June 27, 2023

Composite Measure Phrases and Specialized Measurement

Adam Gobeski & Marcin Morzycki (UBC)

Abstract Purely numerical measure phrases (MPs) that lack a unit term—such as *three* or *two thirds*—are normally treated as denoting degrees on a single shared numerical scale. This paper examines an apparently unrecognized class of complex purely numerical MPs including *two in three* and *six to one*, which we term COMPOSITE MPS because they have other MPs as subconstituents. They demonstrate, we argue, that mathematically equivalent MPs aren't equivalent in their distribution and in their semantics. These differences can be related to restrictions on what sort of objects different MPs can measure. They can be captured by enriching the theoretical notion of degrees with tropes (as advocated in a line of work by Friederike Moltmann). Some MPs, such as sports scores, require taking a further step: embracing tuples of degrees as a sort of degree. We classify composite MPs into three varieties, provide diagnostics for each of them. The inquiry has ramifications for broader questions about the relationship between the facts of the world and how they are encoded in language.

Keywords measure phrases, degrees, tropes, tuple degrees, proportions, relative degrees, odds, mathematical language, composite measure phrases

1 Introduction

There is something deeply odd about sentences such as (1):

(1) $^{??}$The height of Mt. Everest is $\left\{\begin{array}{l}\text{1.6 nautical leagues}\\ \text{44 furlongs}\\ \text{2,090,315 picas}\\ \text{about 25 million points}\\ \text{4,839 fathoms}\end{array}\right\}$.

Whatever that is, it's not that we can't arrive at determinate truth conditions for them. All the claims in (1) are, as a matter of fact and arithmetic, true.

But that is of course a small part of the story. In a normal discourse, these sentences would give rise to the sense that there had been some sort of misunderstanding—either about the restrictions on the felicitous use of words like *nautical league* and *furlong* or else about leagues and furlongs themselves as units of measure. To gloss explicitly what's at issue, nautical leagues are used only to measure distance from the perspective of a seagoing vessel, and might not be familiar to non-sailors at all if not for Jules Verne. Furlongs in contemporary use measure only horse racing tracks. Picas and points measure only typography. To know how unit terms like these work, it is necessary to know more than the physical dimension they measure. One must also know the domain to which they are restricted. It may be a fact of arithmetic and physical measurement that they are equivalent, but it is not a fact of language.

There is an interesting broader linguistic issue here about the nature of scales, but our aim will be slightly more narrow. It is to examine a class of measure phrases that have other measure phrases as subconstituents, which we'll call COMPOSITE MEASURE PHRASES or composite MPs, many of which give rise to a similar insight:

(2)
a. one in four
b. one out of four
c. three to one
d. two by four
e. two inches by four inches
f. from three to six
g. three through six

As in (2), it's often the case that such measure phrases are not interchangeable even when they are arithmetically equivalent. It's natural, for example, to use *one in four* to express a probability, but not *one fourth*:

(3) The probability of winning is $\left\{\begin{array}{l}\text{one in four}\\ \text{one out of four}\\ {}^{\#}\text{one fourth}\\ {}^{\#}\text{a fourth}\\ \text{.25}\\ \text{25\%}\end{array}\right\}$.

We take this fact to be analytically revealing, a reflection of subtle restrictions on the use of measure phrases like the fine-grained scale restrictions in (1). Many composite MPs, we will argue, establish especially that mathematically equivalent measure phrases are not equivalent in their distribution. They will lead us to advocate an enriched understanding of degrees—involving tropes, as advocated in Moltmann (2009) and subsequently—that makes possible a general theory of measure phrase specialization on which certain degrees can be linguistically distinct while being mathematically equivalent.

Other members of the class, such *three to one* in its use as a sports score, establish a distinct point about the ontology of degrees: that it makes sense to construe some degrees as actually composed irreducibly of tuples of values. Ranges such as *from three to six* are related importantly to the grammar of modified numerals, though we will reserve a full analysis of these for other work (Gobeski & Morzycki 2022)

In section 2, we make the case that measure phrases that are mathematically identical do not have the same distribution. In section 3, we propose a taxonomy of composite MPs into three subclasses—proportional, irreducible, and range—and diagnostics to distinguish them. In section 4, we relate the fine-grained distinctions among composite measure phrases explicitly to specialized unit terms such as the ones in (1), and propose a theory of how measure phrases come to be specialized that provides evidence for viewing degrees in terms of tropes. In 5, we propose an analysis of one variety of non-proportional composite MPs—arithmetically irreducible ones—that relies on assuming that some degrees are composed of pairs or other tuples of other degrees. Section 6 concludes.

2 Mathematically equivalent MPs aren't linguistically equivalent

2.1 *Bare numeral measure phrases*

It's natural to assume in English that in the absence of an overt unit term such as *meter* or *pound*, purely numerical MPs like the ones in (4) refer to degrees on the same scale:

(4) a. 6 is greater than 4.

b. 7 plus $\left\{\begin{matrix} \frac{1}{5} \\ .2 \end{matrix}\right\}$ is $\left\{\begin{matrix} 7\frac{1}{5} \\ 7.2 \end{matrix}\right\}$.

c. $\left\{\begin{matrix} 8\frac{1}{4} \\ 8.25 \end{matrix}\right\}$ is between 8 and 9.

d. There are $\left\{\begin{matrix} 3.5 \\ 3\frac{1}{2} \end{matrix}\right\}$ apples on the table.

The key observation in (4) is that non-integer numbers can be freely expressed as either fractions (pronounced, e.g., "one-fifth") or decimals (e.g., "point two"). In fact, there are multiple linguistic strategies for describing arithmetically identical numbers:

(5) a. one fourth
b. 25% (“twenty-five percent”)[1]
c. .25 (“point two five”)
d. 1 $\left\{\begin{array}{l}\text{in}\\ \text{out of}\end{array}\right\}$ 4
e. 1 to 3

Their arithmetic equivalence might lead us to expect that these MPs should be interchangeable across different contexts. The aim in this section is to establish that that expectation is not met, and that mathematically identical MPs aren’t linguistically identical.

Let’s begin by describing the variations in acceptability for these proportional numbers in different contexts. For instance, equatives support percentages and fractions but not *in* MPs, *to* MPs, or decimals:

(6) Bertha is $\left\{\begin{array}{l}\phantom{^{\#}}\text{25\%}\\ \phantom{^{\#}}\text{one fourth}\\ ^{\#}\text{.25}\\ ^{\#}\text{1 in 4}\\ ^{\#}\text{1 to 3}\end{array}\right\}$ as tall as Clyde.

Similarly, comparatives don’t consistently allow fractions either:

(7) Bertha is $\left\{\begin{array}{l}\text{25\%}\\ \text{one fourth}\end{array}\right\}$ taller than Clyde.

Partitives allow percentages, fractions, and *in* MPs but not decimals or other composite MPs:

(8) Let’s disburse $\left\{\begin{array}{l}\phantom{^{\#}}\text{25\%}\\ \phantom{^{\#}}\text{one fourth}\\ ^{\#}\text{.25}\\ \phantom{^{\#}}\text{1 in 4}\\ ^{\#}\text{1 to 3}\end{array}\right\}$ of the donations.

Odds can be described using percentages, decimals, *in*, and *to*, *out of*, but not fractions:

---

[1]We realize, of course, that the reader will know how to pronounce these. We provide the pronunciations to highlight that their symbolic representation in writing doesn’t map trivially to a pronunciation. We address this directly in section 2.3.

(9) Her odds of winning are $\left\{\begin{array}{l}\text{25\%}\\ {}^{\#}\text{one fourth}\\ {}^{\#}\text{.25}\\ \text{1 in 4}\\ \text{1 to 3}\\ \text{1 out of 4}\end{array}\right\}$.

Probabilities work with percentages, *in* MPs, decimals, and, varyingly among speakers, fractions and *to* MPs:

(10) The probability of winning is $\left\{\begin{array}{l}\text{25\%}\\ \text{one fourth}\\ \text{.25}\\ \text{1 in 4}\\ {}^{\%}\text{1 to 3}\\ \text{1 out of 4}\end{array}\right\}$.

Finally, composite MPs—and perhaps percentages—are not in the extension of *number*:

(11) I think $\left\{\begin{array}{l}{}^{??}\text{25\%}\\ \text{one fourth}\\ \text{.25}\\ {}^{\#}\text{1 in 4}\\ {}^{\#}\text{1 to 3}\end{array}\right\}$ is a small number.

How do we explain the differences here?

Because these MPs differ linguistically but not arithmetically, the answer must be supplied on linguistic—and not arithmetic—grounds. One possibility is that this is essentially a syntactic fact about structurally different MPs, perhaps related to some subcategorization idiosyncrasies. But as we'll show in §2.2, this can't ultimately be all about syntax, because there seem to be restrictions on which varieties of numerals can be compared with which others, in positions in which subcategorization is not available as an explanation.

