On the acquisition of scalar and additive inferences

Heredia Murillo, Elizabeth; van Hout, Angeliek ; Demirdache, Hamida

Published in:
Proceedings of the 46th Annual Boston University Conference on Language Development

BUCLD 46 Proceedings
To be published in 2022 by Cascadilla Press
Rights forms signed by all authors

On the Acquisition of Scalar and Additive Inferences: Evidence from Spanish *incluso* 'even' and *ni siquiera* 'not even'

Elizabeth Heredia Murillo, Angeliek van Hout and Hamida Demirdache*

1. Introduction

The meaning of sentences with *incluso* in Spanish (1a), just like *even* in English, involve three components: (1a) asserts that *Juan came* (1b) and conveys two inferences: the scalar inference in (1c) and the additive inference in (1d) (Karttunen & Peters, 1979, Rooth, 1985). The negative counterpart of *incluso* is *ni siquiera* 'not even' which likewise contributes three meaning components. (2a) thus asserts that *Juan didn't come* (2b) and conveys the scalar inference in (2c) and the additive inference in (2d). Note the impact of negation induced by *ni* in (2a), reversing the directionality of the scale associated with the focus particle: while (1a) without negation yields the least-likely scalar inference (1c), (2a) with negation yields the most-likely scalar inference (2c).

Previous studies on the acquisition of *even* in English have focused exclusively on children's comprehension of the scalar inference, leaving the additive inference uninvestigated. The present study seeks to fill this gap by investigating the acquisition of both the scalar and additive inferences. We furthermore extend the investigation of scalar focus particles to Spanish *incluso 'even'* and *ni siquiera 'not even'*. To our knowledge, this is the first experimental study to investigate comprehension of both inferences in tandem.

* Elizabeth Heredia Murillo, Nantes University / CNRS (UMR 6310) and University of Groningen, elizabeth.heredia-murillo@etu.univ-nantes.fr, Angeliek van Hout, Center for Language and Cognition Groningen, University of Groningen, a.m.h.van.hout@rug.nl, and Hamida Demirdache, Nantes University / CNRS (UMR 6310), hamida.demirdache@univ-nantes.fr.
Hamida Demirdache and Angeliek van Hout should be considered joint senior authors on this article. We thank Ana Bosnić, Miguel Santín, the Acquisition Lab and the Semantics and Cognition group in Groningen and the Nantes Linguistics Lab (LLING) for very useful meetings and feedback, Petra Schulz for insightful discussions, Jelle Kisjes for his help programming the experiment, all the children in Bolivia who participated in our experiment, and Caterina Sanchez Cordova who helped us carry it out in Bolivia. Van Hout gratefully acknowledges the Netherlands Institute for Advanced Science (NIAS) for their support.

(1) a. Incluso Juan vino.[1]
even Juan came
'Even Juan came.'
b. Juan came *Assertion*
c. Juan was the least-likely person to come. *Scalar inference*
d. At least someone other than Juan came. *Additive inference*

(2) a. Ni siquiera Juan vino.
N- SIQUIERA Juan came
'Not even Juan came.'
b. Juan did not come *Assertion*
c. Juan was the most-likely person to come. *Scalar inference*
d. At least someone other than Juan did not come. *Additive inference*

We start by briefly reviewing two earlier acquisition studies of the scalar inference induced by *even* in English (section 2). Section 3 presents the experiment and section 4 the results, and section 5 discusses the findings.

2. Theoretical Background and Objectives

Kim (2011) investigated children's understanding of affirmative and negative *even*-sentences using a "guess who game" task. The set-up involved three characters differing in size who must accomplish a given task such as reaching to grab a cookie. Participants were first provided with the contextual scale and asked to identify the endpoints of the scale (e.g., *Who is the shortest / tallest one? Can you point to him?*). They then listened to a story, at the end of which the test sentence was uttered with a picture showing who succeeded in doing the task. For the *even* condition, all the characters succeeded in the task, while for the *even-not*-condition, no character succeeded. Participants had to guess who the agent in the *even* or *even-not*- test sentence referred to (e.g., *Even Larry was (not) able to reach the cookie. Who is Larry?*). On the *even* condition, the target answer was the least-likely character, while on the

[1] Note that while English *even* freely appears in positive and negative contexts, Spanish *incluso* is not allowed under the scope of negation e.g., in object position as in (i). The distribution of *siquiera* on its own is also restricted and subject to dialectal variation: in Iberian Spanish, *siquiera* cannot appear outside the scope of negation (iia), while in Andean Spanish it can, but only under a reading equivalent to that of concessive *at least* (iib). See Alonso-Ovalle & Heredia (2020) for extensive discussion.

