

# *Dutchmen are Good Sailors:* Generics and Gradability

Bernhard Nickel · Harvard University

August 12, 2009

## Abstract

While generics generally tolerate exceptions, they seem to play different roles in different kinds of generics. In paradigmatic generics, such as *ravens are black*, the exceptions are reasonably considered deviations from a norm or default pattern. But there are many generics, including *Dutchmen are good sailors* for which this approach seems problematic: Dutchmen who aren't good sailors intuitively aren't deviations in the way that albino ravens are. This paper presents a new theory on which the interpretation of these apparently troublesome examples is the result of the interaction between the semantics of generics and the semantics of gradable predicates. The theory provides extensionally better truth-conditions than some rivals, predicts generalizations about the phenomenon, and does so while making use of only the most widely shared assumptions about the semantics of generics.

**Keywords** SEMANTICS, QUANTIFICATION, GENERICS, HOMOGENEITY,  
GRADABLE PREDICATES, RELATIVE GENERICS

## 1 Introduction

Paradigmatic generic sentences seem to express quite strong generalizations. Since generics tolerate exceptions, the generalizations do not hold universally, but their exceptions are plausibly deviations of one sort or another, as in (1) and (2).

- (1) Ravens are black.

(2) Tigers have stripes.

A semantic theory for generics starting out from these examples will assign them quite strong truth-conditions, perhaps saying that they are true just in case most or all of the normal members of the kind satisfy the property predicated. The strongest support for this approach comes from generics that are false even though most members of the kind at issue conform to the generalization, such as the examples in (3).

- (3) a. Books are paperbacks.
- b. Humans have dark hair.
- c. Sea-turtles die young.
- d. Prime numbers are odd.

We have a ready explanation for this fact if generics have very strong truth-conditions.

One of the core difficulties generics pose is that there are many examples that cast doubt on the basic workability of this approach. All of these are generics that are true, even though most of the members of the kind do not conform to the generalization, not even over the long haul. Consider (4)-(6).

- (4) Lions have manes.
- (5) Dutchmen are good sailors.
- (6) Sea-turtles are long-lived.

Most lions don't have manes, since most lions are females. Most Dutchmen aren't sailors, and *a fortiori* aren't good ones. And most sea-turtles die almost immediately after they hatch. These facts are completely stable and not mere aberrations from a larger pattern. I'll say that when a generic is compatible with most members' of the kind at issue not conforming to it, not even over the long haul, that generic has *weak* truth-conditions.

Generics with weak truth-conditions are not just problematic, they also promise to be the most illuminating test cases. Very many theories give reasonable interpretations of the paradigms in (1) and (2). Far fewer can do so while simultaneously covering (4)-(6). Clearly, the multitude of interpretive possibilities must be the result of different factors interacting in complex

ways, some purely linguistic, others drawn from non-linguistic cognitive systems, perhaps systems of categorization, induction, or probability. By better understanding which semantic phenomena are due purely to the linguistic component of interpretation, we can gain a foothold for investigating the non-linguistic component, as well: it's responsible for whatever is left over once our purely linguistic resources are exhausted. One way to tell apart the phenomena due to linguistic factors from those due to others is to look not just at particular examples, but at patterns. If we can formulate a generalization about the interpretation of generics purely in terms of linguistic phenomena, such as the interaction of quantification and gradability, that gives strong evidence for believing that generics covered by that generalization aren't particularly revealing of the non-linguistic ingredient.

With this broader context in mind, I present a theory of *some* generics with weak truth-conditions that makes use only of linguistic resources. The examples I'll focus on are generics like (5). These are instances of a generalization about how quantification interacts with some gradable predicates, and thus serve as an illustration of the general methodological outlook I just articulated. They are also interesting for more localized reasons, since the problems they raise can't be solved by simply extending a widely accepted treatment for examples like (4). There, we might account for the proper interpretation by suggesting that the generic isn't about lions generally—in slightly more formal terms, the generic quantifier is suitably restricted, and within the restricted domain, the strong generalization that most lions have manes, or that all normal lions have manes, or what have you, holds.

This quantifier domain restriction only helps us part of the way for examples like (5). These sentences exhibit two problems, one concerning inferential relations, the other concerning their relatively weak truth-conditions, and as I argue in §2, the quantifier domain strategy only helps with the former. The problem of weak truth-conditions persists, which I discuss more explicitly in §3. There, I also criticize some extant treatments. §4 contains my positive proposal.

One preliminary remark: my discussion won't depend on any special assumptions about the interpretation of generics. It doesn't matter how exactly the generic quantifier is interpreted, whether in terms of *most* in a suitable domain or a suitably restricted universal quantifier. It also doesn't matter

where in the LF of generics the generic operator appears, whether as a nominal determiner of a bare plural, as an adverb of quantification, or somewhere else.<sup>1</sup> For that reason, I'll represent the generic operator abstractly as GEN, and I'll represent the interpretation of a generic sentence of the form *As are F* as  $[\text{GEN}(x): A(x)](F(x))$ .

## 2 The Port-Royal Puzzle

(5) has a long provenance. It was first mentioned explicitly by Arnauld and Nicole (1996, originally 1662), which is why it has come to be known as the Port-Royal puzzle. For our purposes, the puzzle is best put as a dilemma. The horns are that either sentences like (5) are treated semantically on a par with other generics or they are not. If they are, then any reasonable quantificational approach predicts wrong entailment relations among generics. If they are not, then quantificational approaches have to recognize a theoretically unattractive ambiguity.

The problematic observation that generates the puzzle is that (5) does not entail (7).

(7) Dutchmen are sailors.

This cannot be accounted for if two extremely plausible assumptions hold. The first is that (8) seems to be true.

(8)  $(\forall x)(\text{Good.Sailor}(x) \rightarrow \text{Sailor}(x))$ .

The other is that GEN is upwards entailing in its nuclear scope, i.e., that it validates the inference in (9).

(9) (i)  $[\text{GEN } x: Ax](Fx)$

(ii)  $(\forall x)(Fx \rightarrow Gx)$

$\therefore$  (iii)  $[\text{GEN } x: Ax](Gx)$

The plausible candidates for the interpretation of GEN such as *most* or *all normal* all satisfy this condition. So if we interpret (5) in terms of GEN, we are

---

<sup>1</sup>For the determiner option, see, e.g., Asher and Morreau (1995); for the adverb of quantification option, see, e.g., Cohen (1999a,b); Schubert and Pelletier (1989); Wilkinson (1991). Elsewhere, I have argued for the "somewhere else" option. See [Reference suppressed for blind review].

committed to an unacceptable entailment pattern. The only way to avoid this is to embrace an unhappy ambiguity.

