

## On the acquisition of *either* and *too*

Naomi Francis, Shuli Jones, Leo Rosenstein, & Martin Hackl\*

**Abstract.** This paper presents an experimental investigation of how English-learning children acquire the additive discourse particles *either* and *too*. In the target grammar these items exhibit near-complementary distribution conditioned on the polarity of their host sentence. The path leading to that grammar appears to be rather intricate. We present comprehension data showing that for an extended period of time (ages 3–5) learners find both items acceptable in both polarity environments, exhibiting only a weak adult-like tendency of preferring *either* in negative and *too* in positive sentences. At age 6 their grammar appears categorical with respect to *either* in that they no longer tolerate it in positive sentences while still exhibiting only a weak dispreference for *too* in negative environments. These findings are even more striking in the context of production data. We find that child-directed speech is essentially categorical, providing unambiguous evidence for the adult grammar. Moreover, we find essentially categorical, adult-like use of *either* and *too* in child production from the earliest stage of development. These observations raise a number of challenges for theories of *either* and *too* and for approaches to learning focus particles more generally. Perhaps most strikingly, the protracted insensitivity of the learner’s grammar to accumulation of unambiguous evidence constitutes a novel argument from the abundance of evidence for encapsulated learning.

**Keywords.** Additivity; focus particles; polarity sensitivity; L1 acquisition

**1. Introduction.** This paper presents an experimental investigation of how English-learning children acquire the additive focus particles *either* and *too*. Just like other non-scalar, additive discourse particles (items such as *also*, *additionally*, *as well*, . . .), their basic function is to signal that the sentence they attach to is part of a sequence of sentences that constitutes a (potentially partial) answer strategy for addressing a question under discussion (see e.g. Beaver & Clark 2008). What is special about *either* and *too* is that they appear to be designated to mark, respectively, negative and positive sequences. That is, they exhibit (near) complementary distribution, which is conditioned on the polarity of the sentence they accompany. This is exemplified in (1).

- (1) a. Sam is eating cake. Sam is eating ice cream **too/\*either**.  
 b. Sam isn’t eating cake. Sam isn’t eating ice cream **either/\*?too**.

One approach to capturing the divergent polarity sensitivity of *either* and *too* while maintaining their shared discourse functionality is to assume that they trigger the same canonical additive presupposition (while being assertorically inert), as shown in (2), and to stipulate that they are

---

\*We are grateful to Yadav Gowda for his help with building the scripts that extracted and coded our CHILDES data, to Anya Keomurjian, Abena Peasah, and Mika Thakkar for constructing stimuli for the adult experiment, and to Steve Worthington for stats guidance. This research was supported by the Social Sciences and Humanities Research Council of Canada (Doctoral Fellowship to the first author). Authors: Naomi Francis, University of Oslo ([n.c.francis@iln.uio.no](mailto:n.c.francis@iln.uio.no)), Shuli Jones, Massachusetts Institute of Technology, Leo Rosenstein, Massachusetts Institute of Technology, & Martin Hackl, Massachusetts Institute of Technology.

(morpho-) syntactically specified for specific environments. *Either* would be marked to occur with negative sentences, taking necessarily wide scope over the negative operator, while *too* would be confined to positive sentences either via designated marking or as the result of competition with its more marked counterpart *either*.<sup>1</sup>

$$(2) \llbracket \textit{either/too} \rrbracket(p) = \lambda w: \exists q \in \text{ALT}(p) \text{ s.t. } q(w)=1. p(w)=1$$

Rullmann (2003) provides compelling arguments against such a treatment, however, and instead proposes an analysis on which *either* and *too* are not synonymous. Concretely, while *too* receives a treatment like any run-of-the-mill additive particle, *either* is assumed to trigger an “anti-additive” presupposition demanding that there is an alternative to its host clause that is false, (3).<sup>2</sup>

$$(3) \llbracket \textit{either} \rrbracket(p) = \lambda w: \exists q \in \text{ALT}(p) \text{ s.t. } q(w)=0. p(w)=1$$

(3) by itself does not prevent *either* from being attached to positive sentences, of course. To make that impossible Rullmann (2003) needs to assume, moreover, that *either* is marked as an NPI, hence confined to occur in the scope of a suitable negative operator.<sup>3</sup> This ensures that the host clause of *either* is negative. Moreover, since *either* presupposes that there is an alternative (antecedent) proposition to *p* that is false results in *either* being able to serve as a signal for a sequence of negated propositions forming an answer strategy. The fact that *too* can mark a positive sequence follows directly. To account for the fact the *too* cannot mark a negated sequence requires, however, a further assumption. Rullmann (2003) proposes that *too* is preferentially attached low in the structure, lower, importantly, than the attachment site of sentence negation. Attaching *too* below negation in (1-b) triggers a positive presupposition which the context does not satisfy, however, giving rise to infelicity.<sup>4</sup> By the same token, this proposal predicts that speakers might find (1-b) with *too* acceptable to the extent to which they are able to access a parse with high attachment of *too*. In that case the host clause happens to express a negated proposition which is licit as long as the context provides a true focus alternative to it. Since this is the case in (1-b) wide scope attachment of *too* would produce a felicitous sequence, offering a perspective on why the unacceptability in (1-b) is typically judged less severe than the unacceptability of (1-a).

