

COMPARISON TO ARBITRARY STANDARDS*

STEPHANIE SOLT

Zentrum für Allgemeine Sprachwissenschaft, Berlin

1 Introduction

This paper is concerned with a set of contrasts in the occurrence of degree modifiers that I will argue relate to restrictions on the sort of degree statements that can be made relative to a standard whose location is in some way arbitrary. Along the way I will explore some implications for the relationship between standards and scale structure.

2 *Slightly* and the nature of distributional standards

2.1 The data

As has been noted by several authors, ‘low degree’ modifiers (LDMs) such as *slightly* and *a bit* occur felicitously with some gradable adjectives, but not others (Rotstein and Winter 2004; Kennedy 2007; Bylinina 2011):

- (1) a. The rug is slightly/a little/a bit dirty.
b. The towel is slightly/a little/a bit wet.
c. That neighborhood is slightly/a little/a bit dangerous.
- (2) a. ??John is slightly/a little/a bit tall.
b. ??The lake is slightly/a little/a bit deep.
c. ??The rod is slightly/a little/a bit long.
- (3) a. ??The rug is slightly/a little/a bit clean.
b. ??The towel is slightly/a little/a bit dry.
c. ??That neighborhood is slightly/a little/a bit safe.

It is not that the adjectives in (2) and (3) do not combine with degree modifiers at all. Those of the *tall* class (2), like those of the *dirty* class (1), are felicitously modified by *very*, while those of the *clean* class (3) occur with *completely* and *perfectly*.

*I would like to thank the reviewers for Sinn und Bedeutung 16, and also the conference audience, for helpful and thought-provoking comments. Support for this research was provided by the European Science Foundation and the Deutsche Forschungsgemeinschaft under the auspices of the EUROCORES programme LogICCC.

- (4) a. The rug is very dirty.
 b. The towel very wet.
 c. That neighborhood is very dangerous.
- (5) a. John is very tall.
 b. The lake is very deep.
 c. The rod is very long.
- (6) a. The rug is completely/perfectly clean.
 b. The towel is completely/perfectly dry.
 c. That neighborhood is completely/perfectly safe.

Nor is it that the dimensions invoked by the adjectives in (2), (3) are in some way incompatible with LDMs: the comparative forms of all three adjective classes occur felicitously with these modifiers.

- (7) a. The rug is slightly dirtier than the floor.
 b. The towel slightly wetter than the rag.
 c. That neighborhood slightly more dangerous than this one.
- (8) a. John is slightly taller than Fred.
 b. The lake is slightly deeper than the pond.
 c. The rod is slightly longer than the box.
- (9) a. The rug is slightly cleaner than the floor.
 b. The towel slightly drier than the rag.
 c. That neighborhood slightly safer than this one.

Interestingly, the examples in (2) do have a reading on which they are acceptable, namely when the adjective is coerced to an excessive reading, such that *slightly tall*, for example, is interpreted as ‘slightly too tall (for the present purposes)’. The same is true, though less consistently, for the examples in (3). The following are corpus examples involving this kind of interpretation:

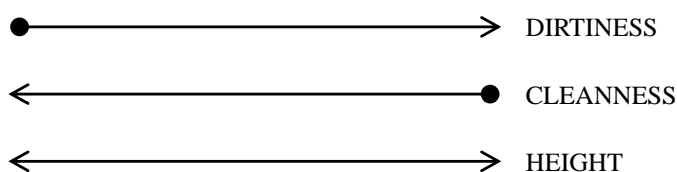
- (10) a. Our rig holds Kriss, P90, and UMP mags perfectly. MP5 mags will sit a little deep, but we can fit pockets with an internal spacer to sit them at the perfect height. (<http://www.originalsoegear.com/smgrig.html>)
 b. However, if you end up with a less-than-perfect joint, you can cope with this situation by recutting the joint (you did cut the board slightly long, right?). (http://www.woodbin.com/misc/cope_molding.htm)
 c. You can make these up to 5 days ahead, and they reheat beautifully. If they’re a bit dry, just stir in water. (Prevention, 61(11), p. 154, 2009; sourced from COCA, Davies 2008-)

Furthermore, this pattern is not limited to English. Kagan and Alexeyenko (2010) report a similar set of facts for the Russian degree modifier *-ovat* ‘slightly’, which occurs felicitously with adjectives of the *dirty* class (e.g. *grijaznovatyj* ‘slightly dirty’), yields an excessive interpretation for those of the *tall* class (e.g. *vysokovatyj* ‘slightly too tall’), and is either infelicitous or has an excessive reading for those of the *clean* class (e.g. **cistovatyj* ‘slightly (too) clean’ but *suxovatyj* ‘slightly too dry’).

