Chapter 1

Splitting atoms in natural language

Andreas Haida
The Hebrew University of Jerusalem

Tue Trinh
ZAS Berlin

The classic Fregean analysis of numerical statements runs into problems with sentences containing non-integers such as **John read 2.5 novels**, since it takes such statements to specify the cardinality of a set which by definition must be a natural number. We propose a semantics for numeral phrases which allows us to count mereological subparts of objects in such a way as to predict several robust linguistic intuitions about these sentences. We also identify a number of open questions which the proposal fails to address and hence must be left to future research.

Keywords: Numerals, Measurement, Density, Scales, Implicatures

1 A new semantics for numeral phrases

1.1. Standard analyses of numerical statements have roots in Frege (1884) and take these to be, essentially, predications of second order properties to concepts, that is specifications of cardinalities. Thus, the sentence

(1)   John read 3 novels

is considered to be a claim about the set of novels that John read, namely that it has three members. The truth condition of (1) is taken to be either (2a) or (2b), depending on whether the 'exact' or the 'at least' meaning is assumed to be basic for numerals.[1]

---

[1] For arguments that numerals have the "at least" meaning as basic see Horn (1972); von Fintel & Heim (1997); von Fintel & Fox (2002); Fox (2007), among others. For arguments that numerals

(2)   a. $|\{x \mid x \text{ is a novel} \wedge \text{John read } x\}| = 3$
       b. $|\{x \mid x \text{ is a novel} \wedge \text{John read } x\}| \geq 3$

Let us now consider (3), which we take to be an expression that is accepted as a well-formed sentence of English.

(3)   John read 2.5 novels

Extending the traditional analysis of numerical statements to this sentence yields absurdity: (4a) is a contradiction, and (4b) is logically equivalent to (2b).[2]

(4)   a. $|\{x \mid x \text{ is a novel that John read}\}| = 2.5$
       b. $|\{x \mid x \text{ is a novel that John read}\}| \geq 2.5$

It is obvious that (3) is neither contradictory nor equivalent to (1). Suppose, for example, that John read *Brothers Karamazov*, *Crime and Punishment*, one-half of *Demons*, and nothing else.[3] In this context, (3) is true and (1) false. The fact that

---

have the 'exact' meaning as basic see Geurts (2006); Breheny (2008), among others. Note that the choice between these two views does not affect what we say in this paper, as will be clear presently.

[2]See Salmon (1997) for related discussion. Note, however, that Salmon is concerned with expressions like **2 ½ oranges**, which he pronounces "two and one-half oranges," while we are dealing with expressions like **2.5 novels** ("two point five novels"). We don't intend to suggest that the two types of expressions should receive the same analysis. In this connection, we should point out that an anonymous reviewer raises the issue of whether some of our judgements might be an artefact of this pronunciation, e.g. of pronouncing **2.5** as "two point five" instead of "two and a half." This issue, we must admit, goes beyond the scope of our paper. We would note, however, that independently of how the issue is settled empirically, the fact that it is raised might be symptomatic of a worry which, as we surmised from various discussions, is shared by a number of colleagues. The worry is that we are not investigating "natural language," but instead, are ruminating on some sort of conventional discourse which has been manufactured for the special purpose of making conversation in mathematics more expedient. A question which we have heard more than once is "what about languages spoken by communities which have no mathematics at this level?" We believe the worry is unfounded. It is true that we have to learn how to write and pronounce decimals, but the linguistic judgements involving these expressions which we present and try to account for below do not come about by way of instruction. In fact, these intuitions should be surprising given the definitions we learn in school. As for the question about languages without expressions for decimals, we would say that our study is similar in kind to one of, say, the Vietnamese pronominal system which can express many distinctions that are not lexically encoded in English. Speech communities may differ, due to historical accidents, in how they lexicalize conceptual space, i.e. in what they can say, but this is of course no reason for assuming that research into language particular phenomena does not inform our understanding of what they *could* say, i.e. of universal grammar.

[3]John is a Dostoyevsky enthusiast.

(3) can be true shows that it is not contradictory, and the fact that it can be true while (1) is false shows that the two sentences are not equivalent.[4]

1.2. This squib proposes an analysis of numeral phrases which can account for intuitions about such sentences as (3). First, we will assume the logical form of **John read 2.5 novels** to be (5), where SOME and MANY are covert (cf. Hackl 2000).[5]

(5)
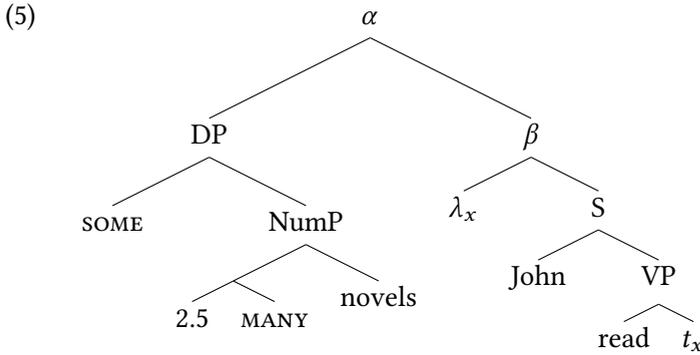
Our proposal will consist in formulating a semantics for MANY, leaving other elements in (5) with their standard meaning.[6] This semantics presupposes the fairly standard view of the domain of individuals, $\mathcal{D}_e$, as a set partially ordered by the part-of relation $\sqsubseteq$ to which we add $\emptyset$ as the least element (cf. Link 1983; Landman 1989; Schwarzschild 1996; Bylinina & Nouwen 2018).[7] The join operation $\sqcup$ and the meet operation $\sqcap$ on $\langle \mathcal{D}_e \cup \{\emptyset\}, \sqsubseteq \rangle$ are defined in the usual way.[8]

---

[4]For Frege, the concept of a "concept" entails, as a matter of logic, that it has sharp boundary: "[...] so wird ein unscharf definirter Begriff mit Unrecht Begriff genannt [...] Ein beliebiger Gegenstand Δ fällt entweder unter den Begriff Φ, oder er fällt nicht unter ihn: tertium non datur" (Frege 1893: §56). Thus, there is no sense in which we can "extend" Frege's theory to include non-integers: the number of objects which fall under a concept must be a whole number. In fact, Frege considers the reals to be of a different metaphysical category from the naturals, and even made the distinction notationally explicit, writing "2" for the real number two and "2" for the natural number two (Snyder 2016; Snyder & Shapiro 2016).

