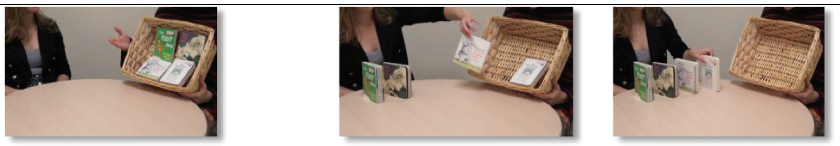


Figure 7. Sample still images from a video corresponding to two versions of a speaker utterance with test item *some* and an infelicitous addressee action in Experiment 2.

differ significantly from each other ($U_A = 540$, $z = -1.32$, $p = .187$), and the monolingual children and adults did not differ from one another ($U_A = 986$, $z = -1.32$, $p = .187$), the bilingual children *did* differ significantly from the adults, being much more likely to accept the target statements ($U_A = 1154$, $z = -2.84$, $p < .005$). The individual responses for the adults and monolingual children reflect a bimodal distribution of acceptance patterns among participants (a pattern also noted by Guasti et al. (2005) for both their Experiments 1 and 2 and the adults in Noveck (2001).

In addition, we note two other points of comparison here. First, consistent with our observation in the previous paragraph about the role of set sizes in assessments, the bilingual children were marginally significantly more likely to accept the infelicitous *some* statements more often than the felicitous ones ($W = -60$, $z = -1.87$, $p = .062$), while the monolingual children displayed the reverse (expected pattern), and were more likely to accept the felicitous *some* statements than the infelicitous ones ($W = 63$, $z = 2.18$, $p = .029$). Second, at the same time, both the monolingual children and bilingual children were more likely to accept the infelicitous *some* statements than the false *all* statements (monolingual: $W = 63$, $z = 2.18$, $p = .0293$; bilingual: $W = 105$, $z = 3.28$, $p = .001$), suggesting a difference between pragmatic implicature calculation and semantic truth evaluation.

3.4. Discussion

Experiment 1 was designed to assess how Heritage Spanish-English bilingual children fare in comparison to English monolingual children in their ability to calculate scalar implicatures with *some*, implementing certain design features that we predicted would support implicature calculation by increasing the salience and relevance of alternatives. Indeed, the rate of rejection we observed in our monolingual group was strikingly similar to the rate observed by Foppolo, Guasti & Chierchia (2012) in their Experiment 3, where an *all* statement preceded a *some* statement. In the current task, neither the monolingual children nor the bilingual children calculated the "some but not all" implicature as frequently as adults. However, the Heritage bilingual children exhibited a more constrained application of *some* to sets of a particular size, a finding that points toward contrastive interpretations of lexical items that denote quantities. We suggest that the fact that this pattern was observed in the Heritage bilingual children but not in the monolingual children may indicate their attempt to establish and differentiate the meanings of multiple lexical entries across the languages they are learning. This pattern may, in turn, be relevant to the first step of the process in calculating implicatures (identifying lexical meaning). Thus, the reason why the bilingual children only slightly diverge from the monolingual children, but not significantly, but diverge from the adults is the additional time needed to sort through the lexical entries they are adding to their lexicons, which induces a slight delay in their ability to calculate these entailment-based implicatures.

Given the results of Experiment 1, the question arises as to how general the observed pattern of results is. Do Heritage bilinguals exhibit challenges calculating *all* pragmatic implicatures relative to monolingual English participants, or is the locus of the difference housed in the lexicon? Will differences wash out when a comparison of entailment-based lexical alternatives is removed from

the calculation? In addition, we might ask if bilingual children are sensitive to the assignment of upper-bounded meaning that is not induced by implicature calculation. We now move to a paradigm that is known to support implicature calculation in Spanish (Syrett et al. 2017a), where we expand the space of linguistic environments involving upper bounds, allowing us to answer these questions.

4. Experiment 2: Communicative context assessment task

4.1. Participants

Participants included the following three groups:

(i) 28 native English-speaking adults, 14 in each of two conditions (1 male, 27 females; 3 additional nonnative speakers excluded)
(ii) 25 English monolingual children (*may take* condition: 11 children [3 boys, 8 girls]; 4;02 to 5;04, M: 4;08, SD: 4.6 months; *please put* condition: 14 children [5 boys, 9 girls]; 4;01 to 5;01, M: 4;08, SD: 3.9 months) (1 additional excluded for "yes" bias)
(iii) 23 Heritage Spanish-English bilingual children (*may take* condition: 9 children [4 boys, 5 girls]; 4;04 to 5;03, M: 4;09, SD: 4.1 months; *please put* condition: 14 children [5 boys, 9 girls]; 3;09 to 5;03, M: 4;08, SD: 4.4 months) (11 additional excluded for "yes" bias, 3 excluded for attentional issues and inability to complete the task).

4.2. Materials and procedure

The experimental design consisted of a series of 17 short video clips (including three training items), lasting approximately 20 seconds each, with a brief transition between each scene, for a total session time of approximately 8 minutes. The videos were filmed using a Sony digital camera and edited and assembled in iMovie. Each scene had the same format and sequence of events. Two young women sat at a table next to each other, with some distance, facing the camera. They began the scene by waving at the camera, smiling, and audibly saying, "Hi!" in a friendly and inviting way. The woman on the right of the screen (consistently the "speaker") then held a basket at torso level above the top of the table at an angle so that the woman on the left (her right) (the "addressee") and the viewer of the video could see the contents. The speaker began by delivering an utterance to the addressee in the form of a request, to which the addressee was to respond by performing an action on the objects in the basket (or not).

4.2.1. Speaker utterance

Because our goal was to investigate the ability of bilingual and monolingual children to assign an upper bound to lexical items beyond instances of the scalar implicature associated with *some*, we structured the speaker's utterance in such a way as to include a frame that either did or did not induce an upper bound with its modal semantics, which then composed with quantificational expression that either semantically or pragmatically licensed maximality, or an upper bound. More specifically, the speaker either uttered an expression of permission/invitation (i.e., *You may take X*) or a request as a directive (i.e., *Please put X on the table*). The first should semantically induce an upper bound via the semantics of the modal *may*. While this upper limit with *may* should not be exceeded, deviation *below* this limit *is* in fact permitted pragmatically—that is, you should always be able to take less. There is suggestive evidence, however, that adults are more willing to allow for deviation below the maximal standard than children are and that children expect the maximal amount to be met (Kennedy & Syrett, under review). The *please put* directive should consistently induce a *lower* bound (and the quantificational expression *X* then perhaps induce an upper bound semantically or pragmatically).

Regardless of the frame, the quantificational expressions were consistently the same. They featured either the universal quantifier *all* in partitive form (*all of the squares*), the existential quantifier in

partitive form (*some of the books*), a definite description of one or more objects (e.g., *the hippo, the yellow dinosaurs*), or a numerical phrase (e.g., *four cars*). In this experiment, as opposed to Experiment 1, the existential quantifier *some* appeared in a partitive construction to better highlight the fact that the interpretation targeted a subset of the discourse-relevant set of items. In addition, the utterances with *some* carried a pitch accent on the lexical item, so that it was neither unstressed nor carried contrastive focus. The target trials were those where an upper bound could be assigned without reliance on the maximal semantics of the embedded lexical item alone (that is, not by the modal in the frame, which served as a control): specifically, those where she took all of the objects when (i) she was instructed to take *some*, (ii) she was instructed to take objects of a particular color, or (iii) she was instructed to take only a certain number of objects. The full set of target responses presented in Appendix B.

4.2.2. Addressee response and actions

The speaker's utterance was followed by a brief pause before the addressee responded. The addressee attempted to comply with the request or expression of permission in the speaker's utterance by putting none, some, or all of the items from the basket on the table. The addressee took the objects from the basket one by one, pausing briefly before each selection, placing the objects in front of her on the table, evenly spaced in a line, clearly away from the basket. These actions were structured in such a way as to highlight cardinality or quantity and to create suspense regarding how many and which objects would be taken. This part of each trial was filmed separately from the scene introduction, filming from the neck down of the addressee to focus on her actions, so that this part of the sequence could be spliced into the video after the speaker's utterance and be the same for each condition. Once she was done, the addressee put her hands on her lap. The two women then looked up at the camera and smiled. Following this exchange, the experimenter turned to the participant and asked, "Is what she did OK?" A sample sequence from a trial showing three still shots from the speaker's utterance, the addressee's initial action, and the completion of the addressee's response appears in Figure 7.

