

# Part-introducing *percent* in English

July 17, 2021

*Barry [the cat] is 99.9 percent dead. Which makes him .1 percent alive.*

– The Practice (late 1990s-early 2000s TV Series)

## Abstract

This paper is dedicated to a detailed empirical investigation of the distribution of *percent* in English, with a focus on uses I call ‘part-introducing’. Two uses of *percent*, called ‘conservative’ and ‘reversed’, have been extensively discussed in the literature. But the ‘reversed’ use is not the only one in which *percent* introduces a predicate that characterizes a part of a larger whole; there are also predicative ones, among others. I develop a catalog of ‘part-introducing’ uses, in which *percent* combines directly with a predicate, using corpus examples. I then consider how existing theories fare in capturing its distribution, and offer two suggestions for improving the empirical coverage with a uniform treatment of the part-introducing uses. First, I propose a type-shift that converts a non-gradable predicate to a gradable one that tracks mereological parthood. This makes any non-gradable predicate eligible for use with a previous analysis of *percent* in constructions like *75% full*. Second, motivated by cumulative-like readings, I sketch an analysis in a dynamic semantics with plurals in which *percent* applies to a cross-assignment sum, evaluated after the rest of the constraints in the clause have been applied to the discourse referent in question.

## 1 Introduction

As originally observed by Ahn (2012) for Korean, there are at least two distinguishable uses of so-called PROPORTIONAL MEASURE NOUNS like *percent*.<sup>1</sup> Sauerland (2014) calls these readings *conservative* and *reversed*, respectively (adapted from Sauerland’s (2014) attested examples):

- (1) The most recent class of NASA consists of 50% of the women. (conservative)
- (2) The most recent class of NASA consists of 50% WOMEN. (reversed)

---

<sup>1</sup>The original observation was made by Ahn (2012) for Korean, and subsequent work has focussed primarily on German and English (Sauerland, 2014; Ahn & Sauerland, 2015, 2017; Sauerland & Pasternak, under review; Li, 2018; Solt, 2018; Spathas, 2019; Pasternak, 2019).

As for why the label ‘conservative’ is used, and for a bit of relevant intellectual background, see Appendix A. Ahn & Sauerland (2015, 2017) show that the alternation is found in German, Korean, French, and English. In the context of the larger cross-linguistic exploration of this special issue, this paper is dedicated to a detailed empirical investigation of the distribution of *percent* in English using plenty of corpus examples, with an eye toward pushing our theoretical understanding of the phenomenon forward.

The two uses of *percent* in (1) and (2) both characterize a situation in which there is a part and a whole, where the part makes up some percent of the whole. In (1) (with *50% of the women*), the whole is denoted by *the women*, and there is a part of that whole making up 50% of it that is denoted by the expression *50% of the women*. On the other hand, in (2), with *50% women*, what comes after *percent* is a description of the part – there is a larger whole and the part making up 50% of it is characterized as being made up of women. Thus, what immediately follows *percent* in the ‘reversed’ case is a predicate that characterizes a part of the whole. In the ‘conservative’ case, what immediately follows *percent* characterizes the whole.

Another place where *percent* can introduce a predicate that characterizes a part of a larger whole is in predicative sentences with adjectives or with nouns:

- (3) A school system that in 1963 had **been** 62 percent white **became** 60 percent black by 1975. (from COCA)
- (4) The solution **is** 30% acid. (adapted from BNCWeb)

In (6-c), a subpart of the student body making up 62% of it in 1963 is characterized as white, and a subpart making up 60% of it in 1975 is characterized as black. In (4), a subpart of the solution making up 30% of it is characterized as acid. Insofar as what comes after *percent* characterizes the part, these uses are like ‘reversed’ uses. Indeed, (4) could be rephrased analogously to (2) above:

- (5) The solution **consists of** 30% acid.

If (2) counts as an example of a reversed use, then so must (5), as it uses the same verb. But because predicative structures are quite different from structures involving transitive verbs, it’s a bit odd to refer to examples like (4) as ‘reversed’ ones.<sup>2</sup> In a predicative structure like (4), we would expect *30% acid* to have type  $\langle e, t \rangle$ , rather than serving as a quantificational noun phrase, as under the Sauerland & Pasternak (under review) analysis of *30% women*. The copula does not normally combine with quantifiers (*\*The class is few students*). I will use the term PART-INTRODUCING to describe both the reversed uses as in (2) and predicative examples like (4) (the idea being, again, that *percent* precedes, or introduces, a description of the part).

After a tour through some corpus examples of part-introducing uses, I show that existing theories do not immediately capture all of them, and offer a uniform treatment, treating the predicative uses as basic. First, I propose a type-shift that converts a non-gradable predicate to a gradable one that tracks mereological parthood. This makes any non-gradable

---

<sup>2</sup>The verb *consists (of)* is arguably not an ordinary transitive verb, expressing a relation between two individuals; it’s rather more like a copula, insofar as it links an individual with a characterization of that individual. I will return to this in §2.1.4.

predicate eligible for use with Pasternak’s (2019) analysis of *percent* in constructions like *75% full*. Second, I sketch a dynamic theory based on Keshet’s (2019) PLural Update Semantics, in which plural discourse referents that summarize the accumulated constraints on a given discourse referent can serve as the ‘whole’ for a *percent* phrase that both introduces and serves as a predicate. This approach has a number of promising features, including the ability to shed light on certain cumulative-like readings that otherwise elude a compositional treatment.

## 2 Part-introducing uses

### 2.1 Predicative constructions

On the view that I will put forward, the most basic configuration in which *percent* has part-introducing uses is in predicative constructions. By ‘predicative constructions’, I mean ones in which a verb selects an expression denoting a predicate of individuals (type  $\langle e, t \rangle$  or, in the case of gradable predicates, type  $\langle d, \langle e, t \rangle \rangle$ ). There are a number of predicate-selecting verbs; these include not only the copula but also verbs like *become*, *remain*, and various others. There are two different types of predicative constructions involving *percent*-modified predicates, which I’ve labeled ‘part-introducing’ and ‘scalar’. These will be exemplified in turn.

#### 2.1.1 Part-introducing predicative uses

Predicative cases that I label ‘part-introducing’ involve a non-gradable adjective serving as the predicate, and carry the implication that  $n\%$  of the subject has that property. These are attested with a range of verbs that accept adjectival complements:<sup>3</sup>

- (6) a. ... playing the Apollo in New York. The audience **was** I’d say 90 percent black.
- b. Towns such as Albertville ... **became** 34 percent Hispanic.
- c. Ole Miss **remains** 90 percent white.
- d. Galileo’s sales have **gone** 30 percent non-military to 94 percent today.
- e. they had thrown out 27,000 votes, 16,000 of them African-Americans that **vote** 98 percent Democratic.

The examples in (6) have entailments that can be expressed using partitives making reference to subparts of the subject:

- (7) a. (6-a)  $\Rightarrow$  90% of the people in the audience were black.
- b. (6-b)  $\Rightarrow$  It became the case that 34% of the people in Albertville were Hispanic.
- c. (6-c)  $\Rightarrow$  It remained the case that 90% of the people at Ole Miss were white.

---

<sup>3</sup>Another interesting case that might fit into this category involves the verb *divide*, which doesn’t usually take adjectival complements but seems to in this example:

- (i) They counted 514 plain M&M’s; in a one-pound bag that **divided** 29 percent brown, 19 percent blue, 17 percent yellow, 15 percent orange, 12 percent red, and 8 percent green.

- d. (6-d)  $\Rightarrow$  94% of the units of Galileo's sales are now non-military.
- e. (6-e)  $\Rightarrow$  98% of (the people among) the African Americans vote Democratic.

In each case, I've made artistic decisions about what sorts of things the parts are made up of. I chose people for 'the audience', 'Albertville', 'Ole Miss', and '(the) African Americans'. I left it open what the units of sales are; they might be transactions, or they might be dollars. Regardless of what the parts are made up of, in every example in (6), the non-gradable predicate following *percent* applies distributively to each individual part in a collection of subparts of a whole making up the indicated percentage of the whole. Furthermore, the whole is rather evenly divided up into these parts. The whole is denoted by the subject of the verb in these predicative constructions.

In other cases where the verb selects a predicate, the predicate is a noun phrase instead of an adjective phrase, and the noun phrase can be 'part-introducing' too. Examples in this category include (4) above as well as the following (again from the COCA corpus):

- (8) After I did Playgirl, my audience **became** 95 percent white women.
- (9) It must be **labeled** 100 percent blue agave

As with the cases above involving adjectives, these examples have entailments involving partitives making reference to subparts of the subject:

- (10) (It became the case that) 95% of the members of my audience were white women.
- (11) (It must be labelled to the effect that) 100% of the parts of it are blue agave

Granted, some of these constructed examples are a bit clunky examples of semanticist-ese, but the intuition seems clear that they are entailed by the corresponding attested sentences. Thus, schematically, predicative part-introducing examples are cases where:

SUBJ VERB NUM% PRED

entails

NUM% of the [parts of] SUBJ VERB PRED

where VERB might have to be fixed up for subject-agreement purposes, PRED can be an adjectival or nominal phrase (possibly other categories as well), and what sorts of things the parts are can vary depending on the nature of SUBJ.

### 2.1.2 Scalar predicative uses

Here are some examples I call 'scalar', taken from the COCA corpus:

- (12) a. the weapon, dating from the 1570s, was **packed** 70 percent full with gunpowder, sealed with mud and stored up on the Wall
- b. Ten minutes of searching your memory to try to pull up what it was you had seen, and after 10 minutes, you said [you **were**] 40 percent sure. Right?
- c. The creators extended the work, while **remaining** 100 percent faithful to the original, and provided an ending that adds a little something extra...

These cases are akin to Pasternak’s (2019) *75% full*: The gradable adjective introduces a scale with maximum and minimum endpoints, and the value that the subject has on that scale is at the  $n\%$  mark, with no implications about the status of any of its mereological parts. Unlike the part-introducing cases below, the examples in (12) do *not* have entailments regarding the subparts of the subject of predication:

- (13) a. (12-a)  $\nrightarrow$  70 percent of the parts of the weapon were full with gunpowder  
 b. (12-b)  $\nrightarrow$  40 percent of the parts comprising you were sure  
 c. (12-c)  $\nrightarrow$  100 percent of the people among the creators were faithful to the original

To say that the weapon is *70% full* does not mean that there is a way of dividing the weapon into equal parts such that 70% of those parts have the property of being full. Being 40% sure is not being such that 40% of the parts constituting you are sure; it is having a confidence level corresponding to 40%. Similarly, remaining 100% faithful to the original does not mean being divisible into parts such that each part of you is faithful to the original. In this case, the parts of the work created might all be faithful to the original, but there’s no implication about the parts of the creators. If the creators remained 99% faithful to the original, then this does not mean that there was a part of the creators that was not faithful to the original, but rather that there was a part of the work that was not. So the faithfulness scale is not tied to parthood of the creators, the subject of the predicate.

There are several cases for which ‘scalar’ and ‘part-introducing’ analyses are both possible, such as the following:

- (14) And I say *we* because I **feel** 100 percent American. I was born and raised here.

This might be said to entail that *I feel that 100% of the parts of me are American*, but it might not be intended that way, implying merely something like *I feel I’m at the 100% mark on the ‘American’ scale*. Indeed, one might wonder whether every part-introducing case could be seen as a scalar case, given that it is always possible to create a scale that tracks the part-whole relation, having a maximum endpoint corresponding to the whole. I’ve chosen to label the cases where a mereological part interpretation is plausible as ‘part-introducing’, and cases where it is implausible as ‘scalar’, because the two classes are associated with different entailments. But under the analysis to be proposed, these uses are linked. Below, I will propose to account for the difference by using the theoretical assumption that some adjectives are fundamentally gradable (like *full*), while others (like *Democratic*) are fundamentally non-gradable and acquire a gradable interpretation through a type-shifting mechanism that tracks part-hood. In fact, I will propose that the same mechanism is at work in the ‘reversed’ uses.