[5]Although we reference Hackl (2000), we should note that existential quantification, i.e. the meaning of SOME, is included in the definition of Hackl's MANY. We thank an anonymous reviewer for reminding us to mention this difference.

[6]In particular, we assume that the covert SOME has the same meaning as its overt counterpart, which is $[\![\text{SOME}]\!] = [\![\textbf{some}]\!] = [\lambda P \in \mathcal{D}_{\langle e,t \rangle} . [\lambda Q \in \mathcal{D}_{\langle e,t \rangle} . \exists x . P(x) = Q(x) = 1]]$.

[7]Note that we don't assume that $\emptyset$ is an element of $\mathcal{D}_e$ itself. However, we also don't exclude this possibility.

[8]Here are the definitions, where ι represents, following standard practice, the function mapping a singleton set to its unique element.

(i) a. $x \sqcup y := \iota\{z \mid x \sqsubseteq z \wedge y \sqsubseteq z \wedge \forall z'(x \sqsubseteq z' \wedge y \sqsubseteq z' \to z \sqsubseteq z')\}$

We assume that plural nouns denote cumulative predicates, i.e. subsets of $\mathcal{D}_e$ which are closed under $\sqcup$ (cf. Krifka 1989; Chierchia 1998; Krifka 2003; Sauerland et al. 2005; Spector 2007; Zweig 2009; Chierchia 2010). For each predicate A, the set of A atoms, A_{at}, is defined as

(6) $A_{at} := \{x \in A \mid \neg \exists y . y \sqsubset x \wedge y \in A\}.$

To illustrate, let b and c be the two novels *Brothers Karamazov* and *Crime and Punishment*, respectively. The individual $b \sqcup c$ has proper parts that are novels, hence $b \sqcup c$ will not be in $[\![\textbf{novels}]\!]_{at}$. In contrast, neither b nor c has proper parts that are novels, hence both of these individuals are in $[\![\textbf{novels}]\!]_{at}$. In other words, $[\![\textbf{novels}]\!]_{at}$ contains things that we can point at and say "that is a novel." The semantics we propose for MANY is (7), where d ranges over degrees.

(7) $[\![\text{MANY}]\!](d)(A) = [\lambda x \in \mathcal{D}_e . \mu_A(x) \geq d]$

We then predict that **John read 2.5 novels** is true iff there exists an individual x such that $\mu_{[\![\textbf{novels}]\!]}(x) \geq 2.5$ and John read x. The term $\mu_{[\![\textbf{novels}]\!]}(x)$ represents "how many novels are in x," so to speak. We want to be able to count novels in such a way that proper subparts of novels, which are not novels, also contribute to the count. To this end, we propose to explicate the measure function μ_A as follows.[9]

(8) $\mu_A(x) = \begin{cases} \mu_A(y) + 1 & \text{if } a \sqsubset x, y \sqcup a = x, \text{ and } y \sqcap a = \emptyset \text{ for some } a \in A_{at} \\ \mu_a(x) & \text{if } x \sqsubseteq a \text{ for some } a \in A_{at} \\ \# & \text{otherwise} \end{cases}$

Thus, each A atom which is a subpart of x will add 1 to $\mu_A(x)$. If x is an A atom or a subpart of an A atom, $\mu_A(x)$ will be $\mu_a(x)$, which represents "how much of the A atom a is in x," so to speak.[10] The measure function μ_a is explicated as

---

b. $x \sqcap y := \iota\{z \mid z \sqsubseteq x \wedge z \sqsubseteq y \wedge \forall z'(z' \sqsubseteq x \wedge z' \sqsubseteq y \rightarrow z' \sqsubseteq z)\}$

[9]Salmon (1997) tentatively suggests to analyze "2½" by means of the quantifier '2.5' in a logical form like '2.5$x(x$ is an F that is G)'. This quantifier is characterized as a "mixed-number quantifier," operating on pluralities, where the quantity of a plurality is measured in such a way that whole Fs count as one and "a part of a whole F counts for part of a whole number." Our proposal can be seen as an order-theoretic specification of such a quantifier.

[10]It follows from the definition in (8) that $\mu_A(x)$ is undefined (for all x) if A_{at} is empty. An anonymous reviewer raises the concern that this definition might exclude the denotation of count nouns like **fence** from being measured by μ, the problem being that fences are homogeneous entities. That is, the concern is that $[\![\textbf{fence}]\!]_{at} = \emptyset$ and, consequently, that $\mu_{[\![\textbf{fence}]\!]}(x) = \#$. We

follows.

(9)   For each $a \in A_{at}$,

         a. μ_a is a surjection from $\{x \in \mathcal{D}_e \mid x \sqsubseteq a\}$ to $(0, 1] \cap \mathbb{Q}$

         b. $\mu_a(x \sqcup y) = \mu_a(x) + \mu_a(y)$ for all $x, y \in \mathrm{dom}(\mu_a)$ such that $x \sqcap y = \emptyset$

         c. $\mu_a(a) = 1$

This definition allows us to use any positive rational numbers smaller or equal to 1 to measure parts of an atom, with 1 being the measure of the whole atom.[11] Furthermore, it guarantees that the measurement of parts of an atom is additive: if x and y are non-overlapping parts of an atom, their mereological sum $x \sqcup y$ measures the arithmetic sum of the measurements of x and y. Thus, two chapters, chapters 1 and 2, of a novel cannot be added to two chapters, chapters 2 and 3, of the same novel to give four chapters of that novel because of the overlap.[12]