4.3. Predictions

We generated a set of predictions about the dependent measure (acceptance of the addressee's actions in response to the speaker's utterance) based on the following factors:

(i) the quantificational lexical items embedded in the utterance and the number of items acted upon
(ii) the semantics of the utterances (a required lower bound for the *please put* request and an obligatory upper bound for the *may take* invitation)
(iii) the pragmatics associated with the utterances (the calculation of an upper bound implicature with *please put* and the permissibility of not meeting the upper bound associated with *may take*)
(iv) the availability of an "and nothing more" particularized conversational implicature associated with the mention of particular objects and their corresponding properties.

First, we predicted that for **control items** where the speaker made a request with *all* and the addressee performed an action with all of the items, where the speaker used a numeral and the corresponding *exact cardinality* was met (|3|), or where all and only the objects named by the speaker were placed on the table by the addressee, participants in *all three groups* would *accept* the addressee's actions (i.e., say that what she did was OK). Likewise, we predicted that when a request with *all* or a numeral was framed with *please put*, and the addressee did not meet the maximal amount, all participants would *reject* the addressee's actions.

Second, we predicted that for ***may take* requests**, participants who were aware of the semantics of the modal would impose a semantically induced upper bound on all associated requests, regardless of

the nature of the lexical item, requiring no more than this maximal specified amount to be put on the table. (Of course, for requests with *all*, this amount cannot be exceeded.) It is an open question how monolingual and bilingual children will fare with such as request, since this has not been previously investigated. Whether or not both adults and children would allow for pragmatically permissible deviation *below* the upper bound for requests with *all* or a numerical item where the addressee put less than this amount or none of the items on the table is also an open question, although independent research leads us to suspect that children might not be as forgiving as adults with an addressee who did not meet the maximum amount (with numbers and *all*) (Syrett & Kennedy, under review).

Third, we predicted that for *may take* requests with *some*, participants should accept the request when some but not all of the objects were on the table. This is because the combination of the upper-bounded modal and the pragmatic upper bound with *some* should conspire to induce an unambiguous upper bound. The open question is whether this prediction holds for *both* groups of child participants. At the same time, given the results of Experiment 1 and independent research cited therein, we predicted that Heritage bilingual children might reject requests for *may take* with *some* when the addressee moves fewer than all of the objects, even though this response is semantically permissible, since the children might think that *some* should not be applied to a numerosity that can more precisely be identified by a numerical term or that *some* should apply to a certain proportion of the total amount.

Fourth, for ***please put* requests**, we predicted that participants who were aware that the modal imposes a lower bound should reject addressee actions when this lower bound is not met. However, the crucial feature is that with these items, the door is opened for a pragmatically generated upper bound (whether generalized in the case of *some* or particularized in the case of color/objects). We therefore predicted that adults would calculate the implicature in both cases but wondered whether children from both groups would. We reasoned that the calculation of the generalized upper-bounded scalar implicature with *some* might be supported in children by the within-task inclusion of felicitous and infelicitous items, weaker and stronger alternative scalemates, and numerals highlighting quantity of objects. At the same time, we predicted that *both monolingual and bilingual children* might be more inclined to accept the actions with *please put* and *some* than adults, since children notoriously calculate implicatures at a reduced rate relative to adults, with levels fluctuating based on the methodology. For the non-*some* items, we predicted success for monolingual children, given independent evidence that monolingual children robustly calculate the "and nothing more" particularized implicature (Papafragou & Tantalou 2004; Stiller, Goodman & Frank 2015; Syrett & Arunachalam 2016). An open question was whether bilingual children would pattern in the same way, given our question of whether their ability to calculate pragmatic implicatures transcends beyond scalar implicatures, which are tied to knowledge of lexical alternatives.

4.4. Results

The mean percentages of participants from the three groups accepting the addressee's actions are presented in the figures that follow. In the legends, we categorize each of the trial types according to how they fare with respect to semantics and pragmatics. If the addressee's action is not semantically permissible given the request/direction, we flag it as S✗ (semantically impermissible); whether or not it is pragmatically (im)permissible is then a moot point, since the response is not licensed semantically. If the response *is* semantically permissible, given the request/direction, we flag it as S✓ (semantically permissible). These trial types can then be further categorized according to whether they are pragmatically felicitous (P✓ for pragmatically permissible) or not (P✗ for pragmatically impermissible). Thus, the participants' responses can shed light on their appeal to semantics and/or pragmatics in their reasoning.

Because what matters is how the groups pattern with respect to each other for each of the trial types, we performed a series of pairwise comparisons between groups, as we did in Experiment 1. These were done for each of the trial types for both of the modal frames separately. In each case, the dependent

measure is the same: acceptance of the addressee's action as permissible. We expect that this will be low in trials where the action is not semantically licensed and high where it is. The main question is how the groups will perform in trials where the addressee's actions are semantically licensed but pragmatically illicit (S✓ P✗).

We begin by presenting the data for the *all* trials in section 4.4.1 (Figures 8–10), followed by the numeral control trials section 4.4.2 (Figures 11–13), all of which patterned precisely as predicted for both modal conditions and all three participant groups and therefore serve as proof-of-concept trials for the study paradigm. In the figures, for each bar, we indicate with a fraction the number of individuals patterning with the aggregate data (since there was only one such trial per participant), which thus indicates that the majority of participants in each group aligned with their respective group's trend. Following these sections, we present the data for the target trials with *some* in section 4.4.3 and the target trials with color or object kind specified (the generalized conversational implicatures and particularized conversational implicatures respectively).

4.4.1. Results for the *all* trials

We begin with the trials where the addressee took all of the items. With the *may take* speaker utterances, the semantics of the modal expression induced an upper bound, so when the addressee took all of them, this should be semantically and pragmatically allowable. Indeed, all participants accepted such actions, and there were no differences among groups (adults vs. monolingual children: $U_A = 82.5$, $z = -0.27$,

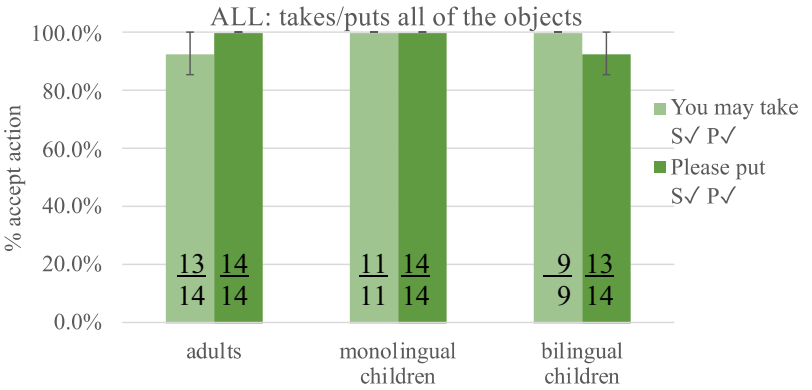


Figure 8. Rate of acceptance for the *all* trial types where the addressee took/put all of the objects in the basket for each of the three participant groups.

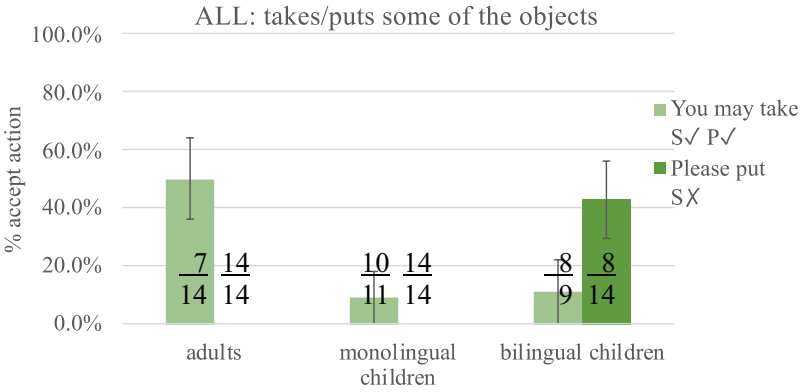


Figure 9. Rate of acceptance for the *all* trial types where the addressee took/put some of the objects in the basket for each of the three participant groups.