Although the proposal in Rullmann (2003) is not without problems and there are competing theories in the literature, notably the alternatives-based analysis of Ahn (2015), we think that it provides a useful framework within which to explore questions about how *either* and *too* might be acquired. In particular, we are interested in finding out whether learners initially adopt the most basic additive semantics for *either* as in (2). Evidence for such a stage in development

<sup>1</sup>This could be implemented via morphological suppletion which is conditioned on the syntactic position of the additive particle, as in König (1991), or via some form of feature probing mechanism.

<sup>2</sup>How exactly alternative propositions are generated is an open question in the literature. For our purpose it will be sufficient to follow Beaver & Clark (2008) and assume that the the focus marking inside the host clause reflects which QuD the sequence is meant to address.

<sup>3</sup>See Rullmann (2003) for details. Since the question of which negative operators can license *either* is not central to our purpose, we will organize our discussion around sentence negation and abstract away from any complications that arise when considering other licensors.

<sup>4</sup>However, in case there is a positive alternative provided by the context as in (i) *too* is predicted to be licensed.

- (i) Mary ate the lasagna, but she couldn't eat the spaghetti, too. (Rullmann 2003; 389)

would come in the form of *either* being fully interchangeable with *too* across positive and negative environments. If such a stage can be identified, we would then hope to trace the time course and the evidence it takes to learn their near-complementary distribution. Of particular interest is whether learners go through a stage predicted by the proposal that Rullmann (2003) ends up rejecting. If so, these learners should expect *either* and *too* to exhibit complementary distribution in basic sentences such as (1) but they would not yet know that *either* is, in fact, an NPI.

To investigate these questions we first conducted a comprehension experiment employing a comparative felicity judgment task. This task is well-suited for probing directly whether speakers prefer *either* over *too* in positive environments and *too* over *either* in negative environments. We find that children do not behave in accordance with Rullmann (2003) or the simpler proposal rejected therein. Indeed, our findings are puzzling under all theories of *either* and *too* we are aware of. We then conducted a corpus study on child-directed speech to see whether their comprehension could be understood in terms of the data available to the learner being noisy, uninformative or even misleading. Interestingly, our corpus study revealed quite the opposite: child-directed speech is in effect categorical with respect to the polarity-dependent distribution of *either* and *too*. Finally, we conducted a corpus study on child speech to examine whether their productions of *either* and *too* reflect their comprehension system. We find, however, that child speech is in fact adult-like even at the earliest stage in development. This puts our comprehension data in even starker relief.

## 2. Child experiment.

2.1. METHOD. To assess children’s understanding of the polarity-conditioned distribution of *either* and *too*, we employed a comparative felicity judgment task (Chierchia et al. 2001, Foppolo et al. 2012). This task presents participants with a forced choice between *either* and *too* in a given environment. For positive sentences, we predict that adults will always choose *too* over *either*, since there is no way for *either* to be licensed there. For negative sentences, a prediction based on Rullmann (2003) is more nuanced; speakers who cannot access a wide-scope LF for *too* should have a categorical preference for *either* over *too*. For speakers who can access a wide-scope LF for *too*, the rate of selecting *too* will depend on them judging the *either* competitor to be nevertheless better. If they are roughly comparable, we might expect non-categorical selection of both items; if they are significantly different, as Rullmann suggests, we might expect a fully categorical pattern.<sup>5</sup>

In each trial, we used a laptop to present a scene like those exemplified in Figure 1 to the participant. One experimenter, playing the role of narrator, introduced the scene and provided an appropriate antecedent for the additive particle. Then two puppets provided competing sentences completing the description; the puppets’ utterances were identical except for the additive item used (*either* vs. *too*). Children were then asked to indicate which puppet “said it better”.

The experiment consisted of eight target items and four filler items, balanced for polarity. Positive and negative versions of each target item were created. For each target item, half of the participants saw the positive version and half saw the negative version; the fillers were seen by all. The ordering of positive and negative trials, the order in which the puppets spoke, and the additive particle used by the first speaker were pseudo-randomized to avoid unwanted patterns.

We recruited 57 participants aged 3–7 years in the Boston area at the Museum of Science and

<sup>5</sup>Since this is an empirical question, we ran an adult control study reported in Section 3.

