Meaning *more* or *most*: evidence from 3-and-a-half year-olds

Alexis Wellwood[*], Justin Halberda[†], Tim Hunter[‡], Darko Odic[†], Paul Pietroski[*], Jeff Lidz[*]
[*]University of Maryland, [†]Johns Hopkins University, and [‡]Cornell University

ABSTRACT

We explore how formal differences in the meanings of the quantifiers *more* and *most* influence how they are evaluated even in cases where they are truth-conditionally equivalent. This influence makes sense as, on our view, meanings can act as a kind of "mental zoom lens", directing perception and attention for the evaluation of sentences in different ways. In our case study, we consider the linguistic properties of *more* and *most* in tandem with the psychophysical properties of different modes of numerical estimation and visual set selection in order to show how this process may be measured. We discuss preliminary results from an ongoing study with preschoolers. Our results suggest that, from the very onset of acquisition, meanings must be understood as strictly richer than format-neutral truth conditions, and that learning language at this level importantly involves hooking up linguistic forms with extralinguistic cognition.

1 Introduction

A wealth of research has investigated the semantics, online comprehension, and acquisition of the quantifier *most* (Hackl 2009, Pietroski *et al.* 2009, Lidz *et al.* 2011, Papafragou & Schwarz 2006, Halberda *et al.* 2008b). Yet, while it is a matter of ongoing debate how and when such quantifiers are learned, and how they are understood in real time, at bottom one faces the question of *what* is learned. In the pursuit of giving a semantics for natural language expressions, it is often thought preferable, or even necessary, to proceed independently of considering cognitive factors (Lewis 1970, c.f. Katz & Fodor 1963). Semanticists routinely take as their goal the translation of sentences into logical formulae that specify the conditions under which the sentence would be true. This paper is part of a larger project taking a different tack (Pietroski *et al.* 2009, Lidz *et al.* 2011).

While we do not abandon the idea that meanings specify truth conditions (Pietroski *et al.* 2009, Pietroski 2010), we pursue the hypothesis that meanings are strictly richer than that. In particular, truth-conditionally equivalent representations of a meaning correspond to different psychological hypotheses. Consequently, the form of the representation contributes to the psychological processes involved in determining whether a given truth condition holds (Lidz *et al.* 2011). On this view, different hypotheses about the way the meaning of an expression is represented make different predictions about what people do when they understand that expression. In this paper, we provide experimental data indicating that adults and children treat certain synonymous sentences containing *most* and *more* differently, and argue that these behavioral differences derive from differences in how the meanings

of those expressions are represented, providing further support for the hypothesis developed in Pietroski *et al.* (2009) and Lidz *et al.* (2011).

Given enough time and motivation, adults evaluate *most* by counting, but without these pressures they use the Approximate Number System (Lidz *et al.* 2011, Tomaszewicz 2011, Odic *et al.* under revision), and research with children has found that acquisition of *most* comprehension is independent of counting ability (Halberda *et al.* 2008b). This paper is an exploration of some of the consequences of such facts, and an advance towards showing that *most* and *more* make differential demands on the visual system in another way: *most* invites a visual set-subset selection procedure, whereas *more* invites visual set-set selection and comparison. We show how these differing demands follow from the way the truth conditions are represented, and that this asymmetry is detectable from the very onset of acquisition. These results suggest that something like the Interface Transparency Thesis (Lidz *et al.* 2011) holds from the beginning, and can not be construed as a byproduct of use or experience.

(1) **Interface Transparency Thesis (ITT)**
The verification procedures employed in understanding a declarative sentence are biased towards algorithms that directly compute the relations and operations expressed by the semantic representation of that sentence. (Lidz *et al.* 2011, 233)

In what follows, we first discuss the relation between meaning and algorithms that is central to the project (§2). Next, we discuss work suggesting that adults utilize representations of numerosity more general than that suggested by appeal to set-theoretic cardinality in the logical forms for *most* and *more*, and provide new preference data that suggest the quantities these two quantifiers compare have measurable consequences (§3). After a brief discussion of previous work with children that motivates our specific hypotheses about how preschoolers understand *most* and *more*, we present the results of a novel study that bolsters the findings from adults with regard to which quantities are compared (§4). We conclude in §5, outlining how the present picture supports the idea that learning the meaning of a word involves learning how to represent a particular truth condition, and linking that representation to particular aspects of cognition.

2 Functions and algorithms

On the sound side of the sound-meaning dichotomy, it is uncontroversial to think that syntax interfaces with systems for articulation and perception via the phonological system. The output of the phonological process is thought to be sets of "instructions" that, however abstract, are interpreted in more or less systematic ways by extralinguistic cognition. In such a model, the adequacy of a phonological theory is determined by how far it goes in explaining what kinds of rules and representations this interface traffics in, and, crucially, what kinds it *doesn't*.

Making the analogous case for the meaning side has proven to be more elusive, in part because it is difficult to know in advance which systems on the other side of this interface are relevant. While there is a rich tradition investigating links between articulatory and acoustic phonetics and linguistic phonology, what are the

"conceptual" or "intentional" systems that imbue syntactic structures with meaning? Pursuing the analogy, if we suppose that meanings (or logical forms) too are instructions, then by examining the systems that are recruited to interpret and use them, we may be able to place new constraints on the adequacy of our semantic theory.