However, this argument is unconvincing at two key steps. The first concerns the generalization (8). As Larson (1998) has pointed out, the generalization need not be true. For example, it may be true that Joan, who only sails once in a while and has a full-time job away from sailing is quite good at it when she does sail. In that case, *Joan is a good sailor* is true, but *Joan is a sailor* is false. Thus, (8) is if true, only a contingent truth, robbing the example of its force.

The argument is also problematic in the inference from the premise that all interpretations of the generic quantifier are upward entailing in their nuclear scope to the conclusion that (5) would have to entail (7) if interpreted using GEN, even if (8) is true. This is easiest to see in the following example.<sup>2</sup>

- (10) a. Lions have manes.  
b. Lions are male.  
c.  $[\forall x: \text{Lion}(x)] (\text{Has.A.Mane}(x) \rightarrow \text{Male}(x))$ .

(10a) doesn't entail (10b), even though (10c) is clearly true. A plausible and common response is to say that the interpretation of GEN, and specifically its restrictor, depends on the predicate of the sentence. The informal justification for this maneuver is that a sentence like (10a) is only about those lions that are even in the business of having sexually selected ornamentation, whereas (10b) is about lions that are in the business of having a gender, a much larger class.

Notice that on this strategy, it's still true that the quantifier used to analyze GEN is upward entailing in its nuclear scope, so long as the restrictor of the quantifier is held constant. But once we change predicates, as we do in moving from (10a) to (10b), that condition is no longer met. Thus, we can acknowledge that the quantifier is upwards entailing in its nuclear scope, that (10c) is true, and still deny that (10a) entails (10b).

The same strategy is appropriate in the case of (5) and (7). While the former quantifies only over those Dutchmen that have some sailing skill, the

---

<sup>2</sup>Examples like this one are already presented in Carlson (1977). He considers the inference from "chickens lay eggs" to "chickens are hens." I use a different example, since most chickens are hens, so that the chicken-hen example doesn't speak against interpretations of generics in terms of *most*. The example about lions makes the point more effectively since most lions don't have manes.

latter might quantify over all Dutchmen that have any avocation at all. Obviously, even if it's true that all Dutchmen in the smaller group are good sailors, and even if they are sailors, nothing at all follows about whether all Dutchmen in the larger group are sailors.<sup>3</sup>

In this context, it is useful to consider the concerns raised by Krifka et al.

we could tamper with the restrictor or with the interpretation of GEN, changing it in such a way that it yields truth conditions which seem to be correct (cf. Geurts, 1985). For example, it seems that the main sentence stress in (5) can be on *good*. Given the correlation between sentence accent and the partitioning between restrictor and matrix [...], we can claim that **good** is part of the matrix, whereas **sailor** is part of the restrictor. We then get a more plausible semantic representation that amounts to 'Dutch sailors are good sailors.'  
[...]

However, we also get this reading with an accent on *sailors* (where we can assume that *good sailors* is in focus), and so just assuming different partitionings will not give us a general solution.<sup>4</sup>

How compelling this objection is depends on what exactly it is an objection to. As I have argued just now, quantificational accounts of generics need to accept on very general grounds that the restrictor of the quantifier depends, at least in part, on the predicate of the generic. So the fact that no particular intonational stress, or for that matter any other obvious marking of a change of the restrictor, is present in a sentence does not speak against this claim. And once this modification of the restrictor by the predicate is accepted, we have a solution to the Port-Royal puzzle.

For some generics with relatively weak truth-conditions, such as (4), *lions have manes*, nothing more than this predicate-induced restriction is required. Within the restricted class, the usual strong generic quantification yields an appropriate interpretation. But in the case of (5), the predicate induced restriction is insufficient to yield the right analysis. Even if we interpret that

---

<sup>3</sup>Geurts (1985) and Cohen (1999b) also hold that the predicate restricts the quantifier used to interpret GEN, and Leslie (2008) suggests using this strategy to solve the Port-Royal Puzzle, building in part on Larson's diagnosis of why (8) is false. But unlike Leslie, my argument in the main text is that the Port-Royal Puzzle is solved by a general feature of the interpretation of GEN that is independent of whether a principle like (8) is true.

<sup>4</sup>Krifka et al. (1995, 82). I've adjusted the numbering to reflect the numbering in this paper; bold-face is used by Krifka et al. to denote the semantic value of the word bold-faced.

sentence as saying roughly that GEN-many Dutch sailors are good sailors, we might still think that the interpretation is too strong. After all, it doesn't seem to be required for (5) to be true that Dutch sailors generally or typically or normally are good sailors. A fair chunk of them may be quite ordinary.

One could, at this point, suggest that the restrictor of the predicate is further restricted in some way in order to give a more reasonable, i.e., weaker interpretation to (5). Perhaps speakers accommodate extra material into the restrictor based on context, so that (5) is interpreted as saying roughly that GEN-many of the best Dutch sailors are good sailors. The criticism that Krifka et al. raise in the passage I just quoted is much more telling against this proposal. In general, generics do not seem to allow for the context to restrict the domain in the way that other quantifiers do. Hence, some special mechanism is required if the restrictor is supposed to be further enriched, and placing focal stress on some material in the sentence certainly is one such special mechanism. But as Krifka et al. rightly point out, focus isn't required. In fact, we might strengthen their observation by pointing out that nothing special is required in order for (5) to have an interpretation that is plausible true, or at least was true at the time Arnauld produced it. And that suggests that we cannot solve the problem raised by the weak reading of (5) by appeal to adding extra material to the restrictor of the generic operator. Let's take a closer look at that problem now.

### 3 Inappropriately Strong Truth-Conditions

The problem of weak truth-conditions is a matter of deriving unacceptable predications about the truth-conditions of (5), or the similar and slightly more contemporary *Bulgarians are good weight-lifters*, from very simple assumptions. All that is required is a hypothesis about the interpretation of the generic quantifier and how that quantifier interacts with the predicate *good sailor*.

