Composite measure phrases: Odds, scores, flavors of scales, and the taxonomy of MPs[1]

Adam GOBESKI — *Unaffiliated*
Marcin MORZYCKI — *University of British Columbia*

Abstract. Purely numerical measure phrases (MPs) like *three* or *two thirds*, which lack a unit term, are often construed as denoting degrees on a single numerical scale. This paper examines an apparently unrecognized class of complex purely numerical MPs such as *two in three* and *six to one*, which we term COMPOSITE MPS. Such MPs demonstrate, we argue, that mathematically equivalent MPs aren't always equivalent linguistically and that different purely mathematical MPs refer to degrees on different and incommensurable scales. Indeed, some, such as sports scores, seem to refer irreducibly to tuples or pluralities of degrees. We classify composite MPs into three varieties, each of which requires a distinct analysis.

Keywords: degrees, tuple degrees, measure phrases, proportion, ontology of scales, odds, mathematical language, composite MPs

1. Introduction

There are many ways to refer linguistically to proportions: forms like *25%, one fourth, a quarter, .25, 1 in 4*, among others. It is perhaps surprising that language provides so many ways of characterizing precisely the same thing. We argue here that this is an illusion, and that underlying this diversity of forms is a diversity of meanings—that mathematically identical measure phrases may be grammatically different. This point emerges especially clearly from the novel data we will focus on: what we'll call COMPOSITE MEASURE PHRASES, which are measure phrases with other measure phrases as parts.

In section 2, we make the case that MPs whose referents are mathematically identical do not have the same distribution, showing that there must be some difference in their denotations. In section 3, we propose a taxonomy of composite MPs and some diagnostics to distinguish them. In section 4, we suggest a way of conceptualizing the fine-grained distinctions among scales necessary to capture the data. In 5, we propose an analysis of what we call irreducible composite MPs that relies on recognizing degrees composed of pairs or other *n*-tuples of other degrees. Section 6 concludes.

2. Mathematically equivalent MPs aren't linguistically equivalent

It's natural to assume that in the absence of an overt unit term such as *meter* or *pound*, numerical measure phrases (MPs) like the ones in (1) refer to degrees on the same scale:

[1] We would like to thank Elizabeth Coppock, Henry Davis, Hotze Rullmann, Marcin Wągiel, Peter Sutton, Rose Underhill, Starr Sandoval, Stephanie Solt, Todd Snider, audiences at SuB 26 and at UBC.

 In: Daniel Gutzmann and Sophie Repp (eds.): Proceedings of Sinn und Bedeutung 26, pp. 322–339. University of Cologne.

(1) 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 (1) 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:

(2) a. one fourth
b. 25% (“twenty-five percent”)
c. .25 (“point two five”)
d. 1 $\left\{\begin{matrix} \text{in} \\ \text{out of} \end{matrix}\right\}$ 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. For instance, equatives support percentages and fractions but not *in* MPs, *to* MPs, and decimals:

(3) Bertha is $\left\{\begin{matrix} \text{25\%} \\ \text{a quarter} \\ {}^{\#}\text{.25} \\ {}^{\#}\text{1 in 4} \\ {}^{\#}\text{1 to 3} \end{matrix}\right\}$ as tall as Clyde.

Similarly, comparatives don’t consistently allow fractions either:

(4) Bertha is $\left\{\begin{matrix} \text{25\%} \\ {}^{?}\text{a quarter} \end{matrix}\right\}$ taller than Clyde.

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

(5) Let’s disburse $\left\{\begin{matrix} \text{25\%} \\ \text{a quarter} \\ {}^{\#}\text{.25} \\ \text{1 in 4} \\ {}^{\#}\text{1 to 3} \end{matrix}\right\}$ of the donations.

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

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

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

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

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

(8) I think $\left\{ \begin{array}{l} {}^{??}\text{25\%} \\ \text{one fourth} \\ {}^{\#}\text{a quarter} \\ \text{.25} \\ {}^{\#}\text{1 in 4} \\ {}^{\#}\text{1 to 3} \end{array} \right\}$ 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 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. These restrictions can be understood, however, in terms of incommensurabilty across scales.