2.2 *Direct comparison*

Direct comparison sentences such as (12) especially can be used to our advantage, because we can use them to test whether various MPs share a scale. The first observation is that a simple bare integer like *2* can be compared with only certain MPs, mainly fractions and decimals:

(12) $$\text{2 is greater than} \left\{\begin{array}{l} {}^{\#?}\text{33\%} \\ \text{one third} \\ {}^{\#}\text{1 in 3} \\ {}^{\#}\text{1 to 2} \\ \text{.33} \end{array}\right\}.$$

Percentages can't be compared with composite MPs (and perhaps decimals), while composite MPs can't be compared to percentages, fractions, or decimals:

(13) $$\text{90\% is greater than} \left\{\begin{array}{l} \text{33\%} \\ \text{one third} \\ {}^{\#?}\text{1 in 3} \\ {}^{\#}\text{1 to 2} \\ {}^{\#?}\text{.33} \end{array}\right\}.$$

(14) $$\text{2 in 3 is} \left\{\begin{array}{l} \text{greater} \\ \text{better odds} \end{array}\right\} \text{than} \left\{\begin{array}{l} {}^{??}\text{33\%} \\ {}^{\#?}\text{one third} \\ \text{1 in 3} \\ \text{1 to 2} \\ {}^{\#?}\text{.33} \end{array}\right\}.$$

Explicitly flagging a decimal as a probability allows comparison with composite MPs, but percentages and fractions are still marginal at best:

(15) $$\text{0.9 is a higher probability than} \left\{\begin{array}{l} {}^{?}\text{33\%} \\ {}^{\#?}\text{one third} \\ \text{1 in 3} \\ \text{1 to 2} \\ \text{.33} \end{array}\right\}.$$

2.3 *Distinct restrictions on symbolic representations*

In the previous sections, we have attempted to finesse an important point that we will now address explicitly. There are linguistic restrictions on the pronunciation of various MPs, and equally there are distinct restrictions on their written symbolic representations. For example, aspect ratios—which express the ratio of the width of a screen to its height—are obligatorily rendered graphically with a colon, as in (16a), and normally pronounced with no overt connective between them, as in (16b):

(16) **Aspect ratios**

a. 16:9 *(symbolic representation)*

b. “sixteen nine” *(pronunciation)*

Other conceivable graphical representations, such as those in (17), are not used (we use * here to indicate orthographic ill-formedness):

(17) **Aspect ratios in symbols**

a. *16/9

b. *16-9 *(this contains a hyphen)*

c. *16—9 *(this contains an em dash)*

d. $*\frac{16}{9}$

e. *16,9

f. *16;9

g. *16x9

h. *16 × 9

Nor are other conceivable pronunciations:

(18) **Aspect ratios as pronounced**

a. ?“sixteen to nine”

b. *“sixteen by nine”

c. *“sixteen on nine”

d. *“sixteen vs. nine”

e. *“sixteen over nine”

Of course, these pronunciations and written representations are used in other contexts, for other flavors of measure phrase. In sports scores, for example—to which we will return in more detail subsequently—the usual symbol in English is a hyphen or en dash, but not a colon as for aspect ratios:

(19) **Scores in symbols**

a. $\left\{\begin{array}{l}\text{16-9}\\ \text{16–9}\end{array}\right\}$

b. *16—9

c. $*\frac{16}{9}$

d. *16:9

e. *16,9

f. *16;9

g. *16::9

h. *16x9

i. *16 × 9

This fact is the typographical counterpart of the linguistic fact that scores are reported in English with *to* and not with other prepositions or conceivable connectives:

(20) The final score was $\left\{\begin{array}{l}\text{ 6 to 2}\\ ^{\#}\text{6 on 2}\\ ^{\#}\text{6 by 2}\\ ^{\#}\text{6 versus 2}\\ ^{\#}\text{six and two}\end{array}\right\}$.

Multidimensional spatial measurement differs yet again, despite its resemblance to aspect ratios. The standard symbol is × or x:

(21) **Spatial measurement in symbols**
- a. 2x4
- b. 2×4
- c. 2×4×2
- d. *2:4
- e. *2/4
- f. *2,4
- g. *2;4

There are many further such restrictions in other domains, and presumably we needn't belabor the point. What's crucial here is that a fine-grained fact about the use to which a measure is put—for example, measuring a screen rather than a living room—imposes corresponding fine-grained restrictions on both linguistic and symbolic representations. One might reasonably wish to separate symbolic and typographical conventions from linguistic restrictions, of course. But doing so only makes the point more strongly. Despite their important differences, they both independently show that mathematically equivalent measure phrases are not interchangeable.

2.4 *Intermediate summary*

So where does this leave us? At the very least, we've established that mathematically identical MPs are not linguistically identical, with differing distributions. This can't be a consequence of subcategorization alone, because even direct comparison sentences behave differently depending on which MPs are compared. The predicate *is greater* is perfectly happy to compare percentages and *in* MPs, but as (13–14) showed, *not to each other*. That's precisely what we would expect if this is about whether degrees share a scale, but not if it's about the subcategorization of *greater*.

It's reasonable to wonder whether whether the facts in this section are the consequence of some particular deeper generalizations. But although we can observe some smaller regularities—such as that bare numerals seem to be on the same scale as fractions and decimals—there are no robust broad generalizations that explain all these facts. Which MP uses which scale seems to be at least in part a lexical matter. We have to recognize the inherent idiosyncrasy at play: numbers, percentages, probabilities, etc., all measure on slightly but consequentially different scales.

On its face, aspects of this pattern of facts aren't too surprising: for instance, a distinction between what might be called ordinary and proportional degrees has been emerging in various contexts, and it can be construed as a sortal difference in scales (versions of this idea occur in Ahn & Sauerland 2015, 2017, Gobeski & Morzycki 2018, and Solt 2018). Ordinary degrees measure along a scale, divided into units (*meters*, *years*, *pounds*, etc.), while proportional degrees measure along a scale that's been proportionally divided: *33%* describes a proportion, not an independent unit, which can vary in size as it applies to different objects (such as *one degree* or *Clyde's height*).

What may be surprising, however, is that we need an even more fine-grained distinction among MP flavors, one that differentiates between, say, odds and decimals (we have more to say on this in section 4). It's also worth noting that the syntax matters: *1 to 2*, for instance, expresses odds specifically, not just any proportion, and that determination is based on the use of *to*. Yet *to* doesn't *have* to represent odds. It also occurs *to* in scores (*lost the game by two to three*) and in ranges (*two to three capybaras*). In order to explore this question further, let's now take a closer look at varieties of composite MPs.

3 Three kinds of composite MPs

3.1 *Two additional varieties*

Having established that numerical MPs in general aren't always commensurable, in this section we turn to composite MPs specifically. The composite MPs that we've considered thus far, such as *3 to 1*, *3 in 4*, and *3 out of 4*, have mostly expressed proportions. But these are by no means the only kinds of composite MPs in English, as (22) illustrates:

(22) a. The score was *6 to 2*.
b. The room is *3 meters by 4 meters*.
c. Floyd saw *(from) four to six* capybaras.

The MPs in (22) are all composite, just as our other examples have been, but they don't seem to be straightforwardly expressing proportions. While (22a), for instance, could possibly be considered a form of proportion, it's certainly not equivalent to *3 to 1*, while even this stretch is unavailable for the other sentences. This suggests that there are at least three varieties of composite MPs.

Consequently, we classify composite MPs into three distinct groups, which we will further characterize in the remainder of this section:

1. **Proportional composite MPs**, including general proportions such as *1 in 3* and odds expressions such as *3 to 1*.
2. **Range composite MPs**, including expressions such as *(from) 3 to 6* and *between 3 and 6*, which intuitively seem to denote a range of possible values.
3. **Irreducible composite MPs**, including score expressions such as *4 to 3* and areas such as *3m by 4m*, which describe neither ranges nor (reducible) proportions.

To distinguish these, we will propose three diagnostics: the Multiplication Diagnostic, the Amount Question Diagnostic, and the Differential Diagnostic.

3.2 *The Multiplication Diagnostic*

The Multiplication Diagnostic is straightforward. Only proportional composite MPs preserve their meaning when multiplied by the same factor; range and irreducible composite MPs do not:

(23) a. **Proportional:**
The probability of winning is 1 in 3.
$\Leftrightarrow$ The probability of winning is 2 in 6.

b. **Range:**
Bertha saw 1 to 3 capybaras. $\not\Leftrightarrow$ Bertha saw 2 to 6 capybaras.

c. **Irreducible:**
Venezuela beat Australia by a score of 3 to 1.
$\not\Leftrightarrow$ Venezuela beat Australia by a score of 6 to 2.

In (23a), if the probability of winning is 1 in 3, it's therefore also 2 in 6. But in (23b), if Bertha saw 1 to 3 capybaras, it doesn't follow that she saw 2 to 6 of them. Likewise, in (23c), if the score was 3 to 1, it's not therefore true that it was 6 to 2. That's especially striking because *to* can in fact occur in proportional composite MPs in other contexts, when characterizing odds (e.g., *odds of 2 to 1* is equivalent to *odds of 4 to 2*).

3.3 *The Amount Question Diagnostic*

The Amount Question Diagnostic tests whether a composite MP can answer an amount question. Proportional and range composite MPs can do so, but irreducible composite MPs cannot:

(24) a. **Proportional:**
A: How many of the donations did they disburse?
B: 2 out of 3.

b. **Range:**
A: How many capybaras did Bertha see?
B: 2 to 3.

c. **Irreducible:**
A: ?How many meters is this room?
B: #3m by 4m.

It is possible, of course, to ask how many *square* meters a room is, and to receive the response the composite MP in (24c) as a response:

(25) A: How many <u>square</u> meters is this room?
B: ?3m by 4m.

This is only mildly odd, as we might expect. It doesn't directly answer the question that has been asked, but rather a larger question-under-discussion ('How big is this room exactly?' or 'What is the area of this room?'). It provides a stronger claim from which an answer to the direct question can be computed. What is crucial, though, is that any direct answer to the question in (25) would have to be a single figure expressed in square meters and not an irreducible composite MP.