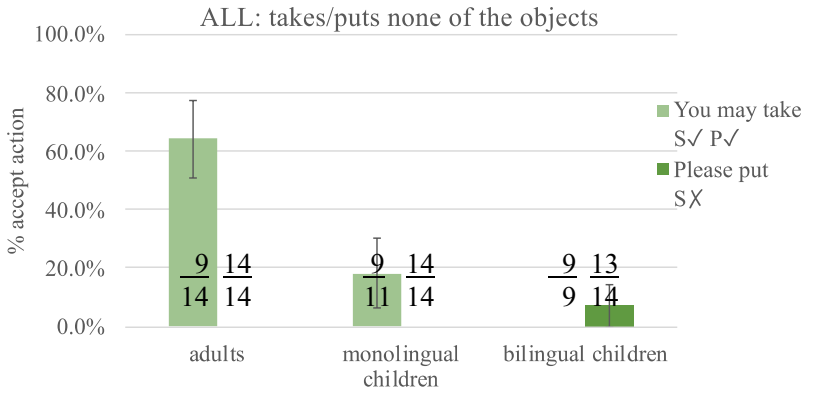


Figure 10. Rate of acceptance for the *all* trial types where the addressee took/put none of the objects in the basket for each of the three participant groups.

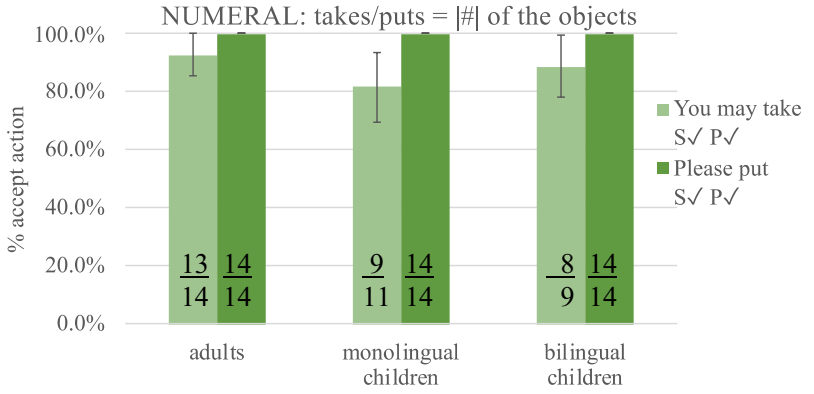


Figure 11. Rate of acceptance for the numeral trial types where the addressee took/put the exact number of the objects in the basket for each of the three participant groups.

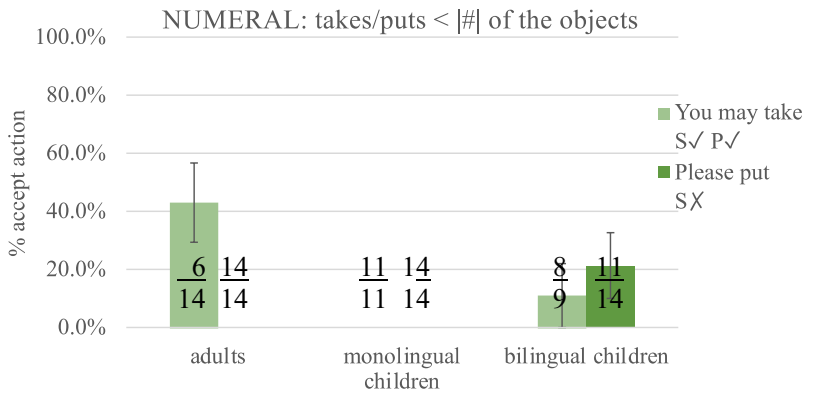


Figure 12. Rate of acceptance for the numeral trial types where the addressee took/put less than the number of the objects in the basket for each of the three participant groups.

p = .787; adults vs. bilingual children: U_A = 67.5, z = −0.25, p = .803; monolingual children vs. bilingual children: U_A = 49.5, z = 0.04, p = .968). With the *please put* requests, only the trials in which the addressee places all of the objects on the table should be permitted, and the others should be ruled out.

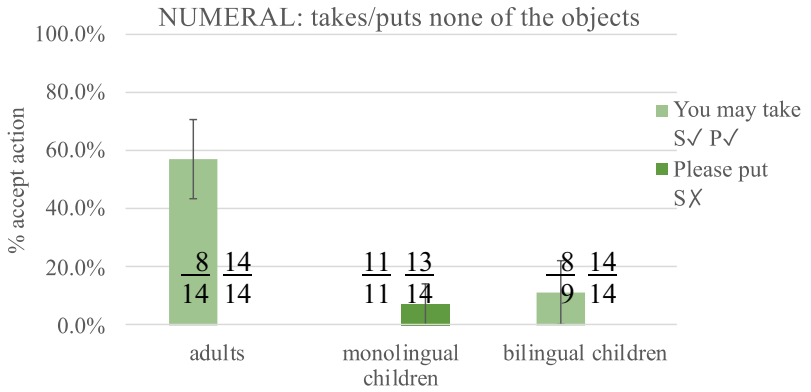


Figure 13. Rate of acceptance for the numeral trial types where the addressee took/put none of the objects in the basket for each of the three participant groups.

Indeed, when the addressee placed *all* of the objects from the basket on the table, all participants accepted the actions (adults vs. monolingual children: $U_A = 98$, $z = 0.02$, $p = .984$; adults vs. bilingual children: $U_A = 91$, $z = 0.3$, $p = .764$; monolingual children vs. bilingual children: $U_A = 91$, $z = 0.3$, $p = .764$).

By contrast, deviation below the maximal limit (where the addressee takes some but not all, or none) displays a different pattern, as shown in Figure 9 where the addressee manipulated some of the objects and Figure 10 with the addressee manipulating none of the objects. With a *may take* utterance, taking some or none is semantically licensed and pragmatically permissible but perhaps a bit odd, so we expect that while such actions are acceptable, adults might be more willing to accept them. In fact, while this was not the case when the addressee took some but not all (adults vs. monolingual children: $U_A = 45.5$, $z = 1.7$, $p = .089$; adults vs. bilingual children: $U_A = 38.5$, $z = 1.51$, $p = .131$; monolingual children vs. bilingual children: $U_A = 50.5$, $z = -0.04$, $p = .968$), it *was* the case when the addressee took none of the objects. Adults were significantly more willing to accept the addressee's actions than the bilingual children ($U_A = 22.5$, $z = 2.52$, $p = .012$), and marginally more so than the monolingual children ($U_A = 41.5$, $z = 1.92$, $p = .055$), but the child groups did not differ from each other ($U_A = 40.5$, $z = 0.65$, $p = .516$).

By contrast, with these scenarios and a *please put* utterance, the addressee's actions of putting some or none of the items on the table are *not* semantically licensed. When the addressee placed some but not all of the items on the table, neither adults not monolingual children accepted the addressee's actions (adults vs. monolingual children: $U_A = 98$, $z = 0.02$, $p = .984$), but bilingual children sometimes did, resulting in a marginally significant difference (adults vs. bilingual children: $U_A = 140$, $z = -1.91$, $p = .056$; monolingual children vs. bilingual children: $U_A = 140$, $z = -1.91$, $p = .056$). This pattern appears to be in line with the results for *todos* 'all' reported in Syrett et al. (2017a, 2017b) for comparable studies run in Spanish with the same population, suggesting that the process of sorting out lexical meaning extends beyond existential quantifiers in both languages. When the addressee placed *none* of the objects on the table with a *please put* utterance, no participants accepted the action (adults vs. monolingual children: $U_A = 98$, $z = 0.02$, $p = .984$; adults vs. bilingual children: $U_A = 105$, $z = -0.3$, $p = .764$; monolingual children vs. bilingual children: $U_A = 105$, $z = -0.3$, $p = .764$).