Thinking of meanings as "mental zoom lenses" that direct perception in particular ways, observing how attention is directed can be used as a measure of differences in meaning. Consider the image in Figure 1. This scene can be described truthfully using either of (2a) or (2b).

Figure 1: *most/more white* is true, *most/more black* is false.

(2)   a.   Most of the dots are white.
       b.   More of the dots are white than (they are) black.

In all cases where there are just a set of white and a set of black dots, the two sentences deliver identical truth values. But how might they focus attention differently?

More concretely, assume that the logical forms for (2a) and (2b) are as in (3a) and (3b), respectively.

(3)   a. $[\![(2a)]\!] = |\text{DOTS} \cap \text{WHITE}| > |\text{DOTS}| - |\text{DOTS} \cap \text{WHITE}|$
       b. $[\![(2b)]\!] = |\text{DOTS} \cap \text{WHITE}| > |\text{DOTS} \cap \text{BLACK}|$

These forms must contain information as to which representations are required to judge whether the sentences they translate are true or false. Can we find evidence that differences in form correspond to a cognitive distinction, despite the lack of *truth-conditional* difference?

In the rest of this section, we first consider a (very simplified) syntactic analysis of (2a) and (2b), and describe its mapping to semantics. Then we show how the resultant logical forms predict differences in the recruitment of different cognitive strategies when they are evaluated. We begin with *most*, as it is somewhat simpler than *more*, at least on the level at which we will discuss it.

The (extremely simplified) syntactic structure assumed for (2a) is as in (4).

(4)

                        most   dots   be   white

most, a generalized quantifier, takes two semantic arguments: its restrictor, *dots*, and scope, the copular AP *be white*. Both of these are interpreted as sets, or predicates of type $\langle e,t \rangle$, which we will represent as in (5).
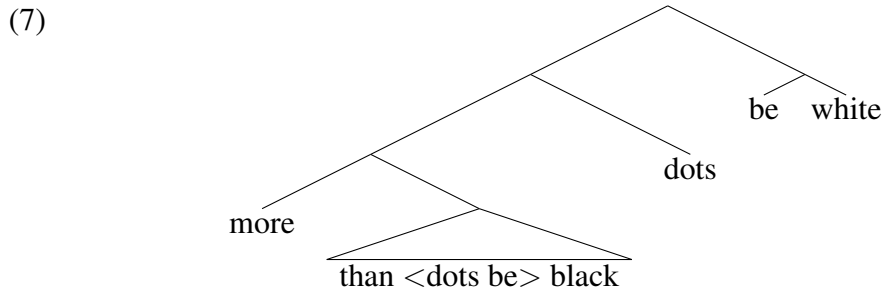
(5) a. ⟦dots⟧ $= \{x|\text{DOTS}(x)\}$
    b. ⟦be white⟧ $= \{x|\text{WHITE}(x)\}$

These sets will usually just be written DOTS and WHITE, respectively. These are combined with the quantifier (6i-6ii) to deliver a complete representation for the sentence as in (6iii).[1]

(6) i. ⟦most⟧ $= \lambda X \lambda Y [\, |X \cap Y| > |X| - |X \cap Y| \,]$
    ii. $⟦(4)⟧ = \lambda X \lambda Y [\, |X \cap Y| > |X| - |X \cap Y| \,](⟦\text{dots}⟧)(⟦\text{be white}⟧)$
    iii. $= |\text{DOTS} \cap \text{WHITE}| > |\text{DOTS}| - |\text{DOTS} \cap \text{WHITE}|$

The resultant expression indicates a numerical comparison between two quantities: that associated with the entities satisfying DOTS ∩ WHITE, and the result of subtracting that quantity from the quantity of entities satisfying the restrictor DOTS itself.

    Contrast this derivation and semantics with that we will assume for *more*. While there are many, many issues in the syntax and semantics of comparatives and superlatives that it would be easy to talk about, for the purposes of the present discussion we must remain at a more surface level. The syntax we assume is that given in (7).[2]

(7)

more is a complex quantifier that takes three arguments. Simplifying considerably, we assume that that denoted by the *than*-clause, which we label C, just picks out the intersection of DOTS and the predicate BLACK. Using the label C will remind us that this argument can be contextually supplied in the absence of an overt *than*-clause.

(8) ⟦C⟧ $= \{x|\text{DOTS}(x)\} \cap \{x|\text{BLACK}(x)\}$

The other two arguments, DOTS and BE WHITE are assumed to have the same denotation as given in (5) above. These arguments combine with *more* as in (9i-9ii), to deliver the resultant interpretation in (9iii).

---

[1]This is the representation arrived at in Lidz *et al.* 2011; cf. Hackl 2009.

[2]This syntax is classical, i.e. *more* combines with the *than*-clause directly, to the exclusion of its evident NP argument: Heim 1985, cf. Bhatt & Pancheva 2004, Wellwood *et al.* 2012. This contrasts with Kennedy 1997, among others. This difference is not crucial here, but is not uncontroversial.

(9) i. $[\![\text{more}]\!]$ $\qquad\qquad\qquad\qquad\qquad = \lambda C \lambda X \lambda Y [\, |X \cap Y| > |C| \,]$

ii. $[\![(7)]\!]$ $\quad = \lambda C \lambda X \lambda Y [\, |X \cap Y| > |C| \,]([\![\text{C}]\!])([\![\text{dots}]\!])([\![\text{be white}]\!])$

iii. $\qquad\qquad\qquad\qquad = |\text{DOTS} \cap \text{WHITE}| > |\text{DOTS} \cap \text{BLACK}|$

It is clear from the surface form of this logical form that it bears a number of similarities with that derived for *most*, in particular that its main operation is a comparison of two quantities.

