More than the sum of its parts: Acquiring semantically complex quantifiers

Cindy Torma (crtorma@mit.edu)

Gabor Brody (gaborbrody@brown.edu)

Athulya Aravind (aravind@mit.edu)

Abstract

How does the acquisition of semantically complex expressions track the acquisition of their constituent meanings? We investigate this question using the English quantifiers *both* and *either*. These quantifiers, while morphologically simplex, are semantically complex, comprising of two pieces: (*i*) universal/existential quantification and (*ii*) a size restriction on the quantificational domain to 2. Across two experiments, we compared the acquisition of these quantifiers with expressions mapping conceptual pieces that contribute to their make-up (*two, all, any*). Our results suggest that having all of the parts is not enough to put together the whole, a finding that could have implications for quantifier learning more broadly.

Keywords: language acquisition; semantics; quantifiers;

Introduction

Unlike noun phrases like *the judge* or *Ruth Bader Ginsburg*, quantificational noun phrases do not refer to individuals — which individuals could expressions like *no student* or *more than two students* possibly pick out? Quantifiers are therefore analyzed not as referring expressions, but as second order predicates that delineate relationships between sets (Frege, 1879; Barwise & Cooper, 1981). Thus, *Some As are B* is true only in the case where the intersection of the set of As and the set of Bs is non-empty; *No As are B* is true only in the case where their intersection is empty.

Learners of all languages must acquire the quantifiers in their language, and in this, they face a formidable challenge. All sorts of set-relations can be articulated using quantificational expressions and constructions — see (1). Quantifier meanings that get lexicalized across languages are, furthermore, varied and rich. Consequently, the hypothesis space the learner must traverse when mapping quantifier meanings is large and complex (Keenan & Stavi, 1986).

(1) a. All the students smiled.
b. Fewer than 3 students smiled.
c. No student but Jack smiled.
d. At least 2 but no more than 5 students smiled.

Consistent with this picture, prior research has revealed considerable variability in the time-course of acquisition of different quantifiers. Whereas children as young as 2 seem to be competent with the quantifier *all*, only at around age 7 do they seem to grasp the meaning of *most* (Barner, Chow & Yang, 2009), a pattern that seems to be robust across languages (Katsos et al., 2016). One hypothesis that has been put forth is that the order of acquisition tracks the relative complexity of the relation expressed by the quantifier (Katsos et al., 2016). To see the import of this hypothesis, compare the standard meaning of *all* with that of *most* (2). To understand sentences with *most*, children need to be able to compare the cardinality of the set of As that are Bs with the cardinalitiy of set of As that are non-Bs.

(2) a. all(A)(B) is True iff $A \subseteq B$
b. most(A)(B) is True iff $|A \cap B| > |A - B|$

The present study turns to another set of quantifiers — the pair *both* and *either* — as a means of studying the acquisition path of semantically complex expressions. This quantifier-pair is special in having conceptual parts that are individually lexicalized using other expressions of English, letting us explore how the acquisition of complex meanings track the acquisition of their constituent meanings.

There are two main meaning components to *both* and *either* (see (3)), which we investigate piecemeal. The first is a cardinality presupposition: these quantifiers restrict the size of the quantificational domain to exactly two elements. Thus, a sentence like *You can have both/either of the toys* can be used felicitously only if there are exactly two toys under discussion. The second key component is quantificational force, which is universal (exhaustive) for *both* and existential for *either*. Thus, whereas *You can have both of the toys* entails giving the entirety of the set of toys, *You can have either of the toys* entails giving a subset.

(3) a. If $|A|=2$, both(A)(B) $\Leftrightarrow A \subseteq B$; else undefined
b. If $|A|=2$, either(A)(B) $\Leftrightarrow A \cap B \neq \emptyset$; else undefined

Our goal in this study is to better understand the relationship between semantic complexity and time-course of acquisition. If children are indeed delayed in mastery of complex quantifier meanings, what is causing the delay? Do children lack command of some of the conceptual structures that the complex meanings are composed of? Or alternatively, is it the process of composition itself, i.e. putting the pieces together the *right* way, that poses a challenge?