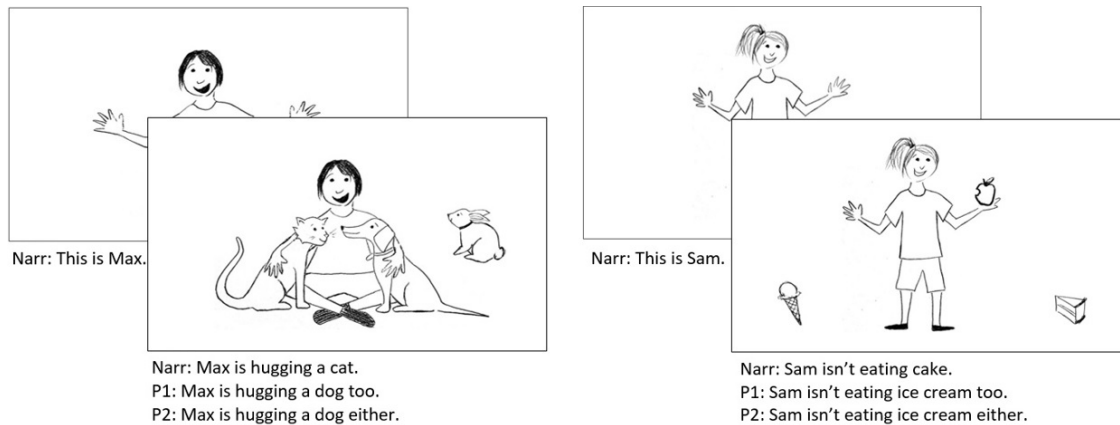


Figure 1: Examples of positive (L) and negative (R) target items

local daycares. Of these, 47 L1 English acquirers completed the experiment: nine 3-year-olds, fourteen 4-year-olds, fifteen 5-year-olds, eight 6-year-olds, and one 7-year-old. The 7-year-old was excluded from our analysis, leaving data from 46 participants. A parent or guardian provided informed consent for each child, and each child received a sticker in thanks for their participation.

2.2. RESULTS. The mean rates at which children selected each item in each environment are shown in Figure 2.

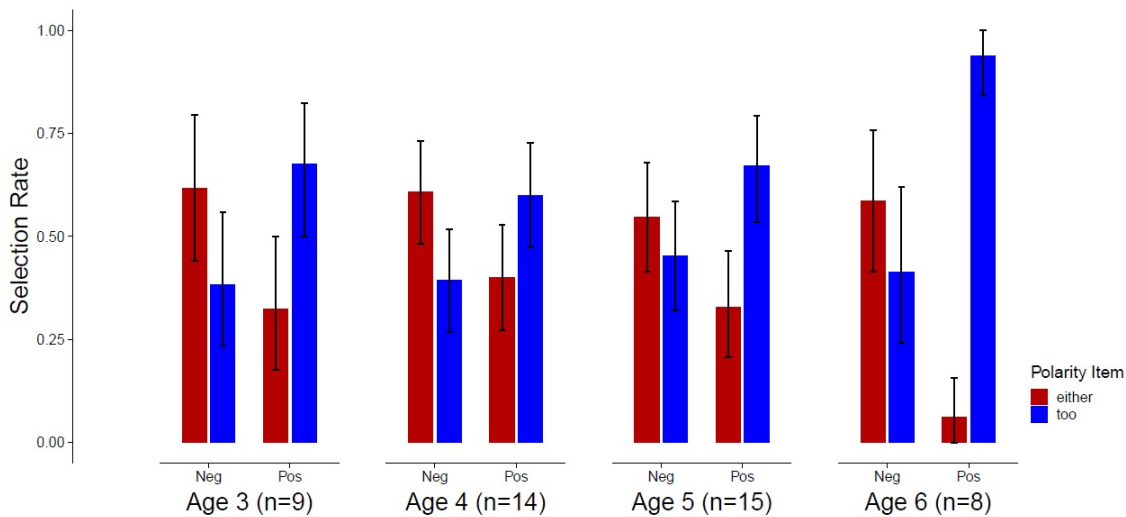


Figure 2: Mean rate of *either/too* selection, by polarity environment and age

To examine the distribution of our responses we constructed a linear mixed effects logit model in R (R Core Team 2019) using the lme4 package (Bates et al. 2015) with fixed effects of polarity (*pos* vs. *neg*) and age (3–5 vs. 6). This model showed main effects of polarity and age; children were more likely to select *either* in negative sentences than in positive sentences ( $\Pr(>|z|) = 0.000171$  \*\*\*), and 6-year-olds were less likely to select *either* overall ( $\Pr(>|z|) = 0.009378$  \*\*). The model also revealed a significant interaction between polarity and age ( $\Pr(>|z|) = 0.036657$  \*); the contribution of this interaction was confirmed with a likelihood ratio test ( $p < 0.05$  \*).

2.3. DISCUSSION. We take the comparative felicity judgment task to reflect comprehension, as it tests participants' responses to utterances of which they are not the author. The results reported above show that children are not adult-like in their treatment of *either* and *too*. From the ages of 3 to 5, children exhibit a non-categorical response pattern, accepting both items in both environments. However, their behavior is not random; they select *either* more often than *too* in negative environments and *too* more often than *either* in positive environments. That is, they trend toward the adult grammar. Importantly, this trending pattern does not reflect a population-level split, with some children being categorical in one direction and others being categorical in the other; the non-categorical trending pattern is attested within subjects.