2.2 Previous accounts

The three classes of adjectives exemplified in (1)-(3) are now standardly taken to correspond to three distinct possibilities for the structure of the measurement scale lexicalized by the adjective (see especially Kennedy and McNally 2005). Those of the *dirty* class relate entities to scales closed at the lower end (i.e. having a minimum point). Those of the *clean* class have scales that are open on the lower end but closed on the upper end (i.e. that have a maximum point). Finally, adjectives of the *tall* class have scales that are open at the lower as well as upper end.

(11)

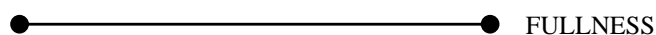


But as yet this has not provided a full explanation for the contrasts exemplified in (1)-(3).

Kennedy (2007) takes occurrence *slightly* to be a diagnostic for the presence of a scalar minimum point. Taken together with *perfectly*, which serves as a diagnostic for the presence of a scalar maximum point, this provides the motivation for the typology depicted in (11). On Kennedy's theory, scale structure in turn determines the standard of comparison for the adjective in its positive form. If the scale has a maximum and/or minimum point, that serves as the standard; only when the scale is totally open (as with the *tall* class) is the standard contextually determined (the principle of interpretive economy).

Kennedy does not, however, address what it is about the semantics of *slightly* that makes it sensitive to this aspect of scale structure. More problematically, his account makes incorrect predictions about adjectives such as *full/empty* and *open/closed*, which lexicalize scales that are closed on both ends (as diagnosed by the compatibility of *perfectly* with both the adjective and its antonym).

(12)



If the occurrence of *slightly* is sensitive simply to the presence of a scalar minimum point, this class should be as felicitous with *slightly* as those of the *dirty* class. But actually, examples such as the following are marginal:

(13) ??The bottle is slightly full / empty

This difference is supported by corpus data. Table 1 provides data from the Corpus of Contemporary American English (COCA; Davies 2008-), showing the frequency of occurrence of the degree modifiers *slightly*, *a little* and *a bit* with adjectives assigned to these four classes. As expected, adjectives that have been diagnosed as having lower closed scales have the highest occurrence with LDMS, while those with upper closed or open scales occur more rarely with these modifiers. But importantly, adjectives with scales closed on both ends occur no more frequently

Table 1: Frequency of LDMs (Source: COCA, 1990-2011)

Scale structure	Adjectives	<i>slightly/a bit/a little Adj</i> per '000 Adj
Lower closed	<i>bent, bumpy, crooked, dangerous, dirty, impure, incomplete, rough, uncertain, wet</i>	11.2
Upper closed	<i>certain, clean, complete, dry, flat, pure, safe, smooth, straight</i>	0.5
Open	<i>big, deep, expensive, fat, large, long, rich, strong, tall, thick, wide, cheap, inexpensive, narrow, poor, shallow, short, small, thin, weak</i>	1.3
Totally closed	<i>open, closed, full, empty, opaque, transparent</i>	1.0

with LDMs than do the latter two classes. Thus the presence of a scalar minimum point is not, in itself, sufficient to license the felicitous occurrence of *slightly* and the like.

In work that predates Kennedy's, Rotstein and Winter (2004) relate contrasts in the acceptability of *slightly* to the nature of the standard of comparison for the positive form of the adjective, in interaction with the semantics of *slightly* itself. Rotstein and Winter propose that the denotation of *slightly Adj* corresponds to an interval of some arbitrary length at the beginning of the denotation of *Adj* on the relevant scale. In the case of adjectives such as *dirty* which are associated with lower-closed scales, the standard for the positive form is a point at or above the scalar minimum S_{\min} , such that *slightly Adj* is associated with the interval immediately above this. Slightly *dirty*, for example, is true of an entity whose degree of dirtiness is just above the relevant minimum point. But in the case of adjectives with upper closed scales, such as *clean*, the standard for the positive form is in the default case the scalar maximum S_{\max} . Since there is no interval above this point, *slightly* is infelicitous in this case.