4.4.2. Results for the numeral trials

All participants accepted the addressee's actions when the exact specified cardinality was placed on the table, in both modal frames (*may take*: adults vs. monolingual children: $U_A = 53$, $z = -0.23$, $p = .818$; adults vs. bilingual children: $U_A = 60.5$, $z = 0.13$, $p = .897$; monolingual children vs. bilingual children: $U_A = 53$, $z = -0.23$, $p = .818$; *please put*: adults vs. monolingual children: $U_A = 98$, $z = 0.02$, $p = .984$; adults vs. bilingual children: $U_A = 98$, $z = 0.02$, $p = .984$; monolingual children vs. bilingual children: $U_A = 98$, $z = 0.02$, $p = .984$).

As with the *all* items, placing less than the upper-bounded numerical amount on the table is semantically valid, and pragmatically permissible, with the *may take* utterances. The percentage of acceptance with the numerical items, as shown in Figure 12, indicates that here too adults are more aware of this, but there is only a marginal difference between them and the monolingual children (U_A = 44, z = 1.78, p = .075), and no difference between the adults and the bilingual children (U_A = 43, z = 1.23, p = .219) or between the monolingual and bilingual children (U_A = 55, z = −0.38, p = .704). When the addressee placed some number less than the specified cardinality on the table, participants also did not accept the addressee's actions with a *please put* utterance (adults vs. monolingual children: U_A = 98, z = 0.02, p = .984; adults vs. bilingual children: U_A = 119, z = −0.94, p = .347; monolingual children vs. bilingual children: U_A = 119, z = −0.94, p = .347).

Finally, we turn to the trials where the addressee placed no objects on the table. Adults were much more likely to accept this action than the monolingual children (U_A = 33, z = 2.38, p = .017) and marginally more so than the bilingual children (U_A = 34, z = 1.8, p = .072), but the children did not differ from each other (U_A = 55, z = −0.38, p = .704). As with the *all* trials, there was no difference between the participant groups when the addressee placed no objects on the table in response to a *please put* utterance (adults vs. monolingual children: U_A = 105, z = −0.3, p = .764; adults vs. bilingual children: U_A = 98, z = 0.02, p = .984; monolingual children vs. bilingual children: U_A = 91, z = 0.3, p = .764).

4.4.3. Results for the *some* trials

We now turn to the target trials, starting with the *some* trials in which the speaker put some, but not all, of the objects on the table. The response patterns in Experiment 1 lead us to anticipate children—bilingual children in particular—might not pattern as expected with these items, rejecting the addressee actions when some but not all of the items were placed on the table. Indeed, this is the pattern we witnessed, as shown in Figure 14.

With both the *please put* and *may take* utterances, adults accepted the addressee placing some but not all of the objects on the table significantly more often than the bilingual children (*please put*: U_A = 210, z = 2.97, p = .003; *may take*: U_A = 126, z = 2.82, p = .005) and than the monolingual children (*please put*: U_A = 280, z = 1.83, p = .067, marginally significant; *may take*: U_A = 196, z = 2.18, p = .029), but children did not differ from each other (*please put*: U_A = 322, z = 1.14, p = .254; *may take*: U_A = 171, z = 0.72, p = .472).

We now move on to the target items where the upper bound indicated was exceeded by the addressee: semantically driven upper bounds with *may take*, generalized conversational implicatures (GCIs) with *some*, and particularized conversational implicatures (PCIs) with color terms, presented in Figure 15.

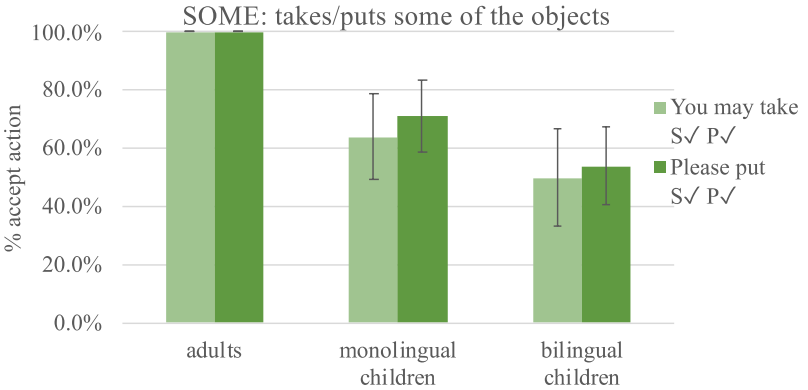


Figure 14. Rate of acceptance for the *some* trial types where the addressee took/put some of the objects in the basket for each of the three participant groups.

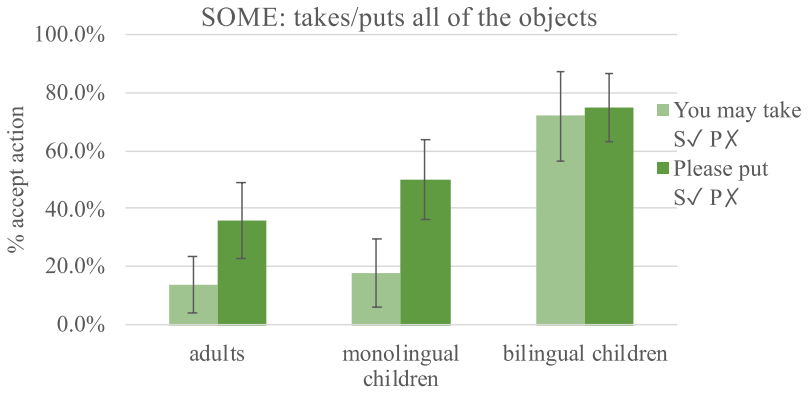


Figure 15. Rate of acceptance for the *some* trial types where the addressee took/put all of the objects in the basket for each of the three participant groups.

With the *some* items with the *may take* utterance where the upper bound was exceeded by the addressee placing *all* of the objects on the table (and was therefore semantically illicit), Heritage bilingual children were significantly more likely to accept the actions than the adults (U_A = 398, z = −3.27, p = .001) and than the monolingual children (U_A = 305, z = −2.9, p = .004), while the monolingual children did not differ from adults (U_A = 320, z = −0.22, p = .826). In fact, both adults and monolingual children were more likely to accept the addressee's actions with the *some*-all trials than the *some*-not all trials with a *may take* speaker utterance (adults: W = 300, z = 4.28, p < .0001; monolingual children: W = 75, z = 2.34, p = .0193), while the Heritage bilingual children were not (n/a).

With the *please put* utterances, which invite a pragmatic implicature of an upper bound but no semantically induced upper bound, adults were the least likely to accept the addressee's actions and differed significantly from the bilingual children (U_A = 546, z = −2.52, p = .012) but not the monolingual children (U_A = 448, z = −0.91, p = .363). The children did not differ from each other (U_A = 490, z = −1.6, p = .110). Indeed, only adults were less likely to accept the addressee's actions with a *please put* utterance when the addressee placed *all* of the items on the table than when she placed only *some* of them on the table (adults: W = 171, z = 3.71, p = .0002; monolingual children: W = 51, z = 1.31, p = .1902; bilingual children: W = −33, z = −1.66, p = .0969).

The contrast between the *some* items where the addressee acted upon some of the objects versus all of the objects and the comparison among the participant groups is further underscored by capturing the individual responses to the two items each trial type in histograms. These are presented in Figure 16 for the *some* "some" trials and in Figure 17 for the *some* "all" trials. The participants were

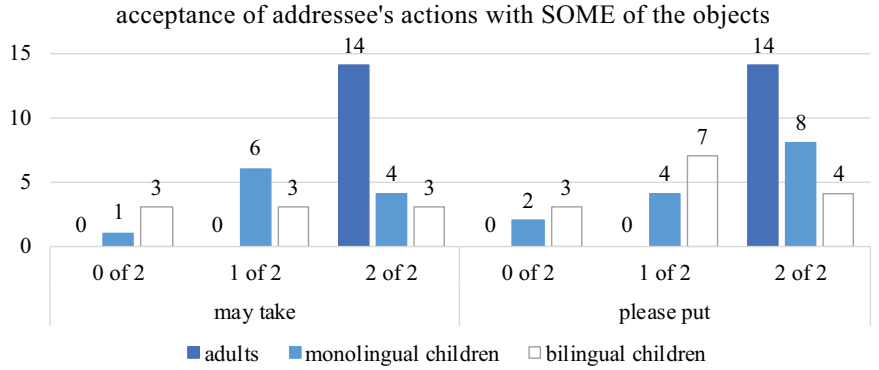


Figure 16. Individual acceptances of the addressee's actions for the *some* "some" trials.

