Linguistic meanings interpreted

Tim Hunter (timhunter@ucla.edu)
Department of Linguistics, 3125 Campbell Hall
Los Angeles, CA 90095 USA

Alexis Wellwood (wellwood@usc.edu)
School of Philosophy, 3709 Trousdale Parkway
Los Angeles, CA 90089 USA

Abstract

A prominent strand of theorizing in linguistics models meaning in language by specifying an "interpretation function" which relates morphosyntactic objects (i.e., those representations whose properties are uncovered by research in morphology and syntax) to elements of non-linguistic experience. Such theorizing has, for the most part, proceeded in relative isolation from developments in the other cognitive sciences. A recent body of experimental work growing out of this tradition has, however, pressed the question of precisely how linguistic representations relate to other faculties of mind. We present the beginnings of a two-step formal proposal for how to do this, specifying: (i) the co-domain of the linguistic interpretation function as the Language of Thought (LoT); (ii) what this mental language is like, (iii) which expression of this language is the semantic value of a sentence like 'Most of the dots are yellow', and (iv) how that LoT expression is interpreted by other cognitive faculties, in ways that produce the choices of verification procedure that have been empirically observed.

Keywords: formal semantics, language of thought, mentalese

Introduction

We pursue a formal framework that can help to contextualise the findings and methodology of a sampling of relevant experimental work, and facilitate closer contact between research in natural language semantics and cognitive science. That experimental work, reviewed below, reveals properties of natural language meaning that (i) cannot be accounted for under semanticists' standard idealization that natural language sentences have mere "truth conditions" as their meanings, and (ii) relate in tractable, identifiable ways to well-understood non-linguistic cognitive systems (e.g. the Approximate Number System, ANS). This provides a clear impetus to formulate concrete alternatives to the standard idealization. We present one such alternative based on "two-step interpretation" via a Language of Thought (LoT) (Fodor, 1975; Rescorla, 2019).

The idea is that semantic theory specifies an interpretation function ("first step") which has as its domain morphosyntactic objects (as described by morphological and syntactic theory) and as its co-domain, we propose, expressions of LoT. Arriving at one of these LoT expressions amounts to understanding the relevant natural language expression in a way that brings it into contact with the rest of cognition. Specifically, we view the relevant non-linguistic systems as providing their own interpretation functions ("second step") with LoT expressions in their domain. Being explicit in this way emphasizes the language-like properties of the intermediate LoT representations, while making it clear that they need not be conflated with representations of natural language.

The experimental case studies we build on all investigate the semantics of sentences like (1). This has proved a particularly fruitful starting point because the design and interpretation of these experiments can draw on extensive literatures both on the mathematical and logical properties of the relation expressed by 'most', and on relevant non-linguistic cognitive systems of number.

(1) Most of the dots are yellow.

We give a hypothesized LoT fragment sufficient for a particular LoT expression resulting from the first step interpretation of (1), and the accompanying interpretations of that expression at the second step that can yield predictions compatible with the experimental results. Many of the details are underdetermined, given the (so far) thin empirical base on which we build. Our main goal is to contribute an overall picture of what such an integrated theory of semantics and cognition might look like, and show how it reveals important theoretical and formal questions that can drive further progress.

Constrained flexibility in verification

In some scenarios, a speaker may seek to learn the truth/falsity of (1) by comparing numbers she arrives at based on a counting procedure: e.g. given a relatively "close call" with just 20 dots, of which 11 are yellow. In other scenarios, she might be satisfied by her ANS representations: e.g. given a display of 173 dots, of which 146 are yellow. Such elementary observations raise serious questions. Since (1) isn't ambiguous,[1] it must have a single meaning that is "flexible" enough to support verification by at least these two methods.[2] However, a series of experimental studies, investigating the verification procedures that speakers adopt for (1) in carefully controlled scenarios, have discovered that there are surprising and fine-grained constraints on this flexibility (Pietroski et al., 2009; Lidz et al., 2011; Knowlton et al., 2021). The conclusion is that certain verification procedures are somehow "better aligned" with the meaning of (1) than other truth-conditionally equivalent procedures. Our account for this is that if a sentence meaning is an LoT expression, then its syntactic structure can align with certain verification procedures more transparently than with others.

[1] A speaker might sometimes verify 'Sam saw the spy with binoculars' in a way that checks whether Sam had binoculars, and sometimes in a way that checks whether the spy had binoculars. But this does not seem analogous to (1).

[2] Of course, these are closely related; see Sella et al. 2021.

