Division vs. Distributivity: Is *per* just like *each*?*

Elizabeth Coppock
Boston University

Abstract This paper argues that there are lexical items that conventionally express the idea of dividing one quantity by another, and *per* is one of them. In particular, the proposal is that there are three ratio-related senses of *per*: (i) a quotient function; (ii) a quotient operator; and (iii) quotient of measure functions. The ratio-based approach, which is built up here in order to handle a wider range of data than previous ratio-based approaches could, is contrasted with an opposing view, one on which *per* is a distributivity marker like *each*. Four types of evidence are used: (i) cases involving measurement of an object or an event whose measure is smaller than the unit given by *per*'s complement; (ii) uses in the differential argument of a comparative; (iii) uses modifying a measure function noun; and and (iv) uses modifying a gradable predicate. All of these are problematic for a distributivity-marker analysis, and support the idea that *per* expresses the concept of ratio. Along the way, we gain diagnostics for whether a given item conventionally expresses the concept of a ratio in a given language.

Keywords: degree semantics, distributivity, ratio markers, quantity division

1 Introduction

Quantities like 10 kg and 2 hours (or as linguists call them, "degrees") can be added, subtracted, multiplied, and divided. Are there lexical items of natural language that conventionally express the notion of quantity division, i.e. ratio? One prominent candidate is *per*, which has been claimed to introduce the denominator of a ratio (Rawlins 2013; Coppock 2021).[1] But is *per* really a ratio marker or is it a distributivity marker like *each*? Consider the following sentence, for example:

1 Other phenomena that have been analyzed using the idea of dividing one quantity by another include (i) the word *average* (Kennedy & Stanley 2009); (ii) proportional readings of quantity words like *many* and *few* (Solt 2009; Bale & Schwarz 2019, 2020); and (iii) proportional measure nouns like *percent* and *thirds* (e.g. Ahn & Sauerland 2017; Sauerland & Pasternak 2022).
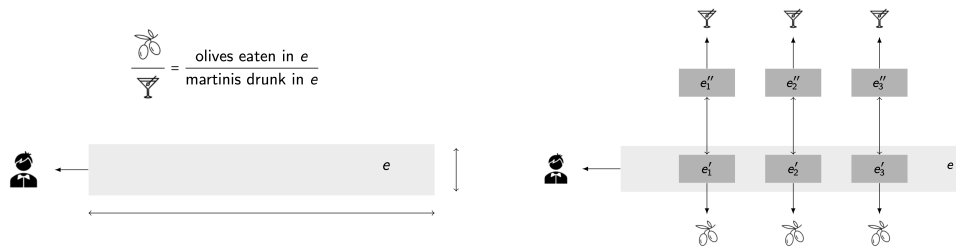
Figure 1   Ratio-based (left) vs. distributivity (right) analyses of *James Bond ate two olives per martini*. Variables e, e'_i, ... stand for events.

(1)   James Bond ate two olives per martini.   (Panaitescu & Tovena 2019)

What sort of event is being described here? On a ratio-based analysis (see Figure 1, left), (1) says: the measure of the event in the 'olives eaten' dimension, divided by its measure on a dimension measured in martinis (presumably a count of the martinis *drunk*), is equal to 'two olives divided by one martini'.

If, on the other hand, *per* is a distributivity marker, then it contributes universal quantification over a 'sorting key' determined by its complement. Indeed, (1) can be paraphrased with the help of universal quantifiers, as in the following:

(2)   a.   For every martini, James Bond ate two olives.
       b.   Every time he drank a martini, he ate two olives.

Sentences of the form *For every A there is a B* are analyzed by Boolos (1981) in terms of a one-to-one correspondence. Rothstein (1995) proposes further that this correspondence is between two sets of events, and applies this idea to *every time* constructions as in (2-b) as well. In their analysis of *per*-like items in Italian and Romanian (*per* and *de*, respectively), Panaitescu & Tovena (2019) (P&T) build on this idea of a one-to-one correspondence among events. A P&T-style analysis of (1) is illustrated in on the righthand panel of Figure 1; it says, 'James Bond is the agent of an event composed of subevents each connected to two olives and another event involving one martini'. Their analysis is presented in more detail in Section 3, after a few touch-ups to the ratio-based approach in Section 2 that are necessary in order to handle (1).

In Section 4, I offer four empirical arguments in favor of a ratio-based approach. Evidence comes from: (i) speed uses, as in *drive 100 mph* (ii) uses as the differential argument of a comparative, as in *2 mph faster* (iii) uses modifying a measure function noun, as in *price per pound*, and (iv) uses modifying a gradable adjective, as in *cheaper per citizen*. I conclude that there is at least one lexical item that

conventionally encodes the concept of ratio, namely *per*. Along the way, we gain diagnostics that can be used to distinguish distributivity markers from ratio markers.

2   Ratio-based approach

2.1   Quantity calculus

The ontological foundations for degree semantics have furnished a relation of comparison among degrees (Cresswell 1977), addition among degrees (Klein 1991), and multiplication of a number by a degree (Sassoon 2010; van Rooij 2011), but not multiplication and division of arbitrary quantities, until recently. *Quantity calculus*, an algebraic branch of the field of metrology, provides foundations for addition and subtraction among degrees ("quantities") of the same kind (e.g. two weights), along with multiplication and division among degrees that may be of different kinds.[2] Following Coppock (2021), to establish foundations for quantity division, let us adopt Raposo's (2019) *dimension-centric* approach to quantity calculus, where one starts with a finite set of basic dimensions $\mathscr{B}$, including for example T (time) and L (distance), which can be multiplied together to form derived dimensions. The full set of dimensions $\mathscr{D}$ forms a group under multiplication ($\cdot$) with identity element $\mathbf{1}_{\mathscr{D}}$. For any dimension D, its *multiplicative inverse*, written D^{-1}, has the property that $D \cdot D^{-1} = \mathbf{1}_{\mathscr{D}}$. In other words, it is possible to divide by any dimension. So the full space of dimensions includes complex ones like $\mathsf{L} \cdot \mathsf{T}^{-1}$ (distance over time).

Every quantity Q has a designated dimension $\dim(\mathsf{Q}) \in \mathscr{D}$. A particular speed, for example, will have $\mathsf{L} \cdot \mathsf{T}^{-1}$ as its dimension. Of particular note are 'dimensionless quantities', which include ratios of two quantities of the same dimension, like a weight divided by a weight, and cardinalities, which are expressed as numbers without an accompanying unit. These quantities have dimension $\mathbf{1}_{\mathscr{D}}$.

In Raposo's (2019) system, the set of quantities forms a *fiber bundle*, with a fiber for each dimension. Within each dimension D, the set of quantities of that dimension can not only be multiplied but also added, and there is an additive identity element $\mathbf{0}_D$. More precisely, $\mathscr{Q}_D$ forms a vector space, like the real numbers. In fact, $\mathscr{Q}_{\mathbf{1}_{\mathscr{D}}}$ is the real numbers itself. While addition is only defined for quantities of the same dimension, any two quantities can be multiplied together (symbolized $\times$), regardless of dimension. The multiplicative identity element, common to all dimensions, is 1, drawn from the real numbers. Non-zero quantities q have multiplicative inverses q^{-1} such that $q \times q^{-1} = 1$. In other words, it is possible to divide by quantities. As an alternative notation, $d \times q^{-1}$ can be written $\frac{d}{q}$.