That this behavior is non-adult-like is transparent for *either*, but the interpretation of the *too* data is more involved. Informal fieldwork with native English speakers and an adult felicity rating study reported in Section 3 suggest that *too* in negative sentences is judged to be quite infelicitous. A small (n=4) replication of the above experiment with adults confirmed that they deliver categorical judgments on this task. Moreover, some children (n=8) explicitly indicated that both items were equally good in the sentences they were presented with, answering "both" or "neither" to the experimenter's question "Who said it better?". We take this to indicate that children's treatment of *too* in our task is non-adult-like.

The non-categorical trending behavior continues virtually unchanged until the age of 6, when children suddenly stop selecting *either* in positive sentences. This is presumably responsible for the statistical effect of age, which indicated that 6-year-olds are less likely to accept *either* overall.

The patterns of behavior exhibited by the 3–5-year-olds on the one hand and the 6-year-olds on the other both raise interesting questions. What knowledge state does the 3–5-year-olds' behavior reflect? Their different response patterns in positive and negative environments suggest they are not ignorant of the role of polarity in conditioning these items' distribution, but neither do they faithfully reproduce its categorical nature. The stability of this trending behavior is also striking; the pattern of responses displayed by the 3-year-olds in our sample is virtually identical to that displayed by the 5-year-olds, despite the fact that the latter group has had two more years' worth of exposure to the distribution of these items in the input. This suggests that whatever is responsible for the change in behavior that occurs between ages 5 and 6 is not merely the result of the steady accumulation of evidence; this constitutes a novel argument for encapsulated learning.

How should we characterize the learning pattern displayed here? The early grammar of additivity seems to have compatibility of *either* and *too* as one of its core properties. At first glance, this might suggest that children initially hypothesize that these items share a basic additive semantics (i.e., (2)). However, on such a view it is difficult to explain the trending behavior, and equally difficult to explain its stability. Why it is so difficult for children to master the restricted distribution of these items? This fact becomes even more striking once we understand the input is robustly categorical (see Section 4). From the perspective of Rullmann's (2003) account, we would ask what it takes to move from a simple additive semantics for *either* (requiring a true antecedent, as in (2)) to the non-standard one (requiring a false antecedent, as in (3)).<sup>6</sup> The answer offered by this framework appears simple enough at first blush: children must realize that *either* is an NPI. What is left open, however, is why a wide-scope theory of *either* is not entertained at all, not even at an

<sup>6</sup>If we take our data seriously, it appears that the children we tested never entertain a simple König-style approach.

intermediate state for which it is plausible to assume that learners have not yet encountered data that decisively argue for the NPI-treatment of *either*.

### 3. Adult control.

3.1. METHOD. To precisify what the target grammar of *too* allows, we ran a variation of the experiment described in the previous section on an adult population. Our goal was to determine whether *too* is acceptable in negative sentences with negative antecedents on the adult grammar. A direct replication of the child experiment may be too blunt a tool to investigate this question, as it does not distinguish between what is dispreferred because it is infelicitous and what is dispreferred because it is merely less felicitous than its competitor. To give participants an opportunity to indicate a difference between what is allowed but dispreferred by their grammar vs. what is not allowed, we conducted a felicity rating experiment.

Our adult study was conducted online. Participants were shown scenes similar to those used in the child experiment and asked to rate the naturalness of positive and negative sentences with *either* and *too* (each following an appropriate antecedent) on a Likert scale with values ranging from 1 (“extremely unnatural”) to 7 (“perfectly natural”).

The experiment consisted of 32 Latin-squared target items and 40 filler items, both balanced for polarity. In each trial the scene and antecedent sentence appeared on the screen; to ensure that participants were actively engaged in the task and had time to read the text, they had to click their mouse or press a key after a 2000 ms embargo period to advance to the target sentence and then to the rating scale. Participants were rejected if they i) failed to complete their rating within 7000 ms of the scale appearing in more than 10% of trials, or ii) rated more than two of eight ungrammatical benchmarking fillers above a 2.

We used Ibex Farm (Drummond 2018) and PennController (Zehr & Schwarz 2018) with Amazon’s Mechanical Turk to recruit participants and deploy the experiment. We collected data from 51 L1 U.S. English-speaking adults. Of these, 6 were excluded due to their performance on the benchmarking fillers, leaving data from 45 individuals. Participants gave informed consent and received 3.85 USD.

3.2. RESULTS. To account for possible variation in how participants used the 7-point scale, we normalized the ratings using z-scores. The resulting mean ratings are plotted in Figure 3.

A maximally-specified convergent linear mixed effects model of z-scored ratings constructed in R using package lme4, with fixed effects of polarity (*pos* vs. *neg*) and particle (*either* vs. *too*), reveals a significant polarity by particle interaction ( $p < 0.001$  \*\*\*). *Either* was judged to be more acceptable in negative than in positive sentences; the opposite was true for *too* ( $t = 76.00$ ).