We go about addressing these questions by comparing children's command of *both* and *either* with their understanding of expressions that are built using the same (or equivalent)

conceptual structures (*two, all, any*). Probing children's understanding of the numeral *two* might lend insight into their understanding of the component of *both/either* that is part of the cardinality presupposition of the expressions: duality. Likewise, command of the universal quantifier *all* and the existential quantifier *any* will provide a baseline against which to compare children's knowledge of the quantificational force of *both/either*. If a command of the relevant conceptual structures is sufficient for the child to build the adult-like denotations for *both* and *either*, success on *two*, *all* and *any* should also yield success on the complex quantifiers. If, on the other hand, acquiring complex meanings also involves a further step of figuring out how their component pieces fit together, we may find the child entertaining alternative modes of composition and in turn, postulating non-adult meanings.

We begin our investigations with the quantifier *both* in Experiment 1. We designed a novel paradigm, the Quantity-Domain Selection Task, in which participants were shown two distinct kinds of unfamiliar objects, each a part of a candidate set of either two or three items. Participants were asked for *both*, *two*, or *all* of a given kind. Crucially, the labeling for the items was left ambiguous, so that the quantifier or numeral was the only aspect of the sentence from which they could base their response. This paradigm allowed us to independently measure children's decision about how many items they gave and their decision about the domain of quantification. We expand on our findings in Experiment 2, in which we probe children's understanding of *either* and the related expressions *two* and *any*.

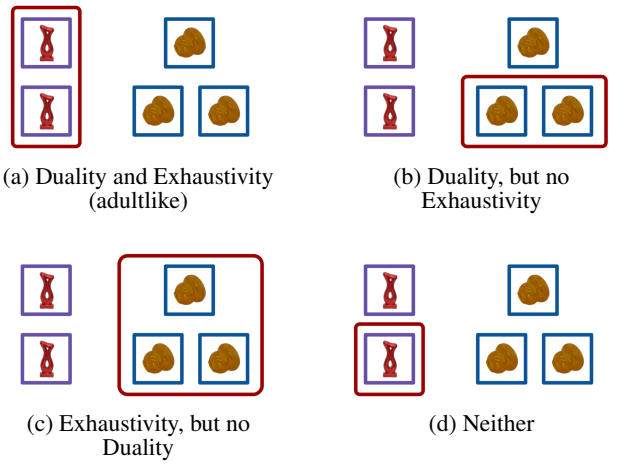
Experiment 1: *Both*

Methods

Participants 28 native English-acquiring children between 2 and 4 years (Age Range: 28-48mos;M = 41mos) in person (pre-pandemic) from childcare centers around the Boston area, and the Boston Museum of Science.[1] An additional 12 children were tested, but excluded for reasons of inattention, non-completion or failing to pass our catch trials (see below). All children were at least 2-knowers, as determined by a Give-N pretest at the time of test. Additionally, we recruited 40 native English-speaking adults online through Amazon Mechanical Turk (MTurk) as controls.

Materials and Design Each trial included five items, each represented by one of two images of real objects, divided into two visually distinct sets of two and three items. The experimental trials involved two sets of novel object kinds, introduced using two unfamiliar words. Crucially, the introduction was ambiguous, e.g. *"Look! There are some tevers and some zavs!*, such that the participant could not know which novel word corresponded to which novel object kind. They were

[1]The wide age-range, during a period of rapid development on children's numerical understanding, was chosen because we aimed to do a direct comparison of two(or more)-knowers and non-two-knowers. The latter comparison could not be carried out due to the pandemic-driven cessation of in-person testing.