3.4 *The Differential Diagnostic*

The Differential Diagnostic tests whether a composite MP can serve as a differential in a comparative construction. Only ranges can robustly do so:[2]

[2] Some people accept (26c), although when asked what the sentence meant they found it difficult to reply. The slipperiness of the judgment suggests that speakers might be attempting to perform two one-dimensional calculations (one per constituent) in an effort to interpret the sentence. Notably, speakers who found the sentence acceptable became uncertain when asked to perform calculations with specific examples, with some saying that there may be an ambiguity depending on the order the constituents are compared to those in the *than* clause, and others attempting to flatten the area into a single dimension, such as square meters. Consequently, it seems reasonable to treat this sentence as semantically anomalous, even though there may be something additional to be said.

(26) a. **Proportional:**
??Bertha disbursed 1 in 4 more donations than Clyde.

b. **Range:**
Bertha saw (from) 2 to 3 more capybaras than Clyde.

c. **Irreducible:**
#This room is 3m by 4m larger than that one.

The following table sums up the results of the three diagnostic tests:

(27)

	Proportional	**Range**	**Irreducible**
Multiplication Diagnostic	✓	✗	✗
Amount Question Diagnostic	✓	✓	✗
Differential Diagnostic	??	✓	✗

No composite MP type passes all the same diagnostics as another type, so these three tests can be used together to determine the type of composite MP being evaluated.

For the remainder of this paper, we'll set aside ranges (but see Gobeski & Morzycki 2022 for more) in favor of proportional and irreducible composite MPs. Semantically, proportional composite MPs work broadly similarly to proportional non-composite MPs, in that they pick out a proportion that maps to the appropriate scale. Irreducible composite MPs, however, have gone unexamined. We'll turn to those now.

3.5 *Irreducible composite MPs*

One important class of irreducible composite MPs is scores, as in (28):

(28) Venezuela beat Australia (by (a score of)) 6 to 2.

Scores could be thought of as a specialized non-proportional form of ratio. After all, a standard way to express a ratio is *N to N*:

(29) Their army outnumbers ours 3 to 1.

Loosely speaking, the ratio in (29) expresses a sort of opposition, comparing two parts of some other whole. Abstractly, scores can be thought of in a similar way, as two proper parts of the total points scored. However, as noted before, ratios express relative proportions: the ratio expressed in (29) isn't sufficient to tell us the actual number of people in each army. By contrast, scores express non-relative amounts and, as noted above, cannot be reduced to relative values: (28) does in fact tell us the actual number of points each team scored.

Scores also systematically entail a corresponding differential:

(30) Venezuela beat Australia (by (a score of)) 6 to 2.
entails: Venezuela beat Australia by 4.

The *by* that marks a score is a distinct lexical item from the *by* that marks the differential of a score. Only the former is optional, as (31) demonstrates:

(31) a. Venezuela beat Australia 6 to 2.
b. *Venezuela beat Australia 4.

That said, the score-marking *by* (or a similar preposition such as *with*) is required if *a score of* is present:

(32) a. Venezuela beat Australia $\left\{\begin{matrix}\text{by} \\ \text{with}\end{matrix}\right\}$ a score of 6 to 2.
b. *Venezuela beat Australia a score of 6 to 2.

Scores aren't the only place in athletics where irreducible composite MPs occur. They also express the results of a series of games:

(33) The Blue Jays are currently **2 for 4** this series.
$\not\Leftrightarrow$ The Blue Jays are currently **4 for 8** this series.

Another kind still expresses defensive coverage, as in basketball:

(34) This offensive drill helps players in a **2-on-1** situation.
$\not\Leftrightarrow$ This offensive drill helps players in a **4-on-2** situation.

In all of these cases, the numbers themselves are the crucial information, not purely their relationship to each other, and thus cannot be scaled up (or down) without losing that information.

Athletics, with its focus on specific numbers, provides a natural place for irreducible composite MPs, but it's not the only environment that we can encounter these. Areas—and relatedly, volumes—are also a common place to find irreducible composite MPs:

(35) This paper measures 8 inches by 10 inches.
$\not\Leftrightarrow$ This paper measures 16 inches by 20 inches.

(36) This box measures 8 inches by 10 inches by 4 inches.
$\not\Leftrightarrow$ This box measures 16 inches by 20 inches by 8 inches.

Once again, the specific length and width (and height, in the case of (36)) are required. The fact that both area and volume can be expressed in one-dimensional terms (namely, square and cubic units of measurement) indicates that what matters in these composite MPs is not the ratio of the individual measures—as that can be ultimately expressed one-dimensionally—but the individual measurements themselves.

Just as different proportional MPs measure vary slightly in what they can measure, different irreducible composite MPs do too:

(37) a. They won 3 $\left\{\begin{array}{c}\text{to}\\ {}^{\#}\text{by}\\ {}^{\#}\text{for}\end{array}\right\}$ 2.

b. This room is 3m $\left\{\begin{array}{c}{}^{\#}\text{to}\\ \text{by}\\ {}^{\#}\text{for}\end{array}\right\}$ 2m.

c. They are 3 $\left\{\begin{array}{c}{}^{\#}\text{to}\\ {}^{\#}\text{by}\\ \text{for}\end{array}\right\}$ 4 in this series.[3]

Clearly the choice of preposition matters: *to*, *by*, and *for* all occur in irreducible composite MPs, but they're not interchangeable. We thus need a more fine-grained way of characterizing these structures, even beyond the three types of composite MPs: a way to discuss the difference between, e.g., scores and win/loss records. To that end, we now turn to the question of how MPs come to be lexically specialized for particular uses.

4 Capturing the specialization of measure phrases

4.1 *Specialized measure terms*

To elaborate a point made fleetingly in the introduction, unit terms are often subject to very fine-grained distinctions. If you were teaching a second-language learner of English the meaning of the term *furlong*, it would be insufficient to simply inform them that it is a unit of measure equivalent to an eighth of a mile. That would invite them to produce deeply odd sentences like (38):[4]

(38) #The Empire State Building is 2.2 furlongs tall.

[3]There is an distinct reading on which *to* is possible here, in which it reflects a win-loss ratio.

[4]To save the effort of Googling, (38) and (39) report the actual height of the Empire State Building more or less accurately.

It's a consequential fact about such sentences that they have determinate truth conditions but that they are nevertheless anomalous. The reason, of course, is that in contemporary English furlongs are only used in horse racing—indeed, even then, only to measure the length of a track and not, say, the height of a horse. For analogous reasons, even in a discourse that explicitly mentioned, say, a 12-point font, (39) remains deeply odd as well:

(39) #The Empire State Building is $\left\{\begin{array}{l}\text{1 1/4 million points}\\ \text{210 thousand picas}\end{array}\right\}$ tall.

A skeptic might suggest that the real difficulty in (38) is that for most English speakers, *furlong* is unfamiliar. But no such objection is possible to (39). In an age in which most of us discuss font size, points—once an obscure unit of measurement known mainly to typesetters—are familiar to all of us. Yet (39) is no better than (38), and anyone inclined to use points and picas to measure linear extent outside of typography would, in a linguistic sense, be using these terms wrong. To really know what *furlong* means, it's necessary to know that it's only used for horse tracks. Likewise for *point* and *pica*, which are used for—and arguably defined only with respect to—typographical measurement. Of course, they're both measures along the same physical dimension: linear extent. But that ontological or physical fact is irrelevant to the linguistic fact that their meaning is more constrained.

A similar point can be made with respect to other conventions of measurement. In Canada, there are complex conventions around the use of metric and imperial units. The height of a human is necessarily in imperial, and it would be strange to express it in metric. The height of a mountain, on the other hand, is obligatorily in metric. Ambient temperature is always in Celsius but oven temperature is always in Fahrenheit. Road distances are in kilometers, but measures of floor space are typically imperial, though apparently not obligatorily. A piece of lumber is called a two-by-four (reflecting its nominal measure in inches), but not a five-by-ten (the equivalent in centimeters). Britain uses *stone* as a measure of human weight equivalent to 14 pounds, but it would be strange to use it to measure the mass of a mountain.

Here we come to the crucial insight. Just as *2.2 furlongs* is a measure phrase that can't be used to measure arbitrary linear extents, so too *3 to 1* and *2 for 2* are measure phrases that can't be used to measure arbitrary proportions. It's a fact of arithmetic that they can both be expressed as fractions, of course, but it is not a fact of language. Both varieties of specialized measure phrase require a common theory of how these linguistic constraints are imposed, despite elementary arithmetic.

4.2 *How specialized measure terms work compositionally*

The next step will be articulate how the specialization of specialized measure terms—and ultimately of the measure phrases they project—is encoded grammatically. This will require some additional compositional and ontological assumptions.

Because the principal use of many specialized measure phrases is alongside degree nominalizations like *the likelihood of losing* and *the odds of winning*, it will be helpful to commit to an assumption about what degree nominalizations denote. These remain studied less extensively than their importance would suggest. One of the most fully articulated proposals, Moltmann (2009), takes them to denote properties of TROPES, which are concrete manifestations of a property. To get an intuitive feel for the notion, it helps to consider one of her go-to examples, the redness of a particular red box. The redness of this box is a trope, and the box is the BEARER of the trope. Crucially, though there are many other boxes that are red, their redness is not the same trope as the redness of this particular box. Some of them may even be precisely the same shade, yet still their redness tropes are different tropes, even if only in virtue of the fact that they are different boxes. This will be important here because we will tie the fine-grained specialized distinctions among measure phrases to the particularized character of tropes.