3.3. DISCUSSION. We found that the ratings depended on a highly significant interaction between the polarity of a sentence and the additive item used. It appears that ratings for negative *too* sentences cluster with those for positive *either* sentences rather than the perfectly acceptable positive *too* and negative *either* sentences; note that this matches Rullmann’s (2003) description of the data rather well. These results suggest that adults indeed treat *too* as degraded in negative environments, albeit less so than *either* in positive environments. We take this to indicate that even at age 6 children are not adult-like in their treatment of *too* in negative environments.

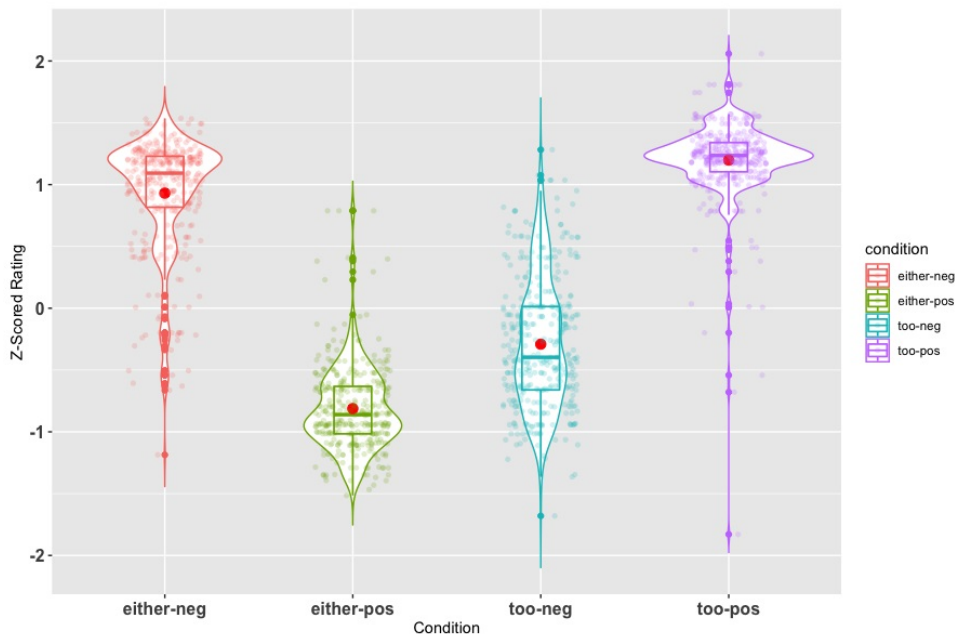


Figure 3: Z-scores of ratings

**4. Corpus study: Child-directed adult speech.** We have observed that children’s comprehension of *either* and *too* is non-adult-like until at least age 6. It is important to know whether this protracted development could be attributed to the input; do children not receive sufficient clear evidence about the distributions of *either* and *too*, or do they have difficulty using the evidence? To determine whether children’s non-categorical trending behavior can be attributed to properties of the input, we conducted a corpus study of child-directed speech.

Using custom R scripts, we extracted every instance of additive *too* and *either* from CHILDES corpora of typically-developing, North American English-acquiring children spoken by Mother, Father, Sister, Brother, Aunt, Uncle, Grandmother, Grandfather, Family Friend, Teacher, or Adult (MacWhinney 2000). Each token was coded for the polarity of the environment in which it appeared. The results are given in the following Table 1. Children hear approximately ten times

Item	Total	Positive	Negative	Unclear
<i>Either</i>	701	28 (3.99%)	670 (95.58%)	3 (0.43%)
<i>Too</i>	7896	7782 (98.56%)	103 (1.30%)	11 (0.14%)

Table 1: Child-directed utterances of additive *either/too* in CHILDES

more occurrences of additive *too* than *either*. The input is overwhelmingly categorical, with very few instances of unlicensed *either* and *too*.<sup>7</sup> It should be noted, however, that if children ask themselves not where to insert *either* or *too* but rather which item to insert in a positive or negative environment, the data available to them in negative environments is somewhat noisier than that

<sup>7</sup>The picture is in fact even clearer than these numbers suggest; 23 of the 28 positive *either* tokens were found in sentences like *Me either!*, uttered in response to a negative antecedent, while 42 of the negative *too* tokens were sentences like *Don’t I get some too?*, following a positive antecedent.

available in positive ones.<sup>8</sup> Even so, the input should provide very few opportunities for confusion about the distribution of *either* and *too*.

**5. Corpus study: Child speech.** In the previous section, we observed clear categorical behavior in child-directed speech that contrasts with the children’s non-categorical behavior profile in the comparative felicity judgment task. This raises a question: What does the production system look like? Does it mimic the input or does it reflect the comprehension system?

We repeated the procedure described above for the target children in each of the corpora investigated. The results are given in Table 2. Children’s productions, like adults’ productions,

Item	Total	Positive	Negative	Unclear
<i>Either</i>	242	22 (9.09%)	210 (86.78%)	10 (4.13%)
<i>Too</i>	5059	5007 (89.95%)	42 (0.83%)	10 (0.20%)

Table 2: Child utterances of additive *either/too* in CHILDES

are categorical; although children were willing to accept the puppets’ productions of *too* in negative environments and (for the younger group) *either* in positive environments, they themselves very rarely use these items in environments where they are not licensed.<sup>9</sup> Children’s productions thus broadly mirror the adult productions that they are exposed to, although the former are somewhat noisier.