(a) Duality and Exhaustivity (adultlike)

(b) Duality, but no Exhaustivity

(c) Exhaustivity, but no Duality

(d) Neither

Figure 1: Possible responses to a *both* trial. Different response patterns might let us infer the presence and absence of different conceptual structures (*exhaustivity* and *duality*) as being represented as part of the meaning.

then asked for *both, two* or *all* of a given kind, e.g. *Can you give me both of the tevers?*. There were 12 trials in total, 3 per quantity word of interest, plus 3 "catch" trials involving the numeral *one*; the latter were used to detect and exclude participants who were not paying attention or did not understand the task. Participants were shown one of four experimental orders, which allowed to control for order factors, including image-label pair, image-domain pair, and pair-side relationships. All preregistration materials and data can be found at: https://tinyurl.com/3sbdb5wc.

For any given experimental trial, participants had to rely on the quantity word (*both, two, all*) to make their decisions. Each participant response involves two decisions — the domain decision, meaning the choice between a 2-set vs. a 3-set, and the quantity decision, or the actual number of items selected from a given set. Figure 1 displays the visual layout of an experimental trial, along with some possible domain-quantity decisions for *both*.

Procedure The child variant of the experiment consisted of two tasks: a standard Give-N task (Wynn 1992) to discern the participant's numerical comprehension, and the main Quantity-Domain Selection task. After their numerical knowledge was assessed, participants were told that they would play a "trading game" with a puppet, who requested some items and would give them stickers in exchange. The experimental trials were preceded by 4 training trials, in which participants were asked to give some quantity of familiar items (e.g. apples), during which they received feedback.

Results

Coding We focus on only the critical trials involving *both, two* and *all*. Each participant response was coded into types based on the number of items given and the choice of set (in

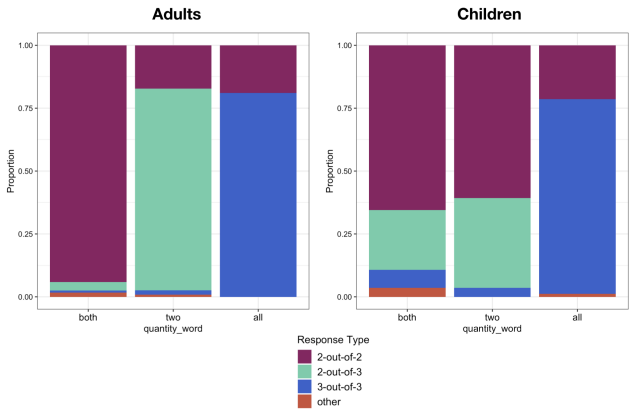


Figure 2: Response Types in Experiment 1 by Quantity Word, Adults and Children

the form *n*-out-of-*n*). There were five such response-types, only three of which were of interest to us, as they express at least one of the component meanings of interest. These were "2-out-of-2", "2-out-of-3", and "3-out-of-3". All other responses were coded as "other". For the child data, We created two additional variables that indicated: (*i*) whether the response-type was logically/semantically consistent given the quantity word ("accuracy"), and (*ii*) whether it was adultlike, i.e. matching the predominant responses of our adult controls. These two measures tend to come apart because adult speakers often pragmatically enrich the literal meaning of quantity-expressions and constructions to arrive at stronger readings (see below). We coded "2-out-of-3" as well as "2-out-of-2" as "accurate" in response *two*, as both responses would indicate an understanding of the literal meaning of the numeral. Analogous reasoning led us to code "2-out-of-2" and "3-out-of-3" as "accurate" for *all*.

Comparison across populations Figure 2 displays the overall pattern of responses from both adults (left) and children (right). We begin with expressions conveying component meanings that are represented in *both*. For *two*, adults had a bias toward the 2-out-of-3 response, although a request for *two* could be satisfied by a selection of 2 items from either the set of 2 or 3. This is likely because in these constructions, adults draw an inference of "proper partitivity" (Hoeksema, 1984, Marty 2019), namely that it's two out of a set with more than two items.

Child responses for *two* were overwhelmingly accurate: 96% of the responses involved giving two objects. However, children's choice of domain was not adultlike, shown by the increased number of 2-out-of-2 responses. A chi-squared test of homogeneity revealed a significant difference between adults and children in their response distributions ($\chi^2(3, 68) = 42.98$, $p < .001$). This pattern of data implies that in our task, children understood the literal meaning of *two*, but failed to draw an adult-like inference of proper partitivity.