Finally, while the dim function maps quanties to dimensions, the unit function goes the other way, designating a unit quantity for every dimension. The unit must be

---

2 See de Boer (1994) for an overview of its two-century history.

non-zero (not $\mathbf{0}_D$ for any dimension D), and the mapping must satisfy the condition that for any dimensions D and D', $\mathrm{unit}(D) \times \mathrm{unit}(D') = \mathrm{unit}(D \cdot D')$.

For the purposes of natural language semantics, these ontological assumptions will be built into the models for our representation language, a typed lambda calculus with basic types e, v, d, and t (Coppock 2021). A model M will determine not only a set of entities $\mathcal{A}$ and a set of events $\mathcal{V}$, but also a set of dimensions $\mathcal{D}_{\mathcal{B}}$ paired with the multiplication operation $\cdot$, a set of quantities $\mathcal{Q}$ paired with the multiplication operation $\times$ and the addition operation $+$, a unit mapping from dimensions to quantities, a dim mapping from quantities to dimensions, and an interpretation function I that maps non-logical constants of the representation language to denotations of the appropriate semantic type, built on these sets. I assume further that the representation language $\mathcal{L}_{\mathcal{Q}}$ is equipped with logical constants denoting the operations and mappings from quantity calculus.

2.2 Quotient function analysis

With these foundations, *per* can be analyzed as a function that takes two quantities (a.k.a. "degrees") and gives back their quotient (Coppock 2021):

(3) **Quotient function analysis of *per***
$per_0 \rightsquigarrow \lambda d \lambda q . \frac{q}{d}$ $\hfill$ (type $\langle d, \langle d, d \rangle \rangle$)

For example, *60 kilometers per hour* would denote the result of dividing the quantity '60 kilometers' by the quantity '1 hour'. Assuming that unit nouns (a.k.a. "measure nouns") are type d lets us derive the analysis in (4) for *60 kilometers per hour*.

(4)

$$
\begin{array}{c}
d \\
\frac{60 \times \mathsf{km}}{\mathsf{hour}}
\end{array}
$$

$$
\begin{array}{cc}
d & \langle d, d \rangle \\
60 \times \mathsf{km} & \lambda q . \frac{q}{\mathsf{hour}} \\
\textit{60 kilometers} &
\end{array}
$$

$$
\begin{array}{cc}
\langle d, \langle d, d \rangle \rangle & d \\
\lambda d \lambda q . \frac{q}{d} & \mathsf{hour} \\
per_0 & \textit{hour}
\end{array}
$$

Note that the result is a complex degree-denoting term, an expression of type d.

As Coppock (2021) shows, this lexical entry can be used in conjunction with some basic assumptions about measurement to derive simple word problem-like inferences, like the following:

(5)   a.   James Bond drove 60 kilometers per hour for a quarter of an hour.
        b.   Therefore, James Bond drove 15 kilometers [in that event].

But to give a ratio-based analysis of (1), it is necessary to remedy two shortcomings of the quotient function analysis.

2.3   Quotient operator analysis

In (1), the object of the verb *eat* is not a particular ratio (two olives divided by one martini); olives can be eaten but olive-to-martini ratios cannot be. Although the *per*-phrase modifies *two olives*, it escapes thematic assignment by the verb. The fact that *per*-phrases can escape thematic role assignment by their local verb can be seen even more clearly in cases where their host is in subject position:

(6)   It's estimated that 150 species per year go extinct.

The predicate *go extinct* does not apply to the result of dividing '150 species' by 'one year'. Ratios don't go extinct; species do. The absurdity of the following argument brings out this fact:[3]

(7)   a.   150 species per year go extinct.
        b.   150 species per year is a high rate.
        c.   #Therefore, a high rate is among those going extinct.

What (6) expresses is that for some event e, the number of species that go extinct in e divided by the duration of e is equal to the result of dividing 150 by one year.

    Truth conditions entailing this ratio claim are obtained from the following **quotient operator** analysis, On this analysis, *per* combines not only with a denominator and a numerator but also a gradable predicate of events, such as: for a given event e, how many species go extinct in e.

(8)   **Quotient operator analysis of *per***

$$per_1 \rightsquigarrow \lambda d_d \lambda q_d . \lambda G_{\langle d,vt \rangle} . \lambda e . \frac{\max(\lambda d . G(d)(e))}{\mu_{\dim(d)}(e)} = \frac{q}{d}$$

In the example under consideration, and in the usual case, this gradable predicate will be constructed through quantifier raising of a complex operator consisting of *per*, its complement (the denominator), and the numerator. After this complex

---

3 The argument in (7-c) bears some structural similarities to Partee's temperature paradox example: *The temperature is 90. The temperature is rising. #Therefore 90 is rising.* The similarities are merely superficial. Example (7-c) does not contain any intensional predicate like *rise*, and while *The temperature is 90* is an equative statement, *150 species per year is a high rate* is a predicative satetment. Thanks to Amy Rose Deal and Enrico Flor for discussion of this point.

operator moves from its base position into a position beside a corresponding lambda abstraction node, it can combine via function application with an argument of the appropriate type, provided that the landing site for QR is right above an event description (type *vt*).

This lexical entry gives appropriate truth conditions for (6), as shown in the following tree. Here I assume that *per day* forms a unit with *150* at LF, even though *per day* is realized on the surface at the right edge of the noun phrase.[4]

(9)

$$\lambda e \, . \, \frac{\max(\lambda d \, . \, \mathsf{ge}(e) \wedge \exists x [\mathsf{th}(e) = x \wedge {}^{*}\mathsf{sp}(x) \wedge \mu_D(x) = d])}{\mu_{\mathsf{dim(day)}}(e)} = \frac{150}{\mathsf{day}}$$

with root type vt, dominating:

left daughter, type $\langle\langle d, vt\rangle, vt\rangle$:
$$\lambda G_{\langle d, vt\rangle} \, . \, \lambda e \, . \, \frac{\max(\lambda d \, . \, G(d)(e))}{\mu_{\mathsf{dim(day)}}(e)} = \frac{150}{\mathsf{day}}$$

right daughter, type $\langle d, vt\rangle$:
$$\lambda d \lambda e \, . \, \mathsf{ge}(e) \wedge \exists x [\mathsf{th}(e) = x \wedge {}^{*}\mathsf{sp}(x) \wedge \mu_D(x) = d]$$

left daughter's daughters:

type d: 150

type $\langle d, \langle\langle d, vt\rangle, vt\rangle\rangle$:
$$\lambda q_d \, . \, \lambda G_{\langle d, vt\rangle} \, . \, \lambda e \, . \, \frac{\max(\lambda d \, . \, G(d)(e))}{\mu_{\mathsf{dim(day)}}(e)} = \frac{q}{\mathsf{day}}$$
dominating: *per*$_1$ *day*