But now suppose we consider what sorts of representations and faculties would be required for a human to do the job of evaluating sentences like (2a) and (2b), given the commitments encoded in their corresponding logical forms. Both quantifiers call for a comparison of two sets, whether simplex (DOTS) or "complex" (DOTS∩BLACK).[3] Both call for numerical information associated with various sets, as indicated by use of set-theoretic cardinality braces, $|\ |$. Both call for a comparison of the numbers so arrived at, represented by the mathematical greater-than symbol, $>$. Finally, both require that the *greater* of the compared quantities be that associated with the set DOTS∩WHITE.

Where the two forms differ is in what the cardinality of DOTS∩WHITE is meant to exceed. The meaning of *more* specifies just that the exceeded quantity is the cardinality of another intersection, DOTS∩BLACK, the set denoted by the *than*-clause when it is overt, or which may be contextually supplied. The meaning of *most*, in contrast, specifies that the exceeded quantity is the result of a subtraction of the cardinality of the first set, DOTS∩WHITE, from the cardinality of its restrictor argument, DOTS.

Thinking algorithmically, this difference in meaning points to a difference in how many and what kind of steps need to be performed to evaluate whether a *more* or *most* sentence is true. The very salient difference is that evaluating *more* requires, in our case example, an enumeration of the black dots, whereas evaluating *most* doesn't. The algorithmic distinctions we have in mind are brought out by considering the pseudocode in (10) and (11), where `enumerate`, `select`, `subtract` and `greater` represent procedures available in extralinguistic cognition. While we don't intend for (10) and (11) to be interpreted as specific claims about the temporal sequencing of the various steps, the idea is simply that there are certain ways speakers might proceed that would count — in a sense we do not aim to specify formally here — as an instance of the pattern indicated by (10), and there are other ways speakers might proceed that would count better as an instance (11). And importantly, there are ways speakers could in principle carry out a truth-conditionally correct evaluation of a *most* sentence that would *not* count as an instance of the pattern in (10), as we will discuss in the next section.

(10) $x := \text{enumerate}(\text{select}(\text{DOTS} \cap \text{WHITE}))$
$y := \text{enumerate}(\text{select}(\text{DOTS}))$
$z := \text{subtract}(y, x)$
return $\text{greater}(x, z)$

(11) $x := \text{enumerate}(\text{select}(\text{DOTS} \cap \text{WHITE}))$

---

[3]"Complex" here is meant to indicate just that the set is derived by performing an intersection of two other sets, but one may ultimately read "complex" as indicating that disparate visual features must be bound together in order to represent the set.

$$y := \texttt{enumerate}(\texttt{select}(\textsc{dots} \cap \textsc{black}))$$
return $\texttt{greater}(x, y)$

In the next section, we discuss in turn the relationship between the semantical reference to different mathematical expressions, i.e. cardinality braces ($|\ |$) and sets (e.g. $\{x|\textsc{dots}(x)\}$), with their executable correspondents (ENUMERATE, SELECT). In particular, we detail how we might measure the different applications of these instructions. This discussion will raise two important questions that the acquisition study reported in the following section addresses.

3   Adult language and cognition

In the previous section, we saw that there was a way of interpreting logical forms as instructions executable by extralinguistic cognition. We now motivate the psychological reality of representations like these. In §3.1, we outline previous work suggesting that *even when a better strategy is available* speakers choose to use an algorithm like ENUMERATE. In §3.2, we discuss the potential relevance of a vision algorithm like SELECT.

3.1   Motivating ENUMERATE

Adult speakers of English, given unlimited time and motivation, answer questions like,

(12)   Are most of the dots black or white?

using their ability to count up the entities in the sets to be compared. This method is the most likely to achieve the correct response with the lowest rate of error.[4]

This method (counting) is consistent with the logical form in (3a)—what better way to find out the cardinality of a set than to count its members? Yet, it is important to ask whether it is necessary to evaluate *most* in terms of the exactness provided by counting. If it turns out that adults don't *fail* to evaluate whether a sentence with *most* is true or false in situations where they can't count, what would this mean for our interpretation of its logical form? In situations where they don't have unlimited time, there are two relevant possibilities to consider: they should use any truth-conditionally equivalent method which will allow them to arrive at the answer, or they should persistently, on every trial, use a (potentially sub-optimal) system that is consistent with the quantifer's meaning.
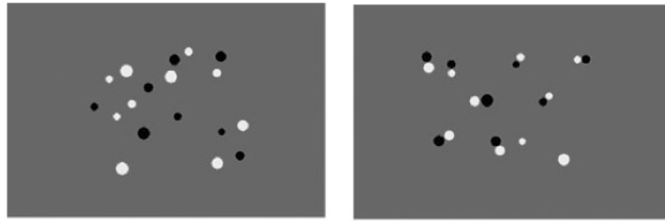
Pietroski *et al.* (2009) tested what adults would do when they only had 200ms to evaluate a *most* question. In this space of time, it is impossible for adults to count up the entities in sets of cardinality greater than 3 or 4. In their experiments, they considered two specific instances of the possibilities given above: (i) adults would make use of any strategy that best suited the case at hand, or (ii) adults would depend on Approximate Number System (ANS) representations exclusively.