The first crucial finding concerns the relationship between cardinality and *one-to-one correspondence*. Given two finite sets A and B, the cardinality of A is greater than that of B if and only if (iff) A has a proper subset whose elements can be put into one-to-one correspondence with the elements of B. We define the relation OneToOnePlus to reflect this.

$$\text{OneToOnePlus}(A,B) \triangleq \exists A'[A' \subset A \text{ and } \text{OneToOne}(A',B)]$$
$$\text{OneToOne}(A,B) \triangleq \text{there is a bijection from } A \text{ to } B$$

The conditions under which (1) is true can therefore be equally well described by (2) or by (3), the latter of which makes no reference to cardinalities.

(2) $|DOT \cap YELLOW| > |DOT \setminus YELLOW|$

(3) $\text{OneToOnePlus}(DOT \cap YELLOW, DOT \setminus YELLOW)$

Is the meaning of (1) flexible in a way that supports verification via detection of one-to-one correspondence, just as readily as it does via ANS computations and precise counting?

Results reported by Pietroski et al. (2009) indicate that the answer is "no". Their participants were shown displays like those in Figs. 1a, 1b for 200ms (a display time that made precise counting impossible) and asked to judge whether (1) was true for each. Analysis of the error rates revealed the signature ratio-sensitivity of the ANS (e.g. more errors when the ratio of yellow to non-yellow dots is 9:10 than when it is 1:2). The surprising result was not that this tell-tale sign of the ANS was found with scenes like Fig. 1a, but that the very same pattern was replicated with scenes like Fig. 1b, which were designed to facilitate verification via detection of one-to-one correspondence relations. Control tasks confirmed that the 200ms display time was sufficient for participants to distinguish the two kinds of layouts, and to identify the color of the un-paired dots in displays like Fig. 1b (both with very high accuracy). So while the task demands did not rule out detecting one-to-one correspondence, participants nonetheless persisted with using their (sometimes less accurate) ANS.

The second crucial finding concerns the specific details of the ANS-based verification procedures. It is possible for the case in Fig. 1a that speakers compare two "primitive" ANS representations of numerosity, one for the yellow dots (there are 10) and one for the blue dots (8). A different possibility is suggested by the intuition that sentence (1) isn't really "about" how yellow and blue dots relate, but about how the dots (18) and their yellow subset (10) relate. Given ANS representations of the latter two numerosities, a speaker might proceed by comparing 10 against $18 - 10$. The relationship between the first, *direct* procedure and this second, *indirect* procedure corresponds roughly to that between (2) and (4).

(4) $|DOT \cap YELLOW| > |DOT| - |DOT \cap YELLOW|$

Importantly, the indirect procedure will yield less accurate results than the direct procedure, because of the additional noise introduced by the subtraction operation.

Converging experimental results indicate that speakers adopt the indirect procedure for (1). First, Lidz et al. (2011) compared performance on displays like those in Fig. 1a and Fig. 1c, where the heterogeneity of the non-yellow dots makes constructing a single, primitive ANS representation of the numerosity of the non-yellow dots all but impossible. If participants used the direct procedure with Fig. 1a, then accuracy there would be higher than on Fig. 1c where the more accurate procedure is not possible; but in fact no difference was found. Second, Knowlton et al. (2021, 139) found that, given displays like Fig. 1a, participants' accuracy verifying a sentence like (1) was lower than their accuracy verifying a different but truth-conditionally equivalent 'more' sentence. This pattern is consistent with participants making use of the direct yellow-vs-blue procedure for 'more' but neglecting it in favor of the less accurate indirect procedure with 'most', even given exactly the same displays.

A full account of the meaning of (1), then, must explain: (i) how ANS-based verification procedures can engage with this meaning but procedures based on one-to-one correspondence cannot (in the same experimental setting); and (ii) why, when people recruit their ANS to verify (1), they use the indirect procedure over the direct one.

Using expressions of LoT as semantic values

We aim to identify a meaning for (1) which is flexible enough to allow for some variety in verification procedures (e.g. both ANS and precise counting), but not so flexible so as to support other logically possible procedures that would yield the same true/false answers (e.g. one-to-one correspondence, or direct comparison with the blue dots). See Fig. 2.