While this explanation nicely captures the contrast between (1) and (3), Rotstein and Winter's account has nothing to say about the infelicity of *slightly* with adjectives of the *tall* class, as in (2). Following a standard view, they take these to involve a contextually derived standard d_{mid} located somewhere in the middle of the scale. It is then not clear why *slightly Adj* in these cases could not correspond to the scalar interval immediately above this point, such that slightly *tall*, for example, would be true of individuals whose heights are just a small amount above the contextually determined standard for what counts as *tall*.

Furthermore, neither Rotstein & Winter nor Kennedy have an explanation for the occurrence of adjectives of the *tall* and *clean* classes with LDMs when they are coerced to an excessive reading.

2.3 Proposal

The account I develop builds on that of Rotstein and Winter, and on Kagan and Alexeyenko's analysis of Russian *-ovat*. I propose that the occurrence of *slightly* and other low degree modifiers is sensitive not to the structure of the scale, but to the nature of the standard of comparison for the positive form. Contra Kennedy, I propose that scalar maximum and minimum points, if they are present, may provide a standard of comparison, but do not necessarily do so. I further will argue that there are more types of standards than are typically assumed in work in this area. Some but not all are compatible with the semantics of *slightly*.

For concreteness, I adopt the framework of Kennedy (2007), in which gradable adjectives denote measure functions, that is, functions from individuals to degrees. For example:

$$(14) \quad \llbracket \text{dirty}_{\langle \text{ed} \rangle} \rrbracket = \lambda x. \mu_{\text{DIRTINESS}}(x)$$

Degree modifiers take gradable adjectives as arguments and return predicates over individuals. In particular, I assume the following for *slightly*, based on Rotstein and Winter (2004). For the present I assume other LDMs have similar semantics, though it is possible that a more in-depth investigation would turn up differences.

$$(15) \quad \llbracket \text{slightly} \rrbracket = \lambda P_{\langle \text{ed} \rangle} \lambda x_e. d_{\text{Std}} < P(x) < d_{\text{Std}} + \delta, \text{ for some small } \delta$$

As defined here, the interpretation of *slightly* is dependent on the standard d_{Std} for the positive form of the adjective. I follow Kagan and Alexeyenko (2010) in assuming at least the following four types of standards:

- (16) a. **Scalar maximum:** $d_{\text{Std}} = S_{\text{max}}$
 b. **Scalar minimum:** $d_{\text{Std}} = S_{\text{min}}$
 c. **Distributional:** $d_{\text{Std}} = f(C)$
 d. **Functional:** $d_{\text{Std}} = \max \{d : \exists w_{\text{Acc}} [P(x)^w = d]\}$

The first two of these are the familiar scalar minimum and maximum assumed by Kennedy; together they can be classified as absolute standards. A distributional standard is a standard determined with reference to the distribution of items in a comparison class C . Distributional standards correspond to the usual reading of adjectives of the *tall* class. For example, *John is tall* is typically interpreted as conveying that John's height exceeds the typical height for members of some set John is a member of; such a set can be made overt with a *for*-phrase (*John is tall for a jockey*). I return to a fuller discussion of distributional standards below. Finally, functional standards – a term introduced by Kagan and Alexeyenko (2010) – correspond to the maximum degree compatible with the requirements of a given situation. Functional standards are similar to those invoked by the overt excessive modifier *too*; for example, *John is too tall* means that his height exceeds the maximum acceptable for some purpose. But the positive form can also reference such a standard; for example, the most natural reading of (17) is that the fourth board was too long for the needs of the situation.

(17) Three of the boards were cut to exactly the right length, but the fourth one was long

See also Bylinina (2011) for a discussion of *for* phrases with functional standards.

I further assume a version of Kennedy's principle of interpretive economy, such that when an absolute standard is available for an adjective, it is the default choice. But I take this principle to be violable, in that contextual or other factors can result in the choice of a different standard even in the case of a scale with a maximum and/or minimum point.

Let us examine how these various standards are compatible with the semantics of *slightly*.