---

hypothesize that measuring this type of noun requires contextual restriction: if C is a syntactic variable and $[\![\mathbf{fence}_C]\!]^g = [\![\mathbf{fence}]\!]^g \cap g(C)$, then $[\![\mathbf{fence}_C]\!]^g_{at} \neq \emptyset$ iff $[\![\mathbf{fence}]\!]^g \cap g(C)_{at} \neq \emptyset$; consequently, $\mu_{[\![\mathbf{fence}_C]\!]}(x)$ is defined if (and only if) $[\![\mathbf{fence}]\!]^g \cap g(C)_{at} \neq \emptyset$ (for certain x). Thus, we surmise that sentences like **Ann passed by 3 fences** or **Ann painted 3.5 fences** presuppose a context in which fences aren't homogeneous entities but maximal stretches of fence, such as the whole stretch of a fence around a property or along a border. Thus, we agree with Wągiel (2018) that counting can involve a notion of 'maximality.' However, we put forth the hypothesis that maximality only comes into play through contextual restriction, in the absence of atoms in the unrestricted extension of a noun.

[11]The need for non-integral counting in natural language has been recognized. Kennedy (2015), for example, says the following about #, the measure function which maps objects to number: "Note that # is not, strictly speaking, a cardinality function, but rather gives a measure of the size of the (plural) individual argument of the noun in "natural units" based on the sense of the noun [...]. If this object is formed entirely of atoms, then # returns a value that is equivalent to a cardinality. But if this object contains parts of atoms, then # returns an appropriate fractional or decimal measure [...]" (Kennedy 2015: footnote 1). However, this is all Kennedy says about the matter. In particular, he does not explicate what he means by "appropriate," and is not concerned with the data that we present below.

The notion of "natural units" refered to by Kennedy in the quote above is due to Krifka (1989), who proposes a function, NU, which maps a predicate P and an object x to the number of natural units of P in x. Like Kennedy, Krifka does not consider the data presented in the next section, and neither does he provide a definition of NU which is explicit enough to relate to them. In fact, Krifka stipulates that NU is an "extensive measure function," on the model of such expressions as **litter of**, which means he actually makes the wrong prediction for the data point presented in 2.2. below. Specifically, Krifka will predict that (11b) must be contradictory as (11a) is.

Thus, what we are doing here is essentially improving upon Kennedy and Krifka, with the improvement being explication in the former and explication as well as correction in the latter case.

[12]Note that overlap is dealt with twice in our definitions, viz. in the first clause of (8), to prevent

2 Some predictions of the proposal

This section presents some intuitions about numerical statements which are predicted by our semantics for MANY. The list is not intended to be exhaustive.

2.1. We predict the observation made at the beginning of this paper, namely that (10a) is neither contradictory nor equivalent to (10b).

(10)   a.   John read 2.5 novels
       b.   John read 3 novels

This is because $\mu_{\llbracket \mathbf{novels} \rrbracket}(x) \geq 2.5$ is neither contradictory nor equivalent to $\mu_{\llbracket \mathbf{novels} \rrbracket}(x) \geq 3$. To see that $\mu_{\llbracket \mathbf{novels} \rrbracket}(x) \geq 2.5$ is not contradictory, let b, c, and d be, again, the three novels *Brothers Karamazov*, *Crime and Punishment*, and *Demons*, respectively, and let d' be a subpart of *Demons* which measures one-half of this novel, so that $\mu_{\llbracket \mathbf{novels} \rrbracket}(d') = \mu_d(d') = 0.5$. Then, $\mu_{\llbracket \mathbf{novels} \rrbracket}(b \sqcup c \sqcup d') = \mu_{\llbracket \mathbf{novels} \rrbracket}(c \sqcup d') + 1 = \mu_{\llbracket \mathbf{novels} \rrbracket}(d') + 1 + 1 = \mu_d(d') + 1 + 1 = 0.5 + 1 + 1 = 2.5$. The non-equivalence follows from the logical truth that $2.5 < 3$ and the fact that there is an x such that $\mu_{\llbracket \mathbf{novels} \rrbracket}(x) = 2.5$ (as shown above).

2.2. We predict that (11a) is a contradiction but (11b) is not.

(11)   a.   #John read 1 Dostoyevsky novel yesterday, and 1 Tolstoy novel today, but he did not read 2 Russian novels in the last two days
       b.   John read 0.5 Dostoyevsky novels yesterday, and 0.25 Tolstoy novels today, but he did not read 0.75 Russian novels in the last two days

The first conjunct of (11a)[13] requires two different novels, say b and c, to have been read by John. As $\mu_{\llbracket \mathbf{novels} \rrbracket}(b \sqcup c) = 2$, the second conjunct of (11a) contradicts the first. On the other hand, suppose John read a subpart of b, call it b', yesterday and read a subpart of c, call it c', today, and suppose that b' measures

---

atoms from being counted more than once, and in (9b), to do the same for subatomic parts. This is in line with the claim that subatomic quantification is subject to the same constraints as quantification over wholes (Wągiel 2018; 2019). However, we are not committed to all aspects of Wągiel's theory. Specifically, we see reason to reject his claim that counting (of atoms and subatomic parts) requires 'topological integrity.' It seems to us that the sentence **John owns 2 cars** can be much more readily accepted as true if John owns (nothing but) a whole car and a car that is sitting disassembled in various places in his garage than the sentence **John owns two cups** if he owns (nothing but) a whole cup and the shards of a shattered cup. While some notion of 'integrity' might play into this contrast, we believe that the way this notion enters is by affecting, dependent on context, what is considered a possible extension of the nouns **car** and **cup** in the actual world. A more thorough comparison of our proposal to Wągiel's theory is beyond the bounds of this article but we believe that the two proposals are largely compatible.