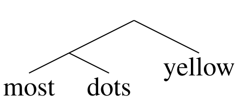
A proposal: interpretation in two steps

We propose to assign LoT expressions as meanings. For (1), this is indicated in (5). The interpretation function $[\![\ldots]\!]^{\text{SEM}}$, then, maps expressions of English to expressions of LoT.[3]

(5) $[\![\text{'Most of the dots are yellow'}]\!]^{\text{SEM}}$
`= |DOT & YELLOW| > |DOT| - |DOT & YELLOW|`

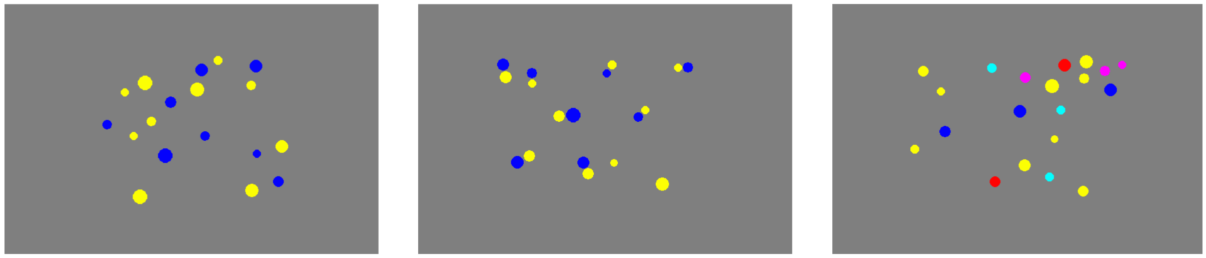
(We will consistently use "typewriter" text to write LoT expressions.) Call the expression on the right-hand side m.[4]

For now at least, we are agnostic about how the "first step" compositional process maps (1) to m. This may arise via a structure like (6), with sisterhood interpreted via function application and (7) as the lexical semantic value for 'most'. (This lexical semantic value is a function that assembles LoT expressions from other LoT expressions.)

(6)


[3]The notation $[\![\cdot]\!]$ is widespread in both linguistics (e.g. Heim & Kratzer, 1998) and computer science (e.g. Winskel, 1993). The superscript serves to distinguish a variety of interpretation functions.

[4]Importantly, m is a hypothesized cognitive representation about which virtually nothing has so far been said. While its symbols and syntax may be familiar and perhaps suggestive, our hypothesized LoT expressions must not be confused e.g. with the mathematical expressions in (2), (3) or (4). Those expressions were helpful for exploring a range of logically possible verification procedures, but they convey only mathematical relationships, not cognitive hypotheses.



(a) (b) (c)

Figure 1: Displays used in the experiments reported in Pietroski et al. (2009) and Lidz et al. (2011).

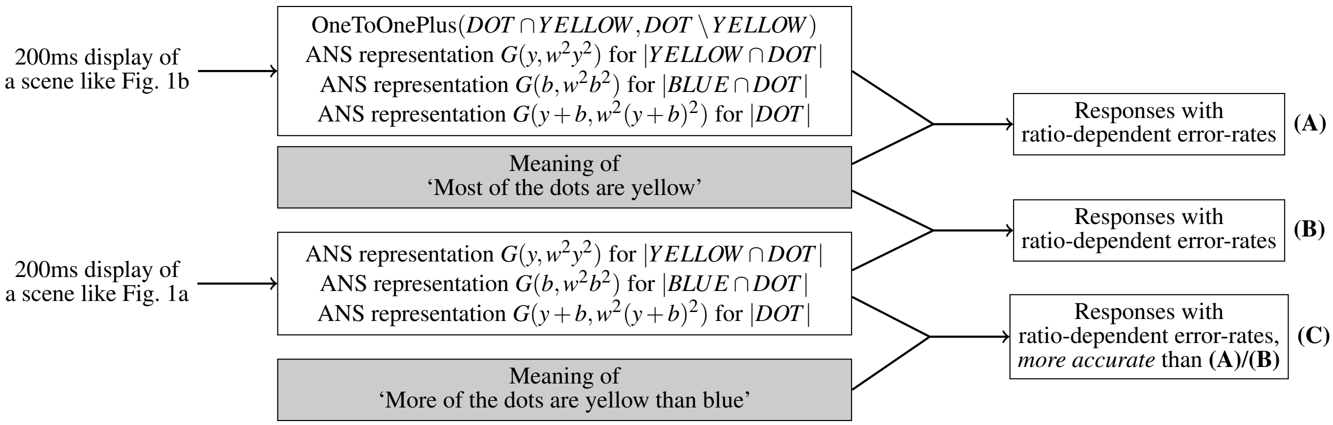


Figure 2: Outline of relevant experimental scenarios. Shaded boxes represent sentence meanings that are under investigation; we triangulate in on these by considering, for various experimentally-manipulated displays that make different kinds of information available (white boxes in the middle), how these meanings interact with the available information to yield observed response patterns (white boxes on the right). We write $G(\mu, \sigma^2)$ for an ANS representation corresponding to a gaussian with mean μ and variance σ^2. The first key finding was that the response patterns at Box **(A)** and Box **(B)** show the same ratio-dependent error rates that are characteristic of the ANS, despite the additional one-to-one correspondence information that is made available by a layout like Fig. 1b. The second key finding was that this common response pattern show more errors than at Box **(C)**, where the different target sentence allows for more direct use of the information made available by a layout like Fig. 1a.