Lower closed scales (*dirty* class): When a scalar minimum point is available to serve as a standard for an adjective *Adj* (per (16a)), *slightly Adj* is associated with a short segment of the scale above this minimum. Thus *slightly dirty* is true of an individual if it has some small positive degree of dirtiness. This is intuitively correct.

$$(18) \quad \llbracket \textit{slightly dirty} \rrbracket = \lambda x_e. d_{\text{Std}} < \mu_{\text{DIRTINESS}}(x) < d_{\text{Std}} + \delta$$

$$d_{\text{Std}} = S_{\text{min}} \quad \mathbf{OK}$$

Upper closed scales (*clean class*): By contrast, when a scale has a maximum point, this does not provide a standard compatible with *slightly*. The reason is that if this point were chosen as the standard (per (16b)), *slightly Adj* would relate entities to a range above this point; but there is of course no range of degrees above the maximum point on the scale.

$$(19) \quad \llbracket \textit{slightly clean} \rrbracket = \lambda x_e. d_{\text{Std}} < \mu_{\text{CLEANNESS}}(x) < d_{\text{Std}} + \delta$$

$$d_{\text{Std}} = S_{\text{max}} \quad \mathbf{BAD}$$

Open scales (*tall class*): In the case of adjectives such as *tall* which correspond to totally open scales, there are no maximum nor minimum points to potentially serve as standards. The infelicity of LDMs with such adjectives on their usual interpretations could be explained if we could show that distributional standards (16c) are in some way incompatible with the semantics of *slightly*. I take this up below. But there is another type of standard that is available in such cases, namely the functional standard (16d). When this standard is plugged into the entry for *slightly*, *slightly tall* is interpreted as a predicate true of an individual if his height is greater by some small amount than the maximum height compatible with the needs of the situation. This corresponds exactly to the felicitous reading of the examples in (2), and to the interpretation of the corpus examples in (10a,b).

$$(20) \quad \llbracket \textit{slightly tall} \rrbracket = \lambda x_e. d_{\text{Std}} < \mu_{\text{HEIGHT}}(x) < d_{\text{Std}} + \delta$$

$$d_{\text{Std}} = f(C) \quad \mathbf{BAD}$$

$$d_{\text{Std}} = \max \{d : \exists w_{\text{Acc}} [\mu_{\text{HEIGHT}}(x)^w = d]\} \quad \mathbf{OK}$$

Note that the choice of a functional rather than scalar maximum standard may also potentially rescue *slightly Adj* with an adjective of the *clean class*. As pointed out by Kagan and Alexeyeko for *-ovat*, whether this is felicitous depends on the meaning of the adjective. For example, it is difficult to imagine a situation in which there is a maximum acceptable degree of cleanliness short of completely clean, and as such *?slightly clean* is awkward. But *slightly dry* on a functional reading is not bad at all, since one can imagine a situation in which there is a maximal acceptable degree of dryness. Such an interpretation seems to be what is involved in (10c).

Totally closed scales (*full class*): Finally, consider adjectives that lexicalize scales closed on both ends, such as *full/empty* and *open/closed*. On the account developed here, it is predicted that they will be acceptable with *slightly* only to the extent that they can be interpreted relative to a standard other than the scalar maximum S_{max} . Kennedy (2007) notes that such adjectives typically prefer maximum standards; this explains the infrequent occurrence of this class with low degree modifiers reported in Table 1. But the present account suggests two other types of standards are possible, namely scalar minimum and functional. Corpus examples such as the following (from COCA) suggest that these are in fact the sorts of standards involved when an items of this class occur with LDMs. (21a-c) seem to involve scalar minimum standards; in (21b), for example, the door deviates slightly from a position of fully closed (i.e. having a zero degree of openness). (21d) on the other hands seems to involve something closer to a functional standard, namely the adequate or expected level of stocking of the shelves.

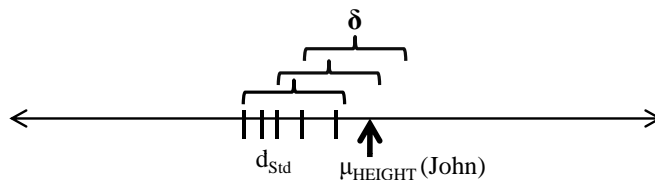
- (21) a. He'd lean his head back, his eyes slightly closed . . . (Ploughshares, Winter97/98, 23/4, p.12)
 b. Helene . . . had been in the bathroom, door cracked. slightly open, peeking out through the small gap (Analog Science Fiction & Fact, 122/10, p. 108)
 c. As soon as the rice is hot and slightly transparent, add about 1 cup of warmed wine. (Chicago Sun Times, 14/2/1999)
 d. The Safeway store . . . hadn't received a truck delivery, and the shelves were looking a little empty. (Denver Post, 22/6/2006)

We are left with the crucial question of why distributional standards should be incompatible with LDMs. What sets such standards apart from scalar minimum or functional standards, such that they do not support the interpretation of *slightly*?