It's not obvious whether the facts we will examine actually *require* using tropes. Alternative notions in this conceptual neighborhood might conceivably deliver similar distinctions, at least if supplemented with additional assumptions. Anderson & Morzycki (2015), for example, use kinds of states to represent degrees, and that might suffice (but see Moltmann 2015 for counterarguments). More general ways of conceptualizing the referents degree nominalizations, such as property concepts (Francez & Koontz-Garboden 2017, 2011), are also potential alternatives, but again significant elaborations would be required. Tropes, on the other hand, come ready-made for our purposes.

Putting this to use, a sentence with a degree nominalization and a measure phrase such as (40) will receive a denotation with two crucial ingredients. First, it contains a definite description of a trope, which corresponds to the iota expression in (40b) (we'll represent tropes with variables $d_t, d_t', d_t'', \ldots$).[5] Second, it makes use of a measure function, $\mu_{\mathbf{feet}}$, that maps tropes to their measure in feet (like the measure functions of Krifka 1989,

[5] For Moltmann (2009), an adjectival nominalization actually denotes something with more structure. It is what she calls an order-constituted trope, which consists of a simple trope and relational tropes—tropes involving a relation rather than a property—built around comparative relations such as 'to be longer than'.

which do the same for ordinary individuals):[6]

(40) a. The length of my arm is 3 feet.
b. $\mu_{\textbf{feet}}(\iota d_t[\textbf{length-of-my-arm}(d_t)]) = 3$

To build this up compositionally, we will assume the measure term *feet* relates a bare numeral (represented here with $n, n', n'', \ldots$) to a property of tropes:

(41) $[\![\, \textit{feet} \,]\!] = \lambda n \lambda d_t [\mu_{\textbf{feet}}(d_t) = n]$ *(elaborated below)*

Of course, not just any trope can be measured in feet:

(42) #The age of my car is 3 feet.

The problem is that $\mu_{\textbf{feet}}$ can't measure an age trope. For the sake of explicitness, we'll encode this as a presupposition in the denotation, as in (43), though one might equally well regard it as being a definedness condition on $\mu_{\textbf{feet}}$ itself:[7]

(43) $[\![\, \textit{feet} \,]\!] = \lambda n \lambda d_t : d_t$ is a trope of linear extent $[\mu_{\textbf{feet}}(d_t) = n]$

This is straightforward, but it's not trivial. Measure terms for linear extent such as *feet* and *meters* are compatible with measurement in any direction—length, width, height—but natural language predicates like *tall* and *wide* typically treat these as different dimensions. Consequently, one might think of degrees of tallness as constituting one scale and degrees of width another. Comparisons across these scales are possible, and further theoretical refinements can reflect this (Kennedy 1997, Bale 2006 among others). But there is a real empirical tension here. On the one hand, there is the scale used by measure terms of linear extent, which includes measures of tallness, width, length, altitude, and so on. On the other hand, there are the scales used by adjectives and nominalizations, which typically distinguish tallness, width, length, altitude, and so on. Indeed, it takes the slightly roundabout term 'linear extent' to express the degree nominalization associated with whatever it is feet and miles measure.

One advantage of working with tropes is that the tension between the dimension-specificity of natural language predicates like *tall* and the relative

[6] We will attempt to make our trope-talk more accessible by echoing more familiar degree-talk. For that reason, we avoid saying e.g. that the trope of the length of my arm is the bearer of a trope of being 3 feet. Moltmann (2009) explores these issues explicitly.

[7] We use n as a variable for real numbers (or possibly rational numbers, given the oddness of e.g. ??*My arm is pi feet*), corresponding to whatever a plain numeral denotes. That side-steps a number of complications that aren't immediately relevant, chief among them that in a framework with tropes, numbers themselves might be better regarded as tropes as well.

dimension-indifference of measure terms like *feet* can be resolved by simply stating more or less restrictive constraints on tropes. There is no need to determine whether tallness or linear extent is the 'real' scale. Rather, different linguistic expressions can simply impose more or less stringent restrictions.

The denotation of *feet* in (43) combines first with the numeral *3*, then with the degree nominalization *the length of my arm* (assuming *is* is vacuous):

(44) a. $[\![\, \textit{3 feet}\,]\!] = [\![\, \textit{feet}\,]\!](3)$
$= \lambda d_t : d_t$ is a trope of linear extent $[\mu_{\textbf{feet}}(d_t) = 3]$

b. $[\![\, \textit{The length of my arm is 3 feet}\,]\!]$
$= [\![\, \textit{3 feet}\,]\!]([\![\, \textit{the length of my arm}\,]\!])$
$= [\mu_{\textbf{feet}}(\iota d_t[\textbf{length-of-my-arm}(d_t)]) = 3]$

The upshot is the full measure phrase *3 feet* denotes a property of tropes.

On this view, a specialized measure term like *furlongs* would differ from a general one like *feet* in how restrictive its presupposition is. In contemporary use, for a trope to be measured in furlongs, its bearer needs to be a racetrack for horses:

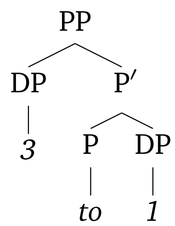
(45) $[\![\, \textit{furlongs}\,]\!] = \lambda n \lambda d_t :$ the bearer of d_t is a horse racetrack $[\mu_{\textbf{furlongs}}(d_t) = n]$

Just as *feet* can't measure age, so too *furlongs* can't measure, say, driveways. The difference between a more and less specialized measure term is just in how narrow its domain is. That's not to say there isn't a deeper conceptual reason that age can't be measured in feet, of course. What's crucial here is that the same linguistic building blocks that encode this linguistically for one measure term can equally do so for the other.

4.3 *Specialization in composite measure phrases*

If some measure phrases come to be narrowly specialized because they inherit this property from their measure terms, how does the specialization of e.g. *3 to 1* arise? It seems to contain no specialized measure term.

The answer, we suggest, lies in the preposition that ties together the two components of the composite measure phrase. The full measure phrase is a PP with DPs in both its complement and specifier position:

(46) 

This is presumably also the case for other prepositional composite MPs, like *1 in 3*, *2 by 4*, and so on. Intriguingly, such a structure may also be what's required for various arithmetic expressions, such as *3 plus 5* (Gobeski 2019).

In PP composite measure phrases, the role of the specialized measure term—the locus of the specialization—is played by the preposition. It maps two bare numerals to a property of tropes. In the case of *3 to 1*, there are actually two homophonous expressions we take to be distinct. One of them is used to report scores in games and perhaps by extension arbitrary oppositions. For the moment, we will focus on the other use, which reports odds. That use involves odds tropes, which we take to be a cousin to tropes of likelihood or probability. Like ordinary likelihood tropes, odds trope instantiate the property of having a certain likelihood. What distinguishes them is that an odds trope is a likelihood trope viewed as the basis of betting or investment. Likelihood is first and foremost a property of events, so the bearers of odds tropes, like likelihood tropes, are events. To represent their mathematical value of an odds trope, we'll make use of a measure function $\mu_{\textbf{proportion}}$, which maps a trope to a proportion representing its value (if such a proportion exists):

(47) a. $[\![\, \textit{to} \,]\!] = \lambda n \lambda n' \lambda d_t : d_t \text{ is an odds trope } \left[\mu_{\textbf{proportion}}(d_t) = \dfrac{n'}{n + n'}\right]$

b. $[\![\, \textit{3 to 1} \,]\!] = \lambda d_t : d_t \text{ is an odds trope } \left[\mu_{\textbf{proportion}}(d_t) = \dfrac{3}{4}\right]$

This means the measure phrase can be straightforwardly predicated of arbitrary odds:

(48) a. The odds of this horse winning the race are 3 to 1.

b. $[\![\, \textit{3 to 1} \,]\!]([\![\, \textit{odds of this horse winning the race} \,]\!])$

$= [\![\, \textit{3 to 1} \,]\!](\iota d_t[\textbf{odds-of-this-horse-winning}(d_t)])$

$= \left[\ \mu_{\textbf{proportion}}(\iota d_t[\textbf{odds-of-this-horse-winning}(d_t)]) = \dfrac{3}{4}\ \right]$

On the other hand, *3 to 1* cannot be predicated of proportions that are not odds—say, the trope of our level of certainty about a particular proposition—because in those cases the presupposition would not be met:

(49) a. $^{\#}$Our level of certainty about this is 3 to 1.

b. $^{\#}$The fraction of dentists who recommend electric toothbrushes is 3 to 1.

Certainty levels and fractions of dentist populations can of course be represented as proportions, but the presupposition of *to* imposes the more stringent requirement of being an odds.

As in the preceding section, one might have considered treating the presupposition as a definedness condition on the measure function $\mu_{\mathbf{proportion}}$ instead, rather than spelling it out independently in the denotation of *to*. But in this instance, the choice has discernible empirical consequences. Many different kinds of tropes can be represented mathematically as proportions, but as we've seen they are not interchangeable. One way to make sense of this situation is to suppose that $\mu_{\mathbf{proportion}}$ can apply to any such tropes, but that particular prepositions impose *additional* restrictions as presuppositions, such as the restriction to odds. The measure function, then, reflects the conceptual similarity of proportional measure phrases, and independent presuppositions reflect their more narrow specialization.