Under normal circumstances, only degrees that are on the same scale can freely be compared (von Stechow 1984; Kennedy 1997 among others; but cf. Bale 2006). Degrees on different scales normally can't be. It's not clear, for example, what it would take to make either version of (9) true:

(9) a. #Floyd is much taller than Bertha is Episcopalian.
b. #The degree to which Floyd is tall is (much) greater than the degree to which Bertha is Episcopalian.

The scales of height and commitment to being Episcopalian are simply incommensurable.

Direct comparison sentences such as (9b) especially can be used to our advantage, because we can use them to test whether various MPs are scale-mate degrees. The first observation is that a simple bare integer like *2* can be compared with only certain MPs, mainly fractions and decimals:

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

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

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

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

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

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

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 the direct comparison sentences behaved differently depending on which MPs were compared. The predicate *is greater* is perfectly happy to compare percentages and *in* MPs, but as (11–12) 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 essential generalizations. But although we can observe some smaller regularities (such as the fact that bare numerals seem to be on the same scale as fractions and decimals), there are no robust broad generalizations that can explain all these facts. Which MP uses which scale seems to be a lexical matter. We have to acknowledge the inherent idiosyncrasy at play here: 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 and Sauerland 2015, 2017; Gobeski and 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 between these "flavors" of scales, one that differentiates between, say, odds and decimals (and we'll have more to say on this particular topic in section 4). It's also worth noting that the syntax matters: *1 to 2*, for instance, expresses specifically odds, not just any proportion, and that determination is based on the use of *to*. Yet *to* doesn't have to represent odds; what do we do with the *to* present in scores (*lost the game by two to three*) or in ranges (*two to three capybaras*)? To that end, let's take a closer look at composite MPs.

3. Three kinds of composite MPs

3.1. Two additional varieties

Having hopefully established that numeric MPs aren't always commensurable, in this section we turn to composite MPs specifically. Again, composite MPs are measure phrases that have other MPs as subconstituents, such as *3 to 1*, *3 in 4*, and *3 out of 4*. These examples, and all we have consider thus far, express proportions. But these are by no means the only kinds of composite MPs in English, as (14) illustrates:

(14) 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 (14) are all composite, just as those in (13) are, but they don't seem to be straightforwardly expressing proportions. While (14a), for instance, could possibly be considered a form of proportion, it's certainly not equivalent to *3 to 1*. Even this stretch is not available for the other sentences. This suggests that there are at least three varieties of composite MPs.

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. We will mostly set these aside here, but see Gobeski and Morzycki (2022).

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.

In order to distinguish between these three kinds of composite MP, we 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:

(15) a. PROPORTIONAL:
The probability of winning is 1 in 3. ⇔ The probability of winning is 2 in 6.

b. RANGE:
Bertha saw 1 to 3 capybaras. ⇎ Bertha saw 2 to 6 capybaras.

c. IRREDUCIBLE:
Venezuela beat Australia by a score of 3 to 1.
⇎ Venezuela beat Australia by a score of 6 to 2.

In (15a), if the probability of winning is 1 in 3, it's therefore also 2 in 6. But in (15b), if Bertha saw 1 to 3 capybaras, it doesn't follow that she saw 2 to 6 of them. Likewise, in (15c), 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.

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 be such answers, but irreducible composite MPs cannot:

(16) 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, of course, possible to ask how many meters *square* a room is, and to receive the response in (16c), though it's not clear that it counts as answering the question directly. Rather, it provides a stronger claim from which an answer can be computed, and crucially, that answer must be a single figure expressed in square meters.

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]

(17) 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:

(18)

	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 largely set aside ranges (but see Gobeski and 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, seem to 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 (19), which systematically entail a corresponding differential:

(19) Venezuela beat Australia (by (a score of)) 6 to 2. $\Rightarrow$ 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 (20) demonstrates:

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

[2]Some people accept (17c), 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 at play.

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

(21) 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 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*:

(22) Their army outnumbers ours 3 to 1.

Loosely speaking, the ratio in (22) 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 (22) 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.

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

(23) 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:

(24) 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 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:

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

(26) 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 (26)) 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 along slightly different flavors of scale, different irreducible composite MPs do too:

(27) 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.