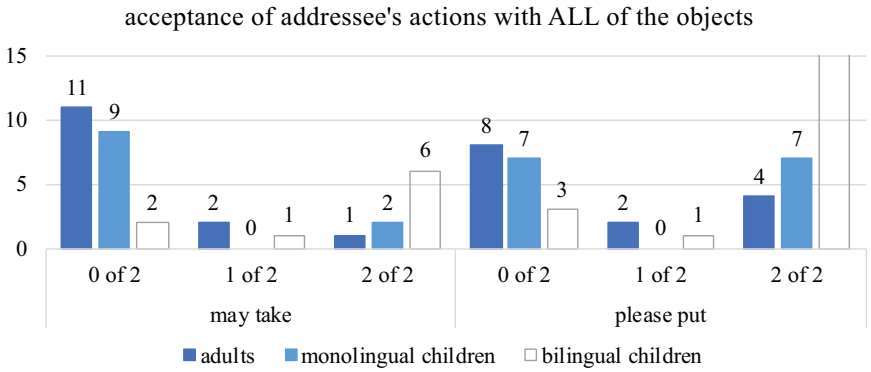


Figure 17. Individual acceptances of the addressee's actions for the *some* "all" trials.

inclined to allow the addressee to act upon some but not all of the objects, regardless of the modal in the speaker's utterance, but—with the exception of the bilingual children—they did *not* allow the addressee to act upon all of the objects with an utterance of permissibility with *may take* (reflecting a semantic upper bound), and—likewise, with the exception of the bilingual children and half of the monolingual children—they also assigned an upper bound with *please put* utterances (reflecting a pragmatic upper bound).

In contrast to these patterns with *some, all* participant groups roundly rejected the addressee's actions when she put more objects on the table than specified by the color indicated by the speaker (particularized conversational implicatures), so acceptance was low, as shown in Figure 18. They did this with the *may take* utterances (adults vs. monolingual children: $U_A = 85.5$, $z = -0.44$, $p = .660$; adults vs. bilingual children: $U_A = 65.5$, $z = -0.13$, $p = .897$; monolingual children vs. bilingual children: $U_A = 46$, $z = 0.23$, $p = .818$) and also with the *please put* utterances (adults vs. monolingual children: $U_A = 63$, $z = 1.59$, $p = .112$; adults vs. bilingual children: $U_A = 91$, $z = 0.3$, $p = .764$; monolingual children vs. bilingual children: $U_A = 126$, $z = -1.26$, $p = .208$). Thus, while the groups diverged with the upper-bounded generalized conversational implicature associated with *some*, they converged with the particularized conversational "and nothing more" implicature associated with the mention of object kinds of a particular color.

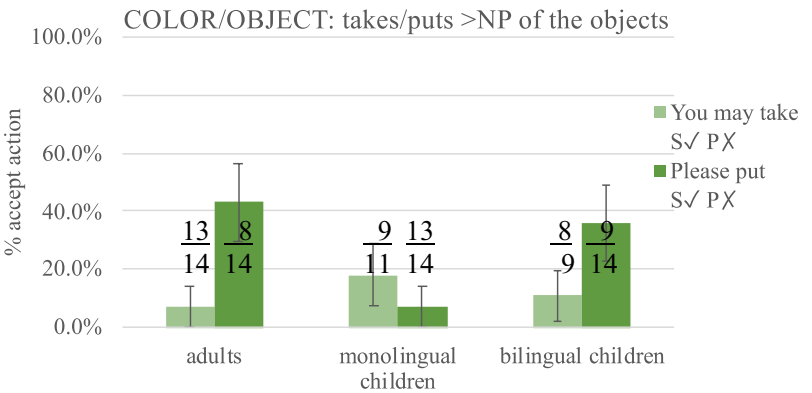


Figure 18. Rate of acceptance for the color trial types for each of the three participant groups where the addressee took/put more than the specified color and NP specified.

4.5. Discussion

The results of the communicative context task reported in Experiment 2 complement the results of the TVJT reported in Experiment 1 to demonstrate that Heritage Spanish-English bilingual children are similar to monolingual English-speaking children in their rate of calculating upper-bound implicatures with *some* but at the same time, diverge significantly more from adults, showing that they calculate implicatures at a notably reduced rate. However, they are not incapable of imposing an upper bound pragmatically, as their responses to the color items demonstrates: The Heritage bilinguals calculate a particularized conversational implicature just as monolingual children and adults do, demonstrating that the difference in behavioral patterns observed with *some* does not have to do with the pragmatic process itself. Rather, it is a result of identifying the meaning of the lexical item and comparing it with its scalemate (within a language) and alternatives acquired across languages, a possibility to which we return in the next section. That the bilingual children did not impose an upper bound with speaker's utterances framed with *may take* indicates that they have not yet mastered the semantic restrictions of this modal expression in English. Additional research would have to be done to further probe this possibility. Taken together, these lines of evidence lead us to propose that the reason for the delay in the bilingual children becoming adultlike is not in the operation of the pragmatic calculation itself (or in sensitivity to context-driven speaker meaning) but is in the acquisition of the lexicon within and across multiple languages. That they can calculate a particularized conversational implicature with the color/NP items (and reject the addressee's actions when they exceed the upper bound) but accept the addressee's actions for the pragmatically infelicitous cases with *some* and the semantically barred cases with *may take* suggests that it is in building up lexical meaning that the bilingual children initially struggle when assigning upper bounds.

5. Conclusions and general discussion

We began this article by presenting two main questions that arise regarding the ability of Spanish-English bilinguals to assign an upper bound: *How general is this ability across languages and lexical items?* and *What is the source of the pattern of implicature calculation in this population?* We presented two experiments in English, comparing their performance to English monolingual children and adults and interspersing *some* with other lexical items, including the stronger scalar alternative *all*. In Experiment 1, we assessed their ability in a TVJT design that yielded a baseline with monolinguals. In Experiment 2, we extended the research to a paradigm that highlighted communicative goals and expanded the range of lexical items and linguistic environments under investigation. The results demonstrate that in some respects, Heritage bilinguals pattern with their monolingual peers in the rate of implicature calculation with *some* but diverge from adults in that they do not reject infelicitous requests as often. They also seem relatively unaware of the semantic upper bound imposed by a *may take* invitation. However, the Heritage bilinguals align with their monolingual peers and adults in their ability to calculate particularized conversational implicatures in which the pragmatic upper bound does not depend on the meaning of individual lexical items. Finally, Heritage bilingual children also exhibited a striking pattern of rejections with *some* whereby they seemed to think it should not apply to sets of a certain size (or should apply to others).

Together, the results strongly suggest that where the Heritage bilingual children in our experiments diverged from the monolinguals in the process of calculating implicatures is in the first two steps: computing the meaning of the target sentence, given the meaning of the lexical item (i.e., *some*), and generating a set of relevant scalar alternatives whose meanings they must also know. This may come as no surprise, especially given the observations we made about bilingual development in the Introduction and the fact that Heritage bilingual children are adding more entries to their lexicon than monolingual children are and, as a consequence, face a more challenging task in development. However, our findings underscore the point that the process of calculating implicatures is just that: a process. The first step of this process involves knowing the meaning of the terms involved, and how

they compare with others. If learners appear to encounter difficulties with entailment-based scalar implicatures but sail through with particularized conversational implicatures, this is evidence that their pragmatic ability is not compromised; rather, they are in the process of establishing and differentiating lexical meanings. What then are the implications of this process in the Heritage bilinguals we tested? Here, we think the fact that they rejected felicitous *some* items is relevant.