(7) $[\![\text{'most'}]\!]^{\text{SEM}} = \lambda X.\lambda Y.\texttt{|}X\ \texttt{\&}\ Y\texttt{|} \texttt{ > } \texttt{|}X\texttt{|} \texttt{ - } \texttt{|}X\ \texttt{\&}\ Y\texttt{|}$

Or, the morphosyntax may be more fine-grained, perhaps with no leaf node corresponding to 'most', but rather some complex composed out of 'many'/'much' and '-est' (Hackl 2009). These are live issues in compositional semantics about which we have nothing to say. Instead, we will focus mainly on the "second step" interpretation of m.[5]

The central idea is that there are different ways for non-linguistic cognitive systems to interpret LoT expressions such as m, which can lead to differences in verification. The language faculty produces m, based on exposure to the string in (1), and then m, in turn, is interpreted by the ANS, or a "precise counting" system, etc; see Fig. 3. Our analytical task is to lay out specific, concrete ways in which LoT expressions may be interpreted by particular cognitive systems, in such a way as to predict the available verification procedures that we know speakers adopt for m, while excluding those that speakers neglect (e.g. one-to-one correspondence).

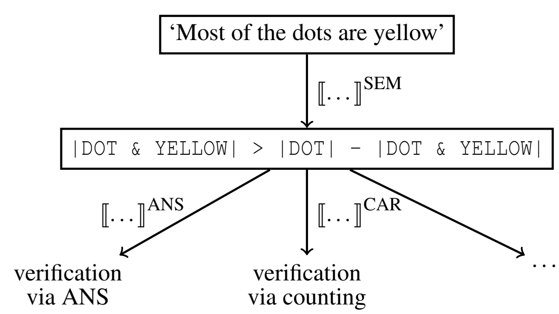


Figure 3: Overview of two-step interpretation of the expression 'Most of the dots are yellow'

Interpretation by the ANS

How does the ANS interpret a LoT expression? Our proposed "ANS interpretation", written $[\![\dots]\!]^{\text{ANS}}$, is given in (8). This

[5] More generally, the project as a whole raises interesting questions about the relationship between the compositional details at each step. But those are empirical questions. Logically, what matters is that the outputs of the first step are the inputs to the second.

interpretation specifies how the ANS will interact with m, as a function of m's parts and how they are put together. We write $G(\mu, \sigma^2)$ for an ANS representation modeled by a gaussian with mean μ and variance σ^2.

(8) a. $[\![\texttt{|DOT|}]\!]^{\text{ANS}}$ is the ANS representation of the numerosity of the dots

b. $[\![\texttt{|DOT \& YELLOW|}]\!]^{\text{ANS}}$ is the ANS representation of the numerosity of the yellow dots

c. $[\![e_1 \texttt{ - } e_2]\!]^{\text{ANS}}$ is $G(n_1 - n_2, \sigma_1^2 + \sigma_2^2)$, where

$$[\![e_1]\!]^{\text{ANS}} = G(n_1, \sigma_1^2) \text{ and } [\![e_2]\!]^{\text{ANS}} = G(n_2, \sigma_2^2)$$

d. $[\![e_1 \texttt{ > } e_2]\!]^{\text{ANS}}$ is $\frac{1}{2}\operatorname{erfc}\left(\frac{n_2 - n_1}{\sqrt{2}\sqrt{\sigma_1^2+\sigma_2^2}}\right)$, where

$$[\![e_1]\!]^{\text{ANS}} = G(n_1, \sigma_1^2) \text{ and } [\![e_2]\!]^{\text{ANS}} = G(n_2, \sigma_2^2)$$

The formula in the last clause of this definition is drawn from a standard model of the ANS (Pica et al., 2004, supplementary materials): it computes the probability of judging a numerosity represented by $G(n_1, \sigma_1^2)$ greater than one represented by $G(n_2, \sigma_2^2)$. The interpretation of m therefore turns out to be a real number in the range from 0 to 1, which is understood as the probability of arriving at a judgement that the sentence is true.

(8) is the glue linking m in (5) to the experimental findings. Consider the scenario that yielded the observed response patterns at Box **(B)** of Figure 2. Since the information made available by a 200ms display of a scene like Figure 1a includes ANS representations of the numerosity of the dots and of the yellow dots, (8) can be used to evaluate m in this scenario, as shown in (9). The first two lines simply draw on the available information indicated in Figure 2 (center white box), and subsequent lines follow the instructions for using this information as specified by (8).