This raises some interesting questions about how this state of affairs would arise, both in acquisition and diachronically. Let's suppose that a particular bookie with little formal education has a keen folk understanding of odds but has somehow failed to discern their relations to other proportions. What would their grammar look like? Would they have multiple measure functions associated with different forms of proportional measurement? Is the process of learning such correspondences in childhood also a process of incrementally revising one's denotations for e.g. *3 to 1* as further mathematical connections are made? Sadly, preschool-aged children are typically not ardent gamblers, and testing this experimentally would require some ingenuity or an unusually accommodating ethics board.

Measure phrases aren't the only expressions that crucially distinguish between odds tropes and likelihood tropes. The adjectives that can be predicated of them differ as well. All likelihood tropes, including odds tropes, can be said to be *high*, but only odds tropes can be said to be *good* or *terrible*:

(50) a. The odds of winning are $\left\{\begin{array}{l}\text{high}\\ \text{good}\\ \text{terrible}\end{array}\right\}$.

b. The $\left\{\begin{array}{l}\text{probability}\\ \text{likelihood}\end{array}\right\}$ of winning is $\left\{\begin{array}{c}\text{high}\\ ^{??}\text{good}\\ ^{\#}\text{terrible}\end{array}\right\}$.

This difference is again something that can be reflected in the semantics of the adjective as a simple presupposition: *good* and *terrible* in the relevant specialized sense presuppose that their argument is an odds trope. This establishes that the fine-grained odds-versus-likelihoods distinction is necessary independently.

Generalizing the proposal in this section, specialization in measure phrases more broadly can be captured via presuppositions about the sort of trope they measure. Composite measure phrases especially tend to be specialized, and in those cases the specialization arises from presuppositions of the preposition that heads them.

4.4 *Degree constructions*

The restrictions on measure phrases imposed by degree constructions such as comparatives and equatives can be understood in a similar spirit, with an additional wrinkle. To confront this issue, we need to return to equatives. They are compatible with percentages and fractional terms like *one third*, but not with various other forms:

(51) Bertha is $\left\{\begin{array}{l} \;\text{33\%} \\ \;\text{one third} \\ {}^{\#}\text{.33} \\ {}^{\#}\text{1 in 3} \\ {}^{\#}\text{2 to 1} \\ {}^{\#}\text{3} \end{array}\right\}$ as tall as Clyde.

One distinction to be made here is between specialized composite MPs like *1 in 3* and *2 to 1* and the more generalized forms. There are at least two reasons these might not be possible in (51). One is that, given the discussion in the previous section, they denote properties of tropes of a particular sort, one incompatible with the measurement of height. Because *2 to 1* imposes the presupposition that what it measures is an expression of odds, it makes sense that it can't measure a height. To make this fully explicit, it will be necessary to provide a denotation for the equative—but as it will turn out, doing so will rule out such measure phrases independently for a different reason.

The crucial consideration is a distinction beyond the specialization we've already explored. It's between ABSOLUTE and RELATIVE degree expressions.[8] The idea that some degree expressions are inherently relative seems to have been regularly reinvented for various purposes over the last decade,

---

[8]This is quite distinct from the distinction between absolute and relative adjectives. It would be desirable to have other equally evocative terms for this.

depending on which aspect of proportional measurement is at issue. It arose in connection with percentages in Ahn & Sauerland (2015, 2017), Gobeski & Morzycki (2018), Gehrke & Wągiel (2023), with proportional readings of comparatives in Solt (2018), with *per* expressions in Coppock (2022a,b), Bale & Schwarz (2022), with rates in the semantics of adverbials (Rawlins 2013), and multipliers like *double* (Wągiel 2019, 2020) and *twice* (Gobeski 2019) . The essence of the idea is that a fraction like one third has a double life: it is both a degree and a way of getting one degree from another. It can express an absolute measurement, like the temperature, or a relative measurement, like the ratio of two prices.

Gobeski & Morzycki (2018) propose that equatives in particular are sensitive to the distinction between relative and absolute degrees. A version of that idea can be implemented in the current framework, with tropes as a background component. The notion that some degree expressions are relational is independent of the role tropes might play. We haven't explicitly used tropes to model degrees themselves in this paper, and that will remain the case—but it bears noting that the distinction between absolute and relative degrees can be recapitulated in the distinction between tropes as instantiations of properties and tropes that are instantiations of relations (Mertz 1996, Moltmann 2009).

One way to approach equatives in English is to suppose that first instance *as*—the degree word rather than the element that introduces a clausal standard—takes as arguments a gradable adjective, a measure phrase that denotes a relation between tropes, a trope expressed by the standard phrase, and an individual, as in (52):

(52) $[\![\text{as}_{\text{EQUATIVE}}]\!] = \lambda G_{\langle e,\langle d^t,t\rangle\rangle} \lambda R_{\langle d^t,\langle d^t,t\rangle\rangle} \lambda d_t \lambda x \,.\, \exists d'_t[\ G(x)(d'_t) \wedge R(d_t)(d'_t)\ ]$

The pieces are assembled so that a sentence like the one in (53) will require that the trope of Bertha's height stand in the 33% relation to Clyde's height:

(53) $[\![\textit{Bertha is 33\%}\ \text{as}_{\text{EQUATIVE}}\ \textit{tall as Clyde}\ \cancel{\textit{is tall}}]\!]$

$$\begin{aligned}
&= [\![\text{as}_{\text{EQUATIVE}}]\!]([\![\textit{tall}]\!])([\![\textit{33\%}]\!])([\![\textit{as Clyde}\ \cancel{\textit{is tall}}]\!])([\![\textit{Bertha}]\!])\\
&= \exists d'_t[\ [\![\textit{tall}]\!]([\![\textit{Bertha}]\!])(d'_t) \wedge [\![\textit{33\%}]\!]([\![\textit{as Clyde}\ \cancel{\textit{is tall}}]\!])(d'_t)\ ]\\
&= \exists d'_t[\ \textbf{tall}(\textbf{Bertha})(d'_t) \wedge \textbf{33\%}(\iota d''_t[\textbf{tall}(\textbf{Clyde})(d''_t)])(d'_t)\ ]
\end{aligned}$$

Thus *33%* differs from other measure phrases not (just) in the details of how its domain is restricted by a presupposition, but also in its type: it denotes a relation between degree tropes rather than a property of them. Its semantics ensures that the measure of the tropes stand in the right ratio to each other:

(54) $[\![\textit{33\%}]\!] = \lambda d_t \lambda d'_t \left[\ \dfrac{\mu(d_t)}{\mu(d'_t)} = \dfrac{33}{100}\ \right]$

The measure phrase *one third* would also have a semantics along these lines, and therefore be able to occur with equatives.

What is less expected, though, is the behavior of *.33*, which is for most speakers unnatural as a measure phrase in an equative. Conceptually, there is no reason why it shouldn't have patterned with *33%*, but as it turns out, it doesn't, which is reason enough to suppose that it is purely an absolute degree expression—and given our assumptions here, that it denotes a property of tropes.

The bare numeral *3* is similar. A priori, one wouldn't have expected it to support relative measurement. But this judgment is misguided. It could easily have been the case that *3* would work exactly the way *300%* does. Yet only the former is possible in equatives:

(55) Bertha is $\left\{\begin{array}{l} {}^{\#}\text{3} \\ \text{300\%} \end{array}\right\}$ as tall as Clyde.

In accord with our theme, the fact that the two expressions are the same as a matter of arithmetic does nothing to diminish the contrast. This means that bare *3* denotes a property of tropes, but not a relation between them, as *300%* does.

Returning to the specialized measure phrases *1 in 3* and *2 to 1*, they could probably be ruled out on the basis of their presuppositions alone, but the issue doesn't arise because they are independently ruled out by their type. But that's not the end of the story. The absolute degree expressions *.33* and *3* can be rescued from their ill-formedness by embedding them in a measure phrase headed by *times*:

(56) Bertha is $\left\{\begin{array}{l} \text{3 times} \\ \text{.33 times} \\ {}^{\#}\text{1 in 3 times} \\ {}^{\#}\text{2 to 1 times} \end{array}\right\}$ as tall as Clyde.

As (56) reflects, this rescue strategy does nothing to improve *1 in 3* and *2 to 1*. *Times* can be taken to denote a function from bare numerals to relative degree expressions:

(57) $[\![\, \textit{times} \,]\!] = \lambda n \lambda d_t \lambda d'_t \,.\, \left[\dfrac{\mu_{\textbf{proportion}}(d_t)}{\mu_{\textbf{proportion}}(d'_t)} = n \right]$

This fails in combination with *1 in 3*, again for simple type-theoretic reasons. But one could imagine reconceptualizing (57) to apply directly to a property rather than a bare numeral and therefore to rely on the presupposition to rule these examples out.

Comparatives introduce additional twists in the story, but we will leave those aside here for brevity.

5 Irreducible composite MPs

5.1 *Finding the right tools*

In the preceding section, we articulated a view of proportional composite MPs that links them to domain-specialized MPs more generally. We captured this specialization in part by supposing that such MPs denote properties of tropes. In this section we turn to irreducible composite MPs such as areas (*3m by 4m*) and sports scores (*3 to 1*), which cannot be expressed as a single proportion. It would be the usual move at this stage to extend the current analysis to these new facts. But as it turns out, that simply adds a layer of complexity without immediate analytical benefit. For that reason, in this section we will build an analysis of irreducible composite MPs essentially from scratch, relying on more standard assumptions about degrees. We will return to tropes only after having done this, at which point it will be easier to distinguish what is empirically necessary to accommodate the irreducible composite MP facts from what is analytically advantageous in view of the full range of data.