[13]Here and below, we refer to the conjuncts of **but**.

one-half of b and c' measures one-quarter of c, i.e. $\mu_{[\![\mathbf{novels}]\!]}(b') = \mu_b(b') = 0.5$ and $\mu_{[\![\mathbf{novels}]\!]}(c') = \mu_c(c') = 0.25$. Then the first conjunct of (11b) is true. However, b' and c', put together, do not make up something which has a subpart that is a novel, or something which is a subpart of a novel. In other words, there is no $a \in [\![\mathbf{novels}]\!]_{at}$ such that $a \sqsubset b' \sqcup c'$ or $b' \sqcup c' \sqsubseteq a$, which means $\mu_{[\![\mathbf{novels}]\!]}(b' \sqcup c') = \#$, which means $\mu_{[\![\mathbf{novels}]\!]}(b' \sqcup c') \ngeq 0.75$, which means the second conjunct of (11b) is true.[14]

2.3. We predict that (12) is a tautology.

(12)   If John read 0.75 novels, and Mary read the rest of the same novel that John was reading, then Mary read 0.25 novels

Suppose John read a portion of b, call it b', which measures three-fourth of b, so that $\mu_{[\![\mathbf{novels}]\!]}(b') = \mu_b(b') = 0.75$. Suppose, furthermore, that Mary read the rest of b, call it b'', which is all of that part of b which John did not read. Then the antecedent is true. Now by hypothesis, $b' \sqcup b'' = b$, and $b \in [\![\mathbf{novels}]\!]_{at}$. This means $\mu_b(b' \sqcup b'') = \mu_b(b) = 1$. Since b' and b'' do not overlap, i.e. $b' \sqcap b'' = \emptyset$, we have $\mu_b(b' \sqcup b'') = \mu_b(b') + \mu_b(b'') = 1$. And because $\mu_b(b') = 0.75$, we have $\mu_b(b'') = 1 - 0.75 = 0.25$, hence $\mu_{[\![\mathbf{novels}]\!]}(b'') = 0.25$, which means the consequent is true.

2.4. We predict that (13) is not a contradiction.

(13)   John read 0.5 novels, and Mary read 0.25 of the same novel that John was reading, but John and Mary together did not read 0.75 novels

Suppose John read b' which measures 0.5 of b, and Mary read b'' which measures 0.25 of b. Thus, $\mu_b(b') = 0.5$ and $\mu_b(b'') = 0.25$. The first conjunct is then true. Now let b' and b'' overlap, so that $b' \sqcap b'' \neq \emptyset$. Furthermore, let o be $b' \sqcap b''$ and

---

[14]Note that our prediction in this case differs from that of Liebesman (2016), who would predict that **John read 0.75 novels** is true in the described context, since Liebesman's proposal, according to our understanding, would allow subparts of different novels to be added, as long as the sum is smaller than 1. Furthermore, judgments might be different for an example like (i), which seems to have a contradictory reading.

(i)   #John ate 0.5 oranges yesterday, and 0.25 oranges today, but he did not eat 0.75 oranges (or more) in the last two days

We believe that the difference between (11b) and (i) comes down to the fact that **orange** can be more easily coerced to a mass interpretation than **novel** (cf. **The smoothie contains orange** vs. **#The shredder bin contains novel**). To accommodate the contradictory reading of (i), we tentatively assume that $[\![\mathbf{oranges}]\!]$ can be contextually extended by sums of subparts of different oranges.

d' and d'' the non-overlapping parts of b' and b'', respectively. Thus, $b' = d' \sqcup o$, $b'' = d'' \sqcup o$, and $b' \sqcup b'' = d' \sqcup d'' \sqcup o$. This means $\mu_b(b' \sqcup b'') = \mu_b(d' \sqcup d'' \sqcup o) = \mu_b(d') + \mu_b(d'') + \mu_b(o) < \mu_b(d') + \mu_b(o) + \mu_b(d'') + \mu_b(o) = \mu_b(d' \sqcup o) + \mu_b(d'' \sqcup o) = \mu_b(b') + \mu_b(b'') = 0.5 + 0.25 = 0.75$, which means $\mu_b(b' \sqcup b'') < 0.75$, which means hence the second conjunct is true.

2.5. We predict that (14a) is coherent, but (14b) is not.[15]

(14)   a.   John read (exactly) 0.5 novels
       b.   #John read (exactly) 0.5 quantities of literature

That (14a) is coherent is, by now, obvious. It will be true if John read, say, half of *Anna Karenina*. What makes (14b) incoherent, then, must lie in the semantics of **quantities of literature**, henceforth **qol** for short. According to the semantics we proposed for MANY, (14b) entails the existence of an individual x such that $\mu_{\llbracket \mathbf{qol} \rrbracket}(x) = 0.5$, which entails the existence of some $a \in \llbracket \mathbf{qol} \rrbracket_{at}$ such that $x \sqsubseteq a$. Given that any subpart of a quantity of literature is itself a quantity of literature, we have $\llbracket \mathbf{qol} \rrbracket_{at} = \{x \in \llbracket \mathbf{qol} \rrbracket \mid \neg \exists y \sqsubset x \wedge y \in \llbracket \mathbf{qol} \rrbracket\} = \emptyset$. Thus, there is no $a \in \llbracket \mathbf{qol} \rrbracket_{at}$, which means there is no x such that $\mu_{\llbracket \mathbf{qol} \rrbracket}(x) = 0.5$, which means (14b) is false. Furthermore, it is analytically false, which is to say false by virtue of the meaning of the word **quantity**. This, we hypothesize, is the reason for its being perceived as deviant. We will come back to this point in the last section.

2.6. We predict (15), which we claim to be a fact about natural language.

(15)   There is no numerical gap in the scale which underlies measurement in natural language

What (15) is intended to say, illustrated by a concrete example, is that to the extent **John read 2.5 novels** is meaningful, **John read 2.55 novels** is too, as well as **John read 2.555 novels**, or any member of $\{$**John read n novels** $\mid \llbracket \mathbf{n} \rrbracket \in \mathbb{Q}^+\}$.[16] This follows from the fact that 0.5, as well as 0.55, as well as 0.555, as well as any other rational number in $(0, 1] \cap \mathbb{Q}$, are all in the range of μ_a, for any $a \in \llbracket \mathbf{novels} \rrbracket_{at}$. This fact, in turn, follows from the fact that μ_a is, by stipulation, a function onto

---

[15]Note that the word **quantity** in (14b) is not intended to mean '200 pages,' or '3000 words,' or any contextually specified quantity of literature. The intended meaning of **quantity** here is the lexical and context-independent one.