(9)

$$\begin{aligned}
[\![\texttt{|DOT \& YELLOW|}]\!]^{\text{ANS}} &= G(y, w^2y^2) \\
[\![\texttt{|DOT|}]\!]^{\text{ANS}} &= G(y+b, w^2(y+b)^2) \\
[\![\texttt{|DOT| - |DOT \& YELLOW|}]\!]^{\text{ANS}} &= G(y+b-y, w^2(y+b)^2 + w^2y^2)) \\
&= G(b, w^2(y+b)^2 + w^2y^2)) \\
[\![\texttt{|DOT \& YELLOW| > |DOT| - |DOT \& YELLOW|}]\!]^{\text{ANS}} & \\
&= \tfrac{1}{2}\operatorname{erfc}\left(\frac{b-y}{\sqrt{2}\sqrt{w^2((y+b)^2+2y^2)}}\right)
\end{aligned}$$

The last line is a prediction (that the studies above have already borne out) about participants' responses as a function of the number of yellows and blues (y and b) and their ANS constant of proportionality (w).

One-to-one information is neglected

Recall the first crucial finding: participants did not recruit one-to-one correspondence information, even in scenarios where control studies confirmed its availability (Box **(A)** of Fig. 2). Why did subjects draw on the available ANS representations rather than draw on this one-to-one correspondence? Our answer is simply that there is no interpretation of m which draws on this information, and so there is no route from one-to-one information to a complete evaluation of m.[6]

Of course, nothing interesting that we have said so far derives this conclusion. For now, our main aim is simply to develop a way of thinking about meaning and verification so that these sorts of facts can be cataloged at all. Once a reasonable body of such facts have been collected and cataloged in a unified framework, we will be able to seek generalisations and principles in the normal way.

Interpretation by precise counting

So far, the only non-linguistic cognitive system to interpret m is the ANS. We split the relevant path into two—from linguistic representation to LoT, then from LoT to ANS—precisely to accommodate alternative interpretations of LoT that capture the fact that other systems can be used instead.

Obviously, such an alternative interpretation of LoT is required to capture the fact that, under more accommodating circumstances, speakers will judge the truth or falsity of (1) via precise counting. This clearly does not invoke the ANS, but rather (we will assume) some other system that deals with cardinal numbers and arithmetic. The interpretation $[\![\ldots]\!]^{\text{CAR}}$ of m into this system might be defined as in (10).[7]

(10) a. $[\![\texttt{DOT}]\!]^{\text{CAR}}$ is the property of being a dot

b. $[\![\texttt{YELLOW}]\!]^{\text{CAR}}$ is the property of being yellow

c. $[\![e_1 \texttt{ \& } e_2]\!]^{\text{CAR}}$ is the property of having both the property $[\![e_1]\!]^{\text{CAR}}$ and the property $[\![e_2]\!]^{\text{CAR}}$

d. $[\![\texttt{|}e\texttt{|}]\!]^{\text{CAR}}$ is the number of things with property $[\![e]\!]^{\text{CAR}}$

e. $[\![e_1 \texttt{ - } e_2]\!]^{\text{CAR}}$ is the difference between $[\![e_1]\!]^{\text{CAR}}$ and $[\![e_2]\!]^{\text{CAR}}$

f. $[\![e_1 \texttt{ > } e_2]\!]^{\text{CAR}}$ is true if $[\![e_1]\!]^{\text{CAR}}$ is greater than $[\![e_2]\!]^{\text{CAR}}$, and false otherwise

Given both of $[\![\ldots]\!]^{\text{CAR}}$ and $[\![\ldots]\!]^{\text{ANS}}$, speakers asked to verify (1) are free, in general, to choose between two "extensionally equivalent" procedures, subject of course to constraints on which primitive pieces of information are available. Given an unlimited amount of time in experiments otherwise like those we've discussed, however, people could take the time to count, or rely on ANS representations that would be automatically triggered in any case.

'Most' neglects the direct-comparison procedure

Consider now the relationship between Box **(B)** and Box **(C)** in Figure 2: the scene displayed, and therefore the available information, is the same, and yet participants responded with lower accuracy with 'most' than with 'more'. To get different response patterns, we must specify some LoT expression

[6] Of course there will be other LoT expressions that can be interpreted in ways that draw on one-to-one information; what we are saying is just that m, the expression which is the meaning of (1), is not one of them.

[7] Since we don't have much to say here about how sentence meanings interact with the specifics of this system, this interpretation boils down to something essentially equivalent to the interpretation of the mathematical set-language used above in (2) and (4).

(distinct from m) as the meaning of the 'more' sentence. Suppose it is that in (11) (compare (5)).