### 2.1.3 Complex attributive modifiers

**Part-introducing.** We turn now to cases involving regular transitive verbs, expressing relations between individuals, instead of selecting a predicate. To set the stage, observe that there is a subtle difference between these two cases:

- (15) Victoria’s Secret... is under pressure to **use 50 percent recycled paper** in its millions of mailed catalogs and shun paper derived from wilderness regions

- (16) tobacco farmers won a provision in the recent budget bill requiring manufacturers to **use 75 percent domestic tobacco**

Both contain a string of the form VERB NUM% ADJ NOUN and can be paraphrased using ‘NUM% ADJ’ as a predicate:

- (17) The paper VS uses must be 50% recycled.  
(18) The tobacco manufacturers use must be 75% domestic.

But the similarities end here. It is more natural to paraphrase (15) by replacing *50 percent recycled paper* with *paper that is 50 percent recycled* than it is to paraphrase (16) by replacing *75 percent domestic tobacco* with *tobacco that is 75% domestic*. Another way of looking at it is that *50% recycled* is recognizable as a type of paper, but *75% domestic* is not recognizable as a type of tobacco. (On the other hand, domestic tobacco is a type of tobacco.) Related: Paper is sold with the label ‘50% recycled’, and tobacco is not sold with the label ‘75% domestic’; it’s sold either as domestic or as imported. Furthermore, the requirement described in (16) is that 75% of the tobacco *used* by manufacturers is domestic. In (15), in contrast, it sounds as if the requirement is that *all* of the paper used (in the catalogs) is 50% recycled. For these reasons, I classify (15) as a case where ‘NUM% ADJ’ serves as an attributive modifier of the the noun:

- (19) use [<sub>NP</sub> [<sub>AP</sub> 50 percent recycled ] paper ]

On the other hand, (16) seems to be a true ‘reversed’ case, the kind that could be analyzed as having focus on *domestic*.

Among these ‘attributive’ cases, a distinction can be drawn between ‘part-introducing’ and ‘scalar’ uses, just as in the realm of predicative uses of adjectives. Example (15) is a ‘part-introducing’ use; the adjective *recycled* is not gradable, and there is a way of dividing this paper up into small, evenly-sized parts such that 50% of those parts have the property of being ‘recycled’. Here are some further examples I would place in this category:

- (20) a. Hey, firefighters can’t afford 100 percent jumbo lump crab meat, OK?  
b. Runners need at least three to six one-ounce servings of whole grains per day, and eating 100 percent whole-grain bread (as opposed to just whole-grain bread, which may contain some refined grains and flours) is an easy way to meet this requirement since one slice equals one serving.

Both cases just given involve mass nouns, which do not require determiners. Below are cases where the modified noun is a count noun. In these, the noun phrase has an indefinite determiner *a* preceding the *percent* phrase:

- (21) a. In one study, drinkers and smokers “spontaneously abstained” after eating a 67 percent raw diet.  
b. In 1980-81, 5,000 white students in Henry schools created a 75 percent white system.

Here is a case in which the modifier of the count noun is itself a noun:

(22) But I found a 23 percent hydrochloric acid cleanser.

All of these can be paraphrased using a relative clause: ‘crab meat that is 100% jumbo lump’, ‘bread that is 100% whole-grain’, ‘diet that is 67% raw’, ‘system that is 75% white’, ‘cleanser that is 23% hydrochloric acid’. This shows that in these strings of the form VERB (a) NUM% ADJ NOUN, the string NUM% ADJ functions as an intersective, attributive modifier of NOUN.

The same types of examples can be found in subject position:

- (23) a. Can a 98% white town attract and keep black, Asian and Hispanic families?  
b. After London, the 100% recyclable sculpture made the rounds on a tour of the U.K.  
c. Actual rehab of the 95% occupied building took 18 months and \$5 million.

These have paraphrases using relative clauses as well: a town that is 98% white; a sculpture that is 100% recyclable; a building that is 95% occupied.

Paraphrases with a relative clause of this kind support the idea that NUM% ADJ serves as an *attributive* modifier of NOUN, but what shows that these cases are *part-introducing* attributive modifiers? The parenthetical clarification in example (20-b) clearly reveals that a mereological scale is involved in this case. A general way of showing this is to use the same kind of transformations on the sentence that we used in the predicative construction:

- (24) (15)  $\Rightarrow$  VS is under pressure to use paper such that 50% of the parts constituting it are recycled.

Again, it sounds a bit clunky, but the entailment is there. Similar observations hold for the other examples in this section. Thus, in these cases,

SUBJ VERB (DET) NUM% ADJ NOUN

can be paraphrased as

SUBJ VERB (DET) NOUN such that NUM% of the parts constituting it are ADJ

**Scalar.** Another class of cases involving attributive modifiers of the form ‘NUM% ADJ’ are cases in which ADJ is a upper- and lower-closed scale gradable adjective like *full*. Here are some constructed examples:

- (25) a. Marie Kondo recommends having 90% full drawers.  
(drawers that are 90% full)  
b. To make this, I create a 50% transparent image in Photoshop.  
(an image that is 50% transparent)

Like the attributive part-introducing cases, these can be paraphrased with relative clauses, as noted in conjunction with the examples. Moreover, these uses carry analogous entailments to the ones observed for scalar predicative cases:

- (26) a. (25-a)  $\Rightarrow$  MK recommends having drawers that are 90% of the way up on the ‘full’ scale.

- b. (25-b)  $\Rightarrow$  I create an image that 50% of the way up on the ‘transparent’ scale.

and do not have entailments analogous to the part-introducing cases:

- (27) a. (25-a)  $\not\Rightarrow$  MK recommends having drawers such that 90% of their parts are full.  
(drawers that are 90% full)  
b. (25-b)  $\not\Rightarrow$  I create an image such that 50% of its parts are transparent.

Here are some attested examples from the COCA corpus in this category:

- (28) With a few tweaks, citrus industry officials say they’ll be able to turn the cameras toward Brazil and get a 95 percent accurate estimate of that country’s orange harvest.  
(29) for three weeks, he spent four to five hours expanding his tough-minded [NP [AP 200 percent nationalistic ] theory ]

Example (28) does not imply that the estimate in question can be divided up into parts such that 95% of them are accurate; rather, 95% expresses the degree of accuracy that the estimate enjoys. (29) does not imply that 200% of the (parts of the) theory is (are) nationalistic (if that even made sense). This example seems to involve hyperbole, as nothing can be more than 100% on any scale (although a quantity can be more than 100% of another quantity, as in 200% zoom), but it is clear how to interpret it: The degree of nationalism was way off the charts.

I did not happen to find any clear examples of scalar attributive uses in subject position in the COCA corpus, but they are not difficult to construct:

- (30) A 95 percent accurate estimate of Brazil’s orange harvest can be obtained using this technique.

Thus, it seems that like part-introducing strings of the form NUM% ADJ, scalar ones can function either predicatively or attributively, and when they function attributively, they can modify a noun in any grammatical position.

To summarize: In scalar attributive cases,

SUBJ VERB (DET) NUM% ADJ NOUN

can be paraphrased

SUBJ VERB (DET) NOUN that is NUM% of the way up on the ADJ scale

and *cannot* be paraphrased as

SUBJ VERB (DET) NOUN such that NUM% of the parts constituting it are ADJ

#### 2.1.4 Reversed readings

We turn now to attested examples of genuine ‘reversed’ readings, like (16), repeated here:

- (31) tobacco farmers won a provision in the recent budget bill requiring manufacturers to **use 75 percent domestic tobacco**

In these cases,

SUBJ VERB NUM% ADJ NOUN

can be paraphrased as

NUM% of what SUBJ VERB was ADJ NOUN

and/or

NUM% of the NOUN SUBJ VERB was ADJ.

These two paraphrases correspond to the two possible placements of focus under Sauerland & Pasternak's (under review) theory. The first corresponds to placement of focus on the 'ADJ NOUN' combination, and the second corresponds to focus on the 'ADJ' combination.

Here are some ambiguous cases:

- (32) The Media Research Center recently documented that Romney **got** 86 percent negative coverage during his trip to England, Israel and Poland last week.  
'86% of what Romney got was negative coverage'  
or '86% of the coverage Romney got was negative'
- (33) he **ran** 95 percent negative ads, he and his PAC, down in Florida  
'95% of what he ran was negative ads'  
or '95% of the ads he ran were negative'
- (34) These locations **receive** 100 percent fresh air from the building's bulk air handling system.  
'100 percent of what these locations receive is fresh air'  
or '100 percent of the air these locations receive is fresh'
- (35) Google and Yahoo both lost ground, with each company **holding** 16 percent pay-per-click shares for the first six months of 2006.  
'16 percent of what each company held was pay-per-click shares'  
or '16 percent of the shares each company held were pay-per-click'

In all of these cases, the second paraphrase is more precise, but the first paraphrase is not out of the question, as long as there is some room for contextual narrowing on the subcategorization requirements of the verb. The verb *get*, for example, is quite polysemous and vague and could potentially take on any number of different more specific senses in context. Given a sense of the verb *get* that is restricted in context to the kinds of things that the media gives—namely coverage—then in context, 'what Romney got' is the same as 'the coverage Romney got'. So these two readings collapse. Analogous observations hold for the other examples in this set.

There are also unambiguous cases, where the ADJ-focus paraphrase is clearly unavailable:

- (36) The only difference was that the first group's food **contained** 1 percent linoleic acid (consistent with the U.S. diet circa 1900) while the second group...  
'1% of what the first group's food contained was linoleic acid'  
(not: '1% of the acid the first group's food contained was linoleic')

- (37) Just as important, the location offered access to raw materials like bog iron ore, which workers collected from the bottom of bogs, swamps, ponds, and riverbeds, yielding a mixture of materials that **contained** 30-50 percent usable iron.  
 ‘30-50% of what the materials contained was usable iron’  
 (not: ‘30-50% of the iron the materials contained was usable’)
- (38) The cosmos **holds** 73 percent dark energy, 23 percent dark matter, and a measly 4 percent ordinary matter  
 ‘73% of what the cosmos holds is dark energy’  
 (not: ‘73% of the energy the cosmos holds is dark’)
- (39) ...campaign directed at Coca-Cola, which made—and failed to keep—a 1990 promise to **use** 25 percent post-consumer waste in its plastic beverage bottles  
 ‘promise that 25% of what they use is post-consumer waste’  
 (not: ‘promise that 25% of the waste they use is post-consumer’)

It may be worth noting that in many of these particular examples, ‘NUM % of what  $x$  VERB’ is equivalent to ‘NUM % of  $x$ ’: ‘1% of what the first group’s food contained’ is equivalent to ‘1% of the first group’s food’. *Hold* is similar to *contain*: ‘73% of what the cosmos holds’ is equivalent to ‘73% of the cosmos’. This equivalence does not always hold, though: ‘25% of what Coca-Cola uses’ is not the same as ‘25% of Coca-Cola’.

The following case involves quantification over the subject position, and each witness for the quantifier corresponds to its own reversed reading.

- (40) Until last year, Beijing required that wind turbines sold in China **contain** 70 percent Chinese-made parts.  
 ‘for all wind turbines sold in China  $x$ , 70% of what  $x$  contains is Chinese-made parts’  
 or ‘... 70% of the parts  $x$  contains are Chinese-made’.

These paraphrases do entail paraphrases of the form ‘NUM% of what SUBJ VERB is ADJ NOUN’ and ‘NUM% of the NOUN SUBJ VERB is ADJ’, respectively (‘... 70% of what wind turbines sold in China contain is Chinese-made parts’; ‘... 70% of the parts wind turbines sold in China contain are Chinese-made’). But they say something more specific too, imposing the same condition on all wind turbines. These observations are entirely expected in light of the fact that quantifiers are scopally mobile, but this case does help to illustrate that the entailment patterns depend on whether SUBJ is quantificational.