For *all*, adults strongly preferred selecting 3-out-of-3, despite the fact that a request for *all* can by definition be satisfied by either a 2-out-of-2 or a 3-out-of-3 response. This preference is consistent with an inference that is typically thought to arise from competition between *all* and *both* (Heim 1991; Sauerland 2008). *All* and *both* express universal quantification. However, *both* has the additional cardinality presupposition that the domain contains exactly 2 elements, so the use of the word *all* instead of *both* can lead one to infer that this condition is not applicable. Reasoning this way, adults settle on the set of 3.

Turning now to children, 99% of their responses to *all* involved exhaustive quantification over the chosen set; 77%, moreover, involved an adult-like choice of the 3-set. A chi-squared test of homogeneity revealed no statistically significant difference between the two populations in the distribution of responses to *all* ($\chi^2(3, 68) = 1.04$, $p = .79$). These results indicate a strong understanding of *all*. Moreover, if the adult preference for the 3-set is due to competition with *both*, children's adultlikeness also raises the possibility that they know and represent the complex quantifier as a competitor; we address this further in the discussion section.

Finally, for *both*, adult responses were highly consistent and in line with expectations: they overwhelmingly chose 2 from the set of 2. 2-out-of-2 was the predominant response for children, as well, but less so (65%) than for adults. The response distributions for *both* pattern significantly differently across populations ($\chi^2(3, 68) = 28.71$, $p < .001$). Looking more closely at children's individual response patterns, we see high rates of within-participant inconsistency in responses to *both*-trials. The largest group (15 out of 28) showed variability in how they responded across trials. The second largest group (10 out of 28) was adult-like, giving 2-out-of-2 responses on all three trials. Finally, there was a small minority (3 out of 28) who consistently responded with 2-out-of-3.

Duality and exhaustivity across quantity-words Next, we analyzed children's representation of the conceptual structures underlying *both* — duality and universal/exhaustive quantification — through pairwise comparisons of the complex quantifier with each of the other two quantity words.

Both vs. *Two*: Our comparison point for duality knowledge is the numeral *two*. For *both* and *two*, we coded each response as "duality-compliant" if it involved giving 2 items (2-out-of-2 and 2-out-of-3). The overwhelming majority of responses to both expressions were duality-compliant (96% for *two*; 89% for *both*). A mixed-effect logistics regression predicting duality-compliance as a function of quantity word (*both/two*) revealed no significant effect of quantity word (model syntax: *duality* $\sim$ *quantity-word* + *(1|participant)*; β=-1.26, SE=0.71, z=-1.77, p=0.08).

Both vs. *All*: *All* is our comparison point for exhaustivity knowledge. For *both* and *all*, we coded each response as "exhaustivity-compliant" if it involved exhaustively/universally quantifying over the chosen set (2-out-of-2 and 3-out-of-3). Exhaustivity-compliance was significantly lower for *both* (72%) compared to *all*

(99%) (model syntax: `exhaustivity ~ quantity-word + (1|participant)`; β=-3.985, SE=1.132, z=-3.521, $p < 0.001$). This effect was driven by a non-trivial proportion (24%) of 2-out-of-3 responses to *both*, which crucially is duality-, but not exhaustivity- compliant. It is also worth noting that this error-type showed no relationship with age in our sample.

Discussion

The goal of Experiment 1 was to chart the acquisition of the semantically complex quantifier *both* in relation to that of the related expressions *all* and *two*. Our results revealed robust understanding of the related expressions, and in turn, the conceptual pieces represented in *both*. For *both*, we found that children were largely adult-like with their quantity decisions, consistently giving 2 items as expected. But our results were less robust when it comes to the domain requirement of the quantifier, i.e. that the cardinality of the domain must be 2.

These findings are consistent with children having an adult-like knowledge of *both*, but some performance issues result in the higher error rates compared to the other two expressions. This possibility gains support from the fact that the predominant response was indeed the adult-like one. As mentioned earlier, children's adultlikeness with *all* might lend further, indirect support: if the 3-out-of-3 bias with *all* is due to competition with *both*, children's adultlikeness with *all*, at least at first blush, implies an understanding of *both*.