Recall that the Heritage bilingual children often—quite surprisingly—rejected felicitous items with *some*, with justifications occasionally appealing to the fact that the cardinality of the set was "two" or "three." It seems that these children did not allow for *some* to map onto the same quantity because there was already a more specific lexical item (the numeral) that did this work. We wish to argue that the Heritage bilingual children, given the increase in input from two languages, have more work to do not only in distinguishing between the words mapping to "some" in their two languages but also between two sets of numerals and other quantifiers. Thus, this increased focus on lexical differentiation within and between the two languages —which also calls upon its own pragmatic process invoking the principle of contrast (Clark 1987)— overshadows the pragmatic comparison between scalar alternatives, which consequently surfaces later. Whereas the monolingual child is able to focus on the contrast between *some* and *all* within a quantificational scale in one lexicon, the bilingual child must make an implicit comparison between multiple quantificational terms both within and across languages and lexicons. With these lexical entries, the process is not simply a matter of translation as it is between, e.g., 'dog' and *perro* mapping to the concept DOG; *some, algunos*, and *unos* do not have lexical entries that align neatly. For bilingual children then a more involved process of pinning down the meaning assigned to quantificational lexical entries within and across languages may precede the onset of the calculation of pragmatic implicatures tied to entailment-based meanings, and this may also be why they diverged from their monolingual peers in their interpretations of *todos* and *all*, and why they, and not the monolingual children, in some cases, differed significantly from adults. The increased amount of comparison among lexical entries prolongs the process.[2]

At first glance, our claim that there is an extended process of lexical feature distinction among child bilinguals may seem surprising, given that "highly proficient" adult bilinguals (i.e., those bilinguals who acquired both languages prior to age 6 and who are exposed to both languages on a daily basis) show very little intrusion of their L1 on L2 processing and production and minimal switch cost between languages compared to "late bilinguals" (Bonfieni et al. 2019; Costa & Santesteban 2004; Poulisse & Bongaerts 1994), and they are adept at selecting and producing words from their lexicon(s). However, we are not focused in this article on proficient adults who (unlike the preschoolers we studied) have relatively stable mental lexicon(s) and many years of experience in switching between and inhibiting their L1 and L2. Rather, we are focused on a process at work in a population of Heritage bilinguals prior to age 6, as they build up their lexicon(s) and sort out lexical entries. Given the right kind of environment and support, these children could grow up to become those highly proficient bilingual adults who appear to rather easily navigate meanings, but they are far from that stage in preschool. Our work shows that during the preschool years, they spend time contrasting and fine-tuning lexical entries, taking into account their communicative role in a discourse context. The studies reported here provide a snapshot of that period of time.

We end on a final note regarding the comparison of monolinguals and Heritage bilinguals in this article. In no way is this comparison meant to privilege one population over another or encourage a view of the Heritage bilingual child as having incomplete competence (see discussion in Grosjean 1985). Rather, the inclusion of both sets of participants within a task presents us with a unique opportunity to evaluate the relationship between the development of the lexicon and pragmatic knowledge linked to a speaker's contextual and conversational goals, as we outlined. The results of the current research demonstrate that an ability to calculate implicatures depends upon both but that the two can be disentangled. Thus, our work highlights the importance of extending research on the development of pragmatics, and its relevance to

[2]We note here that there are other factors that could influence the ability of the two groups of children to calculate generalized conversational implicatures dependent on lexical meaning. The overall amount of exposure to English is, for example, one such difference. There may be others that could only be assessed through independent linguistic and nonlinguistic tasks, which we did not have the opportunity to perform with these populations.

lexical meaning, to a population of participants that is acquiring multiple languages (bilingual children) and the value of bridging L1 and L2 acquisition research.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Kristen Syrett http://orcid.org/0000-0002-3773-3035
Liliana Sanchez http://orcid.org/0000-0002-0991-5392

References

Bach, E., E. Jelinek, A. Kratzer, and B. H. Partee, Eds. 1995. *Quantification in Natural Languages*. Dordrecht: Springer. doi:10.1007/978-94-017-2817-1.

Barner, D., and A. Bachrach. 2010. "Inference and Exact Numerical Representation in Early Language Development." *Cognitive Psychology* 60:40–62. doi:10.1016/j.cogpsych.2009.06.002.

Barner, D., N. Brooks, and A. Bale. 2011. "Accessing the Unsaid: The Role of Scalar Alternatives in Children's Pragmatic Inference." *Cognition* 118:87–96. doi:10.1016/j.cognition.2010.10.010.

Barwise, J., and R. Cooper. 1981. "Generalized Quantifiers and Natural Language." *Linguistics and Philosophy* 4:159–219. doi:10.1007/BF00350139.

Bergen, L., and D. Grodner. 2012. "Speaker Knowledge Influences the Comprehension of Pragmatic Inferences." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 38:1450–60. doi:10.1037/a0027850.

Bialystok, E., G. Luk, K. F. Peets, and S. Yang. 2010. "Receptive Vocabulary Differences in Monolingual and Bilingual Children." *Bilingualism: Language and Cognition* 13:525–31. doi:10.1017/S1366728909990423.

Bonfieni, M., H. P. Branigan, M. J. Pickering, and A. Sorace. 2019. "Language Experience Modulates Bilingual Language Control: The Effect of Proficiency, Age of Acquisition, and Exposure on Language Switching." *Acta Psychologica* 193:160–70. doi:10.1016/j.actpsy.2018.11.004.

Bott, L., and I. Noveck. 2004. "Some Utterances are Underinformative: The Onset and Time Course of Scalar Inferences." *Journal of Memory and Language* 51:437–57. doi:10.1016/j.jml.2004.05.006.

Breheny, R., H. Ferguson, and N. Katsos. 2013. "Investigating the Timecourse of Accessing Conversational Implicatures during Incremental Sentence Interpretation." *Language and Cognitive Processes* 28:443–67. doi:10.1080/01690965.2011.649040.

Clark, E. V. 1987. "The Principle of Contrast: A Constraint on Language Acquisition." In *Mechanisms of Language Acquisition, Mechanisms of Language Acquisition*, edited by B. MacWhinney, 1–33. Hillsdale, NJ: Erlbaum.

Costa, A., and M. Santesteban. 2004. "Lexical Access in Bilingual Speech Production: Evidence from Language Switching in Highly Proficient Bilinguals and L2 Learners." *Journal of Memory and Language* 50:491–511. doi:10.1016/j.jml.2004.02.002.

Crain, S., and C. McKee. 1985. "The Acquisition of Structural Restrictions on Anaphora." In *Proceedings of 16th North Eastern Linguistics Society*, edited by S. Berman, J Choe and J. McDonough, 94–110. Amherst, MA: GLSA.

Crain, S., and R. Thornton. 1998. *Investigations in Universal Grammar*. Cambridge, MA: MIT Press.

Degen, J., and M. K. Tanenhaus. 2015. "Availability of Alternatives and the Processing of Scalar Implicatures: A Visual World Eye-tracking Study." *Cognitive Science* 39: 667-710. doi: 10.1111/cogs.12171.

Fabrégas, A. 2010. "An Argument for Phrasal Spell-out: Indefinites and Interrogatives in Spanish." In *Nordlyd 36.1, Special Issue on Nanosyntax*, edited by P. Svenonius, G. Ramchand, M. Starke and K. T. Taraldsen, 129–68. Tromsø: CASTL.

Flores, C., and P. Barbosa. 2014. "When Reduced Input Leads to Delayed Acquisition: A Study on the Acquisition of Clitic Placement by Portuguese Heritage Speakers." *International Journal of Bilingualism* 18:304–25. doi:10.1177/1367006912448124.

Foppolo, F., T. Guasti, and G. Chierchia. 2012. "Scalar Implicatures in Child Language: Give Children a Chance." *Language Learning and Development* 8:365–94. doi:10.1080/15475441.2011.626386.

Foppolo, Francesca, Maria Teresa Guasti, and Gennaro Chierchia "Scalar implicatures in child language: Give children a chance." *Language learning and development* 8, no. 4 (2012): 365–394.

Gentner, D., and A. B. Markman. 1997. "Structure Mapping in Analogy and Similarity." *American Psychologist* 52:45–56. doi:10.1037/0003-066X.52.1.45.

Gentner, D., and L. L. Namy. 1999. "Comparison in the Development of Categories." *Cognitive Development* 14:487–513. doi:10.1016/S0885-2014(99)00016-7.