A plausible semantics for that predicate relates its argument to an abstract representation of sailing skill, which we can represent as a degree on a scale that allows us to compare various skill levels.<sup>5</sup> The context then supplies an-

---

<sup>5</sup>The view that gradable adjectives should be interpreted in terms of scales is extremely widespread in the literature on these predicates. See, e.g., Bartsch and Vennemann (1973); Bierwisch (1989); Cresswell (1976); Heim (2000); Kennedy (1999, 2007); Kennedy and McNally (2005); Klein (1991); Seuren (1973); von Stechow (1984).

other value on that scale as a standard against which the sailing skill of the argument is measured, and the predicate is true of its argument just in case the degree to which that argument is proficient at sailing exceeds the contextually supplied standard. For example, *John is a good sailor* is true just in case the degree to which John is a good sailor exceeds the contextually supplied standard. When we have a quantified subject, rather than a proper name, we assume the following principle. A quantified claim  $[Qx: Fx](\text{Good.Sailor}(x))$  is true iff  $Q$  many  $F$ s all possess a degree of sailing proficiency that exceeds the contextually given one.

Applied to (5), these assumptions entail that it is true iff GEN-many Dutch sailors exceed a contextually supplied standard. But given that we want GEN to have a fairly strong meaning, such as *most* or a suitably restricted *all*, that in turn entails that (5) has far stronger truth-conditions than are intuitively attested: (5) does not require for its truth that most Dutch sailors exceed a contextually given standard. Something has to give. I'll first discuss the possibility of solving the problem by introducing an alternative generic operator, one that Ariel Cohen has introduced and defended in his (1999b; 2001; 2004). I'll suggest that this proposal does not work, which will motivate my own proposal in §4.

### 3.1 *Absolute and Relative Generics*

To introduce Cohen's system, let me begin with his treatment of the core examples (1) and (2), which he calls *absolute* generics. He interprets the generic quantifier in such a way that its restrictor is determined, in part, by the predicate via its association with a set of alternatives. To interpret *As are F*, we have to compute the set of alternatives  $\text{ALT}(F)$ . In most cases,  $F$  is included in  $\text{ALT}(F)$ , and in most cases, the alternatives are mutually exclusive. For example, to interpret (1), *ravens are black*, we associate the property of being black with alternative colors. With that set in hand, Cohen gives the following truth-conditions.<sup>6</sup>

(11) *As are F* is true iff the probability that a randomly chosen  $A$  that also satisfies at least one of the properties in  $\text{ALT}(F)$  is  $F$  is greater than .5.

---

<sup>6</sup>See Cohen (1999b, 37).

These truth-conditions deal well with the lions/manes problem, because the alternatives to the property of having manes are other ways of having sexually selected ornamentation, and no female lions satisfy any of these alternatives. However, Cohen accepts that these semantics, when applied to (5), yield the truth-conditions that the probability that a randomly picked Dutch sailor is a good one is greater than .5, and that is too strong.

In response, Cohen introduces relative generics, generic sentences that are analyzed in terms of an alternative generic operator. Generics are therefore systematically ambiguous, depending on whether they are analyzed as absolute or relative. The core idea behind relative generics is that they reflect ways in which the members of the kind at issue might be distinguished from members of other kinds that are relevant in the conversation. The relative generic says that the members of the kind at issue are more likely than members of these other kinds to satisfy the predicate of the generic. More formally, when *As are F* is analyzed as a relative generic, we do not just consider the alternatives to *F*, but also the alternatives to *A*,  $\text{ALT}(A)$ . In the case of (5),  $\text{ALT}(A)$  might include other nationalities. Relative generics have the truth-conditions in (12).<sup>7</sup>

- (12) *As are F* is true iff the probability that a randomly chosen *A* that satisfies one of the alternatives in  $\text{ALT}(A)$  is *F* is higher than the probability that an arbitrarily chosen object that satisfies one of the members of  $\text{ALT}(A)$  and one of the members of  $\text{ALT}(F)$  is *F*.

As applied to (5), these semantics yield the truth-conditions that an arbitrarily chosen Dutch sailor is more likely to be a good sailor than an arbitrarily chosen sailor from one of the alternative nations. We can see why this interpretation is aptly called *relative*. A relative generic requires for its truth that the relevant members of the kind be more likely to satisfy the predicate than members of some other kind(s), i.e., we are relating different kinds, such as Dutchmen and Swiss.

So that we may evaluate this proposal, let me say how claims about likelihoods are related to facts “closer to the ground.” For our purposes, we

---

<sup>7</sup>See Cohen (1999b, 55f). One benefit of Cohen’s strategy is that it makes the difference between relative and absolute generics not completely *ad hoc*. As he argues in Cohen (2001), the difference between the two readings can be reduced to a difference in the setting of one parameter, and he argues that we can see the same difference in parameter setting in some non-generic cases involving *many* and *often*.

can simply translate talk of probabilities into talk of ratios. To say that the likelihood that a randomly picked Dutch sailor is good is higher than the likelihood that a randomly picked sailor from some other country is amounts to the claim that the ratio of good Dutch sailors to Dutch sailors of any skill is higher than the ratio of good sailors from other countries to sailors from these countries of any skill.<sup>8</sup>

The introduction of the relative generic quantifier is only half of Cohen's theory. If that was all there was to it, Cohen would predict far more true readings of generics than are intuitively attested. Consider, for example, (13).<sup>9</sup>

(13) Bees are sterile.

It might appear as if the truth-conditions for a relative generic reading of (13) are satisfied, so that (13) ought to be true. After all, the vast majority of bees are worker bees, and worker bees are sterile. Thus, the likelihood that a randomly picked bee is sterile is far higher than the likelihood that a randomly picked member of some alternative kind is.

For this reason, Cohen introduces the *homogeneity* requirement. It is part of his account that when we assess the truth-value of a generic by assessing a ratio within a certain reference class, that reference class has to satisfy certain constraints in order to be eligible to play a role in determining the interpretation of the generic, i.e., in order to be admissible. The constraint he imposes is that such a class must be homogeneous: if the relevant ratio is a certain value in the class as a whole, then it needs to be the same in all cells of a psychologically salient partition. If this homogeneity constraint is not satisfied, then the sentence is unacceptable—perhaps it suffers from presupposition failure, perhaps it is simply false.