I would like to propose that the crucial property of distributional standards is that their location is essentially arbitrary. The other standards in (16) can be reduced to a single non-arbitrary point – an endpoint in the case of scalar maximum and minimum standards, a maximum value compatible with the needs of the situation in the case of functional standards. But this is not so with distributional standards. There is no natural dividing line that separates those entities that count as tall in a given context from those that do not count as tall. This is the central insight behind several theories of vagueness, which has been developed in various ways, e.g through contextualism (Raffman 1996) and supervaluationism (Kamp 1975). While different in their specifics, these approaches agree on the idea that one way of setting a standard is (within some range) as good as another. To put this in terms of comparison classes, Solt (2011) develops an analysis where the comparison class provides a standard in the form of a range calculated as a function of the degree of dispersion in that class. Here arbitrariness reduces to the choice of the comparison class and the width of the range, expressed as a multiple of the standard deviation.

Arbitrariness of standard is problematic for LDMs in particular, because their semantics essentially require precision of measurement. *Slightly* as defined in (15) relates entities to a scalar interval within a small distance δ of the standard. But small distances cannot be measured accurately from a standard whose location is arbitrary in the way described above: on one perfectly acceptable way of setting d_{Std} an entity's measure might be within δ of that standard, while on another equally acceptable choice, its measure might be more than δ beyond d_{Std} , or conversely below d_{Std} . In simple terms, if we cannot say with precision where the lower bound for *tall* is, we cannot say whether an individual's height is within a small distance from that standard. Visually:

- (22) John is slightly tall



As evidence that this is relevant, when the arbitrariness of the standard is reduced, adjectives of the *tall* class become more acceptable with LDMs. For example, (23) is considerably more felicitous than the original (2a) on the distributional reading:

- (23) John is slightly/a little/a bit tall for his age.

What is different here is that the comparison class has been fixed (to boys of a certain age, perhaps in a particular region), and furthermore that we can assume the existence of data that allows a standard to be set relative to this population in some non-arbitrary way, e.g. as the median value.

Note also that the arbitrariness of distributional standards does not affect ‘large degree’ modifiers in the same way. Suppose for example that *very Adj* is true of an individual iff its measure exceeds the standard for *Adj* by a large amount γ . Even if there is indeterminacy in the precise location of d_{Std} (in that there are multiple equally acceptable possibilities), as long as an individual’s measure is large enough we can be sure that that measure exceeds d_{Std} by at least γ . Hence *very* is compatible with distributional standards, and thus may occur with adjectives such as *tall* on their usual meanings (per (5)).

2.4 Summary and further observations

Facts relating to the distribution of LDMs such as *slightly* point to a more complex view of scales and standards than is commonly assumed. In addition to absolute standards, we’ve seen evidence for two types of contextual standards – distributional and functional – which behave quite differently with respect to one aspect of meaning, namely compatibility with *slightly*. Furthermore, even in the case where the scale features a maximum and/or minimum point, these do not necessarily provide the standard of comparison. And it is entirely possible that there are further standard types beyond those discussed here. For example, (24a) appears to involve deviation from an ideal midpoint as standard, whereas (24b) seems to involve a standard based on societal norms (whether this is the same as a functional standard would require further investigation):

- (24) a. John is slightly overweight/underweight.
 b. Costa was in his late forties, and the years showed around his middle but not in the thick dark hair that he left cut just slightly long. (Linda Grant, *Love nor Money*, 2002; from COCA)

While it is beyond the scope of this paper to explore this topic more generally, I’ll suggest here that better characterizing the full array of available standard types has the potential to resolve some open questions (e.g. relating to the apparently imprecise use of *The theater is empty tonight* to describe a situation where there are fewer attendees than usual; Kennedy 2007).