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 flavors of scales.

4. Flavors of scales

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 (28):[3]

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

It's a consequential fact about such sentences that they have determinate truth conditions, more or less, 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.[4] For analogous reasons, even in a discourse that explicitly mentioned, say, a 12-point font, (29) remains deeply odd as well:

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

A skeptic might suggest that the real difficulty in (28) is that for most English speakers, *furlong* is unfamiliar. But no such objection is possible to (29). In an age in which we all discuss font size, points—once an obscure unit of measurement known mainly to typesetters—are familiar to all of us. Yet (29) is no better than (28), 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

[3]To save the effort of Googling: (28) and (29) are approximately correct about the actual height of the Empire State Building.

[4]In some grammars, it might just be that *furlongs* only measure horizontally, but that makes the same point about fine-grained distinctions among scales. A more general case is that in Canadian use, metric doesn't measure the height of a person, but is obligatory for mountains.

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. 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 proportions, of course, but it is not a fact of language. More generally, various specialized measure phrase types can be said to measure different sorts of things or in different ways—which amounts to saying that they measure along slightly different scales.

That's what the observations we have made so far suggest. It will help to have a term for these fine-grained distinctions. We'll call these different FLAVORS of scales. Thus just as the scale of font size is a particular flavor of the scale of linear extent, the scale of gambling odds is a particular flavor of the scale of probabilities, and that scale might in turn be a flavor of the scale of general proportions. These kinds of relations among scales are interesting because, if this line of thought is on the right track, they make it possible to explore the ontology of scales by exploring the lexical semantics of measure phrases. A priori, there are various ways in which one might think about these relations among scales. Perhaps some scales simply contain others as parts. Perhaps the relevant notion is something closer to the taxonomy that kind-subkind relations involve (e.g., POODLE is a subkind of DOG, which is a subkind of MAMMAL). Or perhaps the right way of thinking about it is in terms of mappings from more particular scales to more general ones, in the spirit of Bale (2006). We don't have firm convictions about which of these is the right course, but for the sake of explicitness we'll sketch a version of the first option.

First, a reminder of the analytical intuitions we seek to capture. We need for different flavors of measure phrase to correlate with different flavors of scale. This information needs to be encoded in the semantics of the non-numerical grammatical components of the measure phrase—for example, in a composite measure phrase like *1 in 4*, it has to be the preposition *in* that identifies the whole measure phrase as necessarily a proportion and not e.g. an all-purpose integer. In this respect, *in* and other such non-numerical components resemble unit terms like *inches* or *kilograms*, which also relate a numeral to a particular flavor of scale. *Inches*, for example, relates a numeral on the general scale of numerical measurement to a degree of linear extent; *kilograms* relates a numeral to a degree of weight; and so on.

One simple but non-standard way of thinking about unit terms (cf. Krifka 1989) that reflects this idea is to take them to denote a function, which we'll write δ_S, from a real number to a degree on the appropriate scale S. This means δ_S is of type $\langle d,d \rangle$, and that both *3* and *3 inches* in (30) are degrees:[5]

[5] A more standard approach might treat *inches* as a function from numeral degrees to properties of individuals with a particular measure, as in (30), where $\mu_{\textbf{inches}}$ is a measure function measuring linear extent in inches:

(i) $[\![\textit{inches}]\!] = \lambda n \lambda x[\mu_{\textbf{inches}}(x) = n]$

We're not convinced that the facts at issue here are arguments against such an approach as such, but they make the alternative course we'll take more natural. For reflection on whether *three inches* denotes a degree or a property of individuals, see Schwarzschild (2005).