(i) ??/* No bebió incluso café. 'He didn't even drink coffee'
(ii) a. *Siquiera Juan vino. 'Not even Juan came.' [Iberian Spanish]
b. Siquiera Juan vino. 'At least Juan came.' [Andean Spanish]

negative condition it was the most-likely one. A limitation of this study is that there were only three items per condition which prevented an analysis of mastery and error patterns.

Children chose the target character about half of the time (45.6% for *even*; 46.7% for *even-not* sentences), showing that they had difficulties interpreting *even*. Kim (2011) identified three types of response patterns: choosing the target character, the opposite one (i.e. the most likely instead of the least likely on the *even* condition, and conversely on the *even-not* condition), or always the rightmost/leftmost character. There was a relationship between target responses and age: all children with target-like responses were age 5 while those with non-target-like responses were mainly age 4, suggesting that children below age five are not sensitive to the scalar meaning of *even*.

In contrast, Gowda et al.'s (2020), who also tested affirmative and negative *even*-sentences by adapting Kim's set-up with some modifications, found that English 4-year-olds were already sensitive to the scalar inference of *even*-sentences. They did not ask children to identify the endpoints of the scale but instead presented the characters with the relevant size attribute (i.e., as "the small/ middle /big character"). Target sentences, just as in Kim's set-up, were presented with a picture revealing the outcome, that is, affirmative target sentences were presented with a picture showing that all characters succeeded in the task, and negative target sentences with a picture showing that none had. To avoid children always picking the leftmost or rightmost character, Gowda et al. used two types of stories, which varied the likelihood of the character in succeeding to perform the action (for fitting in a hole for instance, the smallest is the most likely, while for lifting the strongest is the most likely).

Three types of responses were found: target, middle, and (like Kim (2011)) opposite. Target responses increased with age, while middle responses decreased. Opposite responses, however, were stable across age (3 to 6.). The decline of middle responses is argued to be an indication that children were not choosing randomly but instead showed more sensitivity to scalarity with age.

Both studies revealed the same interesting tendency in children's non-adult-like interpretation of *even*-sentences, namely, the selection of the opposite character. Kim (2011) explains this pattern as the result of ignoring *even* in the test sentence, and instead reasoning on the basis of world or practical knowledge to answer the target question. For example, for *Who is Larry?*, Larry is identified as the tallest animal since, as such, he should more easily reach the cookie than the others. In contrast, Gowda et al. (2020) argue that young children are sensitive to the scalar meaning of *even*, but their limited pragmatic skills lead to occasional opposite answers. It is not fully made clear, however, what exactly has not yet been acquired in the children's pragmatic system at this stage.

Summarizing so far, there are only two studies on the acquisition of English *even* and both have focused exclusively on the acquisition of its scalar inference. The present study extends the line of investigation in two ways. First, we investigate the acquisition of the scalar as well as the additive

inference and the relation between them. Second, this study aims to broaden the range of child languages by investigating Spanish learners' acquisition of the scalar focus particles *incluso* 'even' and *ni siquiera* 'not even'.

To investigate the scalar as well as the additive inferences of *incluso* 'even' and *ni siquiera* 'not even', we adapted Kim's (2011) experimental set-up. Just as in Kim's original set-up, a child has acquired the scalar inference triggered by a sentence of the form '*Even/not even* X did VP' if she correctly answered the scalar question '*Who is X*?', where the only clue for identifying X is X's ability or likelihood to perform the relevant action. On the *even* condition, the target answer is the least likely character, while on the *not even* condition, it is the most likely one. To test additivity, we added a question. A child has acquired the additive inference if she correctly answered the question about the other characters *'Did the others do VP?'*. Crucially, for it to be plausible to ask this question, the story could not reveal whether or not any of the characters succeeded in performing the requested action. Here, the current set-up differed crucially from both Kim's (2011) and Gowda's (2020) experiments where the outcome of the story was explicitly provided.