Gentner, D., and L. L. Namy. 2006. "Analogical Processes in Language Learning." *Current Directions in Psychological Science* 15:297–301. doi:10.1111/j.1467-8721.2006.00456.x.
Goldin, M., K. Syrett, and L. Sánchez. (2020). *Perspective-taking with deictic motion verbs in Spanish Heritage child speakers: Context-dependent cross-linguistic influence at the lexical level*. Manuscript submitted for publication.
Grice, H. P. 1989. *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.
Grosjean, F. 1985. "The Bilingual as a Competent but Specific Speaker-hearer." *Journal of Multilingual and Multicultural Development* 6:467–77. doi:10.1080/01434632.1985.9994221.
Guasti, M. T., G. Chierchia, S. Crain, F. Foppolo, A. Gualmini, and L. Meroni. 2005. "Why Children and Adults Sometimes (But Not Always) Compute Implicatures." *Language and Cognitive Processes* 20:667–96. doi:10.1080/01690960444000250.
Gutiérrez-Rexach, J. 2001. "The Semantics of Spanish Plural Existential Determiners and the Dynamics of Judgment Types." *Probus* 13:113–54. doi:10.1515/prbs.13.1.113.
Horn, L.R. 1989. *A Natural History of Negation*. Chicago: University of Chicago Press.
Horn, L.R. 2004. "Implicature." In *The Handbook of Pragmatics*, edited by L. R. Horn and G. Ward, 1–28. Oxford: Blackwell Publishing Ltd.
Hulk, A., and N. Müller. 2000. "Bilingual First Language Acquisition at the Interface between Syntax and Pragmatics." *Bilingualism: Language and Cognition* 3:227–44. doi:10.1017/S1366728900000353.
Katsos, N., C. A. Roqueta, R. A. C. Estevan, and C. Cummins. 2011. "Are Children with Specific Language Impairment Competent with the Pragmatics and Logic of Quantification?" *Cognition* 119:43–57. doi:10.1016/j.cognition.2010.12.004.
Katsos, N., and D. Bishop. 2011. "Pragmatic Tolerance: Implications for the Acquisition of Informativeness and Implicature." *Cognition* 120:67–81. doi:10.1016/j.cognition.2011.02.015.
Keenan, E., and D. Paperno. 2012. *Handbook of Quantifiers in Natural Language*. Dordrecht: Springer.
Kennedy, C., and K. Syrett. (under review). Numerals denote degree quantifiers: evidence from child language. manuscript submitted for publication.
Lidz, J., and J. Musolino. 2002. "Children's Command of Quantification." *Cognition* 84:113–54. doi:10.1016/S0010-0277(02)00013-6.
Markman, A. B., and D. Gentner. 1993. "Structural Alignment during Similarity Comparisons." *Cognitive Psychology* 25:431–67. doi:10.1006/cogp.1993.1011.
Martí, L. 2007. "Restoring Indefinites to Normalcy: An Experimental Study on the Scope of Spanish *Algunos*." *Journal of Semantics* 24:1–25. doi:10.1093/jos/ffl010.
Martí, L. 2008. "The Semantics of Plural Indefinite Noun Phrases in Spanish and Portuguese." *Natural Language Semantics* 16:1–37. doi:10.1007/s11050-007-9023-x.
Martí, L. 2009. "Contextual Restrictions on Indefinites: Spanish *Algunos* Vs. *Unos*." In *Quantification, Definiteness and Nominalization*, edited by A. Giannakidou and M. Rathert, 108–32. Oxford: Oxford University Press, Oxford Studies in Theoretical Linguistics 22.
Marty, P. P., and E. Chemla. 2013. "Scalar Implicatures: Working Memory and a Comparison with *Only*." *Frontiers in Psychology: Language Sciences* 4:1–12. Article 403. doi:10.3389/fpsyg.2013.00403.
Miller, D., D. Giancaspro, M. Iverson, J. Rothman, and R. Slabakova. 2016. "Not Just *Algunos*, but Indeed *Unos* L2ers Can Acquire Scalar Implicatures in L2 Spanish." In *Language Acquisition beyond Parameters: Studies in Honour of Juana Liceras*, edited by A. Alba de la Fuente, E. Valenzuela and C. Martínez Sanz, 125–45. Amsterdam: John Benjamins.
Miller, K., C. Schmitt, H. Chang, and A. Munn. 2005. "Young Children Understand *Some* Implicatures." In *Boston University Conference on Language Development (BUCLD) 29 Proceedings*, edited by A. Brugos, M. R. Clark-Cotton and S. Ha, 389–400. Somerville, MA: Cascadilla Press.
Montrul, S. 2018. "Heritage Language Development: Connecting the Dots." *International Journal of Bilingualism* 22:530–46. doi:10.1177/1367006916654368.
Montrul, S., and N. Sánchez-Walker. 2013. "Differential Object Marking in Child and Adult Spanish Heritage Speakers." *International Journal of Bilingualism* 20:109–32.
Montrul, Silvina. 2004. "Subject and Object Expression in Spanish Heritage Speakers. A Case of Morpho-syntactic Convergence." *Bilingualism: Language and Cognition* 7:125–42. doi:10.1017/S1366728904001464.
Müller, N., and A. Hulk. 2001. "Crosslinguistic Influence in Bilingual Language Acquisition: Italian and French as Recipient Languages." *Bilingualism: Language and Cognition* 4:1–53. doi:10.1017/S1366728901000116.
Namy, L. L., and D. Gentner. 2002. "Making a Silk Purse Out of Two Sows Ears: Young Children's Use of Comparison in Category Learning." *Journal of Experimental Psychology: General* 131:5–15. doi:10.1037/0096-3445.131.1.5.
Noveck, I. A. 2001. "When Children are More Logical than Adults: Experimental Investigations of Scalar Implicature." *Cognition* 78:165–88. doi:10.1016/S0010-0277(00)00114-1.
Oller, D., B. Pearson, and A. Cobo-Lewis. 2007. "Profile Effects in Early Bilingual Language and Literacy." *Applied Psycholinguistics* 28:191–230. doi:10.1017/S0142716407070117.
Papafragou, A., and J. Musolino. 2003. "Scalar Implicatures: Experiments at the Semantics-pragmatics Interface." *Cognition* 86:253–82. doi:10.1016/S0010-0277(02)00179-8.