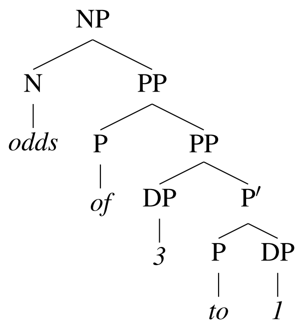
(30) a. $[\![\textit{inches}]\!] = \delta_{\mathbf{inches}}$

b. $[\![\textit{3 inches}]\!] = \delta_{\mathbf{inches}}(3)$

As this reflects, unit terms actually themselves pose questions about the relation among scales. One might take *3 inches* to simply name a degree on the scale of linear extent, the same one as is named by *7.62 centimeters*. The only difference would be that different measure functions are used to get to the same place: in one case, an inches measure function, and in the other, a centimeter measure function. But a conception closer to what we're exploring would take linear-extent-in-inches to be its own flavor of scale, systematically related to the scale linear-extent-in-centimeters. It would be interesting to discover evidence analogous to composite measure phrases that suggests that construing this difference as a difference in scale flavor as well, such as an adjective that only supports measurement in inches.

Next, the syntax. We'll focus on prepositional structures like *3 to 1*. We'll assume *3* is a measure phrase specifier of *to*, and that *1* is the complement of *to*:

(31) a. The bookies are offering odds of 3 to 1.

b.


This could of course be extended to 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 phrases, such as *3 plus 5* (Gobeski 2019). In (31), we've embedded the whole thing under *odds of*, so that the measure phrase PP *3 to 1* is itself the complement of another preposition, *of*.

Combining the semantics and the syntax above, the degree-oriented version of the preposition *to* might apply to a real number, itself a sort of degree (we'll use the variable *n* to represent that, but *d* would also have been appropriate):

(32) a. $[\![\textit{to}_{\textit{degree}}]\!] = \lambda n \lambda n' \,.\, \delta_{\mathbf{odds}}\left(\dfrac{n'}{n+n'}\right)$

b. $[\![\textit{3 to}_{\textit{degree}}\ \textit{1}]\!] = \delta_{\mathbf{odds}}\left(\dfrac{3}{4}\right)$

We represent the scale that contains the degree *3 to 1* with **odds**. We set aside cases like *They outnumbered us 3 to 1*.

In *odds of 3 to 1*, the noun *odds* imposes a presupposition that the degree to which it applies is an odds degree. For simplicity, we'll assume it does nothing more than imposing this presupposition:

(33) $[\![\textit{odds}]\!] = \lambda d : d \in S_{\mathbf{odds}} \,.\, d$

Of course, more ontologically sophisticated alternative assumptions about what *odds* might denote are available, but they aren't immediately relevant.[6]

Importantly, even this simple denotation predicts the ill-formedness of (34), where measure phrases that don't measure on the odds scales are impossible:

(34) #The bookies are offering odds of $\left\{\begin{array}{l}\text{2}\\ \text{25\%}\\ \text{a quarter}\\ \text{one fourth}\end{array}\right\}$.

Although *25%*, *a quarter*, and *one fourth* all name proportional degrees, none of them names an odds degree, and in combining with *odds* these give rise to failure of presupposition. That's also the case for measure phrases that name degrees that aren't even proportional, such as *2*.

Of course, we still need denotations for expressions like *a quarter* as well, and crucially, ones that reflect that they don't measure odds. For the sake of explicitness we'll take these to use a scale of generalized fractional measurement:[7]

(35) $[\![\textit{25\%}]\!] = \delta_{\textbf{fractions}}\left(\dfrac{25}{100}\right)$

(36) $[\![\textit{a quarter}]\!] = \delta_{\textbf{fractions}}\left(\dfrac{1}{4}\right)$

(37) $[\![\textit{one fourth}]\!] = \delta_{\textbf{fractions}}\left(\dfrac{1}{4}\right)$

Interestingly, this seems to be a different scale from the one that decimal proportions like *.25* use. As we've seen, English equatives distinguish these:

(38) Bertha is $\left\{\begin{array}{l}\text{ 25\%}\\ \text{ a quarter}\\ \text{\#.25}\\ \text{\#1 in 4}\\ \text{\#1 to 3}\end{array}\right\}$ as tall as Clyde.

That suggests that the equative itself requires measurement on this fraction scale. We'll sketch how this might work. For the sake of consistency we assume a denotation for equatives in which the degree morpheme and standard phrase are a constituent and this whole MP denotes a generalized quantifier which moves to create a degree property as its argument (Bresnan 1973; Heim 2000; cf. Abney 1987; Kennedy 1997; see Morzycki 2016 for comparative discussion). We also need some way of capturing that the measure phrase in (38) expresses a way of getting from one sort of degree to another. We'll represent that with simple multiplication in (39) (but

[6]One particularly interesting issue: expressions like *3 to 1 odds of winning*, where *3 to 1* seems to be in the specifier position of *odds*, or perhaps a member of a compound, and expressions like *3 to 1 {in favor/against}*. We flag this but have nothing more to say about it at this time.