Our objective is to investigate the acquisition of both the scalar and additive inferences in tandem. We expected to find children who were target-like on both the scalar and additive inferences (like adults), as well as children who were not target-like on either inference (focus particle is not acquired). But we could, in principle, also find systematic asymmetrical non-target patterns of responses showing that one inference, but not the other, is acquired. That is, it could be that children have acquired the scalar inference, but not the additive one, or vice versa. As will be discussed in section 5, Rullmann (1997) contends that there is an entailment relation between the two inferences, arguing that the additive inference does not have an independent status, but arises indirectly as a pragmatic entailment due to the interaction of the scalar inference with the assertion. If Rullmann is correct and the scalar inference combined with the assertion does indeed entail the additive inference, then it is expected that children who have acquired the scalar inference will also have the additive inference. Crucially, one type of asymmetrical pattern (acquisition of scalar but not additive) is **not** expected to occur.

Finally, prior studies found systematic patterns of non-adult responses on the scalar inference where children selected the opposite character (Kim 2011, Gowda et al. 2020). Our study explores whether or not children who show a systematic pattern of non-target responses on the scalar inference, also show a related systematic pattern on the additive inference.

3. Method

3.1 Participants

53 5-to-7-year-old Spanish-speaking children from Cochabamba (Bolivia) were tested. In addition, 21 Spanish-speaking adults from the University of San Simon in Cochabamba-Bolivia participated as a control group. 5 children who

did not pass the controls and 6 who did not understand the task were excluded. The remaining 42 children were divided in two groups: 21 younger children (M_{age}=5;7, range 5;0 – 6;7) and 21 older children (M_{age}=7;1, range 6;8 – 7;11). Children were recruited from several schools in Cochabamba, Bolivia. A parental consent was asked for their participation. Due to the COVID-19 sanitary crisis, the experiment was adapted to test children remotely through Zoom, with the help of a research assistant in Bolivia.

3.2. Materials and Design

The "guess who game" was adapted from Kim (2011) with modifications. Two types of sentences were tested: pre-subject *incluso* (affirmative *even* (3)) and pre-subject *ni siquiera* (negative *even* (4)), with six observations per condition, and eight control items, yielding a total of 20 items. Two lists were created. The *incluso* and *ni siquiera* sentences on list 1 were turned respectively into *ni siquiera* and *incluso* sentences on list 2, with list 2 presented in the upside-down order of list 1. The experiment took about 30 minutes, and was divided in two sessions of ten minutes with a ten-minute break.

(3) Incluso Boris atrapó la pelota.
even Boris reached the ball
'Even Boris reached the ball.'
(4) Ni siquiera Boris atrapó la pelota.
N- SIQUIERA Boris reached the ball
'Not even Boris reached the ball.'

Before the test, participants were presented with three warm-up stories. They then listened to prerecorded short stories while viewing pictures of animals participating in different challenges. For each story, three characters were depicted who varied relative to a scalar attribute such as height: the tallest, the medium, and the shortest one (Stage A in Figure 1). Children were asked to identify the characters on the two extremes of the scale (*Who is the tallest character? Who is the shortest one?*). Participants could not see the outcome of the challenge because a curtain appeared (Stage B, Figure 1). However, Parrot, the announcer, who witnessed the challenge, reported the outcome with a statement containing either *incluso* 'even' or *ni siquiera* 'not even'. The curtain opened and the participant was shown the three protagonists after the challenge (Stage C in Figure 1). Children then answered three test questions.

In this set-up, each question was associated with one of the three meaning components of the scalar focus particles (cf. (1)-(2)). The first question tested children's understanding of the scalar inference (henceforth, the Scalar Q) by asking the participant to guess who the character named in the target sentence refers to (i.e. *Incluso X VP. Who is X?*). Since the experiment was carried out remotely through Zoom, children were instructed to answer by using the color

code given below each animal (red, yellow, green). The second question tested understanding of the assertion (that is, understanding of the truth conditions of the *incluso* 'even/*ni siquiera* 'not even' sentence; henceforth, the Assertion Q) by asking the child whether character *X* performed the action or not (i.e. *Did X VP?*). Positive answers were given by selecting a happy face, negative answers a sad face. Finally, the third question tested children's understanding of the additive inference (henceforth, the Additive Q) by asking the participant to guess whether the other characters performed the action or not (i.e. *Did the others VP?*). A happy or sad face coded *yes* vs. *no* answers.