If (i) was correct, then different types of displays against which adults evaluated *most* would differentially recruit the ANS as well as a strategy that they dubbed

---

[4]The only type of error is double-counting or omission. If one is careful, one nearly never errs with counting.

OneToOnePlus, following Hume's reduction of numerosity to one-to-one pairing: on such a strategy, two sets are determined to be equinumerous just in case the elements of each can be paired up one-to-one, and one set is *greater-than* another in case it alone has any elements unpaired. If adults made use of this instead of or in addition to systems for representing numerosity explicitly, this implies that the relationship between the logical form (assumed to be similar to those in (6iii) and (9iii)) and evaluation procedures is tenuous at best. A test of this would be to see whether adults are better at evaluating *most* against paired displays like that on the right-hand side of Figure 2: pairing the white and black dots makes it easier to determine which color was most.

Figure 2: Scattered (left) and paired (right) displays in Lidz et al 2011.

In contrast, if (ii) was correct, this would suggest that the interpretation of the cardinality braces in logical forms like that in (6iii) are to be understood as stand-ins for a more general notion of cardinality, 'numerosity': the meaning requires only that a system represent numerical information, so a system that (albeit approximately) represents numerosity is "as good a fit" or at least preferential to one that fails to represent number. The ANS, an evolutionarily ancient cognitive system possessed by creatures as disparate as honeybees, rats, pigeons, monkeys, as well as human adults and children, functions by generating "number percepts" whose imprecision is consistently measurable: accuracy is just a function of *ratio*, where the closer the ratio of two sets gets to 1, the harder the set sizes are to discriminate accurately. Evidence for uniform use of the ANS would mean that there is no overall difference in accuracy when adults evaluate *most* against the two figures in Figure 2, but it should always be ratio-dependent.

When Pietroski *et al* asked adults to evaluate *most* against the unpaired and paired displays, they found ratio-dependent performance regardless of that manipulation, supporting (ii). Importantly, in a control task, where participants were told explicitly to "find the leftover dot" over paired displays, they had significantly higher accuracy than when the instruction was to determine which was "most". Participants were not promiscuous with their evaluation procedures during these experiments: while they would have had greater accuracy on paired off displays if they, on those trials, used the "find the lone dot" procedure, nonetheless they didn't: they persisted in using systems that represented numerosity. These findings supported the Interface Transparency Thesis, in particular the idea that the relevant cognitive interpretation of cardinality braces in logical form is ENUMERATE. This procedure may request information from different systems so long as the content of their representations are numerical.

What about *more*? This expression, too, specifies a comparison of numerosities:[5] Odic *et al.* (in-press) found that young 3-year-olds who did not know how to count understood *more* in count noun contexts, and did so clearly using their ANS. Given these similarities, we now turn to a difference in meaning between *most* and *more* that may be detectable in a different way: visual set selection.

3.2 Set SELECTion

As discussed above, on the assumptions presented in this paper the logical forms of *most* and *more* differ only on what appears on the right-hand side of the greater-than symbol. In our case example, *most* compares the numerosity of a focal set (the white dots) to a subtraction over the total set of dots and the focal set. *more*, in contrast, compares the focal set to a contrast set (the black dots). In this section, we set up the question whether we can detect this difference in meaning using available methods.

As we saw in the preceding section, the task is to understand the properties of the systems on the other side of the interface, which these meanings might call up in the course of understanding or evaluating. The relevant processes for the current case are mechanisms of visual set selection. Halberda *et al.* (2006) discovered that, given any array of dots, adults can reliably select three sets of dots to act as input to the ANS: the total set of dots (the superset), and up to two colored subsets (say, white and black). While there was some individual variation, their results were consistent with the 3-item limit observed in other domains, such that certain items may be extracted automatically and effortlessly. As Halberda *et al.* point out, this limit on the number of items that can be automatically attended includes items like spatially bounded groups (e.g., colored set of dots, or flocks of birds; vanMarle & Scholl 2003, Wynn *et al.* 2002).

Early vision can represent up to three sets very early in the presentation of an array of dots — the superset of dots, and up to two colored subsets — just the kinds of sets that are suggested differentially by the meanings of *most* and *more*. As we've seen, the meaning of *most* references the superset of dots, while *more* doesn't; correspondingly, *more* references the set of black dots, while *most* doesn't. Despite the ease of recovering the superset and two subsets, there is a clear difference in performance when the relevant sets are spatially intermixed (SI) or spatially separated (SS) (see e.g. Ly *et al.* 2009, Price *et al.* 2012): for example, accuracy is higher for the selection of a subset of dots if that set is spatially distinguished from other dots. With these facts in hand, we reasoned that manipulating the spatial arrangement of the superset/subsets could differentially affect the accuracy with which sentences containing *most* and *more* are evaluated.
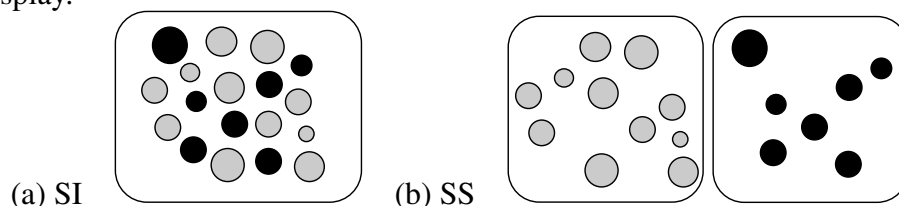
Consider Figure 3, which differ only in whether the black and grey dots are spatially intermixed or separated. The SI array makes the superset of dots immediately available, while the SS array makes the set of white or black dots plain. If *most* requires a representation of the superset of dots, and *more* a representation of the black dots, then participants should be more accurate evaluating a sentence