(11) $[\![$'More of the dots are yellow than blue'$]\!]^{\text{SEM}}$
$= \texttt{|DOT \& YELLOW| > |DOT \& BLUE|}$

Using the interpretation given already in (8), the ANS evaluation of (11) can be described as in (12).

(12)
$$[\![\texttt{|DOT \& YELLOW|}]\!]^{\text{ANS}} = G(y, w^2y^2)$$
$$[\![\texttt{|DOT \& BLUE|}]\!]^{\text{ANS}} = G(b, w^2b^2)$$
$$[\![\texttt{|DOT \& YELLOW| > |DOT \& BLUE|}]\!]^{\text{ANS}} = \tfrac{1}{2}\,\mathrm{erfc}\left(\frac{b-y}{\sqrt{2}\sqrt{w^2y^2+w^2b^2}}\right)$$

The difference between this and the way the ANS is used in (9) is only in the value which is compared with $[\![\texttt{|DOT \& YELLOW|}]\!]^{\text{ANS}}$. Here it is compared with a primitive ANS representation of the numerosity of the blue dots, with variance w^2b^2; in (9) it is a noisier representation of the same numerosity (its mean is still b), with variance $w^2((y+b)^2+y^2)$. Using this representation yields a noisier/less accurate pattern of judgements of truth or falsity, even though the "perfect" response pattern is identical in both cases.

To illustrate the distinct predicted responses more precisely, note that in each case the predicted probability of a 'yes' response still depends only on the ratio of yellow to blue dots: letting $r = \frac{y}{b}$, the relevant formulas from (9) and (12) can be rewritten in terms of only r:

$$[\![\texttt{|DOT \& YELLOW| > |DOT| - |DOT \& YELLOW|}]\!]^{\text{ANS}} = \tfrac{1}{2}\,\mathrm{erfc}\left(\frac{1-r}{\sqrt{2}w\sqrt{(r+1)^2+2r^2}}\right)$$

$$[\![\texttt{|DOT \& YELLOW| > |DOT \& BLUE|}]\!]^{\text{ANS}} = \tfrac{1}{2}\,\mathrm{erfc}\left(\frac{1-r}{\sqrt{2}w\sqrt{r^2+1}}\right)$$

Plotting these two functions of the ratio r then yields a flatter, noisier curve in the case of 'most', as shown in Fig. 4.

There is a seemingly problematic aspect of what we have said so far: if the difference between Box **(B)** and Box **(C)** arises from 'most' (but not 'more') forcing the indirect procedure with subtraction, shouldn't we expect people to perform this subtraction even when precise counting is recruited for evaluating a 'most' statement? This is implausible for an adult carefully inspecting a scene with 51 yellow dots and 49 blue dots. We don't have a concrete account of how it is that speakers can "ignore", in some cases but not others, the verificational slant that comes with a particular sentence meaning. But notice that our proposal is not that meanings simply are verification procedures (nor sets of candidate verification procedures) — the proposal is that meanings are structured enough to allow for certain verification procedures, such as the subtraction procedure in the cast of 'most', to have a distinguished status that sets them apart from truth-conditionally equivalent alternatives. This seems to be necessary for any explanation of the constrained flexibility that we observe, even if there must be room for pragmatics to dictate that what matters is whether the statement is true, rather than how that conclusion is reached. We must, for example, allow speakers to verify 'Most birds can fly' by looking in an encyclopaedia.

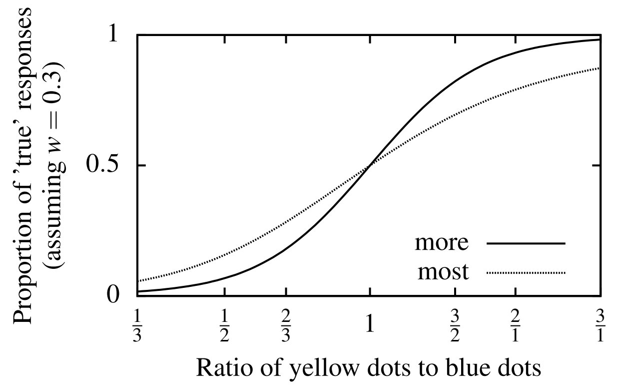


Figure 4: Predicted responses based on (9) ('most') and (12) ('more'). Compare with Fig. 4B of Knowlton et al. (2021)

.

Issues for development

So far, we have specified a kind of cognitive architecture that can support explanations for certain non-obvious findings in "psychosemantics". We hope that precisifying this architecture will generate discussion and debate about the details. Indeed, formal clarity often helps to put further questions into stark relief; we explore three of these in what follows.

Beyond ANS and cardinalities

As linguists are well aware, 'most' (along with many other functional vocabulary items plausibly invoking 'degrees' or 'measures') has many uses that are not plausibly about number, cardinality, or numerosity at all: (13a) is plausibly about area, (13b) is plausibly about volume, and it isn't at all obvious what should be said about the dimensionality of (13c).