[7]We write (35) with a denominator of 100 rather than 4 to more transparently reflect the intended meaning, but of course $\frac{25}{100}$ and $\frac{1}{4}$ are the same thing.

see Ahn and Sauerland 2015, 2017; Solt 2018; Gobeski and Morzycki 2018; Gobeski 2019 for more):

(39) $[\![as_{degree\ word}]\!] = \lambda d \,.\, \lambda d' : d' \in S_{\mathbf{fractions}} \,.\, \lambda D_{\langle d,t \rangle} \,.\, d \geq \mathbf{max}(D) \times d'$

This combines with an adjective, a measure phrase denoting a degree on the scale of fractional measurement, and a pair of individuals to be compared:

(40) a. $[\![\textit{as Clyde}]\!] = \mathbf{max}\{d : \mathbf{tall}(\mathbf{Clyde}, d)\}$

b. $[\![\textit{1 Bertha is } d_1 \textit{ tall}]\!] = \lambda d'' \,.\, \mathbf{tall}(\mathbf{Bertha}, d'')$

c. $[\![as_{degree\ word}]\!]([\![\textit{as Clyde}]\!])([\![\textit{25\%}]\!])([\![\textit{1 Bertha is } d_1 \textit{ tall}]\!])$

$$= \mathbf{max}\{d'' : \mathbf{tall}(\mathbf{Bertha}, d'')\} \geq \mathbf{max}\{d : \mathbf{tall}(\mathbf{Clyde}, d)\} \times \delta_{\mathbf{fractions}}\left(\frac{25}{100}\right)$$

Again, the wrong flavor of measure phrase will result in failure of presupposition. A measure phrase that's not on the fraction scale—even if it's arithmetically equivalent to one that is—will not suffice.

It's worth noting that *2 times* and even *1.29 times* are good in equatives and comparatives, but bare *2* and *1.29* aren't. That suggests that *n times*, *n percent*, and fractions all operate on the same scale, the fractional scale. It is the inventory of English measure phrases that lead to this conclusion, and only further lexical semantic investigation will be necessary to make further such discoveries. It seems very unlikely that our way of representing things here—simply subscripting δ with some vaguely appropriate label—will be adequate to the larger task. We will need a more articulated theory of scale flavors and how they relate. It also seems likely that it will be necessary to impose idiosyncratic restrictions on flavors of scales. For example, it seems likely that *n to n′* measure phrases actually presuppose a scale that represents oppositions of some kind, and a number of distinct scale flavors can meet that requirement. Such an approach would be required if it turns out, as we suspect, the class of possible measure phrases in any one context is a part of the class of possible measure phrases in some suitably more promiscuous alternative context.

One broader issue worth reflecting on is the relation between scale flavor and what it is that's measured. It's in the measured object that we see most clearly the distinctions: whether that is skyscrapers or racetracks or fonts. We take that to be sufficient cause for thinking of each variety of measurement invoking a different scale. Not all theories of degrees deliver that result with equal straightforwardness. One classic view is that degree are literally built up of individuals, so that the degree 'six feet' is the equivalence class of everything that is six feet tall (Cresswell 1976). If a scale is ultimately a set of degrees, this view connects measured objects and scales in just the right way. But even if we assume, as is more standard, that degrees are a distinct type in the model, an abstract representation of measurement, there is still a close connection between measured object and the nature of the measurement. The most essential property of degrees is that they support an ordering, which is typically taken to be total. There is indeed a total ordering in length between horse racetracks, which one might express as the 'is a longer horse racetrack than' relation. This relation orders horse racetracks, and only those. It's not defined for anything else. Of course, it's a special case of the 'is longer than' relation, the one associated with the general scale of length. But this is sufficient to show that, even on more abstract theories of degrees, there is a very close connection between what's measured and the scale.