right daughter's daughters: $\lambda d \, d$ MEAS species go extinct
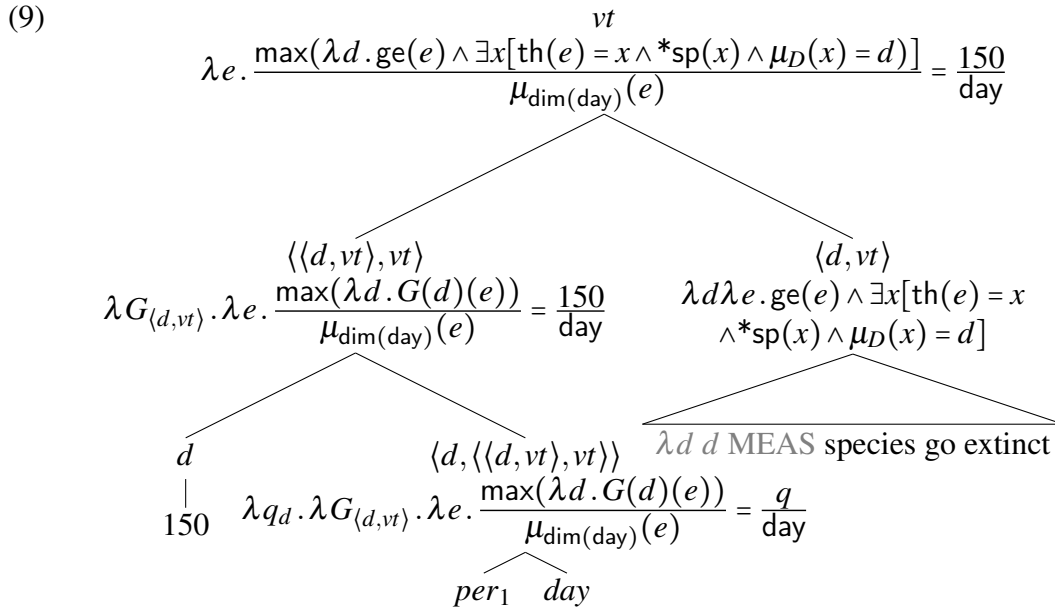
I assume that the trace of *150 per day* combines with *species* in the way that numerals combine with nouns, using a silent head called MEAS as glue (Solt 2009; Rett 2014a). For concreteness, let us assume Rett's (2014a) attributive version of MEAS:

(10)   MEAS $\rightsquigarrow \lambda P \lambda d \lambda x \, . \, P(x) \wedge \mu_D(x) \geq d$

In this lexical entry, μ (type $\langle e, d\rangle$) and D (a dimension) are both free variables set by context.

The formula at the top node can be expressed in English as: "The ratio of how many things go extinct in *e* to the measure of *e* in time (the dimension of the quantity day) is equal to 150 divided by one day." Since two fractions cannot be equal unless their dimensions are equal, it follows that $\mu_D(x)$ must yield a value of the same dimension as 150, which is a plain number, so D to be set to $\mathbf{1}_{\mathscr{D}}$. In other words, the value of D is constrained by other elements in the sentence.

---

4 *Per*-phrases tend to hug the right edge of maximal projections containing their licensor.

2.4 Cardinality dimensions

One more development of the theory is necessary in order to give a satisfactory treatment of (1), because it contains the ordinary sortal noun *martini* in the denominator. *Per* expects its first argument to be a particular quantity, not a property. This is an issue that arises not only with quotient operator uses, but also quotient function uses, i.e., cases where a particular ratio is being referred to, as in:

(11) $20 per martini is a high rate.

To accommodate sortal nouns in the denominator, let us assume that for every predicate denotation S in $D_{\langle e,t\rangle}$, there is a cardinality dimension #S. Assume further that if α is of type $\langle e,t\rangle$ then a model M determines a one-to-one mapping d such that: $[\![\#\alpha]\!]^M = d([\![\alpha]\!]^M)$ and the output is a basic dimension: $d([\![\alpha]\!]^M) \in \mathscr{B}$. This assumption adds $2^{|D_{\langle e,t\rangle}|}$ new basic dimensions to $\mathscr{B}$. As long as the domain is finite, the set of basic dimensions is still finite.[5]

As each dimension D is always associated with its own unit quantity unit(D), there are unit quantities for each flavor of cardinality; for example, the unit quantity for the #martini dimension—unit(#martini)—is '1 martini' as a quantity; unit(#olive) is the quantity '1 olive'.

Let us assume further that a noun like *martini* has, along with its basic denotation of type $\langle e,t\rangle$ as in (12-a), a denotation of type d, the unit for the corresponding dimension:

(12)   a.   *martini* ⤳ martini   (type $\langle e,t\rangle$)
        b.   *martini* ⤳ unit(#martini)   (type d)

So *martini* can denote the unit for the #martini dimension. Let us assume that a unit quantity interpretation is available for any singular NP (including complements and modifiers, excluding articles) with an $\langle e,t\rangle$ denotation.

This allows us to extend our analysis of *per* to cases like *$20 per martini*. The compositional derivation of a ratio term interpretation of that phrase proceeds simply and straightforwardly once a type d meaning for *martini* is established, using the quotient function analysis of *per*.

(13)   *$20 per martini* ⤳ $\dfrac{20 \cdot \text{dollar}}{\text{unit}(\#\text{martini})}$ (type d)

The next section will show an application of the type d lexical entry for *martini* with the quotient operator analysis of *per*.

---

5 As these models are purely extensional, the cardinality dimensions for e.g. 'creature with a kidney' and 'creature with a heart' will be the same in any model where these predicates are co-extensive.

2.5 Two olives per martini

Having in place both the quotient operator analysis of *per* and the ability to handle sortal nouns in the denominator, the denotation in (14) is derived for *eat two olives per martini*.

(14)

$$\lambda e. \dfrac{\max(\lambda d\,.\,\mathsf{eat}(e) \wedge \exists x[\mathsf{th}(e) = x \wedge {}^*\mathsf{olive}(x) \wedge \mu_D(x) = d])}{\mu_{\#\mathsf{martini}}(e)} = \dfrac{2}{\mathsf{unit}(\#\mathsf{martini})} \quad \textit{vt}$$

$$\lambda G_{\langle d,vt\rangle}.\lambda e. \dfrac{\max(\lambda d\,.\,G(d)(e))}{\mu_{\#\mathsf{martini}}(e)} = \dfrac{2}{\mathsf{unit}(\#\mathsf{martini})} \quad \langle\langle d,vt\rangle,vt\rangle \qquad \lambda d\lambda e\,.\,\mathsf{eat}(e) \wedge \exists x[\mathsf{th}(e) = x \atop \wedge {}^*\mathsf{olive}(x) \wedge \mu_D(x) = d] \quad \langle d,vt\rangle$$

$$d \atop \overset{|}{2} \qquad \lambda q_d.\lambda G_{\langle d,vt\rangle}.\lambda e. \dfrac{\max(\lambda d\,.\,G(d)(e))}{\mu_{\#\mathsf{martini}}(e)} = \dfrac{q}{\mathsf{unit}(\#\mathsf{martini})} \quad \langle d,\langle\langle d,vt\rangle,vt\rangle\rangle \qquad \lambda d\ \mathsf{eat}\ [\ \theta_{\mathsf{theme}}\ d\ \mathrm{MEAS}\ \mathsf{olives}\ ]$$

$$\textit{per}_1\ \textit{martini}$$

The formula at the top describes a predicate that holds of an event *e* if the number of olives eaten in *e* divided by a measure of *e* along the 'martinis' dimension is equal to 2 divided by one martini. To give a complete analysis for *James Bond ate two olives per martini*, it suffices to link *James Bond* to this event description via an agent thematic role and existentially close off the event variable, giving:

(15) $\exists e\,.\,\mathsf{agent}(e) = \mathsf{jb} \wedge$
$\dfrac{\max(\lambda d\,.\,\mathsf{eat}(e) \wedge \exists x[\mathsf{th}(e) = x \wedge {}^*\mathsf{olive}(x) \wedge \mu_D(x) = d])}{\mu_{\#\mathsf{martini}}(e)} = \dfrac{2}{\mathsf{unit}(\#\mathsf{martini})}$

So we arrive at a ratio-based analysis of this case, as depicted on the lefthand side of Figure 1.

In this analysis, I have assumed that the *per*-phrase modifies *two* rather than *two olives*. Given the technology we developed for handling cardinality dimensions earlier, another possibility might have been to suppose that the *per*-phrase modifies *two olives*, understood as a degree-denoting expression. This sort of analysis has a great deal of intuitive appeal, as it fits better with the surface constituency. It raises a number of questions, though, given that the theme argument of the verb *eat* should be an individual. If *two olives* is understood as a degree, then the question becomes

how the type requirements of the verb *eat* are satisfied. The answer to this question may lie in Rett's (2014b) 'polysemy of measurement'. I leave it for future research if and how such an analysis could be worked out.[6]

3   Distributivity marker approach

Let us now consider the hypothesis that *per* is not a ratio marker but rather a distributivity marker. As mentioned in the introduction, the intuitive idea behind a P&T-style distributivity-marker analysis is based on the equivalence between (1) and the examples in (2), repeated here as (16) and (17).

(16)   James Bond ate two olives per martini.

(17)   a.   For every martini, James Bond ate two olives.
        b.   Every time he drank a martini, he ate two olives.

P&T take inspiration from Boolos's (1981) and Rothstein's (1995) analyses of these constructions, along with Champollion's (2017) analysis of binominal *each*.
    One piece of evidence bolstering the initial plausibility of this hypothesis is that *per* and binominal *each* have similar distributional requirements. In particular, like adnominal *each*, English *per* can be licensed by a **counting quantifier** but not other determiners (cf. Safir & Stowell 1988).

(18)   a.   They ordered **two/several/many** drinks $\left\{ \begin{array}{l} \text{each} \\ \text{per person} \end{array} \right\}$.

        b. ??They ordered **those/the/most/all** drinks $\left\{ \begin{array}{l} \text{each} \\ \text{per person} \end{array} \right\}$.

These parallels suggest that perhaps *per* is a version of binominal *each* that introduces its own 'sorting key' in the sense of Choe (1987), and is thus a 'share-key relator' (Panaitescu & Tovena 2019).
    On a P&T-style analysis of (16), this sentence describes an event that can be divided into subevents, each corresponding to two olives. The correspondence is via a thematic role, θ_{share}. According to P&T, *per* directly selects both a numeral (e.g. *two*) and a property (e.g. *olives*). (Note that this assumption successfully accounts for the ungrammaticality of (18-b).) For (16), these elements characterize the share participants of subevents thus: $*\text{olive}(*\theta_{\text{share}}(e')) \wedge \mu(*\theta_{\text{share}}(e')) = 2$, which can be read, 'the share participant of e' is made up of olives and has two atomic parts'.[7]
    A contextually determined one-to-one function match associates each subevent

---

6 Thanks to Mitya Privoznov for discussion of this point.

7 The asterisk on predicates denotes the closure operation under sum; *olive holds of any sum of olives. On thematic role functions it signifies that the function

of the main event to another event. In the case of (16), each one of these matching events involves a martini, and is presumably a drinking event. Matching events are linked to their key participants (e.g. martinis) via a thematic role, θ_{key}. P&T assume that the complement of Italian *per* contains a (possibly covert) participial; in (16), the underlying structure is *martini drunk*, where *drunk* is silent. The complement of *per* thus denotes a relation between events and individuals.[8] (19) gives the full truth conditions for (1), where the event-individual relation is represented as drink-martini.

(19) $\exists e . e \in {}^*\lambda e' . {}^*\text{eat}(e') \wedge {}^*\text{olive}({}^*\theta_{\text{share}}(e')) \wedge \mu({}^*\theta_{\text{share}}(e')) = 2 \wedge$
        ${}^*\text{drink-martini}(\text{match}(e'), {}^*\theta_{\text{key}}(\text{match}(e'))) \wedge \mu({}^*\theta_{\text{key}}(\text{match}(e'))) = 1$

The construct $e \in {}^*\lambda e' . \phi[e']$, from Champollion (2017), is a way of saying about an event e that is composed of one or more ϕ-events. The formula thus states that there is an event composed of one or more eating-two-olives events, and each of these corresponds to a drink-one-martini event.

4   Challenges for the distributivity marker approach

4.1   Sub-unit cases

Speed expressions like *100 kilometers per hour* do not fare well under a distributivity-marker analysis à la P&T.

(20)   a.   James Bond drove 100 km per hour.
        b.   ?James Bond drove 100 km for each hour.   (not implied by (20-a))

While (20-a) could be used to describe a driving event that lasted only five minutes, (20-b) could not be. This example illustrates the fact that in general, when *per* expressions are used to measure something, the measurand may have a value that is smaller than *per*'s complement. Let us refer to such cases as **sub-unit cases**.

A P&T-style analysis of (20-a) would falsely predict that the event described lasts at least one hour, because it is composed of subevents lasting one hour each. A ratio-based analysis only implies that the spatial extent of the event, divided by its temporal extent, is 100 kilometers divided by one hour. Such truth conditions can be obtained straightforwardly from the quotient function analysis of *per*.