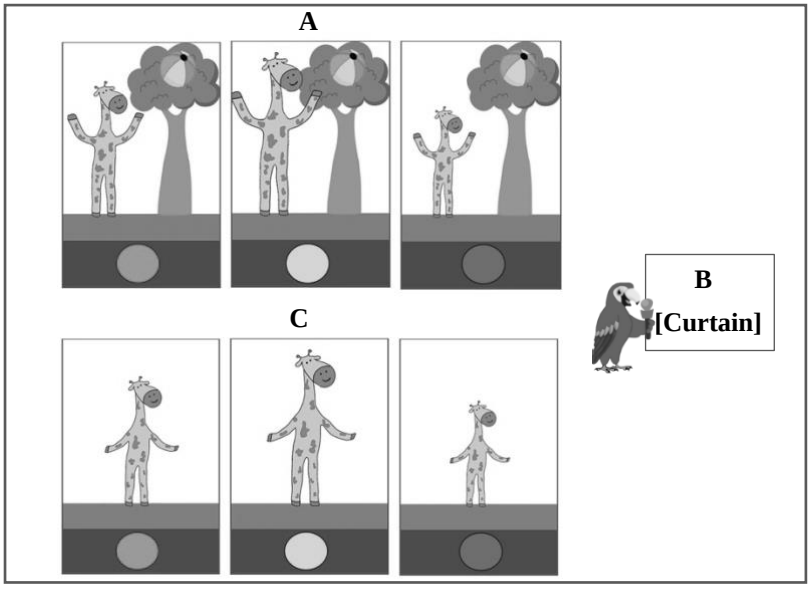

Figure 1: Example of a trial. Stage A shows the opening scene with three characters and their challenge. Stage B shows closed curtains. Stage C shows the final picture, which remained visible when the questions in (5) were asked.

(5) Questions for *Incluso / Ni siquiera Boris reached the ball.*

Scalar Q:	Who is Boris?
Assertion Q:	Did Boris reach the ball?
Additive Q:	Did the others reach the ball?

4. Results

4.1 Group results

Table 1 presents the results for each of the three meaning components for *incluso*/*ni siquiera* sentences. Adults performed as expected on all three meaning components. Both groups of children performed at ceiling with the assertion indicating that they successfully identified the truth conditions of

sentences with scalar focus particles. In contrast, the results for the scalar and additive inferences are low, suggesting difficulties with both inferences.

Table 1: Mean proportion of accuracy of Assertion, Scalar and Additive Q

	Scalar Focus Particle	Assertion	Scalar	Additive
Younger M_{Age} 5;7	*Incluso*	0.93	0.21	0.49
	Ni siquiera	0.90	0.22	0.19
Older M_{Age} 7;2	*Incluso*	0.97	0.54	0.73
	Ni siquiera	0.98	0.36	0.50
Adults	*Incluso*	1.00	0.98	1.00
	Ni siquiera	0.98	0.97	0.98

For each inference, we ran logistic mixed-effect models in R (R core team, 2014). The best model fit was determined per question by adding a variable stepwise and comparing the AIC values. For our models, Age (younger vs. older children) and Focus Particle (*incluso* vs. *ni siquiera*) were the fixed effects, and Participant and Item were the random effects, including Focus Particle as a random slope. The scalar inference responses were analyzed with a binomial *glmer* () function in the *lme4* package (Bates et al., 2015), taking the target character (of the three options) as the correct answer and the other two as incorrect. The model revealed that Age had an effect on the scalar responses. Older children gave more adult-like responses than the younger children in the *incluso* condition (β= 3.73, p=0.007), but not in the *ni siquiera* condition where the older children performed similarly to the younger ones (β= -1.79, p=0.40). For the additive responses the best model fit was again chosen by comparing the AIC values and adding variables in a stepwise manner. In this case, Age did not show any effect, but Focus Particle (*incluso* vs. *ni siquiera*) did within the younger group. Younger children performed worse with *ni siquiera* 'not even' than with *incluso* 'even' (β= -3.55, p= 0.008).

4.2 Response patterns for Scalar Q

Individual answer patterns were analyzed to determine to what extent children were consistent in giving target or reverse responses. Children were classified as Target if they selected the target character at least five out of six times in a given condition, and as Reverse, if they selected the character at the opposite end of the scale at least five out of six times. Children who did not show either of these two consistent answer patterns were classified as Other. Figure 2 presents the distribution of these response patterns for the Scalar Q.