Papafragou, A., and N. Tantalou. 2004. "Children's Computation of Implicatures." *Language Acquisition* 12:71–82. doi:10.1207/s15327817la1201_3.
Poulisse, N., and T. Bongaerts. 1994. "First Language Use in Second Language Production." *Applied Linguistics* 15:36–57. doi:10.1093/applin/15.1.36.
Pouscoulous, N., I. Noveck, G. Politzer, and A. Bastide. 2007. "A Developmental Investigation of Processing Costs in Implicature Production." *Language Acquisition* 14:347–75. doi:10.1080/10489220701600457.
Putnam, M., and L. Sánchez. 2013. "What's so Incomplete about Incomplete Acquisition?" *Linguistic Approaches to Bilingualism* 3:476–506. doi:10.1075/lab.3.4.04put.
Rothman, J. 2007. "Heritage Speaker Competence Differences, Language Change, and Input Type: Inflected Infinitives in Heritage Brazilian Portuguese." *The International Journal of Bilingualism* 11:359–89. doi:10.1177/13670069070110040201.
Serratrice, L. 2007. "Cross-linguistic Influence in the Interpretation of Anaphoric and Cataphoric Pronouns in English–Italian Bilingual Children." *Bilingualism: Language and Cognition* 10:225–38. doi:10.1017/S1366728907003045.
Serratrice, L., A. Sorace, and S. Paoli. 2004. "Crosslinguistic Influence at the Syntax–pragmatics Interface: Subjects and Objects in English–Italian Bilingual and Monolingual Acquisition." *Bilingualism Language and Cognition* 7:183–205. doi:10.1017/S1366728904001610.
Siegal, M., A. Matsuo, and C. Pond. 2007. "Bilingualism and Cognitive Development: Evidence from Scalar Implicatures." In *Proceedings of the Eighth Tokyo Conference on Psycholinguistics*, edited by Y. Otsu, 265–80. Tokyo: Hituzi Syobo.
Siegal, M., L. Iozzi, and L. Surian. 2009. "Bilingualism and Conversational Understanding in Young Children." *Cognition* 110:115–22. doi:10.1016/j.cognition.2008.11.002.
Skordos, D., and A. Papafragou. 2016. "Children's Derivation of Scalar Implicatures: Alternatives and Relevance." *Cognition* 153:6–18. doi:10.1016/j.cognition.2016.04.006.
Slabakova, R. 2010. "Scalar Implicatures in Second Language Acquisition." *Lingua* 120:2444–62. doi:10.1016/j.lingua.2009.06.005.
Stiller, A., N. Goodman, and M. C. Frank. 2015. "Ad-hoc Implicature in Preschool Children." *Language Learning and Development* 11:176–90. doi:10.1080/15475441.2014.927328.
Syrett, K., A. Lingwall, S. Perez-Cortes, J. Austin, L. Sanchez, H. Baker, C. Germak, and A. Arias-Amaya. 2017b. "Differences between Spanish Monolingual and Spanish-English Bilingual Children in Their Calculation of Entailment-based Scalar Implicatures." *Glossa Special Issue: Acquisition of Quantification* 2:1–19.
Syrett, K., J. Austin, L. Sanchez, C. Germak, A. Lingwall, S. Perez-Cortes, A. Arias-Amaya, and H. Baker. 2017a. "The Influence of Conversational Context and the Developing Lexicon on the Calculation of Scalar Implicatures: Insights from Spanish-English Bilingual Children." *Linguistic Approaches to Bilingualism* 7:230–64. doi:10.1075/lab.14019.syr.
Syrett, K., and J. Lidz. 2011. "Competence, Performance, and the Locality of Quantifier Raising: Evidence from 4-year-old Children." *Linguistic Inquiry* 42:305–37. doi:10.1162/LING_a_00043.
Syrett, K., and S. Arunachalam. 2016. "Young Children's Developing Expectations about the Language of Events." In *Proceedings of the 40th Annual Boston University Conference on Language Development*, edited by Jennifer Scott and Deb Waughtal, 375–90. Somerville, MA: Cascadilla Press.
Thordardottir, E. 2011. "The Relationship between Bilingual Exposure and Vocabulary Development." *International Journal of Bilingualism* 15:426–45. doi:10.1177/1367006911403202.
Vargas-Tokuda, M., J. Gutiérrez-Rexach, and J. Grinstead. 2008. "Children's Comprehension of the Spanish Existential Determiners *Unos* and *Algunos*." In *Boston University Conference on Language Development (BUCLD) 32 Proceedings*, edited by H. Chan, H. Jacob and E. Kapia, 516–24. Somerville, MA: Cascadilla Press.
Veenstra, A., B. Hollebrandse, and N. Katsos. 2018. "Why Some Children Accept Under-informative Utterances: Lack of Competence or Pragmatic Tolerance?" *Pragmatics and Cognition* 24:297–313. doi:10.1075/pc.00003.vee.
Yan, S., and E. Nicoladis. 2009. "Finding *le mot juste*: Differences between Bilingual and Monolingual Children's Lexical Access in Comprehension and Production." *Bilingualism: Language and Cognition* 12:323–35. doi:10.1017/S1366728909990101.

Appendix A. Target sentences in Experiment 1

(1) **Universal Quantifier Test Items**
   (1) All of the horses ate an apple. (F; 2 of 3 did)
   (2) All of the kittens ate a cookie. (F; 3 of 4 did)
   (3) All of the penguins ate a popsicle. (T; 3 of 3 did)
   (4) All of the rabbits jumped over the log. (T; 4 of 4 did)

(2) **Existential Quantifier Test Items**
   (1) Some dogs played with a ball. (Felicitous; 3 of 4 did)
   (2) Some cows received a prize. (Felicitous; 2 of 3 did)
   (3) Some boys scored a goal. (Infelicitous; 4 of 4 did)

(4) Some girls tried on a tiara. (Infelicitous4 of 4 did)

(3) **Controls Items**

(1) Boots brought her [1 apple/3 apples].
(2) Blue drew [1 star/3 stars].
(3) Max drew [3 dogs/2 dogs].
(4) Goofy made a [blue car/red car].

Appendix B. Speaker utterances and addressee responses in Experiment 2

1. **Training**

1.1 Objects: three different plastic animals (panda, elephant, cheetah)

Speaker: Please put the elephant on the table./You may take the elephant.
Addressee: [Places only the elephant on table.]
Anticipated response: ok

1.1 Objects: three different plastic animals (cheetah, elephant, zebra)

Speaker: Please put the zebra on the table./You may take the zebra.
Addressee: [Places the cheetah on table.]
Anticipated response: not ok

1.2 Objects: four different plastic animals (zebra, cheetah, elephant, panda)

Speaker: Please put the panda and the cheetah on the table./You may take the panda and the cheetah.
Addressee: [Places all four animals, including the panda and the cheetah, on table on table.]
Anticipated response: not ok

1. **Universal Quantifier Test Items**

1.1 Objects: four spotted circles

Speaker: Please put all of the circles on the table./You may take all of the circles.
Addressee: [Places all of the circles on the table.]
Anticipated response: ok

1.1 Objects: four squares (two green, two blue, interspersed)

Speaker: Please put all of the squares on the table./You may take all of the squares.
Addressee: [Places some but not all of the squares (3/4) on the table.]
Anticipated response: not ok: *please put*; questionable: *may take*

1.2 Objects: four squares (two green, two blue, interspersed)

Speaker: Please put all of the squares on the table./You may take all of the squares.
Addressee: [Places none of the squares on the table.]
Anticipated response: not ok: *please put*; questionable: *may take*

2. **Existential Quantifier *some* Upper Bound Test Items**

2.1 Objects: four different board books

Speaker: Please put some of the books on the table./You may take some of the books.
Addressee: [Places all of the books (4/4) on the table.]
Anticipated response: not ok

1.1 Objects: four different blocks

Speaker: Please put some of the blocks on the table./You may take some of the blocks.
Addressee: [Places all of the blocks (4/4) on the table.]
Anticipated response: not ok

1.2 Objects: five markers

Speaker: Please put some of the markers on the table./You may take some of the markers.
Addressee: [Puts some but not all of the markers (2/5) on table.]
Anticipated response: ok

1.3 Objects: four different plastic animals (hippo, cheeta, panda, elephant)

Speaker: Please put some of the animals on the table./You may take some of the animals.
Addressee: [Puts some but not all of the animals (3/4, all but the elephant) on table.]
Anticipated response: ok

3. **Color Particularized Conversational Implicature Upper Bound Items**
3.1 Objects: four dinosaurs (two pink, two blue, interspersed)

Speaker: Please put the pink dinosaurs on the table./You may take the pink dinosaurs.
Addressee: [Places just the pink dinosaurs on the table.]
Anticipated response: ok

1.1 Objects: four dinosaurs (2 yellow, 2 green, interspersed)

Speaker: Please put the yellow dinosaurs on the table./You may take the yellow dinosaurs.
Addressee: [Places the yellow and the green dinosaurs on the table.]
Anticipated response: not ok

1.2 Objects: four dinosaurs (2 pink, 2 green, interspersed)

Speaker: Please put the green dinosaurs on the table./You may take the green dinosaurs.
Addressee: [Places the two pink but not the green dinosaurs on the table.]
Anticipated response: not ok

4. **Numerical Upper Bound Items**
4.1 Objects: four different cars

Speaker: Please put three cars on the table./You may take three cars.
Addressee: [Places exactly three cars on the table.]
Anticipated response: ok

1.1 Objects: four different cars

Speaker: Please put four cars on the table./You may take four cars.
Addressee: [Places only two cars on the table.]
Anticipated response: not ok: *please put*; questionable: *may take*

1.2 Objects: four different cars

Speaker: Please put two cars on the table./You may take two cars.
Addressee: [Places none of the cars on the table.]

5. **Controls Item**
5.1 Objects: three different plastic animals (cheetah, panda, hippo)

Speaker: Please put the hippo on the table./You may take the hippo.
Addressee: [Places just the hippo on the table.]
Anticipated response: ok