---

[5]If you're Bobaljik in-press or Stateva 2003, this is true as a matter of composition, as *most* is the superlative of *more*. Cf. Bresnan 1973, Hackl 2009, among many others, who take *most* to be the superlative of *many* or *much*.

containing *most* against an SI display, and on a sentence containing *more* against an SS display.

(a) SI   (b) SS

Figure 3: Spatially intermixed (SI) and spatially separated (SS) two-colored arrays of dots.

As a first probe into the question, we informally asked English-speaking conference participants to consider which of the two sentences in (13) better described the pictures in Figure 3.

(13)   a.   Most of the dots are grey.
        b.   More of the dots are grey.

The vast majority of adults preferred (13a) with *most* to describe the spatially intermixed array on the left in Figure 3, while a majority preferred (13b) with *more* to describe the spatially separated array. This pattern held across vision sciences, linguistics, and psychology conferences including many hundreds of participants (Halberda *et al.* 2008a, Halberda *et al.* 2009, Lidz & Halberda 2010). While informal, and certainly not a standardized empirical test, the results are surprising—and predicted, given our understanding of the constraints of visual set selection, the algorithmic specification of the meanings of *most* and *more*, and the ITT.

To conclude this section, we have seen that adults make use of both exact counting and ANS methods to evaluate *most* and *more*, to the exclusion of other possible evaluation procedures even when the environment would make those procedures more accurate. And we have seen the suggestion that adults prefer different visual arrangements of dots to accompany sentences containing *most* and *more*. These findings raise the question of whether these behaviors (explicit or implicit) truly reflect how the meanings of these quantifiers naturally relate to extralinguistic cognition, as opposed to adults' use and experience. Thus, in the next section we consider data from preschoolers who have just acquired competency with these words.

4   Child language and cognition

We saw that adults use the ANS under time pressure, and that they have clear preferences as to the arrangement of sets for the evaluation of *most* and *more*. Are these possibilities and biases present in young children? Acquisition findings suggest that the first meaning children have when they acquire *most* is satisfied by ANS representations, even when given unlimited time, and even when they have access to exact number representations. Here, we present results from a novel study that extends and elaborates this finding. The novelty we report is that we give children *explicit* access to number representations, by asking them to count the sets to be compared, and we test both with *most* and *more*. Finally, novel to this experiment,

we manipulate array type in order to test whether the bias for different arrays (depending on the quantifier) is there from the very beginning.

Halberda *et al.* (2008b) investigated the relationship between developing counting ability and the comprehension of *most*. Given the reference to exact cardinality concepts in the logical form of this quantifier, it might be expected that children failed to understand *most* until at least after they understand the semantic significance of the counting procedure.[6] These two abilities were in fact independent, as they found children between the ages of 3 and 5 instantiated all possibilities: those who could count but didn't understand *most*, those who could and understood it in terms of exact number, those who couldn't count but understood *most*, and those who could neither count nor understand *most*. The intriguing finding was that some of the full counters who understood *most* nevertheless understood it in terms of ANS representations.[7]

These results suggest that there are multiple distinct number-like systems that are equally well-suited for use in evaluating the meaning of *most*. Yet, these results as yet do not apply to *more*, so it remains to be seen whether the same pattern would obtain for that expression. Further, it may be argued that their results are explained by the fact that they did not ask full-counters to count the sets to be compared: perhaps children *would* have answered in terms of exact number if only they had been given contextual support to determine which numbers exactly were at issue.[8] If, however, even after counting the two sets to be compared, full counters persist in understanding *most* and *more* in terms of approximate number, this supports the conclusion above.

4.1 Experiment

We report in this section a study with three goals: first, replicating and extending the finding of Halberda *et al.* (2006) that young full counters at the onset of understanding *most* **or** *more* persist for a time in evaluating them using ANS representations. Second, we want to further extend that finding by having children count all of the sets to be compared. Third, we want to test, amongst these young counters, whether there is evidence for an effect of spatial arrangement depending on which quantifier is being evaluated.

Procedure and materials. For each child, we first use the "What's on this card?" method to determine childrens' level of counting ability (Gelman 1993, Le Corre

---

[6]The developmental trajectory for acquiring the 'cardinality principle' appears unique in the literature on child language acquisition. Children acquire the meaning for *one* around 2-and-a-half, and then *two* and *three* in six month increments following that. Around 3-and-a-half, they appear to grasp the equivalent of the successor function, understanding that the numbers as high as they can count correspond to the cardinality of the set being counted. See (among many others) Wynn 1992, Carey 2009.

[7]That is, their error pattern showed ratio-dependent accuracy, the signature of the ANS. See the appendix in Lidz *et al.* 2011 that explains this in detail.