With *incluso*, there were more older children than younger ones who systematically selected the target character. In contrast, with *ni siquiera*, both

groups of children behaved similarly: in each group only a few children were systematically target-like. Furthermore, quite a number of children systematically gave the reverse response pattern, also found in previous studies, where the character on the reverse end of the scale was consistently selected. In fact, for *incluso*-sentences, more younger children systematically selected the character on the reverse end of the scale than the target one. In contrast, only two of the older children did so too, showing that this pattern has almost disappeared in the older children. With *ni siquiera*, however, the reverse pattern was the most common one in both groups of children.

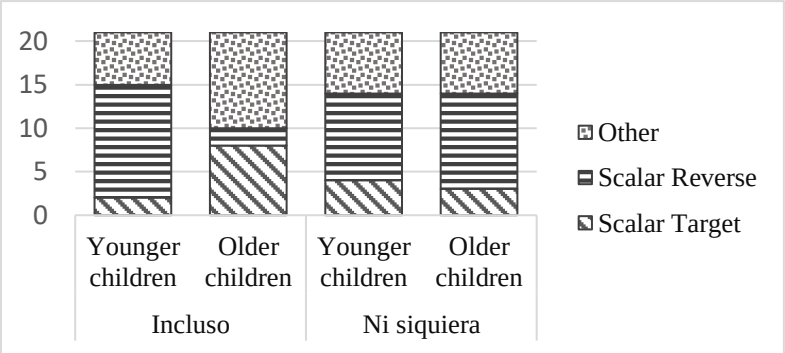

Figure 2: Number of children with a Target, Reverse or Other answer pattern

4.3 Response patterns for Scalar Q and Additive Q together

Given our objective was to investigate the acquisition of the scalar and the additive inferences in tandem, we counted how many children who systematically selected the target character for the Scalar Q, also gave the target response for the Additive Q; and conversely, how many children who systematically selected the reverse character on the Scalar Q, also gave the reverse response on the Additive Q. For *incluso,* a child who selected the least likely character (at least 5 out of 6 times) for the Scalar Q and also answered the Additive Q correctly at least 5 out of 6 times (i.e., "yes, the others did it also"), has acquired both inferences. Likewise, with *ni siquiera,* a child who systematically selected the most likely character and also answered "no, no one else did it" on the Additive Q, has acquired both inferences. Zooming in first on the children with an adult-like interpretation of the Scalar Q, Figure 3 shows the number of children who also consistently answered target-like on the Additive Q. As it turned out, all children who were systematically correct on the Scalar Q were also correct on the Additive Q (except for one younger child on the *ni siquiera* condition, in light gray in Figure 3). So, mastery of the scalar inference goes hand in hand with mastery of the additive inference.