Naturally, it is important conceptually that we as humans can generalize across flavors of scales even when the language doesn't invite us to. But it is not truly a linguistic matter. We know that the odds scale is systematically related to the scale of proportions, but we recognize it only because we are more or less numerate and have embraced the abstractions of arithmetic. It doesn't take too much imagination to call to mind a clever bookmaker or gambler, wily but uneducated, who can reason flawlessly about odds but can't easily express them in terms of fractions. Anyone who has attempted to follow a recipe in imperial units when one is more familiar with metric or vice versa will have a sense of what that might feel like.

Importantly, though, this isn't just a matter of different naming conventions for degrees in different contexts—furlongs for horse tracks, points for fonts, numerical *to* MPs for odds. The relation between odds and proportions seems self-evident because we're so familiar with the necessary conceptual tools. But the same point can be made about the relation between temperature and speed, where the necessary equipment is only slightly more obscure. It's a fact of physics that temperatures can all be expressed as speeds because temperature is ultimately a measure of the speed at which molecules are moving (thermal velocity). But even after we recall this fact, we are no more likely, when asked about the temperature outside, to reply *1038 miles per hour*.[8] That temperatures are a variety of speed is a fact of the world, but that alone doesn't make it a fact of language as well. It does no good to point out that conceptually, facts of the world must 'win' and there's no way to escape the equivalence. That's true of the relation between odds and proportions too. We may know these equivalences are true. But evidently, we don't talk like it.

Of course, we shouldn't lose sight of the fact that conceptually, comparison across flavors of scales can be made sense of readily. One way of thinking about it is that there are conceptual scales—along which we can compare and onto which we can map other scales—that needn't correspond to the lexical scales of any natural language predicate. In some cases, though, a conceptual scale can in fact happen to be the lexical scale associated with some predicate. That may be the case with the general scale of proportions. Any odds degree is conceptually also a proportion, and one might imagine that it is therefore actually a member of the proportion scale too (or mapped onto it). This would, of course, have to be structure preserving, so that if one set of odds is better than another, it must also correspond to a higher proportion:

(41) a. $S_{\mathbf{odds}} \subseteq S_{\mathbf{proportion}}$
b. $\forall d, d' \in S_{\mathbf{odds}}[d \prec_{S_{\mathbf{odds}}} d' \rightarrow d \prec_{S_{\mathbf{proportion}}} d']$

There are more interesting issues in this area, which we don't have the room to address here.

[8]That's room temperature.

5. Irreducible composite MPs and tuple degrees

5.1. Tuple degrees

We've suggested that proportional composite MPs denote relatively ordinary degrees—just proportional ones. But if scores and other irreducible composite MPs denote degrees, as seems plausible, we're left with the question of how those degrees should be represented. We've seen that these behave differently from standard degrees, in that they don't measure along a one-dimensional scale, and any effort to reduce them to a one-dimensional scale would lead to a loss of information.

Consequently, we are faced with the challenge of providing a semantics for an expression that certainly seems like a degree, but can't be thought of one-dimensionally. Unavoidably, this leads to the conclusion that such degrees will need to be made up of multiple other degrees. One possibility is to suppose that they are plural degrees (Dotlačil and Nouwen 2016). That wouldn't provide a way of distinguishing the components of such a degree, as seems to be necessary. A win-loss record of 5 in 7 is different from a win-loss record of 7 in 5. The first is welcome, the second impossible.

Another possibility, which doesn't raise this difficulty, 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 slightly more explicit, one might suppose that tuple degrees are in the domain of degrees D_d. Thus, if d and d' are degrees, then so is $\langle d, d' \rangle$ (and likewise for triples, as in the case of volume, and larger tuples). In order to have this work out, we would use product types to expand the domain of degrees; in other words, the degree domain would include not only standard degrees (D_d) but also tuple degrees ($D_{d \times d}$) as a sort of degree.