There is a further implication for scale structure. I have argued here that *slightly* is sensitive to the nature of the standard of comparison rather than the structure of the measurement scale, and further that scalar endpoints may – but do not necessarily – provide the standard. The corollary to this is that incompatibility with *slightly* cannot be used as a diagnostic for the absence of a scalar minimum point. This is particularly relevant for adjectives of the *tall* class. On the basis of the *slightly* test, Kennedy (2007) concludes, somewhat unintuitively, that such adjectives have scales that are open on the lower end. But an alternate explanation is that such scales have a minimum point, but that this is for some reason unavailable to serve as the standard of comparison for the positive form. Recent work by Toledo and Sassoon (2011) gives a reason to think this might in fact be the case. They point out that adjectives of the *tall* class typically correspond to individual level predicates, while those of the *dirty/clean* classes denote stage level predicates. An entity may, in one of its stages, have a zero level of dirtiness; but no entity of which *tall* could be predicated can have a zero level of tallness. Correspondingly, the scalar minimum point represents a meaningful standard in the first case (*dirty*), but in the second case (*tall*) would be entirely uninformative, since

every relevant individual would have a measure exceeding this minimum.¹ This analysis, if it could be further supported, would result in a simpler and more intuitive view of scale structure.

To return to the main topic of the present paper, facts relating to LDMs suggest that distributional standards are different from other types of standards. I have suggested that the heart of the difference is their essential arbitrariness. While this is by no means a new observation, being the starting point for well known theories of vagueness, the facts discussed here point to another consequence of this, namely incompatibility with degree modifiers whose semantics are based on precise measurement. Large distances can be measured from a standard whose location is imprecisely determined or arbitrary, but small distances cannot. In the next section, I will turn to data from another class of degree modifiers, which I will argue are sensitive to a different sort of arbitrariness.

3 (Not) *exactly*

3.1 The data

The data below illustrate a contrast that on the surface is unrelated to that discussed in the preceding section. Specifically, while *exactly* can occur in positive sentences with point-denoting expressions (25), with vague predicates *exactly* is restricted to occurring in the scope of negation, and is disallowed in the corresponding positive sentences (26).

- (25) a. John is exactly 30 years old.
 b. It's exactly noon.
 c. The circle is exactly in the center of the square.
- (26) a. Red wine is *(not) exactly healthy.
 b. John is *(not) exactly tall.
 c. That's *(not) exactly a mountain.

Sauerland and Stateva (2007) propose that *exactly* is ambiguous between the 'scalar *exactly*' represented in (25) and a separate negative polarity item, 'NPI *exactly*', which is involved in examples such as (26). As evidence, they note that German *genau* and Bulgarian *točno* 'exactly' do not occur with predicates such as *healthy* in either positive or negative contexts (27), nor apparently does English *precisely*, whose meaning is very close to that of *exactly* (28).

- (27) *Der Rotwein ist (nicht) genau gesund.
 (28) ??Red wine is (not) precisely healthy.

This seems to suggest that the data in (25)-(26) reflect an idiosyncrasy of one lexical item.

But one can in fact find corpus examples of *not precisely* with predicates of this sort (29). And while *genau* patterns differently from *exactly*, German has another counterpart of *exactly*, *gerade*, which exhibits the same asymmetry as the English item (30). Hebrew, too, has an equivalent of *exactly* which occurs with point-denoting expressions in positive contexts, but is restricted to negative contexts with vague predicates, just like *exactly* (31).

¹Sassoon (2011) independently reaches a similar conclusion, based on data very similar to those discussed here.

- (29) Delivery pizza is fun but not precisely healthy.
(<http://chiaseedshealthbenefits.com/how-bake-chia-flour-without-any-extra-work/>)
- (30) Der Rotwein ist *(nicht) gerade gesund.
- (31) a. bediyuk esrim anashim
'exactly twenty people'
b. yai'n adom ze *(lo) bedii'uk bari
'Red wine is *(not) exactly healthy'

Thus there seems to be a more general pattern in need of explanation.

3.2 Towards a solution

Note first that *not exactly* has two sorts of readings that might be termed 'literal' and 'strengthened'. In (32), the claim seems to be that the cookies are not, in the strictest sense of the word, healthy, but also not entirely unhealthy (the literal reading). (33), by contrast, implies something stronger, namely that cheeseburgers are quite unhealthy indeed (the strengthened reading).