Now, in some cases, a reversed interpretation is possible, but so is an attributive interpretation, where ‘NUM% ADJ’ serves as a modifier of NOUN. With attributive uses, recall that the sentence can be restated equivalently as ‘SUBJ VERB NOUN that was NUM% ADJ’. For many cases, both of those paraphrases sound about right, so it’s difficult to make a firm decision about which analysis among those is correct. For example, consider the following:

- (41) Up until now, blame for this kind of corrosion has rested squarely on product imported from China, but the Brinkews say they **ordered** 100 percent American dry wall, and they’ve got the contractor receipts to prove it. But is that what they got? They believe either Chinese dry wall is being passed off as American, or contaminated dry wall is being produced in America.

VERB LEMMA	FREQUENCY	EXAMPLE PREDICATES
contain	23	<i>red clover, organic compound, lean protein</i>
use	19	<i>blue agave, post-consumer waste, actual calfskiin</i>
include	6	<i>American content, Pell-eligible students, true topsoil</i>
purchase	3	<i>renewable power, renewable energy, green power</i>
hold	2	<i>dark energy, pay-per-click shares</i>
provide	2	<i>renewable power, renewable energy</i>
show	2	<i>affirmative answers</i>
add	1	<i>green light [to plants]</i>
be fed	1	<i>organic feed</i>
buy	1	<i>clean energy</i>
get	1	<i>negative coverage</i>
go	1	<i>clean energy</i>
harness	1	<i>new oak [in wine production]</i>
incorporate	1	<i>recycled plastic</i>
mix	1	<i>ground-up, dried manure</i>
order	1	<i>American dry wall</i>
receive	1	<i>fresh air</i>
run	1	<i>negative ads</i>
see	1	<i>pure 10s [of women; gross!]</i>
sell	1	<i>Fair Trade coffee</i>
substitute	1	<i>compostible cups</i>

Table 1: Verbs occurring in reversed constructions involving adjectives in the COCA corpus

Did they order dry wall that is 100% American or was 100% of what they ordered American dry wall? Or was 100% of the dry wall they ordered American? All of the above seem to fit reasonably well. This case is thus ambiguous between ‘reversed’ and ‘attributive’ (and between two ‘reversed’ interpretations, corresponding to focus on either ‘ADJ’ or ‘ADJ NOUN’).

**Distribution of verbs.** Even taking the ambiguous cases into account, (potential cases of) reversed readings are restricted to quite a limited set of verbs. Table 1 shows a list of the verb lemmas that figured in examples categorized as (possibly) ‘reversed’, along with their frequency among the search results and example predicates. The frequency information gives a sense of which verbs are most prototypical for a reversed use. There is a striking similarity among them: In almost all cases, the verb can be replaced by ‘incorporate’ without changing the meaning terribly much.

Absence of corpus examples of a particular kind does not prove ungrammaticality, but observe the contrast among the following constructed examples:

- (42) a. The committee hired 75% WOMEN.  
b. ??The committee disliked 75% WOMEN.

To my ears, (42-b) is unrescuably terrible. Why is *dislike* so much worse than *hired* here? An adequate theory of the semantics of *percent* should be able to account for this contrast. It does seem to have something to do with an implication of coming into existence (within a given situation); compare:

- (43) a. The committee hired some women, so there were more women.  
b. #The committee disliked some women, so there were more women.

This is one of the diagnostics that Coppock & Beaver (2015) use for ‘entity-introducing verbs’. Another is the following:

- (44) a. There are seven women. If the committee hires a woman, that will make eight.  
b. There are seven women. #If the committee dislikes a woman, that will make eight.

On the other hand, presentational verbs do not seem very good with *percent* complements:<sup>4</sup>

- (45) #There appeared 75% women.

and yet they are arguably ‘entity-introducing’; although the following sentence is not perfectly natural, it does seem to have an entity-introducing implication.

- (46) There are at least seven women. ?If there appears a woman, that will make eight.

So, being ‘entity-introducing’ may be a necessary, but not sufficient condition.

**Subject uses.** Before moving on, let us briefly address whether there are reversed uses in subject position. It is not easy to find attested uses of *percent* that could be categorized as ‘reversed’, but here is a sentence that could be understood that way, offered by a reviewer:

- (47) Approximately 10% water will remain.

This could certainly be read as a part-introducing use, insofar as what comes after *percent* is a predicate that characterizes the 10% part. It’s clearly not predicative, since *10% water* is not functioning as the complement of a predicate-selecting verb. It’s also clearly not attributive, because 10% does not combine with an adjectival or nominal modifier to create a complex attributive modifier of a noun. The example can be paraphrased analogously to reversed uses:

- (48) Approximately 10% of what will remain is water.

So this looks like a ‘reversed’ use in subject position.

Below, we will review Sauerland & Pasternak’s (under review) analysis of the sentence *30% Westphalian students work here*—another example with a reversed reading in subject position. The original example is in German, and it does not feel perfectly natural in English, but according to my own intuition, it is grammatical and has the relevant reading.

Subject uses can also be obtained by passivizing object uses:

---

<sup>4</sup>This was observed by Ezra Keshet (p.c.).

(49) 25 percent post-consumer waste was supposed to be **used** in Coca-Cola’s plastic beverage bottles.

(50) ?75% women were **hired** by the committee.

My sense is that the slight degradation of (50) relative to (49) is due to a subject-verb agreement problem; *75% women* does not work well as a plural. (It works even less well as a singular; replacing *were* with *was* makes it clearly ungrammatical for me.) But the following sentence strikes me as dramatically worse than (50):

(51) \*?75% women were **disliked** by the committee.

Thus it seems that the same restrictions on the set of verbs that we saw above applies to subject uses with passive verbs.

Overall, it does seem that reversed uses in subject position are a bit degraded and a bit more difficult to find. This may be due to a requirement for focus on the following predicate (which is predicted under some analyses; see below). As is well-known, English subjects tend to be topics, rather than foci (Erteschik-Shir, 2007; Foley, 2007; Lambrecht, 1994).

### 2.1.5 Cumulative readings

The final part-introducing use I will discuss involves something like a cumulative reading (Scha, 1981; Krifka, 1986; Champollion, 2009, i.a.):

(52) In one approach, they **mix** 10 percent ground-up, dried manure with 90 percent coal (by weight) and then...

(52) doesn’t mean ‘10% of what they mixed with 90% coal was ground-up, dried manure’, so it doesn’t pattern exactly like the reversed examples we’ve considered. I therefore place it in a separate category. It also doesn’t mean ‘They mixed manure that was 10% ground-up and dried with 90% coal’, so an attributive reading seems out of the question for this example. It could be paraphrased as ‘10% of what they mixed *in* was ground-up, dried manure and 90% of what they mixed *in* was coal’. This paraphrase does not fit in with any of the categories outlined so far. The example is, however, a part-introducing case, insofar as what immediately follows *percent* is a characterization of the part.

What makes (52) like a cumulative reading is that the quantificational elements do not take scope independently of each other. A classic example of a cumulative reading is in the sentence *600 Dutch firms own 5000 American computers* (Scha, 1981), where 600 Dutch firms and 5000 American computers are involved, but for any given firm, it’s not guaranteed that it owns all 5000 computers, nor is it guaranteed that any of the 500 computers are owned by any of the firms. So neither of the quantificational expressions takes scope over the other. As Champollion (2009, 216) puts it, citing Szabolcsi (1997), “[c]umulative readings express information about the cardinalities of the minimal witness sets associated with the quantifiers involved.” In (52), information is expressed about the proportions of the relevant whole made up by the minimal witness sets associated with the quantifiers involved: They mixed *x* with *y*, and *x* makes up 10% of some whole and *y* makes up 90% of that whole. We’ll see a more detailed rendition of this analysis in section 6.

Here are some additional examples in this category from the COCA corpus, all involving coordination:

- (53) a. it is composed of 55 percent fructose and 45 percent glucose  
b. The 65-gram, 240-calorie breakfast contains 12 percent sodium and 11 percent fat.
- (54) a. E85 is a mixture of 85 percent ethanol and 15 percent gasoline.  
b. Many contractors want a combination of 25 percent compost and 75 percent topsoil for flower beds.

The following cases, using *with*, are a bit more involved:

- (55) Much less biocide is needed per gallon of paint - typically 25 percent along with 5 percent zinc omadine.
- (56) biodiesel runs fine in unmodified diesel engines at up to a 20 percent blend with 80 percent petroleum diesel, a combination known as B20.

In considering these examples, a question that arises is whether and to what extent conjunction (or comitative semantics) plays an important role in cases like (52). In that example, it seems that the two *percent* cases are *not* actually syntactically coordinated:

- (57) Q: #What did they mix?  
A: #10 percent manure with 90 percent coal.
- (58) #What they mixed was 10 percent manure with 90 percent coal.

Rather, the *with*-phrase appears to introduce its own syntactic argument in this case. So while these cumulative-like readings often involve coordination, they need not do so.

## Summary

To summarize, we have identified several categories of predicate-selecting *percent*, each licensing their own paraphrase patterns. A string of the form:

SUBJ VERB NUM% ADJ (DET) (NOUN)

can be paraphrased in number of ways.

- When VERB is predicative, and there is no determiner DET, there are two possibilities, ‘scalar’ and ‘part-introducing’. In the ‘part-introducing’ case, a paraphrase of the form

It VERB the case that NUM% of the parts of SUBJ are ADJ (NOUN)

is possible. In the ‘scalar’ case, that type of paraphrase is not possible, but one of the following form is:

SUBJ VERB NUM% of the way up on the ADJ scale

- When VERB is an ordinary transitive verb, and a NOUN is present, one possibility is that NUM% ADJ acts as an attributive modifier of NOUN. In that case, a paraphrase with a relative clause is possible:

SUBJ VERB (DET) NOUN that is NUM% ADJ

Here again, there are two possibilities, scalar and part-introducing. In the part-introducing case, a paraphrase of the following form is possible

SUBJ VERB NOUN such that NUM% of the parts of it are ADJ

In the scalar case, a paraphrase of that form is not possible, although one of the following form is:

SUBJ VERB NOUN that is NUM% of the way up on the ADJ scale

- There are also reversed readings, when there is no determiner DET and a NOUN is present. Here again there are two possibilities, but in this case it depends on the placement of focus. When focus is on ADJ, the following paraphrase is possible:

NUM% of the NOUN SUBJ VERB is ADJ

When focus is on the ‘ADJ NOUN’ combination, it can be paraphrased:

NUM% of what SUBJ VERB is ADJ NOUN

There are also cumulative-readings, which don’t quite fit into this schema. In the next section, we will apply existing analyses to these cases and see how far we can get.

### 3 Sauerland & Pasternak (under review)

#### 3.1 Conservative uses

At the time of writing, the most recent analysis that has been made publicly available (Sauerland & Pasternak, under review, available on LingBuzz) treats *percent* as in (60). This analysis is analogous to their treatment of absolute measure nouns like *kilo*, which they treat as in (59). Both *percent* and *kilo* combine with two degree predicates  $D$  and  $D'$  (type  $\langle d, t \rangle$ ) in addition to a number  $n$  (which enters into the composition between the two degree predicates).

$$(59) \quad \textit{kilo} \rightsquigarrow \lambda D \lambda n \lambda D'. D' \subseteq D \wedge \max(D') \geq n \cdot \textit{kg}$$

$$(60) \quad \textit{percent} \rightsquigarrow \lambda D \lambda n \lambda D'. D' \subseteq D \wedge \max(D') \geq \frac{n}{100} \cdot \max(D)$$

Note that in (59), I have written  $n \cdot \text{kg}$  to represent the quantity of  $n$  kilograms using the dot ( $\cdot$ ) for multiplication, presupposing a foundation for degree semantics that allows for multiplication among degrees and of numbers with degrees as described by Coppock (2021); I assume  $\text{kg}$  denotes a particular degree, the one corresponding to one kilogram.