It must be noted that competing empirical claims have been made about sub-unit cases. Schwarz and Bale (this volume) claim that there is 'unit sensitivity', whereby

---

8 This choice is motivated by 'recycled individuals' of the kind seen in Krifka's (1990)'s *Four thousand ships passed through the lock*: In a case like *The clerk filed two complaints per lost phone*, the same phone might be lost twice. While the account offered here has nothing in particular to say about recycled individuals, discussions like that of Barker (2010) indicate that this problem is orthogonal to the issue at hand (thanks to Lucas Champollion for discussion of this point).

the two sentences in the following pair give rise to inferences of differing strength:

(21) a. The sample weighs 0.9 grams per milliliter
          implication: mu_{VOL}(the sample) ≥ 1 ml
      b. The sample weighs 0.9 grams per liter
          implication: mu_{VOL}(the sample) ≥ 1 liter

Relatedly, they claim that the following two sentences contrast in acceptability:

(22) a.   The half-liter sample weighs 0.9 grams per milliliter.
      b. #The half-liter sample weighs 0.9 kilograms per liter.

These judgments suggest that at least in some cases, the entity being measured must have a value that is greater than or equal to 'one' of the denominator. Bale & Schwarz propose a meaning for *per* that encodes a presupposition to this effect. I find this contrast quite murky, and informal investigations suggest that it is not felt by everyone.[9] To the extent that there is a contrast, I suggest that it can be derived from a presupposition that the sample is among the types of things that can be measured in liters. This leaves room for contextual and inter-speaker variation, depending on what types of things are taken to be measurable in terms of the relevant unit. But while sub-unit cases remain to be fully understood, their existence still supports a ratio-based analysis.

4.2   Differential arguments

A second type of environment where *per* phrases are not paraphrasable with *for each* phrases is the differential argument of a comparative:

(23) a.   James Bond drove 10 km per hour faster than the speed limit.
      b. ?James Bond drove 10 km for each hour faster than the speed limit.
          (Whatever this means, it's not implied by a.)

The differential argument of a comparative is generally thought to be a position filled by a degree-denoting expression (e.g. Hellan 1981; von Stechow 1984). The differential is not a distance; it is a speed, or in other words, a quantity whose dimension is distance over time. There is no implied subdivision of an event into hour-long events in this case either. The quotient function analysis easily delivers the right sort of truth conditions for this case.

---

9 To be precise, I carried out a Twitter poll, and 53% of the 17 respondents said that both (22-a) and (22-b) were acceptable. 24% said that only (22-b) was good; 17% said that only (22-a) was good; 11% said both sounded weird.

4.3 Licensing without a numeral

4.3.1 Measure functions

As shown by example (18) above, *per* requires a licensor, and to some extent its licensing requirements are similar to those of binominal *each*. But *per* has a wider range of licensors, including nouns that denote measure functions like *cost*:

(24) The guests minimized (the) *cost* $\left\{ \begin{array}{l} ^{??}\textbf{each} \\ \textbf{per person} \end{array} \right\}$.

$\not\Rightarrow$ The guests minimized the cost for each person.

This sort of case is abundant in the Europarl corpus (Koehn 2005):

(25) a. The current *price of oil* **per barrel**, which is more than USD 55, ...
      b. Sweden therefore as the lowest **per capita** *alcohol consumption* of the entire Union.
      c. If they could vote, I am sure these animals would choose to travel in lorries or trailers with proper ventilation, *adequate room* **per animal**, decent lighting ...
      d. ... improving the techniques available for reducing *emissions of greenhouse gases* **per produced energy unit**

Examples like these clearly show that *per* does not select for a numeral, *contra* P&T. But the problem is not so superficial, and the distributivity marker approach cannot be salvaged just by fixing up the types of the arguments. The problem is that this type of example does not imply any one-to-one correspondence between *per*'s complement (people, barrels, etc.) and some other set. While sentences with *per* can be used to express a regular recurrence involving a match-up between two sets—as it is in the case of (1)—this is not an intrinsic feature of its meaning.

4.3.2 Gradable predicates

Gradable predicates can also license *per*; the following is another type of example showing that *per* does not select for a numeral, and behaves differently from binominal *each*:

(26) The guests found it *quite expensive* $\left\{ \begin{array}{l} ^{??}\textbf{each} \\ \textbf{per person} \end{array} \right\}$.

$\not\Rightarrow$??For each person, the guests found it quite expensive.

Such examples are also found in the Europarl corpus:

(27) a.   This Parliament is *far cheaper to operate* **per citizen**.
           ⇏ ??For each citizen, this Parliament is far cheaper to operate.
      b.   Tobacco is the *most heavily subsidised crop* **per hectare** in the EU.
           ⇏ ??For every hectare, tobacco is the most heavily subsidized crop.
      c.   Portugal is the *largest consumer of fish* **per capita** in Europe.
           ⇏ ??For every member of the population, Portugal is the largest consumer of fish.

This type of example reinforces the point that *per* does not select for a numeral. But again, the problem runs deeper than that. A distributivity-marker analysis of (27-a) would require that the citizens stand in a one-to-one relationship with some other set of entities. No such relationship is implied by this sentence.

On a ratio-based view, this case involves comparison along the dimension, '(how cheap x is to operate) divided by (number of citizens x serves)'. This intuition will be made precise in the next section.

5   Triangle equivalences

The analysis I propose for cases like *cost per person* is a ratio-based one. The variant of *per* that is involved in this case, however, is neither the quotient function nor the quotient operator. Rather, it is a third, which I call a *measure function quotient* analysis.

To help motivate this enrichment in our suite of meanings for *per*, let us observe there is a pattern of equivalences involving three expressions: (i) a measure function (e.g. *cost*); (ii) specific quantity (e.g. *$100*); and (iii) a unit, introduced by *per* (e.g. *ton*). Here are two examples based on attested cases in the EuroParl corpus.

(28) a.   The cost (of wheat) is [$100 per ton].
      b.   The [cost (of wheat) per ton] is $100.

(29) a.   a shortfall of [100 billion euros per annum]
      b.   a [per annum shortfall] of 100 billion euros

On one side of the equivalence, the *per*-phrase combines with the specific quantity (e.g. *$100 per ton*). This combination is equated with the measure function (e.g. *cost*), as in (28-a) and (29-a). On the other side, the *per* phrase combines with the measure function (e.g. *cost per ton*), to create an expression that is equated with the specific quantity (e.g. *$100*), as in (28-b) and (29-b). This pattern of equivalences can be depicted as a triangle, as shown in Figure 2, so I refer to them as **triangle equivalences**.
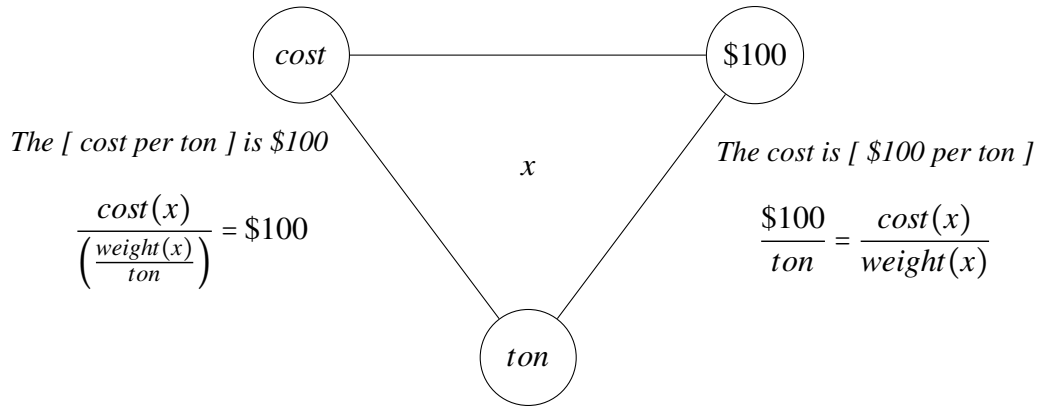
The [cost per ton] is $100

