

Revisiting young children’s understanding of modified numerals

Summary: Few studies have examined the emergence of modified numeral quantifiers in child language. Geurts et al. (2010) identify the following developmental predictions, given the theory presented in Geurts & Nouwen (2007): superlative quantifiers should be acquired later than comparative quantifiers, and upward-entailing (UE) quantifiers should be acquired earlier than downward-entailing (DE) quantifiers. Two existing acquisition studies show that 5-year-old children perform at chance on superlative quantifiers (Musolino 2004) and that 11-year-olds are still mastering the superlative quantifier *at most two* (Geurts et al. 2010). We present two experiments revealing that: (i) some 4-year-olds in fact demonstrate adult-like understanding of the superlative modifier *at least n*, and (ii) even in this younger age range, the two asymmetries reported in Geurts et al. (2010) are already emerging.

Background: According to Geurts & Nouwen (2007), the superlative quantifiers *at least/most n* have a richer semantics (involving modality) and should therefore be acquired later than comparative quantifiers *more/less than n*. In addition, DE quantifiers are expected to be harder than UE ones (Geurts, 2003). To investigate 5-year-old children’s understanding of *exactly two*, *at least two*, *more than two*, and *at most two*, Musolino (2004) used a ‘picky puppet’ task: children were introduced to a puppet who only liked cards that had, e.g., *exactly two* smileys; they would then be shown a series of cards containing different numbers of smiley faces, and judged whether the puppet would want the card. Musolino found that children did not differ from chance on *at least* and *at most*, but were above chance on *more than*. Since Musolino’s participants failed to perform above chance on superlative quantifiers, Geurts et al. (2010) moved to an older group of 11-year-old children. They presented participants with an action-based task in which there was an array of boxes and toys; the experimenter uttered a sentence containing a modified numeral (*At least two of the boxes have a toy*), and participants could add/remove/leave the toys to match the sentence. The 11-year-olds performed significantly better on the UE *more than* and *at least* than on the DE *fewer than* and *at most*, and better on the comparative *fewer than* than on the superlative *at most*; they did not perform significantly differently on *more than* and *at least*. Geurts et al. concluded that the child data are roughly consistent with the predictions of Geurts & Nouwen (2007).

Experimental design: We first introduce the common features of the two experiments, before moving to the specifics of each experiment. We designed a covered box task (Huang et al. 2013) in which participants had to listen to what ingredients a chef needed in order to prepare his/her dish, and then select the box that contained the required ingredients (see Fig.1). Participants could only see the contents of one of the boxes (the visible box, which always contained 3 food items); if the visible box contained what the chef needed, they should select it, otherwise the covered box. In Exp.1, we tested participants’ understanding of *at least n* by presenting visible boxes that could either contain: fewer than the required number of objects (e.g., 1 lime, 2 other food items), exactly the required number (2 limes, 1 other food item), or more than the required number (3 limes). In Exp.2, we used the same paradigm to test: *exactly two*, *less than two*, *more than two*, *at least two*, and *at most two*.

Exp.1: 24 children (3;08-7;00, M=4;11) and 25 adults received 3 training trials and 18 test trials in randomized order: 6 *at least n* trials in which the visible box contained fewer than *n* required items, 6 *at least n* trials in which the visible box contained exactly *n* required items, and 6 *at least n* trials

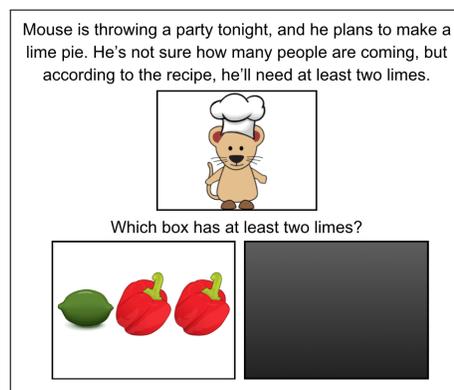


Figure 1: Example

in which the visible box contained more than n of the required items. Fig.2 displays the proportion of visible box selections across the three kinds of contexts. While adults responded as expected, selecting the visible box only when it contained exactly n or more, half of the children responded with an *exactly* reading, selecting the visible box only when it contained exactly n (Group×Picture interaction: $z = -3.9, p < .001$). Nevertheless, the results stand in some contrast to Musolino (2004), who reported that none of his participants correctly interpreted *at least n*.