In partitive, ‘conservative’ uses, these measure nouns will be selected by a function they call MEAS (a bit like Solt’s (2009) MEAS, but compositionally different), defined as follows.

$$(61) \quad \text{MEAS} \rightsquigarrow \lambda M \lambda x \lambda n \lambda y . y \sqsubseteq x \wedge M(\mu(x))(n)(\mu(y))$$

where  $\mu(\alpha)$  is defined as  $\lambda d . d \leq \mu(\alpha)$

Here,  $\mu$  is to be read as a free variable, set by context to an appropriate measure function.<sup>5</sup>

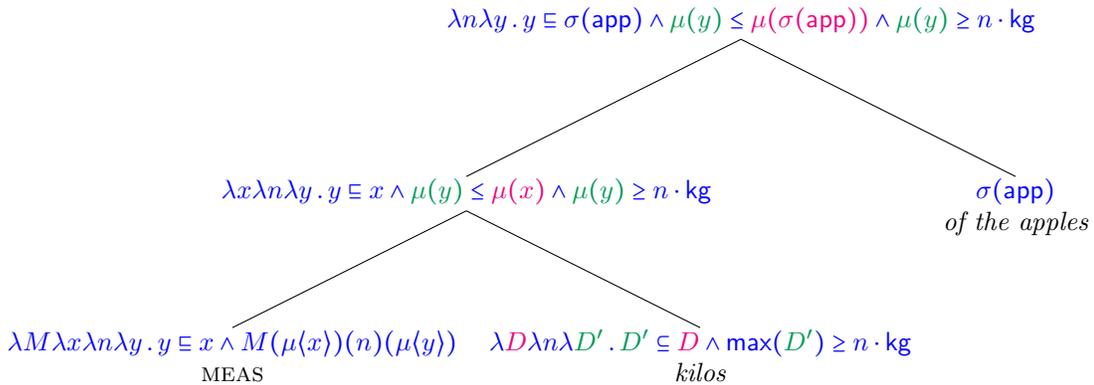
In a partitive example like *30 kilos of the apples*, MEAS combines first with *kilos*, as in the following derivation (in which some simplifications are used, including rewriting  $\max(\mu(x))$  as  $\mu(x)$  and restating a subset relation between two degree predicates as a less-than-or-equal relation among their maxima):<sup>6</sup>

$$(62) \quad \begin{array}{l} \lambda n \lambda y . y \sqsubseteq \sigma(\text{app}) \wedge \mu(y) \leq \mu(\sigma(\text{app})) \wedge \mu(y) \geq n \cdot \text{kg} \\ \left\{ \begin{array}{l} \lambda x \lambda n \lambda y . y \sqsubseteq x \wedge \mu(y) \leq \mu(x) \wedge \mu(y) \geq n \cdot \text{kg} \\ \left\{ \begin{array}{l} \lambda M \lambda x \lambda n \lambda y . y \sqsubseteq x \wedge M(\mu(x))(n)(\mu(y)) \leftarrow \text{MEAS} \\ \lambda D \lambda n \lambda D' . D' \sqsubseteq D \wedge \max(D') \geq n \cdot \text{kg} \leftarrow \text{kilos} \end{array} \right. \\ \sigma(\text{app}) \leftarrow \text{of the apples} \end{array} \right. \end{array}$$

The same happens with MEAS and *percent*, giving the following derivation.

<sup>5</sup>Sauerland & Pasternak (under review) write the context-sensitive  $\mu$  function as  $\mu^c$ , where  $c$  denotes the relevant context. I will not use a superscript  $c$  for  $\mu$  in the representation language. Rather, I will assume that it is an indexical constant of the language like  $i$ , which picks out the speaker of the given context of utterance; just as  $\llbracket i \rrbracket^{M,g,c} = \text{the speaker of } c$ , so  $\llbracket \mu \rrbracket^{M,g,c}$  depends on some feature of context  $c$ . This is a purely stylistic matter.

<sup>6</sup>Here, and throughout, I represent derivation trees using a ‘folder’ style for displaying derivation trees. The tree in (62) in the more familiar tree format would look as follows:



The ‘folder’ style makes it possible to display complex derivations in a more compact yet readable format.



$$(65) \quad \lambda d_1 . \exists x [\text{We}(x) \wedge \text{St}(x) \wedge \mu(x) \geq d_1 \wedge \text{Wo}(x)]$$

$$\begin{array}{l} \lambda_1 \\ \exists x [\text{We}(x) \wedge \text{St}(x) \wedge \mu(x) \geq d_1 \wedge \text{Wo}(x)] \\ \lambda Q . \exists x [\text{We}(x) \wedge \text{St}(x) \wedge \mu(x) \geq d_1 \wedge Q(x)] \\ \exists \rightsquigarrow \lambda P \lambda Q . \exists x [P(x) \wedge Q(x)] \\ \lambda x . \text{We}(x) \wedge \text{St}(x) \wedge \mu(x) \geq d_1 \\ d_1 \rightsquigarrow t_1 \\ \lambda d \lambda x . \text{We}(x) \wedge \text{St}(x) \wedge \mu(x) \geq d \\ \lambda P \lambda d \lambda x . P(x) \wedge \mu(x) \geq d \rightsquigarrow \text{DEG} \\ \lambda x . \text{We}(x) \wedge \text{St}(x) \\ \text{We} \rightsquigarrow \text{Westphalian}_F \\ \text{St} \rightsquigarrow \text{students} \\ \text{Wo} \rightsquigarrow \text{work here} \end{array}$$

Hence the value of  $C$  will be:<sup>8</sup>

$$(66) \quad \{\lambda d . \exists x . P(x) \wedge \text{St}(x) \wedge \mu(x) \wedge d \wedge \text{Wo}(x) \mid P \in \text{ALT}(\text{Westphalian})\}$$

Assuming, following Sauerland & Pasternak (under review), that the trivial property  $\lambda x . \top$  is among the alternatives to *Westphalian*, the 30 percent DIS  $C$  component will have semantics derived as follows:

$$(67) \quad \lambda D' . D' \subseteq \dots \wedge \max(D') \geq \frac{30}{100} \cdot \max(\lambda d . \exists x [\text{St}(x) \wedge \mu(x) \wedge d \wedge \text{Wo}(x)])$$

$$\begin{array}{l} 30 \rightsquigarrow 30 \\ \dots \\ \lambda D \lambda n \lambda D' . D' \subseteq D \wedge \max(D') \geq \frac{n}{100} \cdot \max(D) \rightsquigarrow \text{percent} \\ \lambda d . \exists x [\text{St}(x) \wedge \mu(x) \wedge d \wedge \text{Wo}(x)] \\ \lambda C . \cup C \rightsquigarrow \text{DIS} \\ \{\lambda d . \exists x . P(x) \wedge \text{St}(x) \wedge \mu(x) \wedge d \wedge \text{Wo}(x) \mid P \in \text{ALT}(\text{Westphalian})\} \rightsquigarrow C \end{array}$$

Simplifying a bit, putting these two subtrees together will produce the following:

$$(68) \quad \begin{array}{l} \max(\lambda d_1 . \exists x [\text{We}(x) \wedge \text{St}(x) \wedge \mu(x) \geq d_1 \wedge \text{Wo}(x)]) \\ \geq \frac{30}{100} \cdot [\max(\lambda d . \exists x [\text{St}(x) \wedge \mu(x) \geq d \wedge \text{Wo}(x)])] \end{array}$$

Here is another way of saying the same thing:

$$(69) \quad |\text{We} \cap \text{St} \cap \text{Wo}| \geq \frac{30}{100} \cdot |\text{St} \cap \text{Wo}|$$

<sup>8</sup>I'm using an indirect interpretation style here, which involves translating from English to a version of typed lambda calculus. It's not entirely settled how to deal with focus in an indirect interpretation style, as far as I know. Here I'm assuming that the representation language imports a number of set-theoretic devices and that it contains a context-sensitive function **ALT** that takes a natural language string or parse tree as an argument.

‘(At least) 30% of the students who work here are Westphalian’

So this analysis derives truth conditions for a reversed reading in a focus-sensitive manner, using the same lexical entry that is used to derive conservative truth conditions (given the assumption that the trivial property is always among the focus alternatives, and a silent disjunction operator). Notice that the truth conditions for the reversed case are not conservative, in the sense that *30% Westphalian students work here* does not entail *30% Westphalian students are Westphalian students who work here*.

### 3.3 Predicative uses

Now let us consider predicative uses like the following:

(70) The solution is 30% acid. (adapted from BNCWeb)

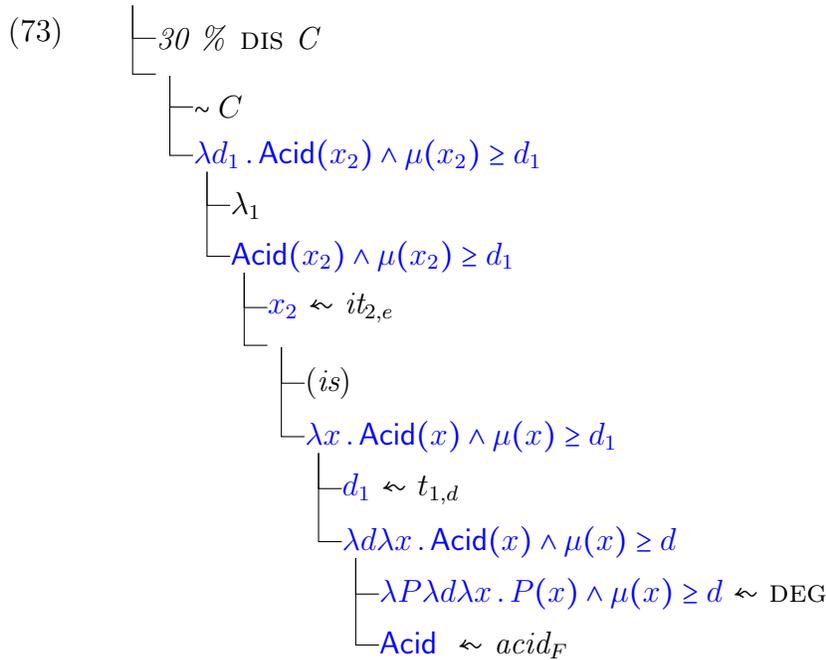
These are like reversed cases in that what follows *percent* is a predicate that holds of a subpart of a whole. Indeed, (70) could be rephrased analogously to (2) above:

(71) The solution consists of 30% acid.

But I assume that the copula takes a predicative complement, while *consists of* does not; for instance, *consists of* can have a quantificational complement:

(72) The solution {consists of, \*is} many different compounds.

If we assume that *30% acid* is type  $\langle e, t \rangle$ , and that *30%* is raised and interpreted as under Sauerland & Pasternak’s (under review) analysis of reversed structures, then we derive the following structure:



Expanding the *30% DIS C* part, we then have:

$$\begin{aligned}
(74) \quad & \lambda d_1 . \text{Acid}(x_2) \wedge \mu(x_2) \geq d_1 \subseteq \lambda d . \mu(x_2) \geq d \\
& \wedge \max(\lambda d_1 . \text{Acid}(x_2) \wedge \mu(x_2) \geq d_1) \geq 30\%[\max(\lambda d . \mu(x_2) \geq d)] \\
& \left| \begin{array}{l}
\lambda D' . D' \subseteq \lambda d . \mu(x_2) \geq d \wedge \max(D') \geq 30\%[\max(\lambda d . \mu(x_2) \geq d)] \\
\quad \left| \begin{array}{l}
30 \\
\lambda n \lambda D' \dots \\
\quad \left| \begin{array}{l}
\lambda D \lambda n \lambda D' . D' \subseteq D \wedge \max(D') \geq n\%[\max(D)] \quad \leftarrow \text{percent} \\
\lambda d . \mu(x_2) \geq d \\
\quad \left| \begin{array}{l}
\lambda C . \cup C \quad \leftarrow \text{DIS} \\
C = \{\lambda d . P(x_2) \wedge \mu(x_2) \geq d \mid P \in \text{ALT}(\text{Acid})\} \quad \leftarrow C
\end{array}
\end{array}
\end{array}
\end{array}
\end{array}
\end{array}
\end{aligned}$$

In words, the truth conditions derived can be expressed as follows: The degree to which  $x_2$  is acid and big (along dimension  $\mu$ ) is greater than or equal to 30% of the degree to which  $x_2$  is big (along dimension  $\mu$ ). Since ‘acid’ is a unary predicate under this treatment, a given object will be in its extension or not; there is no middle ground. If  $x_2$  is not acid, then the first degree predicate will yield an undefined value, because the maximum of the empty set is undefined. So, assuming that something that is 30% acid is not acid, the sentence is predicted to have a truth value of ‘undefined’ when it is in fact true.