The question that now confronts us is what irreducible composite MPs denote. Because it's standard to suppose that MPs in general refer to degrees, we will begin with that as the null hypothesis. But of course a single degree won't suffice. The crucial fact about irreducible composite MPs is precisely that they are irreducible. They don't express a ratio. They are not simply points along only one dimension. So if MPs in general denote degrees, what kind of degrees do irreducible composite MPs denote?

Whatever the answer to this question is, it will require building a multidimensional form of degree from one-dimensional ingredients. There is a conception of degrees made of other degrees for which there is independent evidence: plural degrees (Dotlačil & Nouwen 2016). Pursuing that will be our next step.

5.2 *Plural degrees*

Dotlačil & Nouwen (2016) argue for plural degrees on the basis of sentences such as (58):

(58) John is exactly 3 inches taller than every girl.

They point out that while the intended meaning entails that every girl has the same height, attempting to derive this with standard tools leads only to the assertion that John is 3 inches taller than the tallest girl, with nothing to say about the heights of the other girls. In order to get the intended meaning for (58), they propose (following Beck 2010) treating the *than*-clause as a set of degree pluralities, rather than a set of degree intervals (with a minimality operator then restricting the set of degree plurals to the relevant ones). These plural degrees behave the same way as plurals in other domains do, and Dotlačil & Nouwen show that other properties of plurals, such as cumulative/distributive distinctions, occur in the degree space as well.

Given an independent justification for plural degrees, it might therefore make sense to extend their use to irreducible composite MPs. In most contexts one might think of the area of a room as *10 feet by 12 feet* or *12 feet by 10 feet*. That suggests that an area is simply a plural degree that has only 10 feet and 12 feet as atomic parts. Even in ordinary score reports, it's often the case that the measure phrase doesn't in itself identify which score corresponds to which competitor. If the result of a soccer game between Argentina and Germany is a score of 2 to 3, one can't be sure which of the two is the winner.[9] Even more strikingly, in some languages scores are reported with conjunction, apparently a direct counterpart to conjunction of individual-denoting expressions (as in *Bertha and Clyde*). Henry Davis (personal communication) observes that this is the strategy in St'át'imcets (a Salish language spoken in British Columbia):

(59) t'cún-as i=Canucks-a i=Oilers=a, án'was
beat+DIR-3ERG PL.DET=Canucks=EXIS PL.DET=Oilers=EXIS two.
múta7 pála7
and. one
'The Canucks beat the Oilers 2–1.'
Literally: 'The Canucks beat the Oilers, two and one.'

(60) Gi7i7el' i=Canucks-a é=ki=Oilers=a, án'was
lose PL.DET=Canucks=EXIS. to=PL.DET=Oilers=EXIS two.
múta7 pála7
and. one
'The Canucks lost to the Oilers 2–1.'
Literally: 'The Canucks lost to the Oilers, two and one.'

The measure phrases in (59) and (60) are in the same order, and the assignment of scores to teams is determined by the choice of verb.

[9]This is the case at least in ordinary conversation. There may be extralinguistic conventions in sports journalism around this.

In languages that systematically work this way, analyzing irreducible composite MPs with plural degrees seems promising, although this would of course be a very different application of the concept from the quantification facts that motivated Dotlačil & Nouwen. But at least in English, this won't work for irreducible composite MPs. Win-loss records (e.g. *They are 2 for 3 in the series)* and sales promotions (e.g. *a 2 for 1 sale*) do in fact require that the constituent values be ordered. A 2-for-1 sale on apples is not the same as a 1-for-2 sale on apples—the first is welcome, the second reports a price increase. Similarly, a win-loss record of 5 in 7 is different from a win-loss record of 7 in 5——in this case, the first is again welcome, while the second is impossible. One might, of course, use the implausibility of a reverse sale and the impossibility of winning more games than have been played to one's analytical advantage by supposing that the implausible or impossible order is ruled out independently for pragmatic reasons and an ordered degree plurality would suffice after all. But this would at best distinguish which number plays which role. It wouldn't explain why their relative linear order in the sentence seems to be fixed.

Additionally, individual plurals give rise to patterns of inference that aren't reflected in irreducible composite MPs:

(61) a. Floyd and Clyde sneezed and Bertha and Maude sneezed too.
entails: Floyd, Maude, Clyde, and Bertha sneezed.

b. The score of the first game was 10 to 1, and the score of the second was 5 to 3.
doesn't entail: #The scores of the games were 10, 1, 5, and 3.

When a plurality consisting of two individuals is summed with one consisting of two others, the resulting plurality levels the distinctions among these component parts—it makes no distinction between the individuals that were initially part of the first plurality and those that were part of the second. As (61) reflects, scores don't behave this way. It might be possible to enrich the plural degree analysis with degree groups (in the style of Landman 1989, 1996) to alleviate this by treating this as a case of intermediate distributivity, but such an analysis would, again, be swimming against the current of the language.

Even if we accepted that there were certain restrictions governing the order of constituents in certain composite MPs, it would still be unclear whether plurality would be the right move for these objects. For instance, one of the arguments Dotlačil & Nouwen (2016) offer is the aforementioned cumulative/distributive distinction, as in a sentence such as (62):

(62) The cheetahs are faster than the gazelles.

The distributive reading states that no gazelle was faster than any cheetah (in other words, that each cheetah was faster than each gazelle), but this needn't be the only interpretation. A cumulative reading is available where every cheetah is faster than some gazelle, and that there is some gazelle slower than every cheetah, but that it's not necessarily the case for every individual gazelle that they are slower than every individual cheetah.

But when it comes to irreducible composite MPs, it's not at all clear that such a distinction even exists, let alone is discernible:

(63) 10 to 1 is a bad score.

While a cumulative reading of this is naturally available (where the MP *10 to 1* is treated as a unit), the distributive reading (where 10 is a bad score and 1 is a bad score) is nonsensical. What's more, (63) lacks the flavor of plurality reflected in (62) or even (58), which maintains this sense of plurality despite the fact that *every girl* isn't itself plural—a clear indication, therefore, of a plural degree at work, one that (63) lacks.

What the evidence for plural degrees establishes is that there that degrees may have a richer part structure than ordinarily assumed, and that it is sometimes necessary to package multiple degrees together in the semantics. But at least in English and numerous other languages, irreducible composite MPs need even more than this. The plain fact seems to be that the elements of such measure phrases are not interchangeable, and it is necessary to reflect the difference between them in the model.

5.3 *Tuple degrees*

The alternative to plural degrees is relatively straightforward. It is to bite the bullet and accept that some degrees are actually composed of tuples of other degrees. We'll call such degrees TUPLE DEGREES. To make this more explicit, we will suppose that product types are available among degrees, so that if d and d' are members of the domain of degrees D_d, then $\langle d, d' \rangle$ is a member of $D_{d \times d}$. We will make this recursive by supposing a potentially suspect but convenient variety of recursion: we will assume that any product degree type is itself a member of the domain of degrees D_d. Degree triples are necessary to express volumes. We haven't found examples of reference to larger degree tuples attested outside of scientific language.

To illustrate how the system might work, let's focus on the version of *by* that combines degrees of linear extent. This illustrates the necessary recursion. The tuple degree in (64a) can be combined with an ordinary degree to yield the tuple degree in (64b), and so on in (64c):

(64) a. 3 cm by 4 cm
b. 3 cm by 4 cm by 5 cm
c. 3 cm by 4 cm by 5 cm by 20 minutes
(presumably in a physics context)

The denotations—keeping in mind that we've set aside tropes for in this section—would be as in (65):

(65) a. $[\![\,\textit{3 cm by 4 cm}\,]\!] = \langle \mathbf{3cm}, \mathbf{4cm} \rangle$
b. $[\![\,\textit{3 cm by 4 cm by 5 cm}\,]\!] = \langle \mathbf{3cm}, \mathbf{4cm}, \mathbf{5cm} \rangle$
c. $[\![\,\textit{3 cm by 4 cm by 5 cm by 20 minutes}\,]\!] = \langle \mathbf{3cm}, \mathbf{4cm}, \mathbf{5cm}, \mathbf{20min} \rangle$

How do we build the denotations in (65)? We'll assume that spatial *by* applies to a singleton degree and another degree (singleton or tuple) and yields a degree in which one degree has been appended to the other:

(66) $[\![\,\textit{by}_{\textit{spatial}}\,]\!] = \lambda d\ \lambda \langle d_1, \ldots, d_n \rangle \;.\; \langle d_1, \ldots, d_n, d \rangle$

(67) a. $[\![\,\textit{by}_{\textit{spatial}}\,]\!]([\![\,\textit{4 cm}\,]\!])([\![\,\textit{3 cm}\,]\!]) = \langle \mathbf{3cm}, \mathbf{4cm} \rangle$
b. $[\![\,\textit{by}_{\textit{spatial}}\,]\!]([\![\,\textit{5 cm}\,]\!])([\![\,\textit{3 cm by 4 cm}\,]\!]) = \langle \mathbf{3cm}, \mathbf{4cm}, \mathbf{5cm} \rangle$

This provides an argument for why tuple degrees should in fact be degrees. If singleton and tuple degrees both count as degrees, the generalization about *by* is that it combines degrees to yield new tuple degrees. If singleton and tuple degrees are entirely different types, *by* would have to have a flexible denotation in a way that doesn't correspond to the intuitively simple generalization about its semantic contribution.