It is worth noting that, although children (like adults) produce more utterances containing additive *too* than *either*, both particles appear to follow the same growth curve, Figure 4.

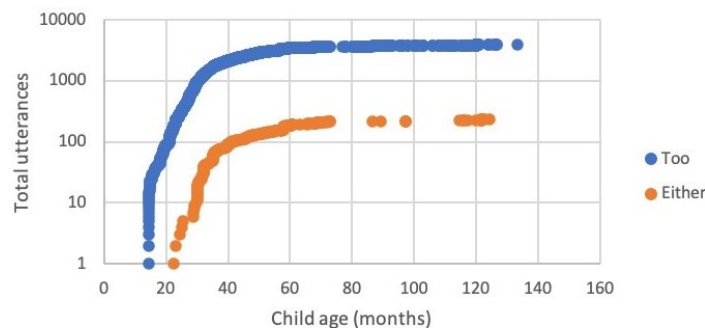


Figure 4: Children’s cumulative productions of *either* and *too*

**6. Discussion.** Our investigation has revealed a number of unexpected and indeed rather noteworthy features of children’s developing understanding of *either* and *too*. The most significant ones are: i) from ages 3–5, children display a stable, non-categorical preference in the direction of the adult grammar, ii) between ages 5 and 6, there is an abrupt transition to adult-like categorical behavior in one half of the paradigm (*either*’s restriction to negative environments) but not the other, and iii) there is an asymmetry between children’s comprehension and production, with the former

<sup>8</sup>The percentages of adult-grammar consistent data would be 86.68% for negative environments and 99.74% for positive ones. We are grateful to Jesse Snedeker for pointing this out to us.

<sup>9</sup>As with the adult production data, most of the positive *either* tokens (18 out of 22) were sentences like *Me either!*, uttered in response to a negative antecedent.



lagging behind the latter. We believe that these findings present serious challenges for all existing theories of (adult) *either* and *too*. In keeping with our previous discussion, we will explicate the nature of these challenges against the backdrop of Rullmann (2003), attempting along the way to distill more general desiderata for any theory which hopes to account not only for the adult pattern but aims to provide a principled basis for explaining how these items are acquired.

From the perspective of the target grammar advocated in Rullmann (2003), we expect the acquisition path for *either* and *too* to unfold roughly along the following lines. Upon first encounter of *either* and *too* learners should entertain a run-of-the-mill additive semantics as in (2) for both items. This is because the treatment in (2) is arguably the simplest way of specifying an additive particle and it is independently needed to account for particles like *also*.<sup>10</sup> The resulting grammar predicts *either* and *too* to be completely interchangeable across all environments and so cannot account for their much more restricted distribution.

Next, as learners begin to record the distributional peculiarity of *either* and *too* in their input data, which, recall, we found to be essentially categorical, they should amend the respective entries in such a way that *either* is confined to negative environments and *too* to positive. The adjustment to the grammar might take the form of specifying each item morpho-syntactically for a suppletion rule along the lines of König (1991), where *either* is simply the spell-out of the particle when it takes scope over a relevant negative operator. Alternatively, *either* might be equipped with a syntactic selection feature that can only be discharged if the host clause contains a negative operator in the scope of *either*. Specifying a lexical item morpho-syntactically for a restricted set of environments is by no means unprecedented in the grammar and so should be a readily available option for the learner. Moreover, since the required morpho-syntactic amendment can take place without any change to the basic additive semantics of our particles we expect a stage in development that is well-characterized by a wide-scope *either* grammar. It would exhibit a fully complementary distribution of our particles for simple sentences like those in (1) and remain in place until the learner encounters more complex sentences that cannot be captured by such a grammar.

One type of data that would provide compelling evidence against the wide-scope *either* grammar is exemplified in (4), where the negative operator is in a higher clause but the presupposition triggered by *either* does not contain the matrix verb.

- (4) John refused to go to church. His parents were so angry that they did not permit him to go to the soccer game either. (Rullmann 2003; 353)

The additive inference in (4) is quite transparently satisfied by the proposition that John didn't go to church. On the wide-scope *either* theory, however, (4) is incorrectly predicted to trigger the presupposition that John's parents did not permit him to go to church. This is because on the wide-scope theory, *either* necessarily outscopes the negative operator of its host sentence. Only in that position is it possible for *either* to generate the correct additive inference in simple cases such as (1). For more complex cases like (4) this assumption fails. To account for the correct presupposition the learner needs to assume that *either*, in fact, takes narrow scope with respect to negation, which in turn requires abandoning the basic additive semantics for *either* in favor of its

<sup>10</sup>Note furthermore that while the proposal in Rullmann (2003) implies a hypothesis space that allows for an “anti-additive” semantics for additive discourse particles, one where the presupposition requires a false antecedent, there are no negative-positive sequences in the data which would unambiguously call for such an entry.

anti-additive counterpart in (3). It also requires marking *either* morpho-syntactically for NPI-status to ensure that it cannot occur with positive hosts. This adjustment to the grammar is clearly more involved than the previous one. Moreover, since the type of evidence that would trigger the change involves more complex constructions which are presumably rarer in the input, it is expected under such an account that the switch to the NPI-grammar of *either* could occur quite late in development.