In the case of (13), the homogeneity constraint is violated because a psychologically salient partition partitions the bees according to sex. And while the workers are sterile, the drones are not. Because (13) would only be true if homogeneity were satisfied, and because homogeneity is violated in this case, (13) is unacceptable.

---

<sup>8</sup>Within the context of Cohen's theory, the initial interpretation of generics in terms of probabilities plays other roles than simply introducing ratios. It also allows him to motivate various constraints on the classes within which the relevant ratios are assessed, what he calls homogeneity constraints, of which far more immediately below. See also Cohen (1999a, 2004).

<sup>9</sup>This example has also been cited by Leslie (2007).

Cohen's theory of weak truth-conditions thus has two interlocking parts: a relative generic quantifier and the homogeneity constraint. Impressionistically, the very weak relative generic quantifier makes weak truth-conditions available for all generics while the homogeneity constraint pares the class of available readings back so that the account as a whole does not overpredict.

### 3.2 *Problems*

There are at least two kinds of concerns about Cohen's proposal. The first is that the interaction of the homogeneity constraint and the relative generic quantifier cannot coherently predict all of the attested cases of weak truth-conditions while ruling out of court cases where weak truth-conditions aren't attested. Here, I will focus on a constellation of data concerning non-gradable predicates and two kinds of gradable predicates, what are sometimes called absolute and relative gradable predicates.<sup>10</sup> The second problem concerns the relative generic reading itself. I will argue that it doesn't get the truth-conditions for generics with relative gradable adjectives right.

Let me begin with a distinction among gradable predicates. I will assume that all gradable predicates are interpreted along the lines I sketched at the beginning of this section: the predicate maps its argument to a point on a scale, intuitively the point that represents the degree to which the argument has the relevant property. The predicate is true of its argument just in case the point to which the argument is mapped is appropriately related to a standard on that scale. For some predicates, it may be required that the argument is mapped to a point above the standard, as for *heavy*, while for others it may have to be mapped to a point below that standard, as for *light*.

Given these basic assumption, we can now define a gradable predicate as *relative* just in case the standard with respect to which the truth of the predication is evaluated can vary from context to context. These include predicates such as *big*, *tall*, *fat*, *expensive*, and *good F*. What counts as *big*, for example, can vary from context to context, which is just to say that the standard value on the size scale can vary. By contrast, *absolute* gradable predicates are predicates for which the standard is obligatorily set to some privileged value, usually an

---

<sup>10</sup>The terminology for the distinction among gradable predicates is unfortunate in this context since it is liable to lead to conflation with relative and absolute readings of the generic. However, I retain both terminological conventions since they are established in the literature on these topics.

extreme value on the associated scale.<sup>11</sup> Absolute gradable predicates include *empty, full, straight, wet, bent*, and color predicates.<sup>12</sup> To complete the typology, genuinely non-gradable predicates are ones that do not tolerate comparative morphology or intensification, such as *locked, political, costs \$100*, and *is 7ft tall*.

Given this typology, we can describe some patterns regarding the availability of weak truth-conditions. First, weak truth-conditions are always available for relative gradable predicates, as the examples in (14) illustrate.

- (14) a. Luxury cars are expensive.  
b. Europeans are tall.  
c. Americans are fat.

However, weak truth-conditions aren't always available for absolute gradable predicates. Consider the examples in (15).

- (15) a. Luxury cars are black.  
b. Buses are empty.  
c. Subways are full.

Finally, weak truth-conditions aren't systematically available for non-gradable predicates, as the examples in (16) show.

- (16) a. Luxury cars cost over \$150,000.  
b. Europeans are 1.95m tall.  
c. Americans weigh 250lbs.

Let me emphasize that the two last observations are crucially not that weak truth-conditions for sentences containing these predicates are *never* available. Obviously, they are. The point is rather that they aren't as systematically available as they are for relative gradable predicates.

The question for Cohen's account is whether he can consistently account for all of these data because these data pull in different directions. The availability of the weak readings for relative gradable predicates requires that when we evaluate (14b)-(14a), the homogeneity constraint is satisfied. The fact that

---

<sup>11</sup>See also Kennedy and McNally (2005); Rusiecki (1985); Unger (1975).

<sup>12</sup>For evidence that the absolute and gradable predicates really are semantically distinct, see Kennedy (2007, §3.2).

weak readings aren't available for the non-gradable predicates in (16b)-(16a) requires that when interpreting them, homogeneity is not satisfied. And as I'll argue below, the fact that weak readings aren't available for (15a)-(15c) speaks against the most plausible way of distinguishing between the (14) and (16) examples.

Consider (14a), *luxury cars are expensive*. Since that sentence is true, homogeneity is satisfied. Now consider the non-gradable counterpart (16a), *luxury cars cost over \$150,000*. It's clear that, if homogeneity were satisfied in this case, it would be true on the relative generic reading. Hence, homogeneity is not satisfied. Importantly, the only difference between (14a) and (16a) is the manner in which we refer to a standard price that luxury cars must exceed in order for the respective sentences to be true. In the former case, we are letting context determine the standard that needs to be exceeded. In the latter, we are mentioning it directly.

What that shows is that whether or not homogeneity is satisfied depends on the manner in which we refer to such a standard. Given that the domain over which we're quantifying is plausibly the same in both (14a) and (16a), the only factor that can account for why homogeneity is satisfied in the one case but not the other is what partition(s) are salient. This, too, is the diagnosis Cohen gives of why examples like (16a) are not acceptable.

whenever a generic predicates a property of some concept, and the predicate contains reference to a value on a scale, the concept is represented as a multidimensional space, with this scale as one of its dimensions. The prototypical case of a value on a scale is, of course, a number word (Horn, 1972). Hence, any sentence whose predicate contains an explicit number word will be ruled out, by failure of homogeneity, so long as there are *any* exceptions to the predicated property.<sup>13</sup>

That is, Cohen's diagnosis runs as follows. A partition along degrees on a scale can be made salient by mentioning one such degree. In the case of (16a), the scale of price is made salient by mentioning a particular price. That means that each cell in the partition contains all of the luxury cars that cost the same dollar amount: one cell contains the cars that cost (say) \$150,001-\$155,000, an-

---

<sup>13</sup>Cohen (2004, 548).

other those that cost \$155,001-160,000, and another contains those that cost \$140,001-\$145,000. Homogeneity is obviously not satisfied, since the likelihood that a randomly picked luxury car in the first cell costs over \$150,000 is 1 (certainty), while the likelihood that a randomly picked luxury car in the last cell costs over \$150,000 is 0 (impossibility).