(13) a. Most of the wall is yellow.
b. Most of the water is brown.
c. Most of the idea is half-baked.

Sticking to one of the relatively concrete cases, if we assume that (13a) has the same LoT interpretation in (5)—just with `WALL` replacing `DOT`—we are immediately presented with the task of working out yet further interpretations of LoT expressions, e.g. ones involving non-countable notions of extent. This will require spelling out some details of how the relevant non-linguistic systems (might) work,[8] just as (8) spells

[8] There is another possibility. We do not have good grounds from linguistic analysis to think that 'most' is ambiguous. However, perhaps we should be thinking about just one approximate magnitude system (Lourenco, 2014) for its meaning to interact with, rather than

out certain details of the ANS and (10) spells out certain details of the precise counting systems.

Inferences and constraints on possible interpretations

Symbols like '>' and '−' that we are using in LoT expressions were introduced as completely fresh symbols whose interpretation was up to us to specify (see fn. 4). This came at a cost: by moving away from the standard logical machinery often used to describe the meanings of natural language sentences, we have lost any account of intuitive entailment relations between sentences. We would like $[\![$'Most of the dots are yellow'$]\!]^{\text{SEM}}$ to bear some relation to $[\![$'Some of the dots are yellow'$]\!]^{\text{SEM}}$ that it doesn't bear to, say, $[\![$'Ten of the dots are yellow'$]\!]^{\text{SEM}}$. Whereas the expressions of first-order logic, for example, have an inference relation already defined on them, no such relation has been defined for expressions of our hypothesized LoT.

A second issue which remains to be resolved is the question of what rules out the possibility of having $[\![\ldots]\!]^{\text{ANS}}$ as defined above in (8) as one interpretation of LoT, and having as another $[\![\ldots]\!]^{\text{CAR*}}$, which is like $[\![\ldots]\!]^{\text{CAR}}$ except that $[\![e_1 \ \texttt{>} \ e_2]\!]^{\text{CAR*}}$ is true iff the integer $[\![e_1]\!]^{\text{CAR*}}$ is *less than* the integer $[\![e_2]\!]^{\text{CAR*}}$. The $[\![\ldots]\!]^{\text{ANS}}$ and $[\![\ldots]\!]^{\text{CAR}}$ interpretations both treat '>' in a way that somehow corresponds to the intuitive notion of "greater than", but the combination of $[\![\ldots]\!]^{\text{ANS}}$ and $[\![\ldots]\!]^{\text{CAR*}}$ breaks this correspondence.[9]

The two issues just raised — the need to account for entailments, and the issue of consistency across interpretations — can be seen as two sides of the same coin. One way to home in on what it means for a non-linguistic system to interpret symbols like '>' and '−' "in the right way" is to identify some of the kinds of inferences that we would like to be valid at the level of LoT expressions. For example, a hypothesized inference schema over LoT expressions such as

(14) $$\frac{e_1 \ \texttt{>} \ e_2 \qquad e_2 \ \texttt{>} \ e_3}{e_1 \ \texttt{>} \ e_3}$$

can also be understood as a requirement that interpretations of LoT must meet: whatever system-specific interpretation is assigned to the symbol '>', it must be some transitive binary relation (albeit perhaps some kind of "noisy" one, as with the ANS). This alone would not rule out the incongruous $[\![\ldots]\!]^{\text{CAR*}}$ interpretation considered above, because "less than" satisfies this requirement; to distinguish $[\![\ldots]\!]^{\text{CAR*}}$ from $[\![\ldots]\!]^{\text{CAR}}$ (and abstracting away from the possibility of equality), we might impose another schema such as (15) (with no premises), constraining the interaction of '>' with '&'.

(15) $$\frac{}{\texttt{|}e_1\texttt{|} \ \texttt{>} \ \texttt{|}e_1 \ \texttt{\&} \ e_2\texttt{|}}$$

A natural goal then would be to assemble inference schemas along these lines, from which it would be possible to prove concrete instances such as (16). The empirically observed entailment relationship between 'Most of the dots are yellow' and 'Some of the dots are yellow' would then follow if the meaning of the latter were as shown in (17).