## 4 Pasternak

Is there a precedent in the literature that we could look to in order to come to grips with these simple predicative cases like *The solution is 30% acid*? Pasternak (2019) provides a theory of some predicative cases involving adjectives like the following:

$$(75) \quad \text{The glass is 75\% full.} \quad (\text{Kennedy \& McNally, 2005; Pasternak, 2019})$$

Pasternak defines *percent* as follows:

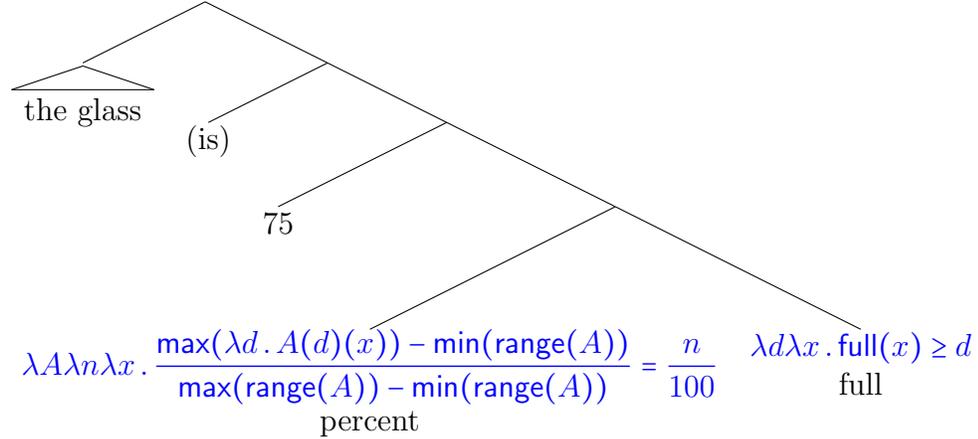
$$(76) \quad \text{percent} \rightsquigarrow \lambda A \lambda n \lambda x . \frac{\max(\lambda d . A(d)(x)) - \min(\text{range}(A))}{\max(\text{range}(A)) - \min(\text{range}(A))} = \frac{n}{100}$$

In this lexical entry,  $A$  is a gradable predicate associating individuals with degrees and  $\text{range}(A)$  refers to the set of degrees that  $A$  associates any individual with. Thus  $\text{range}(A)$  can be written equivalently as  $\lambda d . \exists x . \text{defined}(A(d)(x))$ , where  $\text{defined}(\phi)$  is true if and only if  $\phi$  is true or false.<sup>9</sup>

This lexical entry yields a straightforward analysis of predicative cases like (75):

<sup>9</sup>Note that the fraction appearing to the left of the equals sign in (76) involves division among degrees, implicitly presupposing a foundation for degree semantics that allows one degree to be divided by another, such as the system laid out by Coppock (2021).

(77)



The fraction appearing to the left of the equals sign contributed by the lexical entry for *percent* will have in the numerator the degree to which the glass is full, assuming the **min** term is the zero element for the ‘fullness’ dimension. (Here I am embellishing on Pasternak’s explanation using the system described by Coppock (2021), following Raposo (2018, 2019), where there are many zeroes, one for each dimension.) The denominator will be the degree corresponding to the size of the full range of possible fullnesses, which will be equal to the maximum possible degree of fullness (assuming again that the minimum is the relevant zero element). Whether or not fullness-degrees are themselves proportions, and in whatever units fullness is measured, this fraction can be represented as a number, since the quotient of two quantities associated with the same dimension (e.g. fullness) is a so-called ‘dimensionless quantity’, representable by a number without any accompanying unit (JCGM, 2012; Coppock, 2021). (Here again I am embellishing on Pasternak’s explanation.) The truth conditions expressed by that sentence is that the ratio of these two quantities is equal to the number  $\frac{75}{100}$ .

As Pasternak shows, this lexical entry can elegantly be extended to handle conservative cases, building on a lexical entry for silent MUCH from Wellwood (2015):

$$(78) \quad \text{MUCH} \rightsquigarrow \lambda y \lambda d \lambda x . \partial(\mu(y) \geq d) \wedge x \sqsubseteq y \wedge \mu(x) \geq d$$

where  $\partial(\phi)$  is true if  $\phi$  is true, and undefined otherwise.

I am written the presupposition that  $\mu(y) \geq d$  using the ‘partial operator’  $\partial$  (Beaver & Kraemer, 2001). This presupposition is contributed by Pasternak, and it plays a crucial role in the truth conditions, because it controls the range of the derived gradable predicate.

Let us see how it works with a ‘conservative’ example like (42-a) (*The committee hired 75% women*). This silent MUCH combines with *of the women* to produce a gradable predicate that *percent* then combines with:

$$\begin{aligned}
(79) \quad & \lambda x . \frac{\max(\lambda d . x \sqsubseteq \sigma(\text{women}) \wedge \mu(x) \geq d)}{\mu(\sigma(\text{women}))} = \frac{75}{100} \\
& \left\{ \begin{array}{l} 75 \sim 75 \\ \lambda n \lambda x . \frac{\max(x \sqsubseteq \sigma(\text{women}) \wedge \mu(x) \geq d) - 0_\mu}{\mu(\sigma(\text{women})) - 0_\mu} = \frac{n}{100} \\ \lambda A \lambda n \lambda x . \frac{\max(\lambda d . A(d)(x)) - \min(\text{range}(A))}{\max(\text{range}(A)) - \min(\text{range}(A))} = \frac{n}{100} \sim \text{percent} \\ \lambda d \lambda x . \partial(\mu(\sigma(\text{women}) \geq d)) \wedge x \sqsubseteq \sigma(\text{women}) \wedge \mu(x) \geq d \\ \quad \left\{ \begin{array}{l} \lambda y \lambda d \lambda x . \partial(\mu(y) \geq d) \wedge x \sqsubseteq y \wedge \mu(x) \geq d \sim \text{MUCH} \\ \sigma(\text{women}) \sim \text{of the women} \end{array} \right. \end{array} \right.
\end{aligned}$$

Notice that after *percent* combines with MUCH,  $\max(\text{range}(A))$  becomes  $\mu(\sigma(\text{women}))$  and  $\min(\text{range}(A))$  becomes  $0_\mu$ , which Pasternak uses to represent “the 0-degree of  $\mu$ ” (p. 78). (In Coppock’s (2021) terms, this would be the additive identity element for the dimension along which  $\mu$  measures things.)  $\max(\text{range}(A))$  simplifies to  $\mu(\sigma(\text{women}))$  because of the presupposition, requiring  $d$  to be smaller than or equal to  $\mu(\sigma(\text{women}))$ . That  $0_\mu$  is the bottom element of the range of  $A$  means that  $0_\mu$  is the smallest  $d$  for which  $A(x)(d)$  is ever defined. In this case  $A$  is

$$\lambda d \lambda x . \partial(\mu(\sigma(\text{women}) \geq d)) \wedge x \sqsubseteq \sigma(\text{women}) \wedge \mu(x) \geq d$$

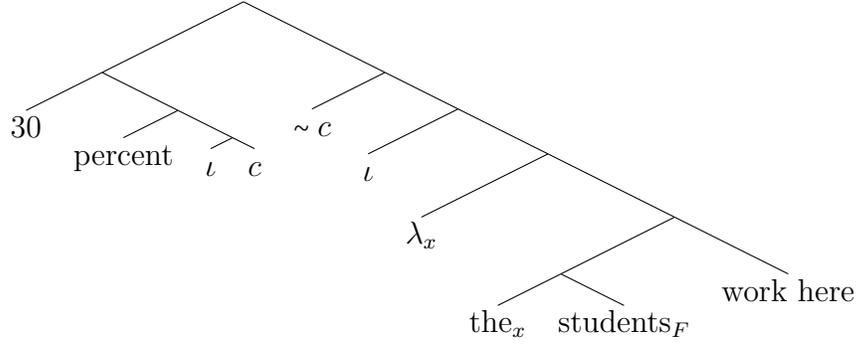
Pasternak is not 100% explicit in the text about why zero is the lowest degree in the range, but this result could be assumed to follow in part from the fact that  $d$  is comparable to  $\mu(x)$  using  $\geq$ , which would mean that it is a degree of the same dimension as  $\mu(x)$ . (I am assuming that every degree/quantity has a corresponding dimension; cf. Raposo’s (2019) *dim* mapping.) The comparability of those degrees, signalling a common dimension, along with an assumption that there are no degrees below zero along that dimension, could derive the result  $0_\mu$  is the lowest  $d$  for which  $A(x)(d)$  is defined for this  $A$  and some  $x$ .

If we add a silent existential quantifier to *75% of the women*, then for (42-a), we derive the proposition that there is a plural individual making up 75% of the women that the company hired. Strictly speaking, this does not rule out the possibility that the company hired more than 75% of the women, but that upper-bounding inference may well be pragmatic.

Pasternak (2019) claims that “this analysis can be extended equally well to Ahn & Sauerland’s (2017) treatment of [reverse cases]”. Under Ahn & Sauerland’s (2017) treatment, *percent* takes an individual as its first argument, representing the ‘whole’. In reversed cases, this argument is saturated by a silent definite description whose descriptive content is (a flattened version of) the focus semantic value of the sentence, abstracting over the position of the *percent* nominal using ‘modified trace conversion’, which inserts an indexed definite determiner  $the_x$  in the would-be trace position. Thus (80-a) has the LF in (80-b).

- (80) a. 30% Studenten<sub>F</sub> arbeiten hier  
30% students work here  
‘30% students work here.’

b.



Ahn & Sauerland use the following lexical entry for *percent*:

$$(81) \quad \textit{percent} \rightsquigarrow \lambda x \lambda n \lambda P . \frac{\mu(x \sqcap \oplus P)}{\mu(x)} = \frac{n}{100}$$

One way to merge their treatment of reverse cases with Pasternak’s assumptions might be to combine Pasternak’s lexical entry for *percent* with Ahn & Sauerland’s (2017) syntactic assumptions, with the exception that a silent MUCH is inserted above  $\iota c$  so as to produce an argument of type  $\langle d, et \rangle$  that *percent* could combine with. I believe this would enable a compositional treatment of relative readings that produces the right truth conditions.

This treatment does not immediately account for cases like *The solution is 30% acid*, though, in part because *acid* is not a gradable adjective, so it is the wrong type to combine. On the other hand, if *acid* could be coerced into a gradable predicate, then perhaps a treatment of such cases could be obtained. The question is what sort of coercion operation would yield the right truth conditions.

## 5 From non-gradable to gradable

### 5.1 Some non-solutions

Not every way of converting a non-gradable predicate like *acid* into a gradable one delivers the right truth conditions. Wellwood’s (2015) MUCH is a device we have already used for converting things into gradable predicates, but it expects an argument of type  $e$ . Suppose that in order to satisfy that type requirement, we coerced *acid* into an individual by making it denote the sum of all (contextually-relevant) acid in order to satisfy the type requirements of MUCH. Then we would derive a meaning for *The solution is 30% acid* that could be rendered back into English as *The solution is 30% of the acid*. This is not a faithful interpretation.

Alternatively, we might try to convert *acid* into a gradable property using M-OP, from Rett (2018, 105), which shifts an  $\langle e, t \rangle$  meaning to a  $\langle d, \langle e, t \rangle \rangle$  meaning:

$$(82) \quad \text{M-OP} \rightsquigarrow \lambda P \lambda d \lambda x . P(x) \wedge \mu(x) \geq d$$

$$(83) \quad \text{M-OP } \textit{acid} \rightsquigarrow \lambda d \lambda x . \text{Acid}(x) \wedge \mu(x) \geq d$$

As a minor variant, we could use Solt’s (2009) MEAS, defined as  $\lambda x \lambda d . \mu(x) \geq d$ , and combine it with *acid* as she proposes using ‘Variable Identification’, which would give the same thing. Whether we derive the meaning in (83) using M-OP or MEAS, the truth conditions for *The*

*solution is 30% acid* would then be ‘The solution is acid and its measure is 30%.’ This is not correct either.