Given that (14a) is true, Cohen is committed to saying that not just any way of referring to a point on a scale is such that referring to it in that way makes the scale salient for purposes of evaluating whether the homogeneity constraint is satisfied. That's because we're referring to such a point in (14a), as well, but homogeneity is nonetheless satisfied. I take it that Cohen would not balk at making this commitment explicit, since he endorses a similar claim about the contrast between (17a) and (17b).

- (17) a. Mammals have a placenta. (Homogeneity satisfied)
- b. Mammals are placental mammals. (Homogeneity violated)

Regarding these examples, Cohen suggests that when we use the predicate *have a placenta* in (17a), we do not make salient a partition that separates the mammals into placental mammals and marsupials. By contrast, we make precisely that partition salient in (17b). Hence, there is a difference in partitions made salient by the predicate, *even though* the predicates that appear in the two sentences are necessarily coextensive.

The concern about this proposal is that it apparently predicts that generics with absolute gradable predicates should systematically have weak readings. In using relative and absolute predicates, we refer to points on a scale. In neither case do we make use of an explicit number word to do so. We should therefore predict that for both kinds of gradable predicates, homogeneity is satisfied. But the unacceptability of the examples in (15) shows that it is not.

This is not a knock-down argument. Rather, it is a challenge to explain what the difference between (14a) and (16a) is in such a way that we can see why the examples in (16) pattern with the non-gradable predicates, rather than the gradable ones. What I've argued for so far is that the most straightforward way of chasing down the details of Cohen's system won't work.

Suppose that a proponent of Cohen's system can find a satisfactory way of drawing the relevant distinctions. She still faces another problem that is independent of concerns about homogeneity: an appeal to relative generics

doesn't get the truth-conditions of generics with relative gradable predicates right.

On the relative generic strategy, the truth-conditions are weakened as compared to absolute generics by requiring that fewer members of the kind have to satisfy a contextually given standard. Rather than it being most, it is enough that relatively more than among other kinds do. That means that the truth of a generic with a relative gradable predicate is insensitive to what the facts are with regards to those members of a kind that fall below the contextually determined standard. Let me put this in terms of the example (5), *Dutchmen are good sailors*. So long as relatively more Dutchmen exceed the contextually determined standard of sailing skill than do members of another kind, the sentence is predicted to be true. It simply doesn't matter how good or bad the Dutchmen are that fall below the standard.

This prediction is not borne out. Suppose that there are some very, very good Dutch sailors, with the kind of sailing skill required to have a maritime colonial empire in the Seventeenth Century. Suppose further that there are more such excellent sailors, relative to the population as a whole, than there are French and German sailors (the relevant alternatives). Suppose finally that all the other Dutch sailors are simply abysmal. In that situation, the generic (5) is false. To help with the intuition, it may be useful to consider (18).

(18) Americans are good athletes.

This generic is clearly false, since the wide swath of Americans are horrible athletes, leading an extremely sedentary lifestyle and consuming a poor diet. The fact that there are also quite a few athletic Americans, the kind who regularly exercise a lot, take part in races or competitions, etc., is insufficient to make (18) true.

The fact that (5) is false in the polarized situation I just described shows that the Dutchmen that aren't particularly excellent are relevant to determining the truth-value of the generic. But the proponent of the relative generic strategy cannot account for this fact, so long as she accepts that a single standard of comparison is relevant in the evaluation of (5). For in that case, she faces two competing demands. On the one hand, she needs to make the standard low enough so that (5) is not predicted to have implausibly strong truth-conditions. On the other hand, she needs to make the standard high

enough so that (5) is not predicted to have too weak truth-conditions, ones that are satisfied if all the Dutch sailors are merely middling.

Let's take these in turn. The only way to make the not-excellent sailors relevant to the truth of (5) is to require that they, too, exceed the contextually salient standard, whatever that may be. If that standard is quite high, then the truth of (5) now requires not just that relatively more Dutch sailors exceed the high standard than do sailors from other nations, but that there are also very many such excellent Dutch sailors. That would make the truth-conditions intuitively too strong. In order to solve this problem, the proponent of the relative generic strategy might evaluate (5) with a relatively low standard of sailing skill. In that case, the truth-conditions might not be exceptionally strong, but they threaten to be too weak. If the standard is sufficiently low so as to not force the interpretation that (5) is true only if an implausibly large number of Dutch sailors are excellent, then (5) is true if a sufficiently large number of Dutch sailors exceed that low standard. And that can be the case even if all Dutch sailors are merely mediocre, i.e., if they just barely exceed the standard. And in that situation, (5) is intuitively false.

### 3.3 *Where We Are*

In the discussion, some empirical desiderata have emerged that I summarize here.

- [E1] Weak readings are systematically available for relative gradable predicates.
- [E2] Weak readings are not systematically available for absolute gradable predicates.
- [E3] Weak readings are not systematically available for non-gradable predicates.
- [E4] The weak truth-conditions of generics with relative gradable predicates are sensitive to facts about members of the kind throughout the distribution of the property.

There are also at least two salient theoretical considerations that speak in favor of an account of the data. First, it's always preferable to posit fewer generic

operators rather than more, so that an account that makes do with a single, strong generic quantifier is preferable to one that requires more. Second, the discussion of the Port-Royal puzzle suggests that we should not account for the data simply by positing contextually determined quantifier domain restriction.

## 4 Binding the Standard

I said at the beginning of the last section that the problem raised by (5) depends on a semantic hypothesis, the hypothesis that a quantified claim  $[Qx: Fx](\text{Good.Sailor}(x))$  is true iff  $Q$  many  $F$ s all possess a degree of sailing proficiency that exceeds the contextually given one. But crucially, that assumption is false, as the following example illustrates.<sup>14</sup>

(19) Everyone in my family is tall.