(16) $$\frac{\texttt{|DOT \& YELLOW| > |DOT| - |DOT \& YELLOW|}}{\texttt{|DOT \& YELLOW| > 0}}$$

(17) $[\![$'Some of the dots are yellow'$]\!]^{\text{SEM}}$
$= \texttt{|DOT \& YELLOW| > 0}$

Compositional grain size across the different interpretations of LoT

A possibly odd difference between the interpretation $[\![\ldots]\!]^{\text{CAR}}$ in (10) and the interpretation $[\![\ldots]\!]^{\text{ANS}}$ in (8) is that the former provides compositional "steps" for $[\![e_1 \ \texttt{\&} \ e_2]\!]^{\text{CAR}}$ and for $[\![\texttt{|}e\texttt{|}]\!]^{\text{CAR}}$, whereas the ANS interpretation does not; instead, it provides only "all at once" steps for larger expressions such as $[\![\texttt{|DOT \& YELLOW|}]\!]^{\text{ANS}}$. It is not obvious to us, yet, what ANS-specific interpretations should be assigned to subexpressions such as `DOT`, `YELLOW` and `DOT & YELLOW` which do not correspond to numerosities.

Conclusion

Careful consideration of sentences of the form 'Most of the dots are yellow' reveals a network of inter-related facts which, taken together, pose interesting questions about the way linguistic meanings interact with the other cognitive systems that can be deployed for the purposes of verification. The meaning of such a sentence must be flexible enough to interface with, at least, two distinct number systems (one precise and one approximate), and yet rigid enough to constrain the particular way in which at least one of these systems is used. And indeed, the meaning must be rigid enough to (at least sometimes) also rule out the use of detected one-to-one correspondences that are deeply related to cardinalities.

Our suggestion is that these facts can be accommodated by postulating that the sentence's meaning is a structured expression which can be independently compositionally interpreted by a number of distinct cognitive systems. Specifying these compositional interpretations has the consequence of specifying what kinds of information must be available to a user of the sentence in order for a particular verification system to be viable. We have shown that it is possible to specify these compositional interpretations in ways that help make sense of speakers' otherwise-surprising choices of which verification system to use under what circumstances.

We have left many issues open. There are important questions about some finer points of how the proposed two-step framework should be set up; in part, these are questions about how the experimental research that has driven this paper can be fully integrated with mainstream semantic research. Other findings concerning the verification profile of expressions like 'most' that can also provide the impetus for further progress along the trajectory set out here.

---

e.g. the ANS, the Approximate Area System (AAS), etc. This is an empirical matter to be determined by cognitive psychology, however; the present model can accommodate either eventuality.

[9] This can be seen as an instance of the very general question of what makes representations "about" things (e.g. Yablo 2014): intuitively, what is wrong with $[\![\ldots]\!]^{\text{CAR*}}$ is that it disregards the way an LoT expression like (5) is about some general notion of number.

Acknowledgments

Thanks to Justin Halberda, Tyler Knowlton, Jeff Lidz, Darko Odic and Paul Pietroski for many valuable discussions, over many years, about the issues addressed in this paper.

References

Fodor, J. A. (1975). *The language of thought*. New York: Thomas Y. Crowell.

Hackl, M. (2009). On the grammar and processing of proportional quantifiers: *most* versus *more than half*. *Natural Language Semantics*, *17*, 63–98.

Heim, I., & Kratzer, A. (1998). *Semantics in Generative Grammar*. Malden, MA: Blackwell.

Knowlton, T., Hunter, T., Odic, D., Wellwood, A., Halberda, J., Pietroski, P., & Lidz, J. (2021). Linguistic meanings as cognitive instructions. *Annals of the New York Academy of Sciences*, *1500*, 134–144.

Lidz, J., Halberda, J., Pietroski, P., & Hunter, T. (2011). Interface transparency and the psychosemantics of 'most'. *Natural Language Semantics*, *19*(3), 227–256.

Lourenco, S. F. (2014). On the relation between numerical and non-numerical magnitudes: Evidence for a general magnitude system. In D. Geary, D. Berch, & K. Mann Koepke (Eds.), *Evolutionary Origins and Early Development of Number Processing* (Vol. 1, p. 145-174). Academic Press.

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, *306*, 499–503.

Pietroski, P., Lidz, J., Hunter, T., & Halberda, J. (2009). The meaning of 'most': Semantics, numerosity and psychology. *Mind and Language*, *24*(5), 554–585.

Rescorla, M. (2019). *The Language of Thought Hypothesis* (Summer 2019 ed.). The Stanford Encyclopedia of Philosophy. Retrieved from `https://plato.stanford.edu/archives/sum2019/entries/language-thought/`

Sella, F., Slusser, E., Odic, D., & Krajcsi, A. (2021). The emergence of children's natural number concepts: Current theoretical challenges. *Child Development Perspectives*, *15*(4), 265-273.

Winskel, G. (1993). *The formal semantics of programming languages*. Cambridge, MA: MIT Press.

Yablo, S. (2014). *Aboutness*. Princeton, NJ: Princeton University Press.