On the other hand, there might be reasons to be cautious. First, observe that children in our study, unlike adults, did not compute the proper partitivity inference with *two*. Not only did they not shy away from the 2-out-of-2 response, it was in fact their predominant response.[2] More pertinently, the pattern of responses for *both* and *two* are close to identical. Furthermore, recall that the most prominent error response with *both* involved giving 2-out-of-3. Together, these facts open up the possibility to an alternative explanation, where children have a non-adult representation of *both* with a meaning analogous to *two* (i.e. lacking the exhaustivity component). If so, it suggests that while children succeed in identifying numerical content associated with *both*, they have a harder time identifying it as a requirement tied to the domain of quantification.

To sum up, we are left with the following questions: Are children opting for the 2-set for the right reasons, namely the cardinality presupposition encoded by *both*, or do they have an initial meaning for the expression that is equivalent or similar to *two*, which the present experiment was unable to detect? More generally, is it the case that having the component pieces is not enough to fully facilitate the acquisition of complex meanings, and more specifically, are learners initially less inclined to posit cardinality requirements on the domain?

In order to address these issues, we turn in Experiment 2 to the complex quantifier *either*. *Either* is like *both* in encoding a cardinality presupposition on its domain, but it has existential, rather than universal, quantificational force. As such, it provides a test environment where the two threads — quantity-selection and domain-selection — come apart. A consistent selection of one item from the 2-set would signal an adult-like representation of *either* as an existential quantifier with a cardinality a requirement on the domain.

Experiment 2: *Either*

In Experiment 2, we extend our Quantity-Domain Selection Task to *either*, and compare its acquisition with the related expressions *two* (encoding duality), and *any* (the corresponding existential quantifier). Our key measure of interest in this experiment is domain-selection, for two reasons. First, the main issue left open by Experiment 1 has to do with whether children encode duality as a condition on the domain for *both*. Second, the weak, existential force of *either* renders quantity-selection a less informative measure. A request for *either* can be satisfied by a *superset* of responses that would satisfy *both*, namely 1-out-of-2 or 2-out-of-2.

These are not the only differences between the two complex quantifiers. *Either*, like its non-presuppositional counterpart *any* (and unlike *both*), is grammatical only in select linguistic environments. More specifically, *either* and *any* require an element like negation (the "Negative Polarity" use, (4-a)) or an existential modal (the "Free Choice" use, (4-b)) to be well-formed. Here, we test children's command of the quantifiers in their free choice uses, as in (4), which allows us to make minimal modifications to our task and prompt. On this reading, (4-b) conveys that all the dishes are fair game, and that the addressee can pick one at will.[3]

(4) a. I didn't eat either/any of the dishes.
b. You can eat either/any of the dishes.

While the free choice component adds a further layer of intricacy to sentences containing *either* and *any*, preschool-aged children have been shown to have little trouble with free-choice (Tieu et al. 2016).

Methods

Participants Data-collection for Experiment 2 was carried out during the COVID-19 pandemic, and thus, using virtual methods. We recruited 32 English-primary children between ages 3 and 5 years (Age Range: 38-70mos; M = 50mos) from a database of parents who expressed interest in having their child participate in a language study. The shift in age range was motivated by both practical and theoretical considerations. The virtual variant of the task was unfeasible with 2-year-olds. The additional meaning complexity associated with *either* and *any* raised the possibility of independent delays in acquisition, and motivated the inclusion of older children. As in Experiment 1, all children were at least two-knowers. An additional 13 children were

[2] We are not sure what is behind this preference; one possibility is that they are inclined to give a complete set, whenever possible.

[3] For analysis, see Kadmon & Landman, 1993; Chierchia, 2013 a.o.

tested but excluded for reasons of inattention, non-completion or caregiver-interference. Additionally, 40 native English-speaking adults were recruited through Amazon Mechanical Turk (MTurk) and served as controls.