[16]Where $\mathbb{Q}^+$ are the positive rationals. Thus, (15) should really be qualified with the phrase "as far as rational numbers are concerned," as pointed out by an anonymous reviewer, who raises the issues of irrational numbers. We refer the reader to subsection 4.5. for more discussion on this point. Here we would only note that by "meaningful," we mean the sentence has non-trivial truth condition, and licenses inferences, as shown for **John read 2.5 novels** in the last section.

$(0, 1] \cap \mathbb{Q}$. Note, importantly, that we cannot guarantee (15) by stipulating, merely, that the set of degrees underlying measurement in natural language is dense. To see that density alone does not exclude gaps, consider the set in (16).

(16) $S := \mathbb{Q}^+ \setminus \{x \in \mathbb{Q} \mid 3 < x \leq 4\}$

This is a dense scale, as between any two elements of S there is an element of S. However, S contains a gap: missing from it are numbers greater than 3 but not greater than 4, for example 3.5. Merely stipulating that the scale is dense, therefore, will not guarantee that **John read 3.5 novels** is meaningful, which we claim is a robust intuition that linguistic theory has to account for.[17]

2.7. On the assumption that overt **many** and the comparative **more** instantiate MANY, we predict that the argument expressed by the sequence in (17) is is invalid.

(17)   John read 2.5 novels and Mary read 2 novels. #Therefore, John read more novels than Mary.

By the definitions in (7) and (8), the scale $[\lambda x \lambda d. \llbracket \text{MANY} \rrbracket (d)(\llbracket \textbf{novels} \rrbracket)(x)]$ is non-monotonic.[18] For instance, if b' is half of *Brothers Karamazov* and c' half of *Crime and Punishment*, then $[\lambda x \lambda d. \llbracket \text{MANY} \rrbracket (d)(\llbracket \textbf{novels} \rrbracket)(x)](b')(0.5) = 1$ but $[\lambda x \lambda d. \llbracket \text{MANY} \rrbracket (d)(\llbracket \textbf{novels} \rrbracket)(x)](b' \sqcup c')(0.5) = 0$. Therefore, (17) is not valid, since it would only be valid if the scale were monotonic, i.e. were a scale of comparison (Wellwood et al. 2012).

This is illustrated in (18). The temparature scale is non-monotonic. Hence, the temperature scale cannot function as the scale of comparison of the comparative in the second sentence of (18a). Therefore, the sequence of the two sentences in (18a) is an invalid argument. The weight scale, in contrast, is monotonic. Hence, the weight scale can function as the scale of comparison of the comparative in the

---

[17]Fox & Hackl (2006), according to our understanding, seems to assume that density of a scale alone guarantees the absense of gaps in it. They claim, for example, that density guarantees that exhaustification of **John has more than 3 children** would negate every element of {**John has more than n children** | $\mathbf{n} \in \mathbb{Q} \wedge \mathbf{n} > 3$}. We quote from page 543 of Fox & Hackl (2006): "Without the UDM [i.e. the assumption that the set of degrees is dense], [...] [t]he set of degrees relevant for evaluation would be, as is standardly assumed, possible cardinalities of children (i.e. 1, 2, 3, ...). The sentence would then assert that John doesn't have more than 4 children [...] If density is assumed, however, [...] the assertion would now not just exclude 4 as a degree exceeded by the number of John's children. It would also exclude any degree between 3 and 4." Taken at face value, this claim is wrong, as is evident from the example in (16).

[18]Let S be a scale, conceived of as a function from entities and degrees to truth values, such that for all x the degree function $S(x)$ is monotonic (i.e. such that $S(x)(d) \rightarrow S(x)(d')$ for all d, d' such that $d' \leq d$). Then, the scale S is monotonic iff $S(x)(d) = 1 \rightarrow S(x')(d) = 1$ for all d and x, x' such that $x \sqsubseteq x'$ (cf. Krifka 1989; Schwarzschild 2002).

second sentence of (18b), as evidenced by the validity of the argument expressed by (18b).

(18)   a.   John ate 90 degree hot spaghetti and Mary 70 degree hot spaghetti. #Therefore, John ate more spaghetti than Mary.
       b.   John ate 500 grams of spaghetti and Mary ate 200 grams of spaghetti. Therefore, John ate more spaghetti than Mary.

To account for the fact that the arguments in (19) are valid, we tentatively assume that MANY can be restricted to atoms and sums of atoms in equatives and comparatives.

(19)   a.   John read 3.5 novels and Mary read 2 novels. Therefore, John read more novels than Mary.
       b.   John read 2.5 novels and Mary read 2 novels. Therefore, John read as many novels as Mary.

This means to say that the scale of comparison of **more than**/**as many as** in (19) is the monotonic scale $[\lambda x \in [\![\textbf{novels}]\!]_{at}^{\sqcup}.\lambda d.\ [\![\text{MANY}]\!](d)([\![\textbf{novels}]\!])(x)]$ (where $A_{at}^{\sqcup}$ is the closure of A_{at} under the join operation).

3 Excursus: conditions on predicates

The semantics we propose for MANY, as presented in (7), (8) and (9), requires that for each atom a of a predicate A the measure function μ_a have $(0,1] \cap \mathbb{Q}$ as its range, and be additive with respect to non-overlapping subparts of atoms.