## 5.2 A solution

My proposal for how to link *75% full* and *30% acid* builds on the following type-shifting operation. It converts a non-gradable predicate into a gradable one tracking parthood of whatever it applies to:

$$(84) \quad \text{PART} \rightsquigarrow \lambda P \lambda d \lambda x . \partial(\mu(x) \geq d) \wedge \exists y[y \sqsubseteq x \wedge \mu(y) = d \wedge P(y)]$$

It’s a lot like MUCH, except that (i) the ‘part’ argument is existentially bound, (ii) it expects a predicate as its first argument, and (iii) the ‘whole’ argument comes after the degree argument rather than before. The partial operator  $\partial$  is used introduce the same presupposition that Pasternak augmented Wellwood’s (2015) MUCH with. We will see below how this presupposition interacts with Pasternak’s semantics of *percent* to give appropriate truth conditions.

Applying the PART shift to *acid* produces the following reasonable-seeming meaning for *30 percent acid*:

$$(85) \quad \lambda x . \frac{\mu(\sigma(\lambda y . y \sqsubseteq x \wedge \text{Acid}(y)))}{\mu(x)} = \frac{30}{100}$$

$$\left\{ \begin{array}{l} \text{— } 30 \rightsquigarrow 30 \\ \text{— } \lambda n \lambda x . \frac{\max(\lambda d . \partial(\mu(x) \geq d) \wedge \exists y[y \sqsubseteq x \wedge \mu(y) = d \wedge \text{Acid}(y)]) - 0_\mu}{\mu(x) - 0_\mu} = \frac{n}{100} \\ \left\{ \begin{array}{l} \text{— } \lambda A \lambda n \lambda x . \frac{\max(\lambda d . A(d)(x)) - \min(\text{range}(A))}{\max(\text{range}(A)) - \min(\text{range}(A))} = \frac{n}{100} \rightsquigarrow \text{percent} \\ \text{— } \lambda d \lambda x . \partial(\mu(x) \geq d) \wedge \exists y[y \sqsubseteq x \wedge \mu(y) = d \wedge \text{Acid}(y)] \\ \left\{ \begin{array}{l} \text{— } \lambda P \lambda d \lambda x . \partial(\mu(x) \geq d) \wedge \exists y[y \sqsubseteq x \wedge \mu(y) = d \wedge P(y)] \rightsquigarrow \text{PART} \\ \text{— } \lambda y . \text{Acid}(y) \rightsquigarrow \text{acid} \end{array} \right. \end{array} \right. \end{array} \right.$$

Two key moments in this derivation are the simplification of  $\max(\text{range}(A))$  to  $\mu(x)$  and the simplification of  $\min(\text{range}(A))$  to  $0_\mu$ . In this case  $A$  is:

$$\lambda d \lambda x . \partial(\mu(x) \geq d) \wedge \exists y[y \sqsubseteq x \wedge \mu(y) = d \wedge \text{Acid}(y)]$$

As we saw above in connection with Pasternak’s analysis of conservative uses, the maximum of the range will be  $\mu(x)$  due to the presupposition, and the minimum will be  $0_\mu$  for analogous reasons. Another important moment is the simplification of

$$\max(\lambda d . \partial(\mu(x) \geq d) \wedge \exists y[y \sqsubseteq x \wedge \mu(y) = d \wedge \text{Acid}(y)])$$

to  $\mu(\sigma(\lambda y . y \sqsubseteq x \wedge \text{Acid}(y)))$ . This is licensed because the greatest degree to which an acid subpart of something measures that degree (on some dimension) is the measure of the mereological sum of the subparts that are acid (on the same dimension).

As a reviewer points out, *The soup is 30% oil* does not mean that the temperature of the oil part of the soup is 30% of the temperature of the soup. This observation is reminiscent of the constraints on the Mon head posited by Schwarzschild (2006). The reviewer thus suggests deriving PART from more basic parts including the Schwarzschild’s (2006) monotonicity head Mon. I will leave this for future work and simply assume here that  $\mu$  must be monotonic.

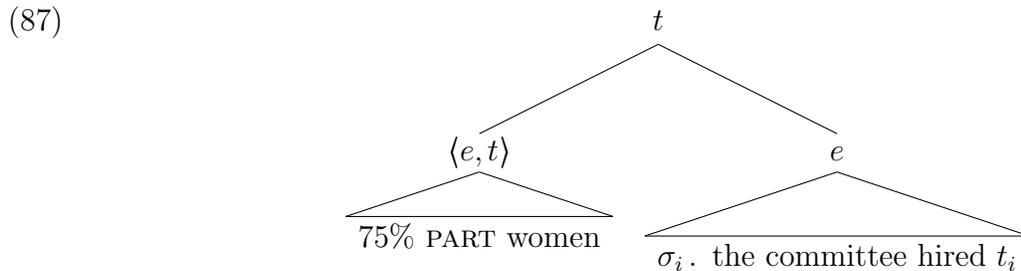
Clearly this analysis works quite unproblematically for the predicative cases. This kind of treatment would also straightforwardly produce the right kind of truth conditions for part-introducing attributive cases like *50% recycled paper*; the modifier *50% recycled* would translate as an expression of type  $\langle e, t \rangle$  in an entirely analogous way to *30% acid*, and this could be combined with *paper* using Predicate Modification.

Following Pasternak (2019), we need only adopt Wellwood’s (2015) MUCH in order to gain a satisfactory compositional account of the ‘conservative’ uses; this option is available under the present proposal as well. (There are no types of uses of *percent* in which M-OP or MEAS plays a crucial role under the present proposal.)

Could we work backwards from this analysis to gain a general account of part-introducing uses, including the reversed cases as in (86)?

(86) The committee hired 75% women.

Suppose that *75% women*, which denotes a predicate on this analysis, could undergo a kind of QR, and that a sum operation took place over the resulting lambda abstract:



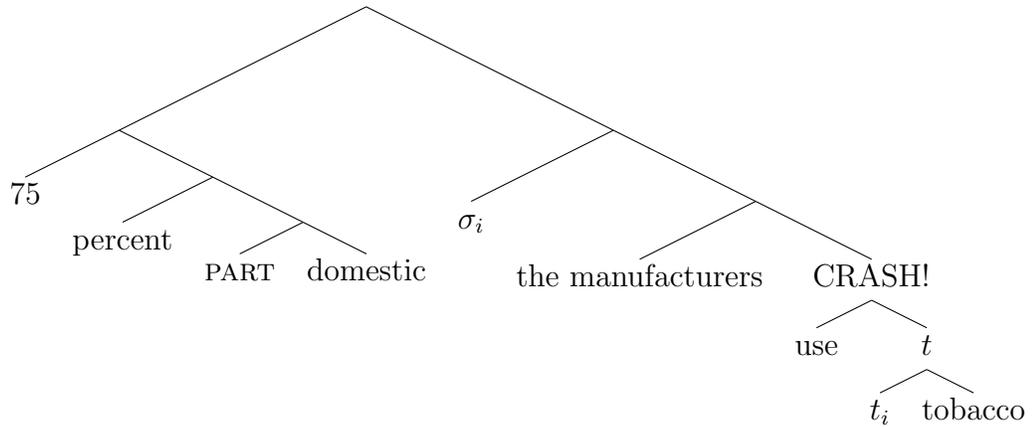
Then the predicate denoted by *75% women* would apply to the sum individual made up of everyone hired. This would produce the right truth conditions; it would imply that 75% of the individuals hired by the committee were women (assuming a context in which  $\mu$  measures cardinality). But can a predicate undergo QR and leave a trace of type  $e$ ?

The following case presents an even more vexing compositional challenge:

(88) The manufacturers use 75% domestic tobacco.

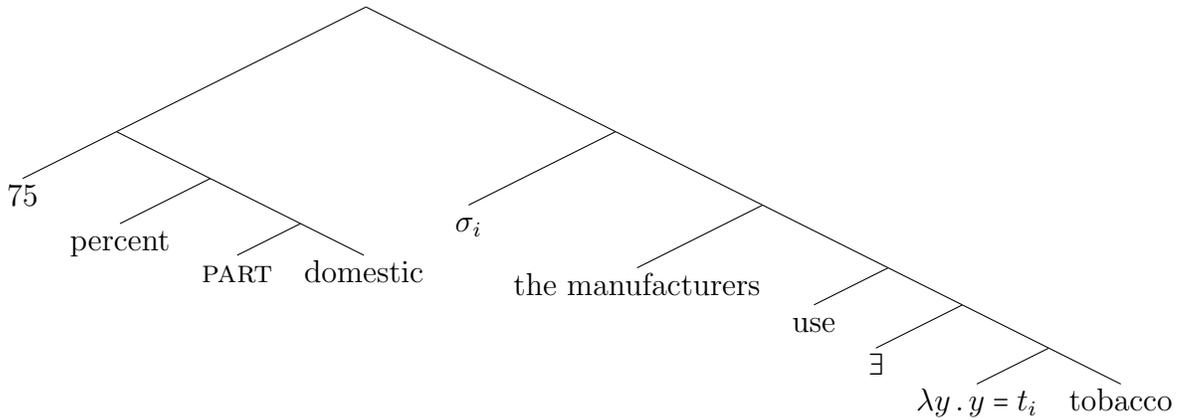
As discussed above, this seems to mean that 75% of the tobacco the manufacturers use is domestic. Sauerland & Pasternak (under review) would get that reading by placing focus on *domestic*. Within the current framework, and pursuing the idea that *75% domestic* forms a unit that acts as a predicate, the same truth conditions could be obtained if it were somehow possible to extract that predicate and sum over the content contributed by the rest of the sentence. It’s not entirely trivial to do that, though, because if extracting *75% domestic* from its position preceding *tobacco* left a trace of type  $e$ , then it wouldn’t be able to combine properly with its surrounding grammatical context:

(89)



An LF that would yield the right result would look something like the following (modulo QR of the object):

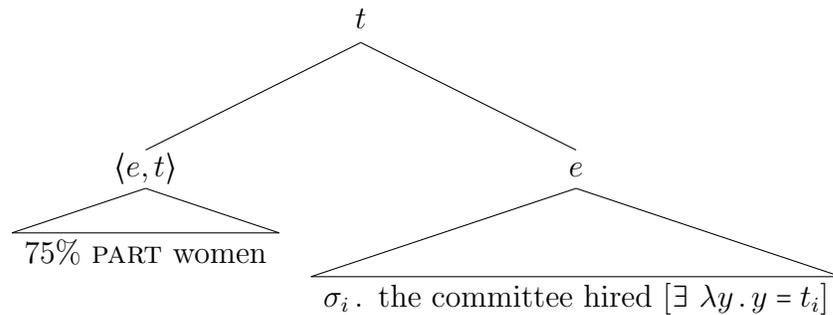
(90)



as if Partee's (1987) IDENT type-shift had applied to the trace;  $\text{IDENT}(x) = \lambda y . y = x$ .

In fact, it would be sensible to assume that the same kind of trace is left in (86), so we have:

(91)



Both of these LFs could be achieved by a covert transformation rule like the following analogue to QR:

(92) **Predicate Raising (PR)**



A first stab at a meaning representation for (94), then, might be the following:

$$(97) \quad [x]; [y]; \text{Mix}(z, x, y); 10\%(\text{Part}(\text{Manure}))(x \oplus y); 90\%(\text{Part}(\text{Coal}))(x \oplus y)$$

Here I assume that  $90\%(\text{Part}(\text{Coal}))$  denotes the predicate that our static system above derives for *90 percent PART coal*, with the PART type-shift.