[8]With this explanation, though, there would be a puzzle as to what non-counters who displayed comprehension of *most* are doing. See Hurewitz *et al.* 2006, who explicitly tested the timecourse of number word and quantifier acquisition: children acquired quantifiers like *some* and *all* earlier than number words, already suggesting that these two classes of words have distinct, and independent, trajectories.

et al. 2006). In this task, children are presented cards with from 1-16 animals, and asked to count the animals. If the child counts correctly, they are prompted to say what's on the card in the format "6 cows", i.e. *Number-Noun*. If they count incorrectly, they are given the opportunity to re-count. If the child makes persistent errors counting a set size greater than 5 (i.e., they do not reach the correct count on recounts, and are unable to produce Number-Noun), they move into a second phase where it is determined whether they are 1-, 2-, or 3-knowers. These "non-counters" did not participate in the task described here.

Children who are determined to be "full counters" by this procedure go on to participate in the Quantifier task (Halberda *et al.* 2008b) with counting. In this task, we present spatially-intermixed (SI) and spatially-separated (SS) cards with two sets of animals, and ask the child to first identify the names of the animals, and then to count each set. If they count incorrectly, a recount is requested. Finally, the experimenter repeats Number-Noun after each count, e.g., "That's right, there are 6 tigers." The child is then asked one of two target questions: half of the participants are asked (14a) and the other half (14b).

(14) a. Are **most** of the animals tigers or elephants?   MOST CONDITION
      b. Are **more** of the animals tigers or elephants?   MORE CONDITION

Use of this question form allows us to control for effects of syntax, as in this yes-no question frame we can swap *more* or *most* without other morphosyntactic changes. The experimenters pronounced these questions with a large prosodic break after uttering the name of the first set of animals, *tigers*, with a sharp rising tone on the last syllable, and normal prosody on *or elephants*.[9]

Each child sees SS and SI cards, with 4 trials each at ratios 7:6 and 6:5, and 2 trials at 5:4 (where 7:6 = 1.17, 6:5 = 1.2, and 5:4 = 1.25). Sample trials are presented in Figure 4.
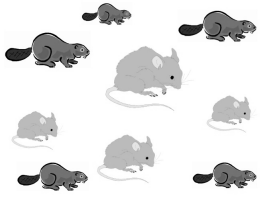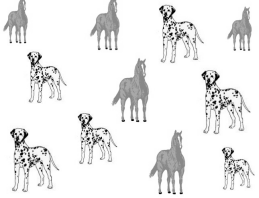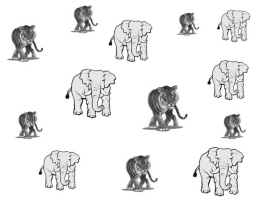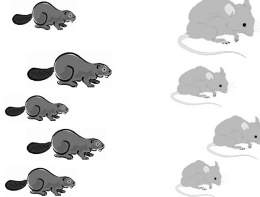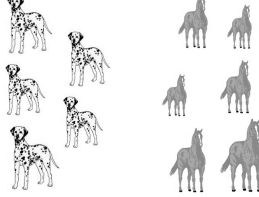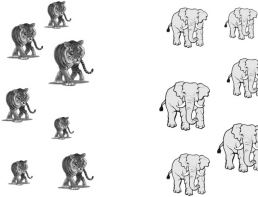
Predictions. Our first prediction is that a number of early full counters will persist in using their ANS to evaluate the target question, regardless of the quantifier. Yet, for these children, we predict an effect of array type will differentiate *more* and *most* as a consequence of their meanings.

We assume that the propositions children would attempt to evaluate given the questions posed in (14) have logical forms as in (15).

(15) a. $[\![(14a)]\!] = |\text{ANIMALS} \cap \text{TIGERS}| > |\text{ANIMALS}| - |\text{ANIMALS} \cap \text{TIGERS}|$
      b. $[\![(14b)]\!] = |\text{ANIMALS} \cap \text{TIGERS}| > |\text{ANIMALS} \cap \text{ELEPHANTS}|$

For those children evaluating the sentence containing *most*, we predict accuracy will be higher on SI than on SS arrays, as SI arrays make the superset (as required by *most*'s logical form) more easily available. There will also be a difference in how quickly performance degrades: as the ratio becomes harder, the difficulty of evaluating *most* on SS arrays will increase, whereas we predict only the usual ratio effect for its evaluation over SI arrays. The opposite is true for *more*: evaluating this quantifier will result in higher accuracy on SS than SI arrays, as SS makes the

---

[9]The stress pattern might best be reflected as: *Are MOST/MORE of the animals TIGERS, or ELEPHANTS?*

Figure 4: Sample experimental materials. Spatially intermixed and spatially separated arrays were matched for difficulty of ratio (easier to harder from left to right).