$$\frac{cost(x)}{\left(\frac{weight(x)}{ton}\right)} = \$100$$

The cost is [$100 per ton]

$$\frac{\$100}{ton} = \frac{cost(x)}{weight(x)}$$

Figure 2   A triangle equivalence

5.1   Measure functions 'equated' with specific quotients

Let us begin with the case of (28-a). On the surface, this sentence equates a certain cost with $100/ton. But a cost is a quantity of dimension 'money', I assume, and $100/ton is a quantity of dimension 'money/weight'. As a dimension is a unique and intrinsic property of a quantity, two quantities cannot be equal unless their dimensions are identical. Only a cost *divided by a weight* can be equal to $100/ton. So to say that the cost of x is $100/ton is really to say that the cost of x, divided by the weight of x, is $100/ton:

(30) $$\frac{\$100}{ton} = \frac{cost(x)}{weight(x)}$$

In other words, although the sentence involves three terms (*cost*, *$100*, and *ton*), the meaning expresses an equation involving four terms.

   The implicit fourth term can be obtained using the quotient operator analysis of *per* from Section 2.3. A derivation is given in (32). Note that this lexical entry requires a bit of type polymorphism for quotient operator *per*; the gradable predicate that it combines with in this case is $\langle d, et \rangle$ rather than $\langle d, vt \rangle$. A revised lexical entry is given in (31).

(31)   **Quotient operator analysis of *per*** (polymorphic version)

$$per_1 \rightsquigarrow \lambda d_d \lambda q_d . \lambda G_{\langle d, \tau t \rangle} . \lambda \alpha_\tau . \frac{\max(\lambda d . G(d)(\alpha))}{\mu_{\dim(d)}(\alpha)} = \frac{q}{d}$$

Division vs. distributivity

(32)

$$\lambda x. \frac{\max(\lambda d. \text{cost}(x) \geq d)}{\mu_{\text{dim(ton)}}(x)} = \frac{\$100}{\text{ton}}$$

Gen wheat$_x$ *et*

$\langle\langle d,et\rangle,et\rangle$

$$\lambda G_{\langle d,et\rangle}.\lambda x. \frac{\max(\lambda d. G(d)(x))}{\mu_{\text{dim(ton)}}(x)} = \frac{\$100}{1\cdot\text{ton}}$$

$\langle d,et\rangle$

$$\lambda d\lambda x. \text{cost}(x) \geq d$$

λd

λx the cost of x is d

d

$\mid$

$\$100$

$\langle d,\langle\langle d,et\rangle,et\rangle\rangle$

$$\lambda q_d.\lambda G_{\langle d,vt\rangle}.\lambda x. \frac{\max(\lambda d. G(d)(x))}{\mu_{\text{dim(ton)}}(x)} = \frac{q}{\text{ton}}$$

per ton
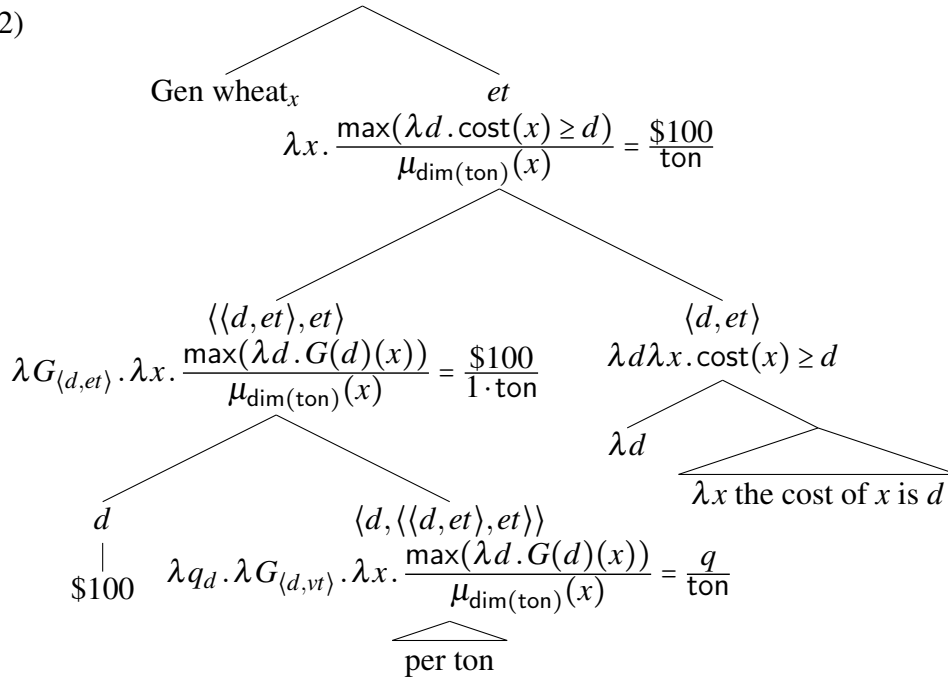
Observe that the fourth term in the equation comes from the lexical entry of *per*, in which the *G*-ness of *x* is divided by by its measure along the dimension given by its complement (*ton*, in this case, whose dimension is weight). The complement of *per* (e.g. *ton*) thus contributes both denominators in the equation. (I assume that in a case like *Wheat costs $100 per ton*, the variable *x* is bound by a generic operator restricted to wheat, as indicated near the top of the tree.)

The same strategy can be employed to account for cases involving 'measurement verbs' like *weigh* and *cost*:

(33)   a.   This sample weighs 0.9 grams per milliliter.
          b.   This wheat costs $100 per ton.

As Schwarz & Bale (this volume) point out, these cases raise exactly the same issue that was encountered with measure function nouns: what wheat costs is not a cost/weight ratio, nor is what samples weigh a weight/volume ratio. Schwarz & Bale give evidence from free relatives that measurement verbs are not "dimension-flexible"; what something weighs is a weight, not a weight divided by a volume:

(34)   #This 1kg cube weighs what that 2kg cube weighs.   (contradictory)

As they point out, the quotient function analysis offered by Coppock (2021) fails to deliver appropriate truth conditions for cases like (33). But the quotient operator analysis straightforwardly resolves the issue, giving truth conditions for (33-b) that

are equivalent to those derived for (28-a).