This sentence has a reading, indeed its most prominent one, on which every member of my family is tall *for the kind of person they are*. Thus, I might be tall for an adult male, my daughter might be tall for a three-year-old, and so on. The sentence does not require that all of us exceed a single, contextually determined standard. The argument for this conclusion is completely parallel to the argument I discussed to show that the truth of (5) is sensitive not only to facts about the sailors that exceed the contextually salient standard required for the evaluation of a sentence of the form *x is a good sailor* (where  $x$  is replaced by a name for a particular person), but also the sailors that fall below that standard. Clearly, (19) is false if I am tall for an adult male, but the rest of my family is short for the kind of person they are. If we wanted to maintain the semantic principle I just mentioned, we would have to analyze (19) in terms of a contextually salient standard that my three-year-old can (and does in fact) exceed. But if we did that, we would predict that the sentence is true, no matter how tall I am, since even a very short adult male will exceed the standard of being tall for a three-year-old (about three-and-a-half feet). So we need to analyze the sentence in such a way that I have to meet a standard that an adult male can reasonably fail to meet, while my daughter has to meet a

---

<sup>14</sup>The example is from Kennedy (2007). The general point that the operative standard used in the evaluation of a gradable adjective can be bound by a higher quantifier is also discussed in Ludlow (1989); Stanley (2002), though the three accounts differ in the particular implementation.

standard that she can reasonably meet. The only way to do this is to analyze the sentence in such a way that we have to meet different standards.

To say that the standard with respect to which an element in the domain of the quantifier is evaluated is bound by the quantifier is not to say that there is no role for context to play. Context can still determine, for each person the quantifier ranges over, just how exacting the standard is. If we're in a conversation discussing the possibility that my offspring has a future as a professional basketball player, we might not count as tall. The crucial point is only that different standards are used as we evaluate different members of the quantifier's domain. We can capture this idea formally by assuming that the context provides a function from members of the domain to standards with respect to which they're evaluated. A rough representation of (19) might be (20).

$$(20) [\forall x: \text{Member.Of.My.Family}(x)] (\text{deg}_{\text{Height}}(x) > f_i(x))$$

In words: each member of my family  $x$  is such that  $x$ 's height-degree exceeds the height-degree the context returns for  $x$ .

I now want to argue that we can have an empirically and theoretically attractive treatment of (5) by directly transferring this theory to generics. That is, (5) is given the semantic representation (21).

$$(21) [\text{GEN}(x): \text{Dutch.Sailor}(x)] (\text{deg}_{\text{Good.Sailor}}(x) > f_i(x))$$

In words: GEN-many Dutch sailors  $x$  are such that  $x$ 's sailing skill exceeds a standard of skill appropriate to  $x$ . The proposal itself is extremely simple, and it does not depend on any particular view of how the generic quantifier is interpreted. All that matters is the general point that quantifiers can bind standards of comparison for the evaluation of gradable predicates that appear in their scope. The key to defending the proposal is therefore to see how well it deals with various problem cases, and how well it does at predicting the generalizations I mentioned in §3.3.

#### 4.1 *Fleshing Out The Account: Objections and Replies*

Before turning to the theoretical benefits of my account, I want to address a number of specific empirical concerns one may have about it. This will also

give me a chance to highlight some of the ways in which we can explain the behavior of sentences like (5) as an interaction effect between the semantics of gradability and the semantics of generics.

First, one might worry that my proposal places undue strain on the context and the background information available to speakers who use (5). To make this problem vivid, let me compare (5) with (19). In the case of (19), it was relatively clear what feature of each of the objects the quantifier ranged over determined the output of the function: being an adult male, an adult female, a three-year-old child, and so on. We should ask what the analogous features are that determine the standard a given Dutch sailor in the scope of GEN has to satisfy in order for (5) to be true. This, in fact, can be turned into an objection: the proposal is implausible because it requires the context to fix on a function that returns a value for a large number of people with unknown features. How could context possibly do that?

There is a feature that is readily available in the context however, that can serve the purpose, and that leads to the proper results. The standard each Dutch sailor has to satisfy in order for (5) to be true is determined by his or her position in the distribution of sailing skill. Suppose, for a minute, that we're considering a relatively rough way of carving up the population of Dutch sailors in the scope of GEN into three pools, the first consisting of the sailors at the top of the distribution, the second consisting of the sailors in the middle, and the third consisting of the sailors at the bottom. We can then interpret (5) as saying that the top Dutch sailors (in the scope of GEN) are good *for being at the top of the country's distribution of sailing skill*, that the middling Dutch sailors are good for being in the middle of the distribution, and that the bottom Dutch sailors are good considering that they're at the bottom. In principle, the distribution of sailing skill could be carved up more finely, but the general strategy remains.

This way of determining the function from elements in the domain of GEN to degrees of comparison has an immediate benefit, since it predicts immediately that a generic can go from true to false if the top of the distribution vanishes. Perhaps the Dutch have abolished the resources to apprentice young sailors to truly excel, or an economic change has occurred to siphon off the top talent. In that case, the sailing skill of those that remain is no different than before, but everybody moves up in the distribution. That is to say, the

sailors who used to be in the middle of the distribution now make up the top, and hence have to satisfy a more stringent standard, and so on down the line.

One might further worry that the truth-conditions I predict for (5) are too weak for all that I have said here.

I predict that (5) is true so long as within each subdivision according to skill, the Dutch sailors in that subdivision reach a skill level that exceeds the standard appropriate for that subdivision. It might appear that the truth-conditions do not require anything in particular about the size of the subdivisions relative to each other. That means that my truth-conditions are insensitive to whether the good Dutch sailors make up only a very small proportion of the Dutch sailors as a whole. Thus, the concern continues, my account predicts that (5) is true even if there are only very few good Dutch sailors. But intuitively, the sentence is false in that kind of situation.

The force of the objection depends on the fact that the situation is underspecified, and filling out the description of the case in different ways leads to different intuitive truth-value judgments. First way of filling out: the few good Dutch sailors are outliers. It is somehow a blip or an accident that there are some exceptional Dutch sailors (much the way that it would be an accident if a Swiss sailor won an important race today). In that case, (5) strikes me as false, but my account predicts that it is false, too. That's because we're only considering what is true generally—or however else one wishes to interpret the generic quantifier GEN. It is a common observation about generics that a generic is not made true by the mere fact that there are *some* members of the kind at issue that conform to the generalization: *ravens are white* is not true, even though there are some albino ravens. Something else is required, though different theories of generics differ over what that something else is. Perhaps what's required is not just that there are some members of the kind, but that members of the kind like them exist over a sufficiently long stretch of time, or across the right set of possible worlds, or that they are normal in some relevant way. Once we take into account this extra requirement, we predict somewhat stronger truth-conditions for generics like (5): not only do there have to be some members at the top of the distribution who are good sailors considering that they're at the top, this also needs to hold generally (in the sense to be spelled out by a theory of generics). To return to the example at hand, if the best sailors are all outliers or for some other reason disqualified

from the scope of the *GENERIC*, the top of the skill distribution relevant for evaluating the truth-value of (5) is not made up of the few outliers, but of the middling sailors. And they do not exceed the relevant level of skill.