But these truth conditions are a bit too weak. All they say is that there is a sum  $x + y$  such that 10% of it was manure and 90% was coal (and  $x$  was mixed with  $y$ ). They don't require that 90% of what was mixed in, total, was coal. To enforce that, we can sum over all candidate values for  $x \oplus y$  using Keshet's  $+$  operator, which gives the sum over all candidate values for a given discourse referent. In PLUS, the only terms are variables, and variables prefixed by  $+$ , whose interpretation is the sum of all current candidate referents for the variable:

(98) **Interpretation of terms in PLUS**

- $\|t\|^{g,\sigma} = g(t)$
- $\|+t\|^{g,\sigma} = \text{SUM}(\{g(t) : g \in \sigma\})$

where  $g$  is an assignment,  $\sigma$  is a set of assignments (a state), and for any set  $S$ ,  $\text{SUM}(S)$  denotes the smallest plural individual that contains every member of  $S$  as a subpart.<sup>11</sup>

With this operator in hand, we can characterize the semantics of (94) as follows:

$$(99) \quad [x]; [y]; \text{Mix}(z, x, y); 10\%(\text{Part}(\text{Manure}))(+(x \oplus y)); 90\%(\text{Part}(\text{Coal}))(+(x \oplus y))$$

Now, *90% coal* applies to the sum of all of the  $x + y$  sums such that  $x$  was mixed with  $y$ .<sup>12</sup>

This sum-operator provides an elegant representation of the meaning of the reversed cases as well. For example, it would allow us to represent *They hired 75% women* as follows:

$$(100) \quad [x]; \text{Hired}(z, x); 75\%(\text{Part}(\text{Women}))(+x)$$

Here  $+x$  denotes the sum of all  $x$  such that they hired  $x$ . For *They used 75% domestic tobacco*, a natural representation would be:

$$(101) \quad [x]; \text{Use}(z, x); \text{Tobacco}(x); 75\%(\text{Part}(\text{Domestic}))(+x)$$

---

•  $\sigma[[x = y]] = (\sigma[[x]])[[x = y]]$

<sup>11</sup>Keshet treats the denotation as the set itself, but treating the denotation as a plural individual works better for the purpose of giving a unified account of *percent*, as 'NUM% ADJ' can clearly function as a property of an individual, as in *The glass is 75% full*. The  $+$  operator is the crucial innovation in PLUS setting it apart from Dynamic Predicate Logic (Groenendijk & Stokhof, 1991).

<sup>12</sup>A reviewer points out that these truth conditions might be too weak. Suppose Jack has two barrels. In one barrel, there is a mixture weighing 50lbs consisting of 10lbs of manure thoroughly mixed with 40lbs of coal. The other barrel contains 50lbs of coal. Jack combines the contents of the two barrels. Did Jack mix 10% manure with 90% coal? The analysis in (99) would predict it is true in the situation described, but it doesn't seem like an appropriate description. In future research, it would be worth probing whether the sentence is really false in this situation or merely pragmatically infelicitous.

Here  $+x$  denotes the sum of all  $x$  such that they used  $x$  and  $x$  is tobacco. I conjecture, therefore, that part-introducing *percent* might always apply to a cross-assignment sum, evaluated after the rest of the constraints in the clause have been applied to the discourse referent in question. I suspect, moreover, that there may be some connection between the use of a cross-assignment sum and the restricted set of verbs that give rise to reversed readings, but I must leave that connection unexplored in the present work. I leave it to future research to develop this in a compositional dynamic framework and thoroughly evaluate its predictions.

## 7 Summary, conclusion, outlook

This paper has explored a wide range of part-introducing uses of *percent*, including not only the ‘reversed’ uses (as in *The committee hired 75% women*), but also predicative uses (as in *The committee is 75% women*), and attributive uses (as in *They used 50% recycled paper*). On the empirical side, this paper also showed that the range of ‘reversed’ uses is restricted to a rather limited set of verbs, and that certain cumulative-like examples can be found (as in *They mix 10% manure with 90% coal*).

In order to give a unified analysis of these part-introducing uses, I made two suggestions. First, I proposed a type-shift called PART that converts a non-gradable predicate to a gradable one that tracks mereological parthood. This makes any non-gradable predicate eligible for use with Pasternak’s (2019) analysis of *percent* in constructions like *75% full*. Pasternak’s (2019) analysis of *percent* provides an adequate analysis of the scalar predicative uses that can easily be extended to the scalar attributive uses. With the PART operator, we gain an adequate treatment of the non-scalar, part-introducing predicative and attributive uses, as well as the reversed uses. Second, motivated in part by slight compositional challenges in the realm of reversed uses but mainly in order to account for cumulative-like readings, I sketched a dynamic theory based on Keshet’s (2019) PLural Update Semantics, in which plural discourse referents that summarize the accumulated constraints on a given discourse referent can serve as the ‘whole’ for a *percent* phrase that both introduces and serves as a predicate. While it is only a sketch, it promises to deliver an elegant account not only of cumulative-like readings but also of reversed uses.

This investigation has raised a number of questions that I hope will be addressed in future work, including: why the range of reversed uses is limited to a small class of verbs, how the PART operator may be related to the monotonicity head Mon, how to capture the cumulative and reversed uses with a fully-explicit and unified compositional analysis, how to explain the infelicity of cumulative-like readings in scenarios like the one given in footnote 12, and how to analyze the wider range of uses of *percent* listed in the appendix.<sup>13</sup> It seems that by making

---

<sup>13</sup>Another open question is the relation between the cumulative cases and other somewhat similar cases like the following (Ezra Keshet, p.c.).

- (i) a. They mixed 1 part manure to 3 parts coal.
- b. They mixed 1 quarter manure with 3 quarters coal.

Note that *part* and *quarter* are not entirely analogous:

- (ii) a. #They mixed 1 part manure into their product.
- b. They mixed 1 quarter manure into their product.

a small amount of progress, we have created much more work for ourselves.

## References

- Ahn, Dorothy. 2012. Reverse quantification in Korean. Term paper, Harvard University.
- Ahn, Dorothy & Uli Sauerland. 2015. The grammar of relative measurement. In *Proceedings of SALT 25*, 125–142.
- Ahn, Dorothy & Uli Sauerland. 2017. Measure constructions with relative measures: Towards a syntax of non-conservative construals. *The Linguistic Review* doi:10.1515/tlr-2017-0001.
- Beaver, David & Emiel Krahmer. 2001. A partial account of presupposition projection. *Journal of Logic, Language and Information* 10. 147–182.
- Champollion, Lucas. 2009. Cumulative readings of *every* do not provide evidence for events and thematic roles. In *Proceedings of the nineteenth amsterdam colloquium*, .
- Cohen, Ariel. 2001. Relative readings of *many*, *often*, and generics. *Natural Language Semantics* 41–67.
- Coppock, Elizabeth. 2021. Challenge problems for a theory of degree multiplication. *Proceedings of Semantics and Linguistic Theory (SALT)* 31.
- Coppock, Elizabeth & David Beaver. 2015. Definiteness and determinacy. *Linguistics and Philosophy* 38(5). 377–435. doi:10.1007/s1098.
- Erteschik-Shir, Nomi. 2007. *Information structure*. Oxford University Press.
- Foley, William A. 2007. A typology of information packaging in the clause. In T. Shopen (ed.), *Language typology and syntactic description*, vol. 1, 362–446. Cambridge: Cambridge University Press.
- Groenendijk, Jeroen & Martin Stokhof. 1991. Dynamic predicate logic. *Linguistics and Philosophy* 14. 39–100.
- Herburger, Elena. 1997. Focus and weak noun phrases. *Natural Language Semantics* 5. 53–78.
- JCGM. 2012. International vocabulary of metrology - basic and general concepts and associated terms (VIM), 3rd edition. Joint Committee for Guides in Metrology (JCGM).
- Kennedy, Christopher & Louise McNally. 2005. Scale structure, degree modification, and the semantics of gradable predicates. *Language* 81(2). 345–381.
- Keshet, Ezra. 2019. Discourse plurals in an update semantics. In M. Teresa Espinal, Elena Castroviejo, Manuel Leonetti, Louise McNally & Cristina Real-Puigdollers (eds.), *Proceedings of Sinn und Bedeutung 23*, 557–571.

---

It seems that *part* is not interpretable without a specification of the total number of parts.

- Krifka, Manfred. 1986. *Nominalreferenz und Zeitkonstitution. Zur Semantik von Massentermen, Pluraltermen und Aspektklassen*. Munich, Germany (published 1989): Fink.
- Lambrecht, Knud. 1994. *Information structure and sentence form: Topic, focus, and the mental representation of discourse referents*. Cambridge: Cambridge University Press.
- Li, Haoze. 2018. Event-related relative measurement. In Chris Cummins Robert Truswell, Caroline Heycock, Brian Rabern & Hannah Rohde (eds.), *Proceedings of Sinn und Bedeutung 21*, vol. 21 2, 801–818. University of Konstanz.
- Partee, Barbara H. 1987. Noun phrase interpretation and type-shifting principles. In Jeroen Groenendijk, Dick de Jongh & Martin Stokhof (eds.), *Studies in Discourse Representation Theory and the theory of generalized quantifiers*, 115–143. Dordrecht: Foris.
- Pasternak, Robert. 2019. Unifying partitive and adjective-modifying *percent*. *Snippets* 37. 77–79.
- Raposo, Álvaro P. 2018. The algebraic structure of quantity calculus. *Measurement Science Review* 18(4). 147–157.
- Raposo, Álvaro P. 2019. The algebraic structure of quantity calculus II: Dimensional analysis and differential and integral calculus. *Measurement Science Review* 19(2). 70–78.
- Rett, Jessica. 2018. The semantics of *many*, *much*, *few*, and *little*. *Language and Linguistics Compass* 12.
- Romero, Maribel. 2020. The many readings of many: Pos in the reverse proportional reading. *Linguistics and Philosophy* .
- Sauerland, Uli. 2014. Surface non-conservativity in German. In Christopher Pi non (ed.), *Empirical Issues in Syntax and Semantics 10*, 125–142. <http://www.cssp.cnrs.fr/eiss10/>.
- Sauerland, Uli & Robert Pasternak. under review. German measurement structures: Case-marking and non-conservativity. Manuscript available on LingBuzz.
- Scha, Remko. 1981. Distributive, collective and cumulative quantification. In Jeroen Groenendijk, Theo Janssen & Martin Stokhof (eds.), *Formal methods in the study of language, part 2*, 483–513. Amsterdam: Mathematical Centre.
- Schwarzschild, Roger. 2006. The role of dimensions in the syntax of noun phrases. *Syntax* 9(1). 67–110.
- Solt, Stephanie. 2009. *The semantics of adjectives of quantity*. New York, NY: The City University of New York PhD dissertation.
- Solt, Stephanie. 2018. Proportional comparatives and relative scales. In Robert Truswell, Chris Cummins, Caroline Heycock, Brian Rabern & Hannah Rohde (eds.), *Proceedings of Sinn und Bedeutung 21*, 1123–1140.

Spathas, Giorgios. 2019. Plural anaphoric reference and non-conservativity. *Snippets* 37. 95–96.

Szabolcsi, Anna. 1997. *Ways of scope taking*. Dordrecht: Kluwer.

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

Westerståhl, Dag. 1985. Logical constants in quantifier languages. *Linguistics and Philosophy* 8. 387–413.

## A Conservativity

The so-called ‘reversed’ uses are analogous, at least in some ways, to ‘reverse proportional’ uses of *many* and *few*:<sup>14</sup>

(102) Many Scandinavians have won the Nobel Prize (Westerståhl, 1985)

*Proportional:*

$$\frac{|\text{SCANDINAVIAN} \cap \text{NOBEL}|}{|\text{SCANDINAVIAN}|} \text{ is high}$$

‘The ratio of Scandinavian Nobel Prize winners to Scandinavians is high’  
or ‘Many of the people of Scandinavia have won the Nobel Prize’

*Reverse proportional:*

$$\frac{|\text{SCANDINAVIAN} \cap \text{NOBEL}|}{|\text{NOBEL}|} \text{ is high}$$

‘The ratio of Scandinavian Nobel Prize winners to Nobel Prize winners is high.’  
or ‘Many of the Nobel Prize winners are Scandinavian’

(103) Few cooks applied. (Herburger, 1997)

*Proportional:*

The ratio of cook-applicants to cooks is low (few of the cooks)

*Reverse proportional:*

The ratio of cook-applicants to applicants is low (few of the applicants)

Both types of readings have to do with the proportion that a part makes up of a whole, but they differ in which part of the sentence describes the whole. In the proportional case, the description immediately following the quantity word (*Scandinavians* in (102), *cooks* in (103)) is the whole, and in the reverse proportional case, the rest of the sentence (*have won*

---

<sup>14</sup>See Cohen (2001) and Romero (2020) for more updated views on the precise semantics of examples like (102) and (103).