To illustrate how the system might work, let's focus on the version of *by* that combines degrees of linear measure to build more complex degrees. This seems to allow degree recursion, so that the tuple degree in (42a) can be combined with an ordinary degree to yield the tuple degree in (42b), and so on in (42c):

(42) a. 3 inches by 4 inches
b. 3 inches by 4 inches by 5 inches
c. 3 inches by 4 inches by 5 inches by 20 minutes *(presumably in a physics context)*

The denotations would be as in (43):

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

How do we build the denotations in (43)? The natural move is to construct these tuple degrees recursively. 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:

(44) $[\![by_{spatial}]\!] = \lambda d\ \lambda \langle d_1, \ldots, d_n \rangle\ .\ \langle d_1, \ldots, d_n, d \rangle$

(45) a. $[\![by_{spatial}]\!]([\![\textit{4 inches}]\!])([\![\textit{3 inches}]\!]) = \langle \mathbf{3in}, \mathbf{4in} \rangle$
b. $[\![by_{spatial}]\!]([\![\textit{5 inches}]\!])([\![\textit{3 inches by 4 inches}]\!]) = \langle \mathbf{3in}, \mathbf{4in}, \mathbf{5in} \rangle$

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.

5.2. A case study of tuple degrees in action: Scores

Having covered spatial irreducible composite MPs, let's now return to scores. Just 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*:

(46) a. $[\![to_{score}]\!] = \lambda d \lambda d' \,.\, \langle d', d \rangle$
b. $[\![[_{PP}\ 6\ [_{P'}\ to_{score}\ 4]]]\!] = \langle \mathbf{6}, \mathbf{4} \rangle$

Beat applies to an individual and the resulting tuple degree:

(47) a. $[\![beat]\!] = \lambda x \lambda d_{d \times d} \lambda y \,.\, \mathbf{beat}(y, x, d)$
b. $[\![Colombia\ beat\ Uruguay\ 6\ to_{score}\ 4]\!] = \mathbf{beat}(\mathbf{Colombia}, \mathbf{Uruguay}, \langle \mathbf{6}, \mathbf{4} \rangle)$

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:

(48) Colombia beat Uruguay by_{diff} 2.

This *by* maps a degree-pair to a degree generalized quantifier (von Stechow 1984; Heim 2000, a.o.), which Quantifier Raises in the standard fashion (**diff** is a difference function, and $d_{\mathbf{score}}$ is a fancy variable name for the score degree):

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

This predicts correctly that *by* is obligatory only in its differential use, as noted in (20). In its absence, there would be a sort clash. The full sentence denotation is built up in (50):

(50) a. Colombia beat Uruguay by_{diff} 2.
b. [by_{diff} 2] [1 Colombia beat Uruguay d_1]
c. $[\![1\ Colombia\ beat\ Uruguay\ d_1]\!] = \lambda D_{d \times d} \,.\, \mathbf{beat}(\mathbf{Colombia}, \mathbf{Uruguay}, D)$
d. $[\![by_{diff}]\!]([\![2]\!])([\![1\ Colombia\ beat\ Uruguay\ d_1]\!])$
$= \exists d_{\mathbf{score}} \in D_{d \times d}\, [[\![1\ Colombia\ beat\ Uruguay\ d_1]\!](d_{\mathbf{score}}) \wedge \mathbf{diff}(d_{\mathbf{score}}) = d]$
$= \exists d_{\mathbf{score}} \in D_{d \times d}\, [\mathbf{beat}(\mathbf{Colombia}, \mathbf{Uruguay}, d_{\mathbf{score}}) \wedge \mathbf{diff}(d_{\mathbf{score}}) = d]$

This correctly predicts that (50) 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 like those discovered by Heim (2000) when score differentials occur with modals. Suppose the requirement in (51) is imposed:

(51) Colombia must beat Uruguay by exactly 2.

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

(52) WIDE-SCOPE MODAL READING
a. must [[by$_{\text{diff}}$ exactly 2] [1 Colombia beat Uruguay d_1]]
b. $\Box\exists d_{d\times d}[\textbf{beat}(\textbf{Colombia},\textbf{Uruguay},d) \wedge \textbf{diff}(d) = \textbf{2}]$
c. 'It is required that Colombia beat Uruguay by a score whose difference is exactly 2.'