Second way of filling out: generally (in a sense to be spelled out by a theory of generics), there are lots of quite good Dutch sailors, it just so happens that right now, there aren't very many—perhaps there was a war in which many lost their lives. In that case, the few remaining good Dutch sailors do qualify for being in the scope of *GEN*, and hence the top of the skill distribution is made up of them, not their less skillful compatriots. The relatively small size of the top of the distribution is a temporary aberration. And in that case, (5) strikes me as true, as predicted.

Finally, one might worry that I predict truth-conditions that are too strong, not too weak, as the following two examples are supposed to illustrate.

- (22) a. Brazilians are good soccer players.
- b. Cubans are good baseball players.

There are two ways of cashing out the argument, one directly the semantic, the other epistemic. The semantic argument suggests that the truth of (22a) depends only on facts about the very top of the soccer skill distribution, that is, on the fact that the Brazilian national team is usually among the very best in the world, and that Brazilian players are among the most dominant players in club soccer.

But I doubt that this is really how our intuitions go on this example. It's important to distinguish a correct observation about this example from a false conclusion one might draw from that observation. The correct observation is that (there are at least some very natural contexts in which) it is *necessary* for the truth of (22a) that the top Brazilian soccer players are world class. That is the kind of context in which the corresponding *Americans are good soccer players* is false, simply on the grounds that the best American players aren't as good, even though they are very good players indeed. But from that, it does not follow that it is also *sufficient* for the truth of (22a) that the top Brazilian soccer players are world class, and considering a (counterfactual) situation in which soccer is very unpopular in Brazil, but where there is a specifically designed academy aimed at producing twenty or so top players for the World Cup and club play shows that this is so. In this kind of situation, I take it that (22a)

is clearly false. And my account can well account for the true observation that the top of the distribution has to be absolutely exceptional. All that is required is that the soccer skill distribution is carved up extremely finely, and that the standard in play for the top of the distribution is very high indeed, high enough that the top US players do not meet that standard. Thus, I do not think that a straightforward semantic challenge to my proposal that takes off from the examples in (22) succeeds.

However, one might mount a slightly more indirect challenge that focuses on epistemic considerations. One might think that (22) is problematic for my view, since we are happy to accept it even if all we know about Brazilian soccer players is that the best ones are the best soccer players in the world. It might therefore seem as if we can reasonably accept (22) without knowing how well Brazilians play who aren't great soccer players, and thus without knowing whether the truth-conditions I predict for (22) are satisfied. However, this seems to be a case in which we take the skill of the internationally known players to be representative of the country as a whole.

The fact that we often draw inferences from a small sample about a larger population also helps to account for a different phenomenon, exhibited by (22b), *Cubans are good baseball players*. We might accept it on the basis of the consideration that there are an awful lot of good baseball players considering how small the country is. This reasoning, of course, is precisely the reasoning we would predict to lead to accepting (22b) if Cohen's relative generic account of the sentence was correct. After all, this bit of reasoning simply says that the pool of Cuban baseball players is quite small, making the number of good players relatively large. Since my account makes the truth of (22b) depend not just on the relative plenty of good baseball players, but on players throughout the skill distribution, it might seem as if the acceptance of (22b) on the basis of the evidence I just mentioned is unjustified. If true, that would count against the theory I've put forward.

However, my truth-conditions can also accept this reasoning. Suppose we believe that the distribution of skill among Cuban baseball players follows roughly a normal distribution. Given that assumption, the claim that there's a relatively large number of good players entails that the distribution is shifted towards the highly-skilled side of the spectrum as compared to other nations, so that my truth-conditions are satisfied, as well. So for both (22) and (22b),

my semantics can accommodate the reasonableness of accepting the sentences on the basis of evidence that initially seemed too weak to warrant such acceptance.

#### 4.2 *Generalizations and Theoretical Benefits*

My analysis derives the distinctively weak truth-conditions of (5) by exploiting the following aspect of the interpretation of the predicate ‘good sailor.’ The operative standard with respect to which that predicate is interpreted in a context can be bound by some higher quantifier, specifically, the generic operator GEN. My account therefore predicts that such distinctively weak truth-conditions are available in those generics that contain a predicate that is interpreted with respect to an operative standard *that can be bound by a higher quantifier*. That is true of all relative gradable predicates, and this immediately predicts that E1 is true: weak truth-conditions are systematically available for generics with relative gradable predicates.

The same feature explains why weak truth-conditions aren’t systematically available for either generics with absolute gradable predicates or with non-gradable predicates, i.e., it explains why E2 and E3 hold. This is more immediately obvious for adjectives like *sterile*, *cost \$150,000*, or *live to be 95 years old*. These adjectives aren’t interpreted with respect to a standard at all, and hence cannot possibly have a reading where an operative standard is bound by a quantifier. Hence, the mechanism I identified for relative gradable predicates cannot get started here. If that is the only general semantic mechanism that allows weak readings, we have an explanation for why we don’t systematically find them here.

Similarly, though absolute gradable predicates such as *empty* or *yellow* are interpreted with respect to a standard of comparison, that standard does not vary from context to context. And for that reason, the standard cannot be bound by a higher quantifier. Just as in the case of non-gradable predicates, the mechanism for generating weak readings cannot get started. And if it is the only systematic mechanism for generating weak truth-conditions, we have an explanation for why we don’t systematically find weak truth-conditions with absolute gradable predicates.

These predictions about when weak truth-conditions are available and when they aren’t depend on no particular assumptions about how to inter-

pret generics. They also do not rely on any mechanism by which the restrictor of the generic quantifier is enriched with information available in the context, a mechanism that would be ill-understood and most likely underconstrained. Rather, the predictions are derived from completely general principles about the interaction between quantification and various kinds of predicates. Theoretically, this account is therefore the most parsimonious possible.<sup>15</sup>

## 5 Conclusion

I have argued that one of the persisting problem cases for quantificational accounts of generics, *Dutchmen are good sailors*, can be addressed using only the barest resources to account for generics—a strong interpretation of the generic quantifier—and exploiting how other elements of the sentence interact with quantification. Along the way, I’ve tried to establish and explain some generalizations about when weak-truth conditions are systematically available.