subsets of animals more easily available, and accuracy will decrease more abruptly when evaluating *more* on SI than SS arrays, as the difficulty of extracting the sets will be compounded with the usual ratio effect.
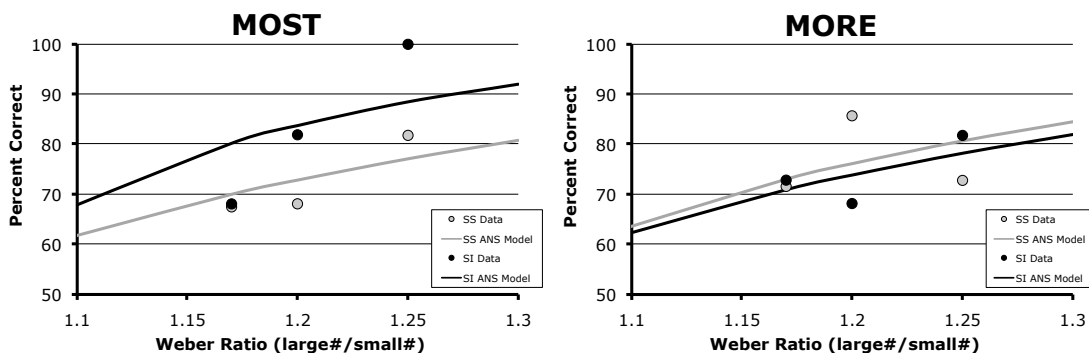
Data analysis. We questioned whether (a) there existed a population of young full-counters who would understood quantifiers like *most* and *more* in terms of approximate number, and (b) if amongst this group there was differential accuracy for array type. If children understand *more* or *most* in terms of ANS representations, accuracy should be ratio-dependent, and well-modeled by the psychometric function of the ANS (Halberda & Feigenson 2008; Pica *et al.* 2004; Pietroski *et al.* 2009; Lidz *et al.* 2011, Odic *et al.* in-press). We used the standard ANS model to check whether a child's performance obeyed Weber's law using a least-squares method (see Halberda *et al.* 2006 for details).

Consequent to the experiment design, we included in the analysis only the results of children whose performance was consistent with this model. We excluded two groups: (a) children who were 100% accurate on the task, suggesting they used their exact cardinality representations to answer the question; and (b) children whose performance could not be fit to the psychophysical model, suggesting they have not yet acquired a meaning for the test quantifier (these children typically had around 50% accuracy). Of 108 children recruited, 55 were non-counters by the counting titration procedure described above (mean age 3 years;6 months, range 3;0-4;0), and 53 were full-counters as determined by the same method. Of the 53 full-counters, 9 children were excluded (5 for fussiness, 1 for displaying a clear preference to name their favorite animal, and 3 for experimenter error).

We attempted to fit the data from the remaining 44 full-counters. Using this method, we successfully fit 11 kids in the *more* condition (mean age 3;9, range 3;6-4;0), and 11 kids in the *most* condition (mean 3;9, range 3;5-4;1). Of the remaining 22 full counters, 11 *more* kids (mean 3;6, range 3;3-3;9) and 5 *most* kids (mean 3;8, 3;5-3;9) could not be fit. Finally, 6 *more* kids (mean 3;9, range 3;8-3;11) performed

100% accurately.

Results. We plot the results from 11 *more* and 11 *most* kids in Figure 5. Accuracy is plotted on the y axis, and ratio on the x axis in terms of increasing ease (i.e., the ratios get further from 1.0 as you move right on the graph).

Figure 5: Preliminary results for *more* and *most*, with accuracy plotted as a function of ratio and array type. Lines represent ANS model fit. The y axis in both graphs begins at chance (50%).

While preliminary,[10] these results show a clear trend in the predicted direction. Accuracy for *most* was better for SI arrays, with accuracy at its lowest point in the middle ratio for SS but not until the hardest ratio for SI. In contrast, accuracy for *more* was better for SS, with accuracy at its lowest point in the middle ratio for SI but the hardest ratio for SS.

5 Conclusions and prospects

We presented experimental results with children, that augment how we understand the acquisition of the quantifiers *most* and *more*. The overall decline in accuracy for young full counters as a function of ratio replicates previous work on *most* and extends it to *more*. These results go beyond previous work in that, here, children counted the sets to be compared before answering a *most/more* question. Importantly, our sample consisted entirely of young children who understand the semantic significance of the counting procedure, yet insist on using ANS representations to answer the question; this suggests a privileged role of this cognitive system for understanding comparison in the number domain.

Given the very narrow age window in which it is possible to find young counters who yet do not understand quantificational expressions in terms of exact counting, the current dataset is limited. However, considering that the data trends in the same direction as the adult preference data which suggests a bias in favor of spatially intermixed displays for communicating a *most* thought, and the opposite bias towards spatially separated displays for communicating a *more* thought, we take the child data to support the conclusion that it is the *form* of the meaning that underlies the bias in preference (adults) and in accuracy (children).

---

[10]The study design ultimately requires 16 'approximate' children per quantifier to definitively test the array type manipulation.

Given that our experimental materials were such that they made a *most* and an equivalent *more* sentence true in exactly the same cases, any observed differences must derive from differences in the formal specification of the meaning. Truth-conditionally, the answer a participant would arrive at in any one of our trials is identical regardless of the quantifier one was exposed to. Appearances, however, suggest that the path one takes to get there differs as a function of the meaning.

We take these results to support the Interface Transparency Thesis of Lidz *et al.* (2011), demonstrating that meanings are strictly richer than format-neutral specifications of truth conditions from the point of acquisition. The patterns of behavior we see with adults (i.e., ratio-dependent accuracy, array type preference) cannot just be the result of language and extralinguistic experience, but points to deeper facts about how linguistic meanings and other systems of human cognition interact. In particular, the process of linking linguistic forms with extralinguistic systems is a subtle affair, but one that may display uniformities across development.

References

Bhatt, Rajesh, & Roumyana Pancheva. 2004. Late merger of degree clauses. *Linguistic Inquiry* 35.1–46.