Our findings are only partially consistent with these expectations. To begin with, we do not see a stage in development where both items are completely interchangeable across positive and negative sentences. Of course, it is possible that such a stage occurs earlier, before the youngest age at which we were able to gather production and comprehension data. Looking at the data we do have, we see that even our youngest participants exhibit behavior that is not fully compatible with this prediction. This is most clearly documented by our production data where we find that even the earliest productions of *either* and *too* are adult-like in their polarity sensitivity. Turning to our comprehension data, the picture gets more nuanced and more interesting. Recall that we find that 3–5-year-olds judged both particles to be contextually appropriate with both sentence types. By itself, this would be expected if they hypothesized the basic additive semantics in (2) for both particles. However, what is not expected is that these children also exhibit a weak adult-like trend towards preferring *either* with negative hosts and *too* with positive hosts, and that they do so essentially unchanged for an extended period of time (three years of development). Both aspects of this development are quite puzzling from the perspective of Rullmann (2003) and even more so from the perspective of the simpler grammar in König (1991). The fact that these children exhibit some sensitivity as evidenced by their trending behavior shows that they have caught on to the fact that polarity matters somehow for the distribution of *either* and *too*. Even more striking, we know from their productions that they view the effect of polarity on the choice of the particle for the purpose of production as decisive. Why then are they not able to execute a rather simple adjustment to the lexical entries of these items that would result in a grammar that matches their input as well as their own productions? Put differently, we have strong indications from our learner's production and comprehension that the polarity of the host sentence matters for the status of *either* and *too*. Nevertheless, our learners seem to view them as compatible with both environments in comprehension, attributing to them the ability to trigger a contextually suitable additive inference irrespective of the polarity of the host sentence. This is of course starkly non-adult-like. And it is entirely mysterious from the perspective of the wide-scope grammar of *either* advocated e.g. in König (1991) but also from the perspective of the NPI-grammar proposed in Rullmann (2003) since that theory offers no reasons why learners shouldn't entertain a wide-scope grammar at an intermediate stage of acquisition when they may have taken note of the polarity effect on the distribution of *either* and *too* in basic sentences but have not yet encountered data such as (4) that are transparently incompatible with a wide-scope grammar of *either*.

Arguably the most striking aspect of this pattern is that it persists over three years of development in the face of categorical and readily available data to the contrary. For some reason, learners seem to not make use of this evidence; the only reason for this seems to be that the wide-scope *either* grammar is never entertained by learners to begin with. That is, it is not part of the hypothesis space for additive discourse particles.<sup>11</sup> If this diagnosis is correct, we have here a novel case of

<sup>11</sup>This would predict that there are no such items in the world's languages. To our knowledge, this has not been

encapsulated learning, one that is based on the abundance rather than the poverty of the stimulus.<sup>12</sup>

Next, we turn to the change that occurs between ages 5 and 6, which affects *either* in that it is now excluded from positive environments but does not seem to affect *too*, which is still readily tolerated in negative environments. Rullmann (2003) offers a perspective for the timing of this change since it is quite plausible that informative data such as (4) are rare in the input and, furthermore, require a well-developed grammar to be fully appreciated by a learner.<sup>13</sup> It is less clear, however, whether 6-year-olds judging *too* in negative sentences to be fully felicitous (albeit at a slightly lower rate than *either*) fits as well. On Rullmann (2003) *too* is disallowed in negative sentences because the required high-scope LF is only marginally accessible due to a parsing preference to attach sentence final particles low and blocking by *either* which can convey the same presupposition without violating the parsing preference. Note that our adult control data bear out this prediction quite nicely. The behavior of our 6 years olds in the comparative felicity judgment task seems to indicate, however, that *too* does not compete with *either* in negative sentences even though *either* is already understood to be an NPI at that age. The only way this could be explained within Rullmann (2003) is to assume that 6-year-olds have not yet adopted the low-attachment preference for sentence-final particles.<sup>14</sup>

The non-categorical (non-adult-like) behavior observed in the 3–5-year-olds' comprehension performance contrasts sharply with their fully categorical (adult-like) productions. While asymmetries between comprehension and production in language acquisition are not uncommon, such asymmetries usually go in the opposite direction, with comprehension maturing first. One attractive way of deriving the appearance of comprehension lagging behind production is to exploit the speaker's advantage of knowing the message;<sup>15</sup> however, it is not clear that our case is amenable to such an explanation, as the stimuli in the comprehension experiment should have left no doubt about the intended message and the context provides no invitation to infer an enriched meaning. We suspect that our asymmetry is instead a reflection of formal complexity.<sup>16</sup> If two structures can express the same meaning, with one being more complex than the other, we might expect that children would systematically choose to produce the simpler structure without necessarily insisting on

systematically investigated.