5.2 Measure function quotients

Now let us turn to (28-a), *The cost per ton is $100*. As given in Figure 2, I propose that what is expressed in this example can be written as follows:

(35) $\quad \dfrac{cost(x)}{\left(\frac{weight(x)}{ton}\right)} = \$100$

The lefthand side of the equation represents the meaning of *cost per ton* (of *x*). The numerator of that fraction is the cost of the object in question. The denominator is the number of tons that object weighs, or in other words, the weight of *x* divided by 'one ton'. Since the weight of an object has dimension 'weight' and 'one ton' does too, $\frac{weight(x)}{ton}$ is a quotient whose numerator and denominator are of the same dimension. That makes it a "dimensionless" quantity, one of dimension $\mathbf{1}_{\mathscr{D}}$. Dividing a quantity by a dimensionless quantity does not change the dimension, so *cost per ton* of *x* is a quantity of dimension 'money'. So it can be equated with $100, as it is in (35).

To derive this meaning compositionally, let us start the more basic, quotient function analysis of *per* given in Section 2.2. Next, both of the arguments of type *d* are lifted to type $\langle e, d \rangle$ (the type of measure functions), and those measure functions are given something to apply to:

(36) **Measure function quotient analysis of *per***

$per_2 \leadsto \lambda g_{\langle e,d \rangle} \lambda f_{\langle e,d \rangle} \lambda x. \frac{f(x)}{g(x)}$ $\qquad\qquad\qquad \langle ed, \langle ed, ed \rangle \rangle$

This version of *per* could be seen as derived from the quotient function *per* via a binary version of the Geach shift, lifting the first two arguments at once, and then introducing an argument for both of them to apply to.[10]
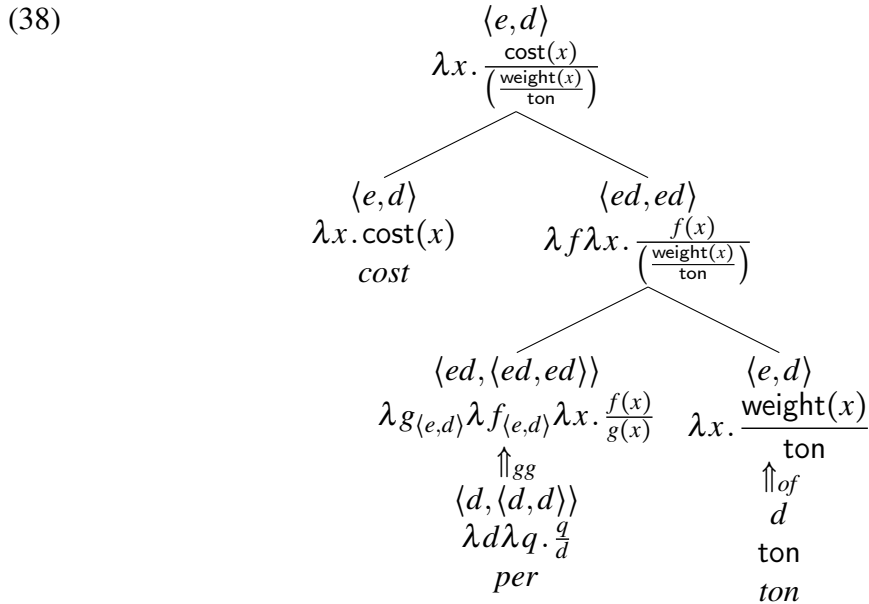
The second step is to make *ton* into a measure function, so that it can be fed as an argument to per_2. To do so, let us posit a type-shifting operation that converts a particular unit quantity (like one ton) to a function that measures an object in terms of that unit. For example, it converts 'ton' (a weight) to a measure function, one which applies to an object *x* and gives back 'the number of tons *x* weighs' (which is not a weight but rather a dimensionless quantity). In general, for any unit *u*, it can be shifted that way:

(37) **_of_-shift**

$u \mapsto_{of} \lambda x. \dfrac{\mu_{\mathsf{dim}(u)}(x)}{u}$

---

10 See also McBride & Paterson (2008) for a precedent (Patrick Elliot, p.c.).

In fact, this shift gives us the more commonly accepted treatment of unit nouns (a.k.a. "measure nouns", not to be confused with "measure function nouns") like *ton*. Champollion (2017), for example, treats *ton* as a function from an individual to a number, the number of tons the individual weighs.
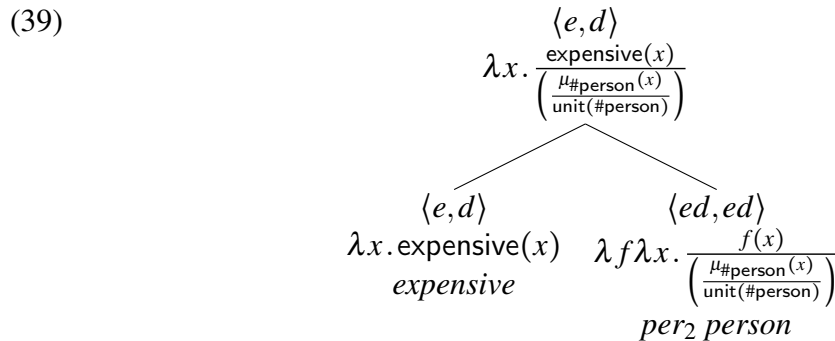
These two tools yield an analysis of *cost per ton*. The unit noun *ton* undergoes the unit shift, to become an expression of type $\langle e, d \rangle$, yielding a measure function that applies to an object and gives its weight in tons. Then *per* undergoes the binary Geach shift, so instead of $\langle d, \langle d, d \rangle \rangle$, it becomes $\langle ed, \langle ed, ed \rangle \rangle$:

(38)

$$\langle e, d \rangle$$
$$\lambda x. \frac{\text{cost}(x)}{\left(\frac{\text{weight}(x)}{\text{ton}} \right)}$$

$$\langle e, d \rangle \qquad \langle ed, ed \rangle$$
$$\lambda x. \text{cost}(x) \qquad \lambda f \lambda x. \frac{f(x)}{\left(\frac{\text{weight}(x)}{\text{ton}} \right)}$$
$$cost$$

$$\langle ed, \langle ed, ed \rangle \rangle \qquad \langle e, d \rangle$$
$$\lambda g_{\langle e,d \rangle} \lambda f_{\langle e,d \rangle} \lambda x. \frac{f(x)}{g(x)} \qquad \lambda x. \frac{\text{weight}(x)}{\text{ton}}$$
$$\Uparrow_{gg} \qquad\qquad \Uparrow_{of}$$
$$\langle d, \langle d, d \rangle \rangle \qquad d$$
$$\lambda d \lambda q. \frac{q}{d} \qquad \text{ton}$$
$$per \qquad\qquad ton$$

The measure function derived at the top node, once applied to an object *x*, returns an amount of money (a function of how much *x* costs and weighs). That amount can be equated to $100. The truth conditions derived for (28-b), then, are equivalent to the truth conditions derived for (28-a). Thus, the triangle equivalence is explained.

To summarize, several key assumptions were used in order to capture triangle equivalences. On the right hand side (*cost is $100 per ton*) the quotient operator analysis of *per* is used in order to get the fourth term. On the left hand side (*cost per ton is $100*), two tricks are used: Along with a type shifting operation that converts unit quantities to measure functions, the quotient function *per* is type-lifted into a function that carries out division at the level of the measure function.