5.4 *A case study of tuple degrees in action: scores*

Having addressed spatial irreducible composite MPs in general, let's linger briefly on scores. Like areas, scores are also tuple degrees, which verbs such as *beat* take as arguments. First, a tuple degree like *6 to 4* is constructed from singleton degrees by *to*:

(68) a. $[\![\,\textit{to}_{\textsc{score}}\,]\!] = \lambda d \lambda d' \;.\; \langle d', d \rangle$
b. $[\![\,[_{PP}\ 6\ [_{P'}\ \textit{to}_{\textsc{score}}\ 4]]\,]\!] = \langle \mathbf{6}, \mathbf{4} \rangle$

Beat applies to an individual and the resulting tuple degree:

(69) a. $[\![\, beat\,]\!] = \lambda x \lambda d_{d\times d}\, \lambda y\ .\ \mathbf{beat}(y, x, d)$

b. $[\![\, \textit{Colombia beat Uruguay 6 to}_{\text{SCORE}}\ \textit{4}\,]\!]$
$= \mathbf{beat}(\mathbf{Colombia}, \mathbf{Uruguay}, \langle \mathbf{6}, \mathbf{4}\rangle)$

To explain how *beat* works, it would be useful to capture the crucial entailment that if Colombia beat Uruguay 6 to 4, it is also true that Colombia beat Uruguay by 2 points. That's expressed with a differential form of *by*:

(70) Colombia beat Uruguay by$_{\text{DIFF}}$ 2.

This *by* maps a degree-pair to a degree generalized quantifier (von Stechow 1984, Heim 2000, a.o.), which raises by Quantifier Raising in the standard fashion. To express this, we'll use a difference function **diff**, and to improve clarity we'll use $d_{\mathbf{score}}$ as a more transparent variable name for the score degree:

(71) $[\![\, by_{\text{DIFF}}\,]\!] = \lambda d\, \lambda P_{\langle d\times d, t\rangle}\ .\ \exists d_{\mathbf{score}} \in D_{d\times d}\, [P(d_{\mathbf{score}}) \wedge \mathbf{diff}(d_{\mathbf{score}}) = d\,]$

This predicts correctly that *by* is obligatory only in its differential use, as noted in (31). In its absence, there would be a sort clash. The full sentence denotation is built up in (72):

(72) a. Colombia beat Uruguay by$_{\text{DIFF}}$ 2.

b.
- TP, t
  - PP, $\langle\langle d\times d, t\rangle, t\rangle$
    - P, $\langle d, \langle\langle d\times d, t\rangle, t\rangle\rangle$
      - *by*$_{\text{DIFF}}$
    - DP, d
      - *2*
  - TP, $\langle d\times d, t\rangle$
    - λd_1
    - TP, t
      - DP, e
        - *Colombia*
      - VP, $\langle e, t\rangle$
        - VP, $\langle d\times d, et\rangle$
          - V, $\langle e, \langle d\times d, et\rangle\rangle$
            - *beat*
          - DP, e
            - *Uruguay*
        - PP, $d\times d$
          - d_1

c. $[\![\, \lambda d_1 \textit{ Colombia beat Uruguay } d_1 \,]\!]$

$$= \lambda D_{d\times d}\,.\,\mathbf{beat}(\mathbf{Colombia}, \mathbf{Uruguay}, D)$$

$[\![\, \textit{by}_{\text{DIFF}} \,]\!]([\![\, \textit{2} \,]\!])([\![\, \lambda d_1 \textit{ Colombia beat Uruguay } d_1 \,]\!])$

$$= \exists d_{\mathbf{score}} \in D_{d\times d} \left[\begin{array}{l} [\![\, \lambda d_1 \textit{ Colombia beat Uruguay } d_1 \,]\!](d_{\mathbf{score}}) \wedge \\ \mathbf{diff}(d_{\mathbf{score}}) = 2 \end{array} \right]$$

$$= \exists d_{\mathbf{score}} \in D_{d\times d} \left[\begin{array}{l} \mathbf{beat}(\mathbf{Colombia}, \mathbf{Uruguay}, d_{\mathbf{score}}) \wedge \\ \mathbf{diff}(d_{\mathbf{score}}) = 2 \end{array} \right]$$

This correctly predicts that (72) requires that Colombia beat Uruguay by some score with a difference between its members of 2.

A nice consequence of the QR strategy is that it also predicts scope ambiguities of the sort discovered by Heim (2000) when score differentials interact with modals. Suppose that an organized crime syndicate has imposed the requirement in (73):

(73) Colombia must beat Uruguay by exactly 2.

The interaction of the differential and the modal will give rise to two readings:

(74) **Wide-scope modal reading**

a. must [[by$_{\text{DIFF}}$ exactly 2] [λd_1 Colombia beat Uruguay d_1]]

b. $\Box \exists d_{d\times d}[\mathbf{beat}(\mathbf{Colombia}, \mathbf{Uruguay}, d) \wedge \mathbf{diff}(d) = \mathbf{2}]$

c. 'It is required that Colombia beat Uruguay by a score whose difference is exactly 2.'

(75) **Wide-scope existential reading**

a. [by$_{\text{DIFF}}$ exactly 2] [λd_1 must Colombia beat Uruguay d_1]

b. $\exists d_{d\times d}[\Box\, \mathbf{beat}(\mathbf{Colombia}, \mathbf{Uruguay}, d) \wedge \mathbf{diff}(d) = \mathbf{2}]$

c. 'There's a particular score by which Colombia is required to beat Uruguay, and its difference is exactly 2.'

Both describe scenarios in which the match is fixed, but with different varieties of fixing. In the first, any score will do so long as at the end, Colombia be ahead by 2. In the second, a particular score is required, and it happens to be true of that score that it has a differential of 2.

5.5 *Multidimensionality and the One-Dimensional Degree Construction Conjecture*

A striking aspect of the view that degrees may come in tuples is that it makes possible what might be called a multidimensional degree semantics. 'Dimension' here is to be taken very literally. An area is literally a measure across two dimensions, and a volume across three.

This is worth recognizing because it brings to light an intriguing gap. If degrees can represent multiple dimensions in a single package, and if predicates like verbs can take such packages as arguments, one might imagine that degree constructions in a language could likewise operate across multiple dimensions simultaneously. Suppose, for example, that the living room is 20 feet by 18 feet and the dining room is 15 feet by 10 feet. One can compare their sizes straightforwardly with comparatives like the ones in (76):

(76) a. The living room is 5 feet longer than the dining room.

b. The living room is 8 feet wider than the dining room.

Both, in this scenario, are true. And *20 feet by 18 feet* is a measure phrase that can manifestly occur in argument positions that involve comparisons of size:

(77) 20 feet by 18 feet isn't $\left\{\begin{matrix}\text{big enough}\\ \text{too big in the Midwest}\end{matrix}\right\}$.

In light of that, there is every reason to expect expect that it could occur as the differential in a comparative, and yet it seems not to:

(78) ?The living room is 5 feet by 8 feet bigger than the dining room.

The sentence is grammatical, and it doesn't immediately strike one as semantically anomalous either. But what does it mean? Like grammatical illusions of Phillips et al. (2011)—also known as Russia sentences after a favorite example, *More people have been to Russia than I have*—(78) seems just fine until one reflects on what it means and realizes it has no determinate truth conditions. At least in the case of (78), we know exactly what the sentence is trying to mean: something like the conjunction of the sentences in (76). And yet, it doesn't quite manage to mean that. Perhaps the difference in room sizes should be reported instead as *3 feet by 10 feet*, swapping which dimensional measure is compared to which? But using this differential results in no improvement.

At most, it *might* be possible to judge (78) true on a reading in which the dining room and living room are perfectly congruent, precisely the same size and shape except that the living room has a single addition whose area is 5 feet by 8 feet.

The only way to build the differential comparative we seek is to simply compute the square footage of each room and compare those directly:

(79) The living room is 210 square feet bigger than the dining room.

That's clearly true. This seems to reflect that the English comparative construction is restricted to one-dimensional measurement, even though it's easy to imagine otherwise. There are concerns about which dimension is compared to which, of course, but that's no obstacle in principle—one might have thought that the differential comparative could be ambiguous, judged true on any combination of dimensions. Yet we have found no English degree construction that works this way, and some informal queries have yielded no examples in other languages.

Further research is of course required, but for the sake of encapsulating things in a suitably provocative way, (80) suggests a possible crosslinguistic universal:

(80) **The One-Dimensional Degree Construction Conjecture**
In any language with differential comparative or excessive constructions, differential measure phrases can express comparison across only one dimension.

This seems plausible. But to be clear, nothing said here explains why this should be the case—even just for English, much less across languages.

5.6 *A momentary return to tropes*

What, then, of tropes? They have played no role in the analysis of irreducible composite MPs offered here. That seems fair enough—we don't expect that any class of constructions make use of all available semantic tools. But minimally, the analysis of irreducible composite MPs must be compatible with the analysis of specialization in MPs offered in the previous section. We believe that it is, once a few simplifying assumptions are set aside.

There are two ways in which this section has differed from the one in the previous section. In the previous section, we treated MPs as denoting properties of degree tropes (or in some cases, relations). Here, we treated them as denoting degrees directly. Second, degree tropes were crucial in the previous section, and apparently extraneous in this one.