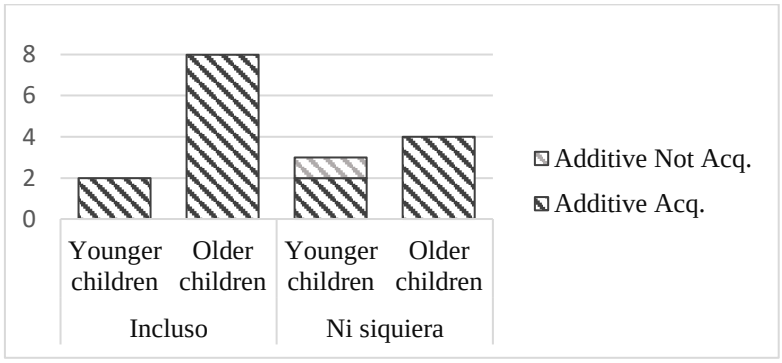

Figure 3: Number of adult-like children on scalar & additive Q

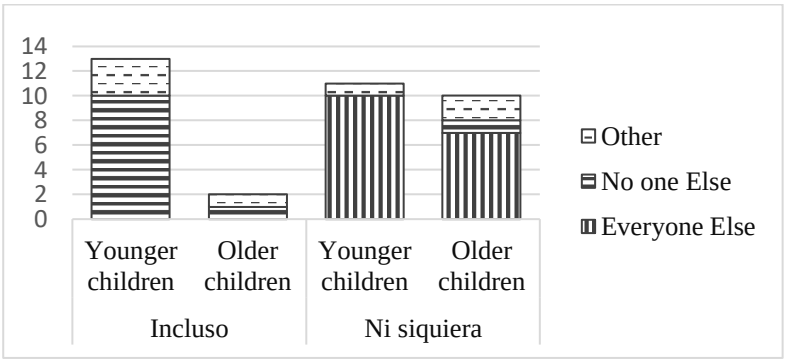

Figure 4: Number of children with reverse responses on scalar & additive Q

Figure 4 zooms in on children who consistently gave the reverse answer on the Scalar Q, that is, who consistently picked the most-likely character for *incluso* sentences, and conversely the least-likely character for *ni siquiera*. Figure 4 shows that these children did not answer the Additive Q randomly, but had a systematic pattern. For *incluso,* most children who reversed the scalar inference systematically answered, "no one else did it" and, conversely, for *ni siquiera*, "everyone else did it". In other words, children who consistently picked the character on the reverse end of the scale, also consistently reversed the additive inference. There thus appeared to be a systematic relation between the two inferences for the non-target response patterns: children who reversed the scalar inference also reversed the additive inference.

4.4 The distribution of scalar and additive response pairs

To further compare the relation between the two inferences, we considered all data points for which the Assertion Q was correctly answered, i.e., 239 responses for *incluso* 'even' and 236 responses for *ni siquiera* 'not even', and

tabulated the accuracy of the answers to the Scalar Q and Additive Q for these response pairs. Table 2 presents the distribution of these pairs for target responses on the Scalar Q, and Table 3 for reverse responses on the Scalar Q.

Table 2: Distribution of target responses on Scalar Q with related target and non-target responses on Additive Q

	Incluso **Scalar Target**	***Ni siquiera*** **Scalar target**	**Target on both**
Additive Target	84	57	141
Additive Non-Target	0	9	

Table 3: Distribution of reverse responses on Scalar Q with related reversal and non-reversal responses on additive Q

	Incluso **Scalar Reversal**	***Ni siquiera*** **Scalar Reversal**	**Reversal on both**
Additive Reversal	79	138	217
Additive Non-Reversal	23	16	

(6) Reversal pattern for Scalar Q with consistent reasoning on Assertion Q and Additive Q for "*Incluso* / *Ni siquiera X* reached the ball"

	Incluso	*Ni siquiera*
Scalar Q Who is X?	X is the tallest	X is the shortest
Assertion Q Wait, did X reach the ball?	Yes, he did it	No, he didn't do it
Additive Q Did the others reach the ball?	No, no one else	Yes, everyone else

Table 2 shows that all target responses on the Scalar Q for *incluso* and most for *ni siquiera* were also accurate on the Additive Q. There were no scalar target responses for *incluso* that were additive non-target and only nine for *ni* siquiera. Table 3 shows that most reversal responses on the Scalar Q for *incluso* as well as for *ni siquiera* were also reversed on the Additive Q.

An example of reversal response pairs is illustrated for *incluso* and *ni siquiera* in (6). We classified this answer pattern as "additive consistent" (in Tables 4 and 5). Although for these reverse responses, children incorrectly picked the tallest character on the *incluso* condition, they correctly reasoned that the other two characters are shorter and as such less likely to have

performed the requested action, leading to "*No, no one else did it*". Conversely on the *ni siquiera* condition, for these reverse responses, children incorrectly picked the shortest character, but correctly reasoned that the other two characters are taller and as such more likely to have performed the requested action, leading to, "*Yes, everyone else did it*".

There was a third answer pattern that has not yet been reported. It was attested for only 53 responses out of 239 for *incluso*, and 17 out of 236 for *ni siquiera*. Here, the medium character was selected for the Scalar Q. For this response pattern the corresponding *yes/no*-response to the Additive Q was hard to interpret. *Yes* and *no* answers sometimes received the same justification, as the answers in (7) with spontaneously provided justifications show. Since children answered either *yes* or *no* with the same reasoning, we could only analyze yes/no responses with a justification as consistent or not. Responses without justification were labeled as uninterpretable in Table 4.