*the Nobel Prize* in (102), *applied* in (103)) describes the whole. The proportional reading is *conservative*, in the following sense:

- (104) A relation  $Q$  between two sets  $A$  and  $B$  is CONSERVATIVE if and only if  $Q(A)(B)$  is equivalent to  $Q(A)(A \cap B)$ .

For instance, the proportional reading of (102) – ‘Many of the people of Scandinavia have won the Nobel Prize’ is conservative because it is equivalent to ‘Many of the people of Scandinavia are Scandinavians who won the Nobel Prize’. Here  $A$  is ‘Scandinavian’ and  $B$  is ‘won the Nobel Prize’, and  $Q(A)(B)$  says that  $\frac{|A \cap B|}{|A|}$  is high. Replacing  $B$  with  $(A \cap B)$  in this formula gives  $\frac{|A \cap (A \cap B)|}{|A|}$ , which is equivalent to  $\frac{|A \cap B|}{|A|}$ , since  $A \cap (A \cap B)$  is the same thing as  $A \cap B$ .

The reverse proportional reading of *many* is not conservative, at least on the surface (although as most recently Romero (2020) discusses, the reverse proportional reading of *many/few* need not be analyzed as a violation of the hypothesized universal that determiners of category are conservative). In this case,  $Q(A)(B)$  says that  $\frac{|A \cap B|}{|B|}$  is high. Replacing  $B$  with  $(A \cap B)$  in this formula would give something very different:  $\frac{|A \cap (A \cap B)|}{|(A \cap B)|}$ , or equivalently  $\frac{|A \cap B|}{|A \cap B|}$ , where the numerator and the denominator are the same. Rather than ‘Many of the Nobel Prize winners are Scandinavian’, this would say ‘Many of the Scandinavian Nobel Prize winners are Scandinavian Nobel Prize winners’. Unlike this sentence, the reverse proportional reading is not tautologous, so the reading is not conservative.

The sense in which (1) (*The most recent class of NASA consists of 50% of the women*) is conservative is a bit trickier to explain because the quantifier is not in subject position. In this case,  $A$  is *of the women* and  $B$  is something like *who the most recent class of NASA consists of*. The use is conservative because it is equivalent to *The women in the most recent class of NASA consist of 50% of the women*. In (2) (*The most recent class of NASA consists of 50% women*),  $A$  is *women* and  $B$  is again *who the most recent class of NASA consists of*. This sentence is not equivalent to *The women in the most recent class of NASA consist of 50% women*; in fact, the latter would be tautologous, just as in the case of *many*. So, both *percent* and *many* give rise to both conservative and non-conservative readings. The non-conservative readings of *many* are called ‘reverse proportional’, and the non-conservative readings of *percent* are called ‘reversed’, alluding to that terminology.

## B Corpus study

In order to develop a picture of the combinatorial potential of *percent*, I carried out a corpus study on the Corpus Of Contemporary American English (COCA), which comprises American speech from a variety of genres (ca. one billion words, 1990–2019, drawn from spoken, fiction, popular magazines, newspapers, academic texts, TV and movie subtitles, blogs, and other webpages).

I searched for strings matching the following pattern:

VERB NUM[BER] percent ADJ[ECTIVE]

This search was intended to capture examples in which ‘NUM% ADJ’ serves as (an initial substring of) the complement of VERB, either with ADJ as the head of an adjectival phrase

serving as the complement, or with ‘NUM% ADJ’ at the left edge of a nominal complement to VERB. With 570 unique results, the search returned many such cases, and some additional ones. I sorted these into several categories.<sup>15</sup> In the main text of the paper, I discussed categories of part-introducing usages. Here, I review the types of examples that I set aside. These types of cases should ultimately be accounted for in a theory of *percent*, so I list them here for the purposes of future research.

**ADJ not an adjective.** In the dataset resulting from the search, there are a number of cases in which the string matching ADJ is not actually an adjective, but rather a verb in the past tense (e.g. in *recent Quinnipiac poll showed 63 percent favored execution for the Aurora theater shooter*) or a preposition (*Total costs at the company fell 4 percent due to lower expenses for compensation and promotion...*), or an adverb modifying the verb, not forming a constituent with NUM PERCENT ( *in the first week after Hurricane Sandy hit, the number of burglaries increased 11 percent citywide – not three*). These can safely be ignored.

**Small clauses and depictives.** There are also cases where NUM PERCENT functions on its own as a noun phrase, and ADJ serves as another argument of the verb, as a predicate of a small clause, or as a depictive modifier of NUM PERCENT as a noun phrase.

VERB [ NUM% ] [ ADJ ... ]

Examples include the following:

- (105) Both of you **got** 96 percent correct (small clause)  
 (106) we **dump** 40 percent untreated into the Vistula (depictive)

These can be paraphrased with an *of them* or *of it* following *percent*, e.g. ‘Both of you got 96 percent of them correct.’ Thus ‘NUM%’ here functions independently as a referential noun phrase. These cases may be interesting for some purposes but they are not particularly germane to our investigation, which is concerned with the semantics of *percent* when it combines directly with a predicate of some sort.

**Comparative and comparative-*ish*.** There’s a large class of examples in which the ADJ is comparative, and these include both predicative and attributive uses:

- (107) The moon will **appear** 14 percent bigger and 30 percent brighter than when it’s at its farthest from us.  
 (108) In a properly tuned engine, natural gas combustion **delivers** 20 percent lower carbon emissions ...  
 (109) Residential customers are **guaranteed** 10 percent better rates than whatever they are currently paying.

---

<sup>15</sup>The COCA corpus did not contain any matches for ‘VERB NUM [thirds|fourths|fifths|sixths] ADJ’, so *percent* was the only proportional measure noun that I investigated.

(110) As the government is keen to point out, Islamic banking is **growing** 50 percent faster than conventional banking.

(Note that the last example is actually adverbial; this makes up a fairly large proportion of the comparative examples returned in the search.) I have left these aside in the present investigation.

One interesting thing to note in relation to comparatives, however, is that there are examples that are not exactly comparative but seem to have a comparative-like semantics, so that what's being described is some kind of change. For example, just as you have *15% bigger*, where the result is 115% of the size of the starting state, you have *15% extra*, where the result is 115% more than some starting state. *Extra* is one of the words that is somehow comparative-like, and I've coded cases involving *extra* and similar words as 'comparative-ish'.

(111) and professionals said it was striking in this case that the testing companies granted 100 percent extra time for students who were suddenly diagnosed with a serious learning disability during SAT...

(112) we've **got** 35 percent increased revenues.

(113) The shares have **marched** 19 percent upward since their initial public offering in May.

(114) New Dell PowerEdge R410 Server **Brings** 80 Percent Improved Performance<sup>1</sup> to Technical and High Performance Computing

(115) San Diego **fell** 8 percent short of its goal.

(116) The children he treats are **averaging** 80 percent overweight, meaning that if the average weight for a child at a certain age and height is 100 pounds, the average overweight child treated by Dr. Robinson would weigh 180.

The context surrounding the 'overweight' case helps to clarify that *80 percent overweight* can be paraphrased as '180% of average weight', or '80% more than average weight'.

**Quantity uses.** Finally, there are a number of other uses that are not of great interest for the present study, although they do make explicit the very wide range of quantity types whose values are percentages, and may be interesting for investigations related to that issue. A very common use of 'NUM%' – at least in this dataset, where it linearly precedes an adjective – is as a characterization of the value that a particular type of quantity takes on. These examples frequently concern finance:

(117) Since 1995, the country has often **achieved** 7 percent annual growth.

(118) A newly **instituted** 11 percent value-added tax has been condemned by everyone from labor unionists to businesspersons.

But they can concern other subjects, such as politics or science:

(119) The state's goals, according to its application, **include** 100 percent grade-level proficiency by 2014

- (120) Once I **reach** 96 percent synaptic connection as measured by this device all I need to do is ...

In these cases:

SUBJ VERB NUM% ADJ NOUN

can be paraphrased as

SUBJ VERB a(n) ADJ NOUN value of NUM%

For example, (120) can be paraphrased, ‘Once I reach a synaptic connection value of 96 percent as measured by this device all I need to do is....’.

These kinds of ‘quantity’ uses, as I call them, are very frequent among cases in which ADJ is followed by a noun. Here is a possible member of the ‘quantity’ category lacking an overt noun:

- (121) Clinton’s unfavorables nationally are not all that far behind. She’s got 49 percent unfavorable nationally among all voters out there, 44 percent favorable.

The sentence in question can be paraphrased as: ‘Clinton has an *unfavorable* value of 49%.’ To me this case sounds like politician slang, and not fully grammatical, probably because *get* requires a nominal complement, and therefore *49% unfavorable* must be coerced into a nominal phrase in order to work in this grammatical context. Indeed, the pluralized use of *unfavorable* in the preceding sentence suggests that this genre easily allows for phrases headed by *unfavorable* to be construed as nominal.

Here are some ‘quantity’ uses with an indefinite article and a count noun:

- (122) you have got a 14 percent black unemployment rate  
(123) We get a 20 percent military discount.  
(124) When somebody has got a 92 percent liberal rating and a 0 percent conservative rating, he’s a liberal.

**Proposition-denoting complements.** Finally, there is one part-introducing use that I did not include in the survey above. Transitive verbs that expect proposition- or situation-denoting complements give rise to slightly different entailment patterns from the other part-introducing uses with transitive verbs. For example, in (125) the noun phrase seems to denote a state of being: ‘100 percent of the electricity that there is is carbon-free’, so ‘there is no electricity other than carbon-free electricity’.

- (125) Cuomo says his state [New York] hopes to **achieve** 100 percent carbon-free electricity by 2040.

This is not a ‘reversed’ case like (16)/(31) (... *requiring manufacturers to use 75% domestic tobacco*). Whereas in (16) the requirement is about a percentage of the tobacco that the manufacturers *use*, here the hope is not about a percentage of the electricity that New York achieves. The hope is that New York will achieve a certain state of being, one where 100 percent of the electricity is carbon-free. The verb *achieve* is very different from the verb

*use*; *achieve* expects a complement-denoting proposition. More concretely, (125) can be paraphrased as ‘... achieve a situation where 100% of the electricity is carbon-free’.

This case is similar to some of the attributive cases we’ve seen, insofar as ‘carbon free’ characterizes the part and ‘electricity’ characterizes the whole. Yet the relative clause paraphrase of this case is mildly odd:

(126) ?Cuomo... hopes to achieve electricity that is 100 percent carbon-free.

A better paraphrase is as follows:

(127) Cuomo... hopes to achieve a situation where 100 percent of electricity is carbon-free.

In general, this class of usages licenses entailments of the form ‘... VERB a situation where NUM PERCENT of (...) NOUN is ADJ’. Some further examples in this category:

(128) Aspen, Colo. [is] potentially the second city in the country to **achieve** 100 percent clean power  
‘... achieve a situation where 100% of the power is clean’

(129) Hutchison’s amendment would have **authorized** 100 percent federal funding for the Northline and Southeast extensions in exchange for Metro funding...  
‘...authorized a situation where 100% of the funding... is federal...’

(130) It imposed a freeze on all new foreign investment in mining and **required** fifty percent Australian equity in all mining ventures by 1975.  
‘...required a situation where 50% of equity... is Australian’

Insofar as the adjective that appears after *percent* characterizes the part, as it were, these cases can broadly be characterized as ‘part-introducing’, along with the part-introducing predicative uses, the part-introducing attributive uses, and the ‘reversed’ uses, but I consider them their own subclass within that broader category.