Exp.2: 25 children (3;03–6;07, $M=4;07$) and 25 adults received 3 training trials and 30 test trials (5 quantifiers x 3 contexts x 2 repetitions), corresponding to *exactly two*, *less than two*, *more than two*, *at least two*, and *at most two* paired with three kinds of visible boxes ($< n$, exactly n , $> n$). Fig.3 displays the results. Adults didn't interpret *at most* as expected (see discussion), so we excluded it from analysis. Responses were coded as correct/incorrect according to the naive semantics associated with each modifier. A GLMM with DE-ness, Category (sum-coded), subject Group (treatment-coded), and Age (log, neutralized for adults) revealed no effect of DE-ness or Category for adults, a main effect of Group ($z = -3, p = .003$), and interactions between Group and both DE-ness ($z = -2.1, p = .03$) and Category ($z = -2.1, p = .04$). Children's performance improved with age ($z = 2.2, p = .03$), but age did not interact with other factors ($|z| < 1.5, p > .13$).

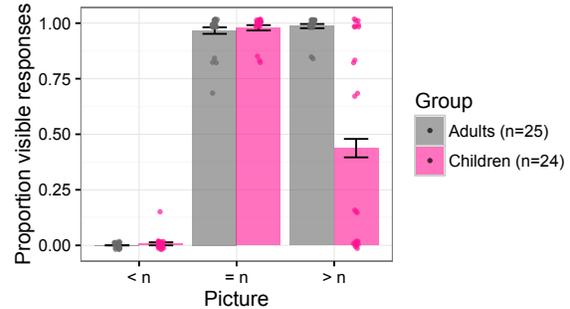


Figure 2: Experiment 1 results (*at least n*).

A GLMM with DE-ness, Category (sum-coded), subject Group (treatment-coded), and Age (log, neutralized for adults) revealed no effect of DE-ness or Category for adults, a main effect of Group ($z = -3, p = .003$), and interactions between Group and both DE-ness ($z = -2.1, p = .03$) and Category ($z = -2.1, p = .04$). Children's performance improved with age ($z = 2.2, p = .03$), but age did not interact with other factors ($|z| < 1.5, p > .13$).

Discussion: The present results are only in part consistent with previous studies. On the one hand, using a different methodology, we observe that half of our sample of 4-year-old children were able to access an adult-like ' $\geq n$ '-meaning for the superlative quantifier *at least n*. This is in contrast to the results reported in Musolino (2004). Many adults also seemed to access a ' $\geq n$ '-reading for 'at most'. This may have been an artifact of the task: if the context specified that the chef "needs at most two apples" and the visible box contains three apples, then it is true that it contains what the chef needs, even though it is false that it contains at most two apples. We plan to run a follow-up where modals are absent from both contexts and prompts. Nevertheless, the results of the present experiments are striking in that they reveal that asymmetries in modified numeral quantifiers (comparative vs. superlative, UE vs. DE) are already emerging in early childhood; such findings indicate the need for a theory that can capture such basic distinctions.

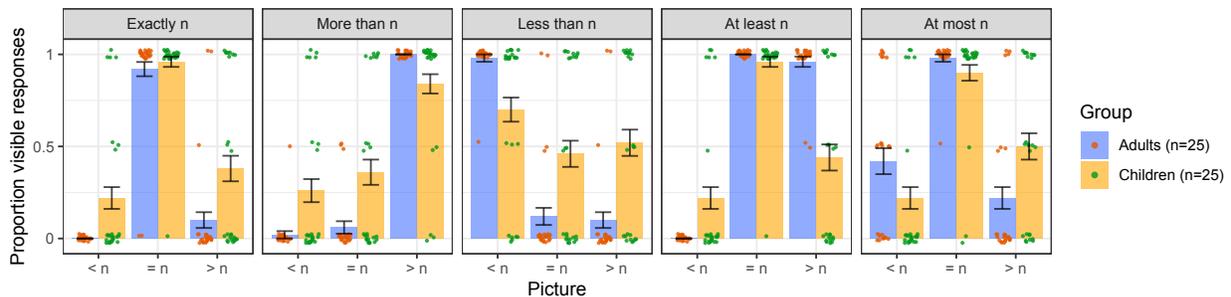


Figure 3: Experiment 2 results (across quantifiers). Bars show mean and SE; dots correspond to individuals.

References: Geurts, B. (2003). Reasoning with quantifiers. *Cognition* 86(3). 223-251. Geurts, B., N. Katsos, C. Cummins, J. Moons & L. Noordman. (2010). Scalar quantifiers: Logic, acquisition, and processing. *Language and Cognitive Processes* 25(1). 130-148. Geurts, B. & R. Nouwen. (2007). At Least et al.: The semantics of scalar modifiers. *Language* 83(3). 533-559. Huang, Y.T., E. Spelke & J. Snedeker. (2013). What exactly do numbers mean? *Language Learning and Development* 9. 105-129. Musolino, J. (2004). The semantics and acquisition of number words: integrating linguistic and developmental perspectives. *Cognition* 93. 1-41.