I began this paper by suggesting that generics lie at the intersection of our linguistic capacities and some other aspect of cognition, and that by learning about generics, we might be able to learn about this other aspect, as well. If the account I’ve offered in this paper is successful, we have a better idea of where to look. We can learn the most about the other system or systems involved in generics by looking at where the purely linguistic mechanisms that can account for their interpretation seem to get these interpretations wrong. In those cases, the interpretation must be more of an interaction effect between language and the other system. I’ve argued that the weak readings that generics with relative gradable predicates are not the place to look. They can be accounted for using purely linguistic resources. The best evidence for this claim is that they are completely systematic in the interpretations they make available. A better place to look are the examples where weak readings aren’t systematically available. These promise to hold more evidence.

---

<sup>15</sup>There is another theoretical benefit that goes beyond the scope of the present paper. The account I have presented here very naturally extends to explicitly comparative generics such as *girls do better than boys in grade school* or *dogs are bigger than cats*. In a companion piece to this paper, Nickel (2009), I argue that other proposals cannot make a similar claim, along with giving a positive theory that makes use of the same mechanisms I introduce here.

## References

- Arnauld, A. and P. Nicole. *Logic or the Art of Thinking* (Cambridge, UK: Cambridge UP, 1996).
- Asher, N. and M. Morreau. "What Some Generic Sentences Mean". In G. N. Carlson and F. J. Pelletier, eds., *The Generic Book*, 300–339 (Chicago: University of Chicago Press, 1995).
- Bartsch, R. and T. Vennemann. *Semantics Structures: A Study in the Relation between Syntax and Semantics* (Frankfurt am Main: Athäneum Verlag, 1973).
- Bierwisch, M. "The Semantics of Gradation". In M. Bierwisch and E. Lang, eds., *Dimensional Adjectives*, 71–262 (Berlin: Springer, 1989).
- Carlson, G. N. *Reference to Kinds in English*. Ph.D. thesis, University of Massachusetts, Amherst (1977).
- Cohen, A. "Generics, Frequency Adverbs, and Probability". *Linguistics and Philosophy*, 22, (1999a), 221–253.
- . *Think Generic!* (Stanford, CA: CSLI Publications, 1999b).
- . "Relative Readings of *Many*, *Often*, and Generics". *Natural Language Semantics*, 9(1), (2001), 41–67.
- . "Generics and Mental Representation". *Linguistics and Philosophy*, 27(5), (2004), 529–556.
- Cresswell, M. "The Semantics of Degree". In B. Partee, ed., *Montague Grammar*, 262–292 (New York: Academic Press, 1976).
- Geurts, B. "Generics". *Journal of Semantics*, 4, (1985), 247–255.
- Heim, I. "Degree Operators and Scope". In B. Jackson and T. Matthews, eds., *Proceedings of Semantics and Linguistic Theory (SALT) 10*, 40–64 (Ithaca, NY: CLC Publication, 2000).
- Horn, L. R. *On the Semantic Properties of Logical Operators in English*. Ph.D. thesis, University of California, Los Angeles (1972).
- Kennedy, C. *Projecting the Adjective* (New York: Garland, 1999).
- . "Vagueness and Grammar: The Semantics of Relative and Absolute Gradable Adjectives". *Linguistics and Philosophy*, 30(1), (2007), 1–45.
- Kennedy, C. and L. McNally. "Scale Structure and Semantic Typology of Gradable Predicates". *Language*, 81(2), (2005), 345–381.
- Klein, E. "Comparatives". In A. v. Stechow and D. Wunderlich, eds., *Semantik: Ein Internationales Handbuch der Zeitgenössischen Forschung*, 673–691 (Berlin: Walter de Gruyter, 1991).

- Krifka, M., F. J. Pelletier, G. N. Carlson, A. ter Meulen, G. Chierchia, and G. Link. "Genericity: An Introduction". In G. N. Carlson and F. J. Pelletier, eds., *The Generic Book*, 1–124 (Chicago: University of Chicago Press, 1995).
- Larson, R. "Events and Modification in Nominals". In D. Strolovitch and A. Lawson, eds., *Proceedings from Semantics and Linguistic Theory (SALT) 8*, 145–168 (Ithaca, NY: Cornell UP, 1998).
- Leslie, S.-J. "Generics and the Structure of the Mind". In J. Hawthorne, ed., *Philosophical Perspectives 21, Philosophy of Mind*, 375–403 (Malden, MA: Blackwell, 2007).
- . "Generics: Cognition and Acquisition". *Philosophical Review*, 117(1), (2008), 1–47.
- Ludlow, P. "Implicit Comparison Classes". *Linguistics and Philosophy*, 12(4), (1989), 519–533.
- Nickel, B. "Generic Comparisons" (2009). Ms., Harvard University.
- Rusiecki, J. *On Adjectives and Comparisons in English* (New York, NY: Longman Linguistics Library, 1985).
- Schubert, L. K. and F. J. Pelletier. "Generically Speaking, or, Using Discourse Representation Theory to Interpret Generics". In G. Chierchia, B. H. Partee, and R. Turner, eds., *Properties, Types, and Meaning, Vol. II*, 193–268 (Dordrecht: Kluwer Academic Publishers, 1989).
- Seuren, P. A. "The Comparative". In F. Kiefer and N. Ruwet, eds., *Generative Grammar in Europe*, 528–564 (Dordrecht: Reidel, 1973).
- Stanley, J. C. "Nominal Restriction". In G. Preyer and G. Peter, eds., *Logical Form and Language*, 365–388 (Oxford: Oxford UP, 2002).
- Unger, P. *Ignorance* (Oxford: Clarendon Press, 1975).
- von Stechow, A. "Comparing Semantic Theories of Comparison". *Journal of Semantics*, 3(1-2), (1984), 1–77.
- Wilkinson, K. *Studies in the Semantics of Generic Noun Phrases*. Ph.D. thesis, University of Massachusetts, Amherst (1991).