Bobaljik, Jonathan David. in-press. *Universals in comparative morphology: Suppletion, superlatives, and the structure of words*. Boston MA: MIT Press.

Bresnan, Joan. 1973. Syntax of the comparative clause construction in English. *Linguistic Inquiry* 4.275–343.

Carey, Susan. 2009. Where our number concepts come from. *Journal of Philosophy* 106.220–254.

Gelman, Rochel. 1993. A rational-constructivist account of early learning about numbers and objects. *The psychology of learning and motivation: Advances in research theory* 30.

Hackl, Martin. 2009. On the grammar and processing of proportional quantifers: *most* versus *more than half. Natural Language Semantics* 17.63–98.

Halberda, Justin, & Lisa Feigenson. 2008. Developmental change in the acuity of the "number sense": The approximate number system in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology* 44.1457–1465.

——, Tim Hunter, Paul Pietroski, & Jeffrey Lidz. 2008a. An interface between language and vision: quantifier words and set-based processing. Talk presented at VSS, The Vision Sciences Society, May 9-14, Naples Florida.

——, Jeffrey Lidz, Paul Pietroski, & Tim Hunter. 2009. Set based visual processing in the acquisition of 'most'. Talk presented at LSA, the Linguistic Society of America, January 8-11 San Francisco, CA.

——, Sean F. Sires, & Lisa Feigenson. 2006. Multiply spatially overlapping sets can be enumerated in parallel. *Psychological Science* 17.572–576.

——, Len Taing, & Jeffrey Lidz. 2008b. The development of "most" comprehension and its potential dependence on counting ability in preschoolers. *Language Learning and Development* 4.99–121.

Heim, Irene, 1985. Notes on comparatives and related matters. Unpublished manuscript, University of Texas, Austin.

Hurewitz, F., A. Papafragou, L. Gleitman, & R. Gelman. 2006. Asymmetries in the acquisition of numbers and quantifiers. *Language Learning and Development* 2.77–96.

Katz, Jerrold J., & Jerry A. Fodor. 1963. The structure of a semantic theory. *Language* 39.170–210.

Kennedy, Christopher, 1997. *Projecting the Adjective: The Syntax and Semantics of Gradability and Comparison*. University of California, Santa Cruz dissertation.

Le Corre, M., G. Van de Walle, E.M. Brannon, & S. Carey. 2006. Re-visiting the competence/performance debate in the acquisition of the counting principles. *Cognitive Psychology* 52.130–169.

Lewis, David. 1970. General semantics. *Synthese* 22.18–67.

Lidz, Jeffrey, & Justin Halberda. 2010. Interface transparency in the acquisition of number words. Invited lunch symposium at BUCLD 35, the Boston University Conference on Language Development, November 6-8, Boston MA.

——, ——, Paul Pietroski, & Tim Hunter. 2011. Interface transparency and the psychosemantics of *most*. *Natural Language Semantics* 1–30.

Ly, R., H. Y. Im., & J. Halberda. 2009. Spatial overlap of collections affects the resolution of ensemble features. Poster presented at 2009 Vision Sciences Meeting, Naples, Florida.

Odic, Darko, Paul Pietroski, Tim Hunter, Jeffrey Lidz, & Justin Halberda. in-press. Children's understanding of "more" and discrimination of number and surface area. *Journal of Experimental Psychology: Learning, Memory, and Cognition* .

——, ——, ——, ——, & ——, under revision. Interface transparency and the count/mass distinction. under revision.

Papafragou, Anna, & Naomi Schwarz. 2006. Most Wanted. *Language Acquisition* 13.207–251.

Pica, Pierre, Cathy Lemer, Véronique. Izard, & Stanislas Dehaene. 2004. Exact and approximate arithmetic in an Amazonian indigene group. *Science* 306.499–503.

Pietroski, Paul. 2010. Concepts, meanings, and truth: First nature, second nature, and hard work. *Mind & Language* 25.247–278.

——, Jeffrey Lidz, Tim Hunter, & Justin Halberda. 2009. The meaning of *most*: semantics, numerosity, and psychology. *Mind & Language* 24.554–585.

Price, G.R., D. Palmer, C. Battista, & D. Ansari. 2012. Nonsymbolic numerical magnitude comparison: Reliability and validity of different task variants and outcome measures, and their relationship to arithmetic achievement in adults. *Acta Psychologica* 140.50–57.

Stateva, Penka. 2003. Superlative more. In *Proceedings of SALT XIII*, ed. by Robert B. Young & Yuping Zhou, 276–291, Cornell University, Ithaca, NY. CLC Publications.

Tomaszewicz, Barbara. 2011. Verification strategies for two majority quantifiers in Polish. In *Proceedings of Sinn und Bedeutung 15*, ed. by Ingo et al. Reich, Saarbrücken, Germany. Saarland University Press.

vanMarle, K, & Brian J. Scholl. 2003. Attentive tracking of objects versus substances. *Psychological Science* 14.498–504.

Wellwood, Alexis, Valentine Hacquard, & Roumyana Pancheva. 2012. Measuring and comparing individuals and events. *Journal of Semantics* 29.207–228.

Wynn, K., P. Bloom, & W.-C. Chiang. 2002. Enumeration of collections by 5-month-old infants. *Cognition* 83.

Wynn, Karen. 1992. Children's acquisition of the number words and the counting system. *Cognitive Psychology* 24.220–251.