The shift from degree-denoting to property-denoting MPs is straightforward, being essentially a matter of type shifting. The score *6 to 4*, for example, can be treated as denoting a tuple degree, as in (81a), or as denoting the property of being that tuple degree, as in (81b):

(81) a. $[\![\textit{6 to 4}]\!] = \langle \mathbf{6}, \mathbf{4} \rangle$

b. $[\![\textit{6 to 4}]\!] = \lambda d \in D_{\langle d \times d, t \rangle}[d = \langle \mathbf{6}, \mathbf{4} \rangle]$

Adding tropes back into the system is likewise formally a relatively small step. We need to recognize that our argument for tuple degrees was really an argument that tuples need to be added to any theory of degrees, whatever it is. A theory in which degrees are represented with tropes is no different, and the change can be represented straightforwardly by replacing the pairs of degrees in (81b) with pairs of tropes—that is, replacing $d \times d$ in (81b) with $d^t \times d^t$.

But it's certainly not the case that our trope-based analysis of MP specialization is irrelevant to irreducible composite MPs. There is specialization there, too, and as with their reducible counterparts, these MPs express their specialization in English with a preposition. The score *4 to 6* can't equally well be #*4 from 6*, #*4 against 6*, #*4 above 6*, or #*4 in 6*. On the other hand, the record of wins in a series of games can be expressed as *4 in 6*, but not #*4 to 6* or anything else. Yet, without making additional assumptions, both *4 to 6* and *4 in 6* would involve exactly the same degree: $\langle 4, 6 \rangle$. This is, of course, a special case of the puzzles around MP specialization, and the same analytical path beckons. A score involves two degree tropes, which can be referred to with trope-denoting expressions like *the number of goals scored by Colombia*, a counterpart to e.g. *the length of my hand*. A record of wins in a series likewise is composed of not of bare numbers but a count of games won and games played. Like any measure phrases, irreducible ones can be specialized, and so like any measure phrases they may require appealing to tropes.

6 Concluding remarks

There are three empirical points to highlight. The first: arithmetically identical measure phrases like *50%* and *1 in 2* do not have identical syntactic distributions, and this has important implications for the relationship between the facts of the world and how they are encoded in language. Second, we brought attention to a previously unrecognized natural class of expressions, composite measure phrases, composed of other measure phrases. These can be divided into three varieties: proportional, irreducible, and range. Third, the idiosyncratic restrictions on composite MPs can be viewed as part of a broader picture of such specialized restrictions in MPs more generally, itself a phenomenon not well understood theoretically.

Proportional composite MPs (like *3 in 4* in reference to probability), the first of our three classes, are interesting in part because they provide an excellent proving ground for capturing fine-grained lexical restrictions. To do this, we adopted a version of the theory pioneered by Moltmann in which tropes play a crucial role in the semantics of degree expressions. Tropes allowed us to capture distinctions among measure phrases that ordinary degrees couldn't deliver, making it possible to make sense of how language makes finer-grained distinctions than arithmetic does. One might imagine alternatives to tropes as a means of capturing this, but it's clear that any adequate explanation will require a richer concept of degree than is commonly assumed. Another respect in which proportional composite MPs are potentially important is the very fact that they are proportional, and therefore associated with a sort of degree that has recently become a focus of increasingly intense interest.

Irreducible composite MPs (like *2 to 1* as a sports score) are unlike proportional ones in just the way reflected by the name: they are irreducibly a matter of grouping together several distinct degrees. We entertained the idea that pluralities of degrees may suffice for this, and in some languages there is evidence that this is the right path. But English isn't one of them. In English, as in many other widely spoken languages, such MPs show, we argued, that ordered tuples of degrees should be recognized as themselves a complex sort of degree.

Range composite MPs (like *3 through 5*) are a topic we set aside here for the sake of brevity, but have developed elsewhere Gobeski & Morzycki 2022.

All of these break some new empirical ground. Our data here has been chiefly English, but such expressions seem to occur in many languages and seem to be a promising area for inquiry crosslinguistically.

References

Ahn, Dorothy & Uli Sauerland. 2015. 'The grammar of relative measurement'. In Sarah D'Antonio, Mary Moroney, & Carol Rose Little (eds.), *Proceedings of Proceedings of Semantics and Linguistic Theory (SALT) 25*, pp. 125–142.

Ahn, Dorothy & Uli Sauerland. 2017. 'Measure constructions with relative measures: Towards a syntax of non-conservative construals'. *The Linguistic Review* **34**(2), 215–248.

Anderson, Curt & Marcin Morzycki. 2015. 'Degrees as kinds'. *Natural Language and Linguistic Theory* **33**(3), 791–828.

Bale, Alan. 2006. *The Universal Scale and the Semantics of Comparison*. Doctoral dissertation, McGill University.

Bale, Alan & Bernhard Schwarz. 2022. 'Measurements from per without complex dimensions'. In Starr et al. (2022), pp. 543–560.

Beck, Sigrid. 2010. 'Quantifiers in *than*-clauses'. *Semantics and Pragmatics* **3**(1), 1–72.

Coppock, Elizabeth. 2022a. 'Challenge problems for a theory of degree multiplication (with partial answer key)'. In Nicole Dreier, Chloe Kwon, Thomas Darnell, & John Starr (eds.), *Proceedings of Semantics and Linguistic Theory (SALT) 31*. LSA.

Coppock, Elizabeth. 2022b. 'Division vs. distributivity: Is per just like each?' In Starr et al. (2022), pp. 384—403.

Dotlačil, Jakub & Rick Nouwen. 2016. 'The comparative and degree pluralities'. *Natural Language Semantics* **24**, 45–78.

Francez, Itamar & Andrew Koontz-Garboden. 2011. 'Semantic variation and the grammar of property concepts'. Ms., University of Chicago and University of Manchester.

Francez, Itamar & Andrew Koontz-Garboden. 2017. *Semantics and Morphosyntactic Variation: Qualities and the Grammar of Property Concepts*. Cambridge University Press, Cambridge.

Gehrke, Berit & Marcin Wągiel. 2023. 'On word order and non-conservative percentage quantification in Slavic and German'. *Glossa: A Journal of General Linguistics* **43**(1).

Gobeski, Adam. 2019. *Factor phrases: The semantics of multiplicative modification of events, degrees, and nominals, and the grammar of arithmetic*. Doctoral dissertation, Michigan State University.

Gobeski, Adam & Marcin Morzycki. 2018. 'Percentages, relational degrees, and degree constructions'. In Dan Burgdorf, Jacob Collard, Sireemas Maspong, & Brynhildur Stefánsdóttir (eds.), *Proceedings of Proceedings of Semantics and Linguistic Theory (SALT) 27*. eLanguage.

Gobeski, Adam & Marcin Morzycki. 2022. 'Ranges'. Ms., UBC.

Heim, Irene. 2000. 'Degree operators and scope'. In Brendan Jackson & Tanya Matthews (eds.), *Proceedings of Semantics and Linguistic Theory (SALT) 10*, pp. 40–64. CLC Publications, Ithaca, NY.

Kennedy, Christopher. 1997. *Projecting the Adjective: The Syntax and Semantics of Gradability and Comparison*. Doctoral dissertation, UC Santa Cruz. Published in 1999 by Garland, New York.

Krifka, Manfred. 1989. 'Nominal reference, temporal constitution, and quantification in event semantics'. In Renate Bartsch, Johan van Benthem, & Peter van Emde Boas (eds.), *Semantics and Contextual Expression*, pp. 75–115. Foris, Dordrecht.

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

Landman, Fred. 1996. 'Plurality'. In Shalom Lappin (ed.), *The Handbook of Contemporary Semantic Theory*, pp. 425–457. Blackwell Publishers, Oxford.

Mertz, D. W. 1996. *Moderate realism and its logic*. Yale University Press, New Haven.

Moltmann, Friederike. 2009. 'Degree structure as trope structure: a trope-based analysis of positive and comparative adjectives'. *Linguistics and Philosophy* **32**(1), 51–94.

Moltmann, Friederike. 2015. 'States versus tropes. comments on curt anderson and marcin morzycki: 'degrees as kinds''. *Natural Language and Linguistic Theory* **33**(3), 829–841.

Phillips, Colin, Matthew W. Wagers, & Ellen F. Lau. 2011. 'Grammatical illusions and selective fallibility in real-time language comprehension'. In Jeffrey T. Runner (ed.), *Experiments at the Interfaces*, *Syntax and Semantics*, vol. 37, pp. 147–180. Emerald Group Publishing Limited.

Rawlins, Kyle. 2013. 'On adverbs of (space and) time'. In Boban Arsenijević (ed.), *Studies in the Composition and Decomposition of Event Predicates*, pp. 153–197. Springer.

Solt, Stephanie. 2018. 'Proportional comparatives and relative scales'. In Robert Truswell, Chris Cummins, Caroline Heycock, Brian Rabern, & Hannah Rohd (eds.), *Proceedings of Proceedings of Sinn und Bedeutung 21*.

Starr, John R., Juhyae Kim, & Burak Öney (eds.). 2022. *Proceedings of Semantics and Linguistic Theory (SALT) 32*. LSA.

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

Wągiel, Marcin. 2019. 'Partitives, multipliers and subatomic quantification'. In *Proceedings of Sinn und Bedeutung 2019*, vol. 23, pp. 443–460.

Wągiel, Marcin. 2020. 'Entities, events, and their parts: The semantics of multipliers in slavic'. *Current developments in Slavic linguistics: Twenty years after* pp. 105–130.