(7) Medium pattern for "*Incluso / Ni siquiera X* reached the ball"

	Incluso	*Ni siquiera*
Scalar Q Who is X?	X is the medium	X is the medium
Assertion Q Wait, did X reach the ball?	Yes, he did it	No, he didn't it
Additive Q Did the others reach the ball?	No, the shortest one no, and the tallest one yes	Yes, the tallest reached it, the shortest no

Table 4: Distribution of scalar medium responses with consistent and non-consistent responses to Additive Q

	Incluso **Scalar Medium**	***Ni siquiera*** **Scalar Medium**
Additive Consistent	24	9
Additive Non-consistent	2	3
Uninterpretable	27	6

Although some children incorrectly picked the medium character on the Scalar Q, Table 4 shows that their reasoning about the other characters on the Additive Q was consistent in most cases. For *incluso* children assumed that the shortest character was less likely to have reached the ball, while the taller one was more likely to have done so, and said that the tallest one also did it, but the shortest one did not. And conversely for *ni siquiera*.

In summary, Table 5 presents the consistency of the additive responses for children who consistently chose the target, reverse or medium character, combining the two focus particles. Table 5 shows that the vast majority

involved additive responses that were consistent with the scalar response. Interestingly, although there were many non-adult-like responses for the Scalar Q, these were for the most part paired with consistent responses on the Additive Q. We argue below that these findings show children's sensitivity to a systematic relation between the scalar and additive inferences.

Table 5: Scalar and additive responses for *incluso* and *ni siquiera* together

	Additive consistent	**Additive inconsistent**
Scalar Target	141	9
Scalar Reversal	217	39
Scalar Medium	33	3
Total	391	51

5. Discussion

Our research objective was to investigate in tandem the acquisition of both the scalar and additive inferences induced by *incluso* and *ni siquiera*, extending the empirical scope of prior studies on the acquisition of *even* in English, which have focused exclusively on the scalar inference. Age played a role for *incluso*: there were more older children who were adult-like on the scalar inference in comparison to the younger children, and vice versa, there were more younger children who gave a reverse interpretation for *incluso* 'even' in comparison to the older children (Figure 2). There was no similar age effect for *ni siquiera*, however. Another finding that we share with the two previous acquisition studies on English *even*, was that children with non-adult-like comprehension tended to systematically give reverse answers. As for our main focus, the relation between the scalar and additive inference, our study revealed two major findings. First, mastery of the scalar inference entails mastery of the additive inference: children who have acquired the scalar inference (responded correctly on the Scalar Q at least five out of six times) have also acquired the additive inference (Figure 3). Second, the comparison of response pairs for the scalar and additive inference indicates that there was a dependency between these two inferences: reverse responses on the Scalar Q were systematically given together with reverse responses on the Additive Q (Table 3). More generally, the additive inference was in most cases consistent with the scalar inference, independent of the accuracy of the scalar inference (Table 5). We now offer an account for these findings.

Our study was explicitly designed to investigate the relation between the scalar and the additive inference by asking questions about both, different from earlier studies that focused on the scalar inference only. Moreover, using a design with six items per condition allowed us to analyze mastery of the two inferences together. Interestingly, a dependency between the two inferences

emerged in the following two findings. All children who behaved adult-like with the scalar inference also behaved adult-like with the additive inference (Figure 3). Moreover, looking at the 475 responses where the Assertion Q was correctly answered, a correct response to the Scalar Q always implied a correct response to the Additive Q (Table 2). We take this interpretation pattern across both inferences to support Rullmann's (1997) claim that the additive inference of *even* is entailed by the scalar inference combined with the assertion. In particular, Rullmann argues that the additive inference associated with *even* in a sentence such as *Even Boris reached the ball* (which for Rullmann is the existential presupposition that there is at least another *x* such that *x* reached the ball) "does not have independent status, but instead arises indirectly as a pragmatic entailment of the scalar presupposition of *even* combined with the assertion expressed by the sentence in which *even* occurs…" (Rullmann, 1997: p. 58). A child who correctly answered the assertion Q for the sentence *Incluso Boris reached the ball*, understood that the proposition that *Boris reached the ball* is true. If, furthermore, she correctly answered the Scalar Q, she also knew that the proposition that *Boris reached the ball* was the least likely of the alternative propositions. Now, on the basis of this, this child, as Rullmann himself (p. 59) puts it, "will most certainly be inclined to conclude that all the more likely propositions in the set of alternatives are also true, on the basis of the default assumption that, if p is less likely than q and p is true, then (in all likelihood) q is also true." This way of reasoning thus gives the child the answer that *Everyone else did it also* for free.