Materials and Design For adult participants, all visual materials were the same as in Experiment 1, though the linguistic stimuli changed in accordance with the change in quantifiers. Thus, the prompts were changed to the form: *Can you give me either/two/any of the tevers?*. The child experiment was adapted to an online format; otherwise, all materials and design elements remained the same, modulo the prompts.

Procedure For the adult control, all procedures were the same as for Experiment 1. For child participants, the in-person task from Experiment 1 was adapted for testing through video-conferencing, with as few changes to the design and participant interaction scheme as possible. Experimental sessions were conducted over Zoom with a researcher, who shared their screen displaying the study materials in PowerPoint. Participants underwent a brief calibration phase to ensure a proper technical setup, and to determine the best interaction method for the participant (i.e. allowing the child to use the mouse without assistance, or asking parents to place the child on their lap and click on the item(s) to which their child points.) Participants assisted an on-screen cartoon character by completing the same two tasks as for Experiment 1 (Give-N and Quantity-Domain Selection). Children were given a certificate of completion at the end of the session.

Results

We again focus on only the critical trials, which in Experiment 2 involved *either, two* and *any*. As before, participant responses were coded into types based on the number of items given and the choice of set (in the form *n*-out-of-*n*). All five response-types are theoretically meaningful in this experiment, given the weak, existential quantificational force of *either* and *any*. Given that our hypothesis-driven question concerns the choice of domain, we also created a variable, "dual-domain selection", indicating whether or not the participant chose from the domain of 2, irrespective of the number of items given. Only *either* encodes the duality presupposition on its domain, and as such, we expect the rate of 2-set selection to be highest for this expression, at least for adults.

Figure 3 displays adults' and children's responses to each quantity word. For *two*, we largely replicated the pattern of responses from Experiment 1. Both populations overwhelmingly gave 2 items (93% adults; 78% children). As before, children's response distribution differed significantly from adults ($\chi^2(3, 72) = 30.76$, $p < .001$), due to a failure to compute the proper partitivity inference.

For *any*, adults most often selected 1-out-of-3. The preference for the 3-set is likely due to competition with the *either*, analogously to *all* and *both*. Children were like adults in consistently giving 1 item (82%), but they did not consistently choose from the 3-set. The result was a significantly different

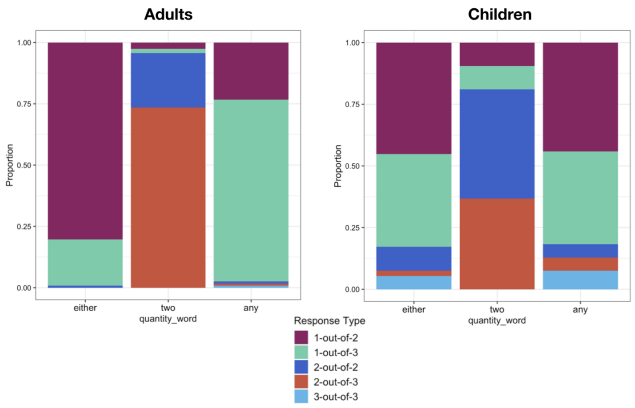


Figure 3: Response Types in Experiment 2 by Quantity Word, Adults and Children

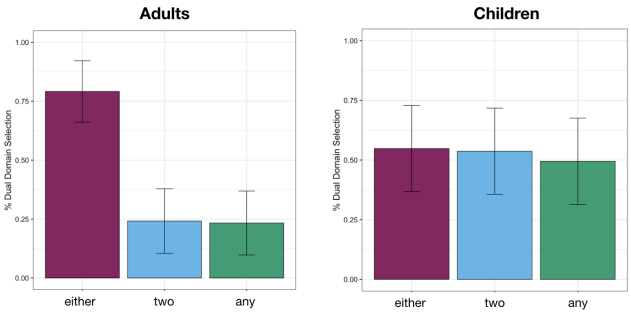


Figure 4: Rate of selecting the 2-set, Adults and Children

response pattern from adults ($\chi^2(3, 72) = 32.07$, $p < .001$).