(20)   Conditions on μ_a
       a. $\text{ran}(\mu_a) = (0,1] \cap \mathbb{Q}$
       b. $\mu_a(x \sqcup y) = \mu_a(x) + \mu_a(y)$ if $x, y \sqsubseteq a$ and $x \sqcap y = \emptyset$

This section details the conditions under which such measure functions μ_a exist. While it is possible to derive empirical predictions from these conditions (see footnote 20 below), which could have been added to section 2, the main purpose of the current section is to tie in our proposal with a general theory of measurement. Conditions on the existence of measure functions μ_a of the right kind are conditions on subsets A of $\mathcal{D}_e$ with $A_{at} \neq \emptyset$ such that for each $a \in A_{at}$ there is a function μ_a that satisfies (20a) and (20b). Call such subsets of $\mathcal{D}_e$ "measurable predicates."

Let A be an arbitrary subset of $\mathcal{D}_e$ such that $A_{at} \neq \emptyset$. The first assumption

we need to make for A to be a measurable predicate is that all of its atoms are divisible into arbitrarily many discrete parts.[19, 20] This is stated in (21), where $\mathcal{P}_a := \{x \in \mathcal{D}_e \mid x \sqsubseteq a\}$.

(21)   For all $a \in A_{at}$ and $n \in \mathbb{N}$, there is a set $S \subseteq \mathcal{P}_a$ such that $|S| = n$, $\bigsqcup S = a$, and $\bigsqcap S' = \emptyset$ for all $S' \subseteq S$ with $|S'| > 1$

It follows from (21) that no A atom a has a smallest part, and also, that there is no smallest difference between two parts of a. This condition is necessary to guarantee that the range of a measure function μ_a can be the rational interval $(0, 1] \cap \mathbb{Q}$, as demanded in (20a).

The second and final assumption we need to make about a measurable predicate A is that its atoms satisfy the condition in (22).

(22)   For all $a \in A_{at}$, $\langle \mathcal{P}_a, \sqsubseteq \rangle$ is a σ-algebra on $\langle \mathcal{D}_e \cup \{\emptyset\}, \sqsubseteq \rangle$[21]

σ-algebras are well-known structures of measure theory (see e.g. Cohn 1980) which guarantee, in our case, that the parts of an entity a are measurable in the sense of there being a function μ_a that satisfies (20a) and (20b). In simple words, what we require with (22) is that each $a \in A_{at}$ satisfy the following conditions: (i) the set of parts of a contains a greatest element (trivially satisfied, since a is a part of itself); (ii) for every (proper) part of a, there is another part of a, discrete from the first, such that the two parts together are a; and (iii) countably many parts of a joined together are a part of a. We add another condition to make sure that counting the atoms of a member x of a measurable predicate A is consistent

---

[19]It seems that a stricter condition might be desirable, viz. that every entity is arbitrarily divisible into discrete parts. However, such a condition would afford a notion of *possible division* of an entity and it is doubtful whether such a notion can be defined independently of the partial order $\langle \mathcal{D}_e \cup \{\emptyset\}, \sqsubseteq \rangle$.

[20]There are predicates whose members withstand being conceived of as being (arbitrarily) divisible. For example, it is hard to conceive of partial results of an achievement. Correspondingly, combining nominalizations of achievement verbs with non-integer nominals leads to deviance:

(i)   a.   #Ann fired 3.5 shots.
       b.   #Bob witnessed 1.5 arrivals.

[21]A partial order $\langle A, \sqsubseteq \rangle$ is a σ**-algebra** on a lower bounded partial order $\langle B, \sqsubseteq \rangle$, with $A \subseteq B$, iff (i) it is upper bounded, (ii) closed under complementation, and (iii) closed under countable joins, where $\langle B, \sqsubseteq \rangle$ is **lower bounded** iff $\bigsqcap B \in B$, and $\langle A, \sqsubseteq \rangle$ is **upper bounded** iff $\bigsqcup A \in A$, **closed under complementation** iff for all $x \in A$ there is a $y \in A$ such that $x \sqcup y = \bigsqcup A$ and $x \sqcap y = \bigsqcap B$, and **closed under countable joins** iff for all countable subsets S of A it holds that $\bigsqcup S \in A$.

with measuring all of its subatomic parts. For this to be the case, the atoms of A must be pairwise discrete from each other, as stated in (23).

(23)   For all $a, b \in A_{at}$, if $a \sqcap b \neq \emptyset$ then $a = b$

4 Open questions

We end with some open questions for future research. Again, the list below is not intended to be exhaustive.

4.1. Concepts – The semantics we propose for MANY predicts the contrast between (14a) and (14b), repeated in (26a) and (26b) below, because it entails that to be half an A is to be half an A atom. This semantics, as it is, makes the wrong prediction that (24) is false.[22]

(24)   The *Unvollendete* is 0.5 symphonies

Let u be the *Unvollendete*. From (8) and (9), it follows that $\mu_{[\![\textbf{symphonies}]\!]}(u) \neq 0.5$, as there is no $a \in [\![\textbf{symphonies}]\!]_{at}$ such that $u \sqsubseteq a$. Obviously, modality is involved: while there is no singular symphony s such that $\mu_s(u) = 0.5$, there could be one, since the last two movements could have been completed. Thus, counting symphonies seems to be about what could be a symphony, not what is actually a symphony. In other words, it is concepts, not predicates, that seem to be at play. This means we should, perhaps, revise our semantics so as to predict that to be half an A is to be half of something which is an A atom in some possible world. There is a possible world, say one where Schubert died at 41 instead of 31, in which the *Unvollendete* is part of a whole symphony, and this is what makes (24) true. However, we do not want to predict, incorrectly, that (25) is true, for example.

(25)   *Crime and Punishment* is 0.5 symphonies

Thus, while there certainly is a possible world w in which *Crime and Punishment* is a subpart of a symphony, we want w to be inaccessible from the world of evaluation. Plausibly, specifying the relevant accessibility relation in this particular

---

[22]According to an anonymous reviewer, this prediction is not wrong. Specifically, the reviewer says that s/he sees the Unvollendete not as half of a symphony, but as a symphony, hence finds (24) to be false. We are not sure to what extent this opinion of the Unvollendete can be accounted for within a semantic theory of numerals. Our point concerns the problems faced by our account given the understanding that the Unvollendete is not a whole symphony, i.e. is "unvollendet." That there is a different understanding is orthogonal to the discussion.

case amounts to fleshing out the concept of "symphony," and specifying it in the general case, to fleshing out the concept of "concept." We leave this task to future work.