Conversely, take a child who correctly answered both the additive Q and the Scalar Q for the sentence *Ni siquiera Boris reached the ball*. Since she correctly answered the Assertion Q, she has understood that the proposition that *Boris reached the ball* was false, and since she correctly answered the Scalar Q, she also knew that the proposition that *Boris reached the ball* is the most likely of the alternative propositions. This child is then warranted to conclude that all less likely alternative propositions are false too, thus answering that *No one else did it either*.

Importantly, the line of reasoning just sketched also explains the second major finding, namely, that children who behaved non-adult-like with the scalar inference (by giving either a reversal or a medium response) gave additive-consistent inferences (Tables 3 and 4). As the non-target answer patterns illustrated in (6)-(7) show, although these children incorrectly answered the Scalar Q, they nonetheless correctly reasoned that the shorter character(s) was (or were) less likely to have performed the challenge, while the taller character(s) was (or were) more likely to have also performed it.

A question remains, however: why do children pick the reverse character in the first place? It could be, as Kim (2011) suggests, that they ignore the focus particle, relying on world knowledge to associate the ability of the characters under discussion with their likelihood to succeed in the task. But children's answer patterns for both inferences combined showed that they nevertheless have scalar knowledge: reversal children know that there is a scale and

understand its directionality. More importantly, on the basis of the least/most likely inferences they draw, they correctly infer the other characters' likelihood to have also succeeded, given that their answers on the Additive Q were consistent with their answers on the Scalar Q. This was independent of whether their scalar answer was target-like or not. This dependency points to children's sensitivity to scalarity and makes Kim's explanation unlikely.

The question remains though: why do younger children systematically start reasoning from the opposite end of the scale. We close this paper by suggesting an alternative explanation: children recognize that the particle has scalar force, but have not yet acquired the full range of meanings that scalar particles have, and thus assign the particle a default scalar meaning: that of ONLY, the covert exhaustivity operator associated with scalar items (Chierchia et al. 2011). That is to say, on the reversal pattern, the child interprets *Incluso Boris reached the ball* as *Only Boris reached the ball*, reasoning that if *Boris* is the only one to reach the ball, he is the most likely (here, tallest) character (Scalar Q) and that no one else reached it (Additive Q). Conversely, the child interprets *Ni siquiera Boris reached the ball* as *Only Boris didn't reach the ball* reasoning that if *Boris* is the only one not to reach the ball, he must be the least-likely (here, shortest) character and that everyone else reached the ball.

References

Alonso-Ovalle, Luis, Heredia-Murillo, Elizabeth. (2020). Concessive scalar items: The case of Spanish *Siquiera*. Presentation at SENSUS: Constructing Meaning in Romance, UMass.

Bates, Douglas, Mächler, Martin, Bolker, Ben and Walker, Steve. (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software 67(1).1-48.

Chierchia, Gennaro, Danny Fox & Benjamin Spector. 2011. The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In Paul Portner, Claudia Maienborn & Klaus von Heusinger (eds.), Handbook of semantics, Mouton de Gruyter.

Gowda, Yadav, Newman, Elise, Rosenstein, Leo, and Hackl, Martin. (2020). Scalar inferences in the acquisition of *even*. *Frontiers in Communication*. Vol.5. https://doi.org/10.3389/fcomm.2020.593634

Karttunen, Lauri and Peters, Stanley. (1979). Conventional implicature. In Choon-Kyu Oh and David A. Dinneen (Eds.), *Syntax and Semantics* Vol.11, pp. 1-56.

Kim, SoYoung. (2011). *Focus particles at syntactic, semantic, and pragmatic interfaces: the acquisition of only and even in English*. Doctoral Dissertation, University of Hawaii.

Rullmann, Hotze. (1997). *Even*, polarity, and scope. In Martha Gibson, Grace Wiebe, and Gary Libben (Eds.), *Papers in Experimental and Theoretical Linguistics*. Vol. 4, pp. 40-64.

R Core Team. (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.

Rooth, Mats. (1985). *Association with Focus*. Doctoral Dissertation, University of Massachusetts, Amherst.