Child and adult response patterns differed significantly for *either*, as well ($\chi^2(3, 72) = 33.95$, $p < .001$). Adults' main response when asked for *either* was 1-out-of-2, but children vacillated between 1-out-of-2 and 1-out-of-3 responses.

Given our focus on domain-selection, we directly compared the rates of choosing the 2-set across quantifiers. As shown in Figure 4, the patterns are vastly different across populations. We fit a pair of mixed-effects logistic regressions to statistically analyze these patterns. For adults, quantity word had a strong effect on dual-domain selection, with *either* being significantly more likely to elicit such choices (β=1.668, SE=0.181, z=9.225, $p < .001$). However, a comparable model on the child data revealed no statistically significant effect of quantity word.

Discussion

Findings from Experiment 2 seem to suggest that children understand the existential quantificational force of *either*, but not the domain presupposition. Indeed, a large number of response patterns are in principle compatible with understanding an expression as existential. Here, our conclusion of competence comes from noting the striking similarity in children's quantity-selection behavior on *either* and *any*.

On the other hand, children were not adult-like in their

domain-selection. Adults distinguished *either* and *any* in their choice of domain, in two ways. First, they consistently opted for the 2-set with *either*. Second, likely due to competition with *either*, they consistently opted for the 3-set with *any*. In neither respect were children adult-like. In fact, they did not differentiate between *either* and *any* at all.

Altogether, these findings suggest that having the conceptual parts (duality, existential quantification) does not by itself lead to an adult-like meaning for the complex quantifier. It also bolsters the hypothesis raised after Experiment 1, that children's initial hypotheses about *both* encodes a numerical component, though not as a condition on the domain.

General Discussion

Our goal in this study was to understand the developmental relationship between complex quantifiers and the conceptual structures they are composed of. We did so by comparing the acquisition of *both* and *either* with expressions that encode the sub-constituent concepts. Children showed good understanding of these expressions (*two, all, any*) across our two experiments, which we take to provide a critical baseline: children have the necessary conceptual structures needed to build the complex quantifier meanings.

Was this sufficient for an understanding of the complex quantifiers themselves? *Prima facie*, the results of Experiment 1 suggests a positive answer. Children mostly behaved like adults, giving exhaustively from the set of two. Moreover, their behavior with *all* was consistent with a treatment of *both* as its competitor (a strong bias toward the 3-set). But other aspects of the data urged us towards caution. In particular, children's responses to *both* and *two* were similar enough to raise the possibility that their apparent adultlike behavior with *both* was only apparent. We considered an alternative hypothesis on which children associate *both* with duality, but crucially not as a size restriction on the domain. In other words, they may initially posit a "two-like" meaning for *both*.

To further explore this latter hypothesis, we turned to *either*, where quantity-selection and domain-selection come apart. We found that while children treated *either* as having existential quantificational force (at least to a similar extent as for *any*), they clearly failed to recognize the duality requirement on the domain. This failure lends validity to the alternative interpretation of the results of Experiment 1. Caveats are in order, however: both the age ranges and the procedure (in-person vs. online testing) are different across the two experiments.

All in all, then, our findings suggest that even when in command of the component parts, children may have difficulties composing them in the right way. What might be behind this compositional trouble? One possibility is that the specific form of the complexity in these quantifiers — a cardinality presupposition on the domain — is at the heart of the issue. This would be consistent with prior proposals that children do not represent the uniqueness presupposition on the domain as part of their initial meanings for definites (Wexler, 2003).

Another possibility is that arriving at these complex meanings presents a more general learning problem. *Both* and *either* instantiate the subset problem in learning (Manzini & Wexler 1987; Piantadosi, et al. 2008). To wit, whenever a *both*-sentence is true, so are corresponding *two-* and *all*-sentences. Rasin & Aravind (2020) has shown that for simplex quantifiers like *every*, the subset problem does not arise so long as pragmatic evidence — in particular, informativity considerations — are taken into account. Crucially, the key piece that makes *both/either* "stronger" than candidate alternatives like *all/any* is packaged as a presupposition, i.e. information that is shared background knowledge. Consequently, informativity considerations cannot help, because there won't be many situations where *both* is true, felicitious and informative, but where *all* is not also equally true and informative. This might make induction to the complex meaning non-trivial in the case of these quantifiers. In fact, our data is consistent with children initially hypothesizing such subset meanings, where only one of the sub-components are encoded.