4.2. Analyticities – Suppose John read one quarter of *Brothers Karamazov* and one quarter of *Crime and Punishment*, our semantics of MANY predicts, correctly, that neither (26a) nor (26b) is true.

(26)   a.   John read 0.5 novels
       b.   #John read 0.5 quantities of literature

Both sentences claim of something, which does not exist, that John read one-half of it: in the case of (26a), a novel which contains parts of both *Brothers Karamazov* and *Crime and Punishment*, and in the case of (14b), an quantity of literature which contains no subpart that is also an quantity of literature. Our semantics, however, does not predict the contrast in acceptability between (26a) and (26b): while the former is perceived as false, the latter is perceived as deviant. In §2.5, we said that this contrast has to do with analyticity: it lies in the meaning of the word **quantity** that any subquantity is a quantity, while nothing in the meaning of **novel** rules out a novel which contains parts of both *Brothers Karamazov* and *Crime and Punishment*. Analyticity has been appealed to in explanations of deviance (cf. Barwise & Cooper 1981; von Fintel 1993; Krifka 1995; Abrusán 2007). However, it has been pointed out that all analyticities are not equal: both (27a) and (27b) are analytically false, but only the latter is deviant.[23]

(27)   a.   Some bachelor is married
       b.   #Some student but John smoked

Gajewski (2003) proposes that the kind of analyticity which leads to deviance is "L-analyticity." Thus, while (27a) is analytically false, (27b) is L-analytically false, and therefore is deviant. Discussing Gajewski's notion of L-analyticity will take us beyond the scope of this squib. Hence, we will leave to future research the question whether, and if yes how, sentences such as (26b) can be considered L-analytical.

---

[23]Assuming that (27a) has the truth condition in (ia) (cf. Heim & Kratzer 1998) and (27b) the truth condition in (ib) (cf. von Fintel 1993).

(i)   a. $\{x \mid x \in [\![\textbf{bachelor}]\!] \land x \in [\![\textbf{married}]\!]\} \neq \emptyset$
      b. $\{x \mid x \in [\![\textbf{student}]\!] \land x \notin \{\text{John}\} \land x \in [\![\textbf{smoked}]\!]\} \neq \emptyset \land$
           $\forall P(\{x \mid x \in [\![\textbf{student}]\!] \land x \notin P \land x \in [\![\textbf{smoked}]\!]\} \neq \emptyset \rightarrow \{\text{John}\} \subseteq P)$

4.3. Countabilities – Words such as **quantity** have been analyzed as a sort of "classifier" which turns a [−count] noun into a [+count] one (cf. Chierchia 2010). This analysis is motivated by such contrasts as that in (28).

(28)   a.   #The vampire drank 2 bloods
     b.   The vampire drank 2 quantities of blood

Since **blood** is a [−count], it cannot be counted. On the other hand, **quantity of blood** is [+count], therefore it can be. However, such contrasts as that between (28b) and (29), to the best of our knowledge, has not been paid attention to.

(29)   #The vampire drank 2.3 quantities of blood

The semantics we proposed for MANY, unfortunately, makes no distinction between (28b) and (29): both are predicted to be analytically false. The proposal thus shares with several others the shortcoming of not being able to differentiate between subtypes of [+count] noun phrases. The task remains, therefore, of refining the semantics of MANY as to predict the contrast in question.

It should be noted, in addition, that words like **quantity**[24] may pose a challenge for the theory of measurement proposed in Fox & Hackl (2006). These authors derive the fact that (30a) does not license the scalar implicature (30b)

(30)   a.   The vampire drank more than 2 quantities of blood
     b.   ¬The vampire drank more than 3 quantities of blood

from the assumption that the scale mates of **2**, for the deductive system (DS) which computes scalar implicatures, are not the set of natural numbers, but the set of rational numbers. The proposal, therefore, claims that (29) is a scalar alternative of (30a) (see note 17). To the extent that the deviance of (29) is due to this sentence being deemed ill-formed by the DS itself (see Gajewski 2003; Fox & Hackl 2006, and the discussion in the previous subsection), the question arises as to whether DS uses a sentence which it deems ill-formed in its computation. Again, we leave this topic to future work.

4.4. Morphology – The plural vs. singular distinction in number marking languages has usually been considered to mirror the bare vs. classified distinction

---

[24]These include **amount** and **fraction**, among possibly others.

(i)   a.   #The vampire drank 2.3 amounts of blood
     b.   #The vampire ate 2.3 fractions of the apple

in classifier languages (cf. Chierchia 1998; Cheng & Sybesma 1999). Specifically, plural/bare nouns have been analyzed as denoting "number-neutral" predicates, i.e. sets containing both singularities and pluralities, while singular/classified nouns have been analyzed as denoting "atomic" predicates, i.e. sets containing only singularities. However, with respect to numerical statements involving non-integers in English, a number marking language, and Vietnamese, a classifier language, the correlation falls apart: what is obligatory is a plural noun in English and a classified noun in Vietnamese.

(31)   a.   John ate 0.5 cake-*(s)
       b.   John ăn 0.5 *(cái) bánh
            John ate 0.5 *(CL) cake

We know of no account for this fact, and leave an investigation of it for future research.

4.5. Reals – We have been assuming that the set of numbers underlying measurement in natural language is $\mathbb{Q}$, the set of rationals. But what prevents us from assuming that it is in fact $\mathbb{R}$, the set of reals? Clearly, that assumption will be true to the extent that sentences containing reals which are not rationals are meaningful. Is (32) meaningful?

(32)   John ate π (many) cakes

We have no clear intuition about (32). A confounding factor for such examples as (32) might be that "π" is too "artificial" to be perceived as part of natural language. One might, then, imagine an experiment along the following lines. Let *ABC* be a cirle on which lie the three points *A*, *B*, and *C*. Let *AB* be the diameter of *ABC*. Now suppose a mathematican, say Euclid, uttering the sentence in (33).