<sup>12</sup>See e.g. Babyonyshev et al. (2001) for a similar case.

<sup>13</sup>Our own corpus work on child-directed speech is not fine-grained enough to examine this prediction and we will have to leave this issue for future research.

<sup>14</sup>Whether this is plausible has, to our knowledge, not been investigated in literature.

<sup>15</sup>See Hendriks (2014) for discussion of a variety of different relevant cases and an Optimality Theory-based perspective under which the advantage for production is a consequence of a non-adult ranking of grammatical constraints.

<sup>16</sup>We do not present a detailed comparison with Ahn's (2015) account here, but we do not believe that it has a significant advantage in capturing our data. In a nutshell, Ahn proposes that *either* and *too* assert disjunction and conjunction, respectively, between the prejacent (p) and a covert propositional anaphor (q) that is presupposed to be a focus alternative of the prejacent. She derives the NPI behavior of *either* via exhaustification of obligatorily-activated subdomain alternatives in the style of Chierchia (2006). An alternatives-based approach along these lines might seem to provide an explanation for the children's behavior, as children under the age of 6 have been found to be non-adult-like in their ability to access certain alternatives; indeed, it has been claimed that this restricted set of alternatives allows at least some children in this age group to systematically strengthen overt disjunction to a meaning that is fully equivalent to conjunction, though the structure that produces this meaning is more complex (Singh et al. 2016). However, our experimental setup is known to circumvent this kind of problem, with children systematically selecting the more informative of two options presented to them (Chierchia et al. 2001, Foppolo et al. 2012).

the same degree of economy from other speakers. For example, suppose that children's developmental path takes them through a stage where a sentence can have the same meaning regardless of which additive item is inserted, but the *either*-sentence has a more complex LF than *too*-sentence if it is positive and the reverse holds in negative environments. If children always produce the simplest structure that will convey their intended message, we should expect them to systematically produce *too* in positive sentences and *too* in negative ones. Importantly, this leaves open the possibility that children will accept other speakers' productions of both items in both environments, as we observed in the comprehension task, without necessarily inferring an enriched meaning in case the speaker chooses the less economical option.

## References

- Ahn, Dorothy. 2015. The semantics of additive *either*. In Eva Csipak & Hedde Zeijlstra (eds.), *Proceedings of Sinn und Bedeutung 19*, 20–35.
- Babyonyshev, Maria, Jennifer Ganger, David Pesetsky & Kenneth Wexler. 2001. The maturation of grammatical principles: Evidence from Russian unaccusatives. *Linguistic Inquiry* 32(1). 1–44. <https://doi.org/10.1162/002438901554577>.
- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. <https://doi.org/10.18637/jss.v067.i01>.
- Beaver, David I. & Brady Z. Clark. 2008. *Sense and sensitivity: How focus determines meaning*. Explorations in Semantics. Oxford, UK: Wiley-Blackwell.
- Chierchia, Gennaro. 2006. Broaden your views: Implications of domain widening and the “logicality” of language. *Linguistic Inquiry* 37(4). 535–590. <https://doi.org/10.1162/ling.2006.37.4.535>.
- Chierchia, Gennaro, Stephen Crain, Maria Teresa Guasti, Andrea Gualmini & Luisa Meroni. 2001. The acquisition of disjunction: Evidence for a grammatical view of scalar implicatures. In *BUCLD 25 Proceedings*, 157–168. Somerville, MA: Cascadilla Press.
- Drummond, Alex. 2018. Ibex farm. Web. <https://spellout.net/ibexfarm/>.
- Foppolo, Francesca, Maria Teresa Guasti & Gennaro Chierchia. 2012. Scalar implicatures in child language: Give children a chance. *Language Learning and Development* 8. 365–394. <https://doi.org/10.1080/15475441.2011.626386>.
- Hendriks, Petra. 2014. *Asymmetries between language production and comprehension*. Dordrecht: Springer.
- König, Ekkehard. 1991. *The meaning of focus particles: A comparative perspective*. London and New York: Routledge.
- MacWhinney, Brian. 2000. *The CHILDES project: Tools for analyzing talk*. Mahwah, NJ: Lawrence Erlbaum Associates 3rd edn.
- R Core Team. 2019. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna, Austria. <https://www.R-project.org/>.
- Rullmann, Hotze. 2003. Additive particles and polarity. *Journal of Semantics* 20. 329–401. <https://doi.org/10.1093/jos/20.4.329>.
- Singh, Raj, Ken Wexler, Andrea Astle-Rahim, Deepthi Kamawar & Danny Fox. 2016. Children

interpret disjunction as conjunction: Consequences for the theories of implicature and child development. *Natural Language Semantics* 24(4). 305–352. <https://doi.org/10.1007/s11050-016-9126-3>.

Zehr, Jérémy & Florian Schwarz. 2018. PennController for internet based experiments (IBEX). Web. <https://doi.org/10.17605/OSF.IO/MD832>.