Some questions remain open. If children do not have an adultlike meaning for *both*, what drives the competition effects we found with *all*? A promising idea to take into account is the idea that alternatives entering into pragmatic competition need not be lexicalized so long as they are conceptually represented (Chemla, 2007). Another question concerns the relationship between semantic and morphological complexity. The morphological make-up of English *both/either* is not revealing of their meaning complexity, but this does not hold cross-linguistically; consider the Dutch *allebeide* or the Hungarian *mindkettő* ('all-two'). Further research could explore if and how the acquisition trajectory of similar expressions varies across languages.

Acknowledgements

We would like to thank Dr. Martin Hackl (MIT) and Leo Rosenstein (MIT) for their critiques and insight, and our research assistants, Katie Dretler, Katherine Chan, Ella Apostoaie, Toluwalase Asade, Suji Balfe, Ceren Temizyurek, and Stella Ho, without whom data collection would not have been possible.

References

Barner, D., Chow, K., & Yang, S.-J. (2009). Finding one's meaning: A test of the relation between quantifiers and integers in language development. *Cognitive Psychology*, *58*, 195-29.

Barwise, J., & Cooper, R. (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy*, *4*, 159–219.

Chemla, E. (2007). French 'both': a gap in the theory of antipresupposition. *Snippets*, *15*, 4-5.

Frege, G. (1879). *Begriffsschrift. eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. Halle, Germany: Neubert.

Heim, I. (1991). Artikel und definitheit. In *Semantik: Ein internationales Handbuch der zeitgenossischen Forschung* (pp. 487–535). Berlin: Mouton de Gruyter.

Hoeksema, J. (1984). *Partitives.* (Ms. University of Groningen)

Kadmon, N., & Landman, F. (1993). Any. *Linguistics and Philosophy*, *16*(4), 353–422.

Katsos, N., Cummins, C., Ezeizabarrena, M.-J., Gavarró, A., Kuvač Kraljević, J., Hrzica, G., ... Noveck, I. (2016). Cross-linguistic patterns in the acquisition of quantifiers. *Proceedings of the National Academy of Sciences*, *113*(33), 9244-9249.

Keenan, E. L., & Stavi, J. (1986). A semantic characterization of natural language determiners. *Linguistics and Philosophy*, *9*, 253–326.

Marty, P. (2019). On the source of proper partitivity. *Proceedings of Sinn und Bedeutung 23*, *23*(2).

Piantadosi, S., Goodman, N., Ellis, B., & Tenenbaum, J. (2008). A Bayesian model of the acquisition of compositional semantics. *Proceedings of the thirtieth annual conference of the cognitive science society*, 1620-1625.

Rasin, E., & Aravind, A. (2020). The nature of the semantic stimulus: the acquisition of every as a case study. *Natural Language Semantics*, 1-37.

Sauerland, U. (2008). Implicated presuppositions. In A. Steube (Ed.), *The discourse potential of underspecified structures* (pp. 581–600). Berlin, Germany: Mouton de Gruyter.

Tieu, L., Romoli, J., Zhou, P., & Crain, S. (2016). Children's knowledge of free choice inferences and scalar implicatures. *Journal of semantics*, *33*(2), 269-298.

Wexler, K. (2003). *Maximal trouble.* Paper presented at the CUNY Human Sentence Processing Conference.

Wexler, K., & Manzini, R. M. (1987). Parameters and learnability in binding theory. In T. Roeper & E. Williams (Eds.), *Parameter setting* (pp. 41–76). Dordrecht, The Netherlands: D. Reidel Publishing Company.

Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology*, *24*(2), 220-251.