(53) WIDE-SCOPE EXISTENTIAL READING
a. [by$_{\text{diff}}$ exactly 2] [1 must Colombia beat Uruguay d_1]
b. $\exists d_{d\times d}[\Box\textbf{beat}(\textbf{Colombia},\textbf{Uruguay},d) \wedge \textbf{diff}(d) = \textbf{2}]$
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.

6. Conclusion

Empirically, there are two major points here. First, we've demonstrated that arithmetically identical measure phrases like *50%* and *1 in 2* do not have identical syntactic distributions. Second, we've examined a previously unrecognized class of measure phrases, which we've called composite MPs. These composite MPs are divided into three categories: proportional composite MPs, range composite MPs, and irreducible composite MPs. Proportional composite MPs behave much like other proportional degrees. And much like other proportional degrees, they are subject to fine-grained lexical restrictions on their particular flavor of proportional scale, with language making finer-grained distinctions than arithmetic does. Irreducible composite MPs, by contrast, prioritize not the relationship between the members of the composite MP but instead the information each constituent MP provides. We set aside range composite MPs.

Theoretically, our contributions follow from the distinctions we've made. We propose that scales come in fine-grained flavors to which linguistic expressions are sensitive. For proportional composite MPs, we suggest that they denote proportional degrees. For irreducible composite MPs, we capture their multidimensional meaning by treating tuple degrees as a new variety of degree. These degrees interact with predicates that specifically require them, such as *beat*, and maintain the information that each constituent of the composite MP provides.

References

Abney, S. (1987). *The English Noun Phrase in its Sentential Aspect*. Ph. D. thesis, MIT.
Ahn, D. and U. Sauerland (2015). The grammar of relative measurement. In S. D'Antonio, M. Moroney, and C. R. Little (Eds.), *Proceedings of Proceedings of Semantics and Linguistic Theory (SALT) 25*, pp. 125–142.

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

Bale, A. (2006). *The Universal Scale and the Semantics of Comparison*. Ph. D. thesis, McGill University.

Bresnan, J. (1973). Syntax of the comparative clause construction in English. *Linguistic Inquiry 4*(3), 275–343.

Cresswell, M. J. (1976). The semantics of degree. In B. H. Partee (Ed.), *Montague Grammar*, pp. 261–292. New York: Academic Press.

Dotlačil, J. and R. Nouwen (2016). The comparative and degree pluralities. *Natural Language Semantics 24*, 45–78.

Gobeski, A. (2019). *Factor Phrases: The Semantics of Multiplicative Modification of Events, Degrees, and Nominals, and the Grammar of Arithmetic*. Ph. D. thesis, Michigan State University.

Gobeski, A. and M. Morzycki (2018). Percentages, relational degrees, and degree constructions. In D. Burgdorf, J. Collard, S. Maspong, and B. Stefánsdóttir (Eds.), *Proceedings of SALT 27*. eLanguage.

Gobeski, A. and M. Morzycki (2022). *Ranges*. Ms., UBC.

Heim, I. (2000). Degree operators and scope. In B. Jackson and T. Matthews (Eds.), *Proceedings of Semantics and Linguistic Theory (SALT) 10*, Ithaca, NY, pp. 40–64. CLC Publications.

Kennedy, C. (1997). *Projecting the Adjective: The Syntax and Semantics of Gradability and Comparison*. Ph. D. thesis, UC Santa Cruz. Published in 1999 by Garland, New York.

Krifka, M. (1989). Nominal reference, temporal constitution, and quantification in event semantics. In R. Bartsch, J. van Benthem, and P. van Emde Boas (Eds.), *Semantics and Contextual Expression*, pp. 75–115. Dordrecht: Foris.

Morzycki, M. (2016). *Modification*. Key Topics in Semantics and Pragmatics. Cambridge: Cambridge University Press.

Schwarzschild, R. (2005). Measure phrases as modifiers of adjectives. In *L'adjectif*, Volume 34 of *Recherches Linguistiques de Vincennes*, pp. 207–228. Paris: Presses universitaires de Vincennes.

Solt, S. (2018). Proportional comparatives and relative scales. In R. Truswell, C. Cummins, C. Heycock, B. Rabern, and H. Rohd (Eds.), *Proceedings of Sinn und Bedeutung 21*.

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