The measure function quotient analysis can be applied to cases involving gradable predicates as well, given an analysis of gradable predicates as measure functions:

(39)

$$\lambda x.\dfrac{\text{expensive}(x)}{\left(\dfrac{\mu_{\#\text{person}}(x)}{\text{unit}(\#\text{person})}\right)} \quad \langle e,d\rangle$$

$$\langle e,d\rangle \qquad\qquad \langle ed,ed\rangle$$
$$\lambda x.\text{expensive}(x) \qquad \lambda f\lambda x.\dfrac{f(x)}{\left(\dfrac{\mu_{\#\text{person}}(x)}{\text{unit}(\#\text{person})}\right)}$$
$$\textit{expensive} \qquad\qquad per_2\ person$$

For a case like *cheaper per citizen* in (27-a), a comparative operator that expects a measure function (as in e.g. Wellwood 2015) could be incorporated.

6 Conclusion

There *are* lexical items that conventionally express the concept of ratio, and *per* is one of them. Four types of evidence were used to argue for this view over a ditributivity marker analysis: sub-unit cases (where the measurand measures less than the unit given by *per*'s complement), uses in the differential argument of a comparatives, uses modifying a measure function noun, and uses modifying a gradable predicate. I have proposed that there are three ratio-related senses: (i) a quotient function; (ii) a quotient operator; and (iii) a quotient of measure functions.

As a side effect, some methodological tools have emerged; the empirical arguments presented here indicate a potential methodology for deciding whether a given item conventionally expresses the concept of a ratio in a given language. Returning to Italian *per*, for example, we find that it actually does show signs of being a ratio marker; (40) and (41) show examples of Italian *per* being hosted by a measure function noun and a gradable predicate, respectively (examples from Europarl):

(40) l'aumento   della *superficie* minima **per gallina**
      the:increasing of:the floor.area minimum per bird
      'the increase of the minimum floor area per bird'

(41) *maggior contribuente* netto **per cittadino**
      greatest contributor   net   per citizen
      'largest net contributor per citizen'

The Europarl corpus reveals a host of other items in European languages that are strong candidates for being ratio markers, including French *per*, Swedish *om*, Polish *na* and *za*, and Hungarian *-nként*, among others. Ratio markers are almost certainly not limited to English, and there is much to discover about how the concept of ratio is expressed in European and non-European languages.

References

Ahn, Dorothy & Uli Sauerland. 2017. Measure constructions with relative measures: Towards a syntax of non-conservative construals. *The Linguistic Review* doi:10.1515/tlr-2017-0001.

Bale, Alan & Bernhard Schwarz. 2019. Reverse proportionality without context dependent standards. In M. Teresa Espinal et al. (ed.), *Sinn und Bedeutung 23*, 93–107. Universitat Autònoma de Barcelona.

Bale, Alan & Bernhard Schwarz. 2020. Proportional readings of *many* and *few*: the case for an underspecified measure function. *Linguistics and Philosophy* 43. 673–699. doi:10.1007/s10988-019-09284-5.

Barker, Chris. 2010. Nominals don't provide criteria of identity. In Artemis Alexiadou & Monika Rathert (eds.), *The semantics of nominalizations across languages and frameworks*, 9–24. De Gruyter.

de Boer, J. 1994. On the history of quantity calculus and the international system. *Metrologia* 31. 405–429. doi:10.1088/0026-1394/31/6/001.

Boolos, George. 1981. For every A there is a B. *Linguistic Inquiry* 12. 465–467.

Champollion, Lucas. 2017. *Parts of a whole: Distributivity as a bridge between aspect and measurement.* Oxford University Press. doi:10.1093/oso/9780198755128.001.0001.

Choe, Jae-Woong. 1987. *Anti-quantifiers and a theory of distributivity*: University of Massachusetts at Amherst PhD dissertation.

Coppock, Elizabeth. 2021. Challenge problems for a theory of degree multiplication. Proceedings of Semantics and Linguistic Theory (SALT) 31.

Cresswell, Max J. 1977. The semantics of degree. In Barbara Partee (ed.), *Montague Grammar*, 261–292. New York: Academic Press.

Hellan, Lars. 1981. *Towards an Integrated Analysis of Comparatives*. Tübingen: Narr.

Kennedy, Chris & Jason Stanley. 2009. On 'average'. *Mind* 118. 583–646.

Klein, Ewan. 1991. Comparatives. In Arnim von Stechow & Dieter Wunderlich (eds.), *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung / Semantics: An international handbook of contemporary research*, 673–691. Berlin: Walter de Gruyter.

Koehn, Phillip. 2005. Europarl: A parallel corpus for statistical machine translation. MT Summit.

Krifka, Manfred. 1990. Four thousand ships passed through the lock: Object-induced measure functions on events. *Linguistics and Philosophy* 13(5). 487–520.

McBride, Conor & Ross Paterson. 2008. Applicative programming with effects. *Journal of Functional Programming* 18(1). 1–13.

Panaitescu, Mara & Lucia M. Tovena. 2019. Distributivity over pairs of events and

entities. In M. Teresa Espinal et al (ed.), *Sinn und Bedeutung 23*, 225–235. Universitat Autònoma de Barcelona.

Raposo, Álvaro P. 2019. The algebraic structure of quantity calculus II: Dimensional analysis and differential and integral calculus. *Measurement Science Review* 19(2). 70–78. doi:10.2478/msr-2019-0012.

Rawlins, Kyle. 2013. On adverbs of (space and) time. In B. Arsenijević et al. (ed.), *Studies in the Composition and Decomposition of Event Predicates*, 153–197. Springer. doi:10.1007/978-94-007-5983-1_7.

Rett, Jessica. 2014a. Measure phrase equatives and modified numerals. *Journal of Semantics* 32(3). 425–475. doi:10.1093/jos/ffu004.

Rett, Jessica. 2014b. The polysemy of measurement. *Lingua* 143. 242–266.

van Rooij, Robert. 2011. Measurement and interadjective comparisons. *Journal of Semantics* 28(3). 335–358. doi:10.1093/jos/ffq018.

Rothstein, Susan. 1995. Adverbial quantification over events. *Natural Language Semantics* 3. 1–31.

Sassoon, Galit Weidman. 2010. Measurement theory in linguistics. *Synthese* 174. 151–180. doi:10.1007/s11229-009-9687-5.

Sauerland, Uli & Robert Pasternak. 2022. Non-conservativity with precise proportions. Special collection of *Glossa*.

Solt, Stephanie. 2009. *The semantics of adjectives of quantity*. New York, NY: The City University of New York PhD dissertation.

von Stechow, Arnim. 1984. Comparing semantic theories of comparison. *Journal of Semantics* 3. 1–77.

Wellwood, Alexis. 2015. On the semantics of comparison across categories. *Linguistics and Philosophy* 38(1). 67–101.

Elizabeth Coppock
621 Commonwealth Avenue
Boston, MA 02215
ecoppock@bu.edu