(33)   If *AB* is one novel, then *ABC* is how many novels John read

Obviously, there is no natural language numeral **n** such that Euclid's thought can be expressed as **John read n novels**. The question is whether this thought is, nevertheless, representable by grammar, or more specifically DS, and thus plays a role in inferences such as scalar implicatures (see subsection 4.3). We leave this question to future research.

Andreas Haida & Tue Trinh

Acknowledgements

References

Abrusán, Martha. 2007. *Contradiction and Grammar: the Case of Weak Islands*. MIT dissertation.

Barwise, Jon & Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4(2). 159–219.

Breheny, Richard. 2008. A new look at the semantics and pragmatics of numerically quantified noun phrases. *Journal of Semantics* 25(2). 93–139.

Bylinina, Lisa & Rick Nouwen. 2018. On "zero" and semantic plurality. *Glossa: a journal of general linguistics* 3(1). 1–23.

Cheng, Lisa Lai Shen & Rint Sybesma. 1999. Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry* 30(4). 509–542.

Chierchia, Gennaro. 1998. Reference to kinds across languages. *Natural Language Semantics* 6(4). 339–405.

Chierchia, Gennaro. 2010. Meaning as an inferential system: polarity and free choice phenomena. Harvard University.

Cohn, David L. 1980. *Measure theory*. Boston: Birkhäuser.

Fox, Danny. 2007. *Pragmatics in Linguistic Theory*. MIT classnotes.

Fox, Danny & Martin Hackl. 2006. The universal density of measurement. *Linguistics and Philosophy* 29. 537–586.

Frege, Gottlob. 1884. *Die Grundlagen der Arithmetik*. Verlage Wilhelm Koebner.

Frege, Gottlob. 1893. *Grundgesetze der Arithmetik, Band 2*. Verlag Hermann Pohle.

Gajewski, Jon. 2003. *L-analyticity in natural language*. Unpublished manuscript.

Geurts, Bart. 2006. Take 'five'. In Svetlana Vogeleer & Liliane Tasmowski (eds.), *Non-definiteness and Plurality*, 311–329. Benjamins.

Hackl, Martin. 2000. *Comparative quantifiers*. Massachusetts Institute of Technology dissertation.

Heim, Irene & Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Oxford: Blackwell.

Horn, Laurence. 1972. *On the semantic properties of the logical operators in English.* UCLA dissertation.

Kennedy, Christopher. 2015. A "de-Fregean" semantics (and neo-Gricean pragmatics) for modified and unmodified numerals. *Semantics and Pragmatics* 8(10). 1–44.

Krifka, Manfred. 1989. Nominal reference, temporal constitution and quantification in event semantics. In R. Bartsch, J. van Benthem & P. von Emde Boas (eds.), *Semantics and contextual expression*, 75–115. Dordrecht: Foris Publication.

Krifka, Manfred. 1995. The semantics and pragmatics of polarity items. *Linguistic Analysis* 25(3-4). 209–257.

Krifka, Manfred. 2003. Bare NPs: kind-referring, indefinites, both, or neither? *Proceedings of SALT* 13. 180–203.

Landman, Fred. 1989. Groups, I. *Linguistics and philosophy* 12(5). 559–605.

Liebesman, David. 2016. Counting as a type of measuring. *Philosopher's Imprint* 16(12).

Link, Godehard. 1983. The logical analysis of plurals and mass terms: a lattice-theoretical approach. In Rainer Baeuerle, Christoph Schwarze & Arnim von Stechow (eds.), *Meaning, Use, and the Interpretation of Language*, 302–323. Berlin: de Gruyter.

Salmon, Nathan. 1997. Wholes, parts, and numbers. *Noûs Supplement: Philosophical Perspectives, 11, Mind, Causation, and World* 31. 1–15.

Sauerland, Uli, Jan Anderssen & Kazuko Yatsushiro. 2005. The plural is semantically unmarked. In Stephan Kesper & Marga Reis (eds.), *Linguistic evidence.* Mouton de Gruyter.

Schwarzschild, Roger. 1996. *Pluralities.* Springer.

Schwarzschild, Roger. 2002. Singleton indefinites. *Journal of Semantics* 19(3).

Snyder, Eric. 2016. *Counting and Other Forms of Measurement.* The Ohio State University dissertation.

Snyder, Eric & Stewart Shapiro. 2016. Frege on the Real Numbers. In P Ebert & M Rossberg (eds.), *Essays in Frege's Basic Laws of Arithmetic.* Oxford University Press.

Spector, Benjamin. 2007. Aspects of the pragmatics of plural morphology: On higher-order implicatures. In Uli Sauerland & Penka Stateva (eds.), *Presupposition and Implicature in Compositional Semantics*, 243–281. Palgrave-Macmillan.

von Fintel, Kai. 1993. Exceptive constructions. *Natural Language Semantics* 1(2). 123–148.

von Fintel, Kai & Danny Fox. 2002. *Pragmatics in Linguistic Theory*. MIT Class-notes.

von Fintel, Kai & Irene Heim. 1997. *Pragmatics in Linguistic Theory*. MIT class-notes.

Wągiel, Marcin. 2018. *Subatomic quantification*. Masaryk University in Brno dissertation.

Wągiel, Marcin. 2019. Partitives, multipliers and subatomic quantification. In M.Teresa Espinal, Elena Castroviejo, Manuel Leonetti, Louise McNally & Cristina Real-Puigdollers (eds.), *Proceedings of Sinn und Bedeutung 23*, vol. 2, 445–462. Universitat Autònoma de Barcelona.

Wellwood, A., V. Hacquard & R. Pancheva. 2012. Measuring and comparing individuals and events. *Journal of Semantics* 29. 207–228.

Zweig, Eytan. 2009. Number-neutral bare plurals and the multiplicity implicature. *Linguistics and philosophy* 32(4). 353–407.