- (32) Sure, they're not exactly healthy, but they're much healthier than your typical cookie, have a soft but slightly chewy texture that I adore, and they are vegan
(<http://catesworldkitchen.com/2010/09/maple-almond-butter-cookies/>)
- (33) Being comfort food, grilled cheese burgers are not exactly healthy. And that's okay, you're not eating them everyday. So you'll want a side dish that's equally comforting ... and unhealthy. (<http://www.squidoo.com/Grilled-Cheese-Burger>)

The strengthening observed in examples such as (33) is part of a broader pattern in which sequences of the form *not+INTENSIFIER+Adj* take on a strengthened meaning (Horn 1989). For example, all of the variants of (34) imply that Fred is quite dumb.

- (34) Fred is not exceedingly / very / too / overly bright.

But the intensifiers that participate in this pattern are not consistently NPIs, and as such this parallel does not in itself provide insight into why this *exactly* is restricted to negative contexts. In any case, I take the literal reading to be the one to be accounted for in the semantics, with the strengthened reading derived via a pragmatic process that is also involved in examples such as (34).

The cross-linguistic data discussed above suggest that something about the semantics of *exactly* (and its counterparts in other languages) predisposes it to develop a second interpretation as an NPI. This would not be exceptional. It is well known that minimizers such as *a bit*, *a drop*, *a wink*, *a red cent* and so forth have a tendency cross-linguistically to evolve into NPIs (Horn 1989). But importantly, NPI *exactly* is not a minimizer. By way of illustration, *a drop* on its NPI interpretation can be understood as a minimal quantity of liquid. As such, *not drink a drop* is a stronger claim than alternatives such as *not drink a glass/a bottle/etc.*; that is, the entailment in (35a) holds. But the equivalent entailment in (35b) does not hold; if anything, the entailment is in the opposite direction. In other words, *not exactly Adj* makes a weaker claim than potential alternatives.

- (35) a. Sue didn't drink a drop → Sue didn't drink a glass / a bottle / a lot / etc.
b. Red wine isn't exactly healthy → Red wine isn't healthy at all / in the slightest / etc.

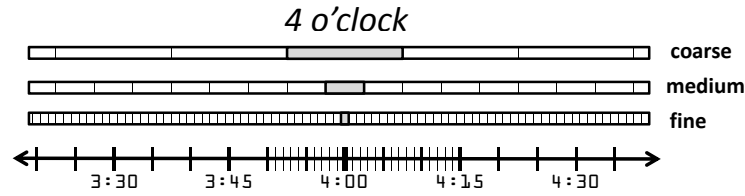


Figure 1: Multiple granularity levels

This difference has as a consequence that analyses developed for NPI minimizers (e.g. Krifka 1995) cannot readily be extended to NPI *exactly*.

In a taxonomy of polarity items developed by Israel (1996), minimizers of the *(drink) a drop* type fall into the class of emphatic NPIs, characterized by low quantitative value and emphatic informative value. Israel does not consider NPI *exactly*, but its behavior allows it be assigned to the opposing class of understating NPIs, characterized by high scalar value and understating informative value. Other members of this class are claimed to include *much* and *all that*. For example, *Sue didn't drink much* makes a weaker statement than *Sue didn't drink anything / a glass / a drop* (cf. (35b)). Israel proposes that items of this class conventionally encode understating informative value, and thus are restricted to negative contexts where the resulting statement is minimally informative. This view is however somewhat problematic in the case of *exactly*, since on its scalar use *exactly* isn't in any obvious way understating. Thus here, too, we find no explanation for the cross-linguistic tendency for items with the meaning *exactly* to acquire a second use as NPIs.

3.3 Proposal

I propose that NPI *exactly* derives from scalar *exactly*, and has the same basic meaning.

Following Sauerland and Stateva (2007), I take scalar *exactly* to play the role of regulating the granularity level at which scalar expressions are interpreted. I assume granularity to be a contextual parameter of interpretation. Any context is associated with multiple possible granularity levels, which may be associated with granularity functions. For example, the expression *4 o'clock* might be interpreted in any of the ways depicted in Figure 1: as a 15-minute interval around the point 4:00 (coarse granularity), as a 5-minute interval around that point (medium granularity), or as a 1-minute interval around that point (fine granularity).

The function of *exactly* is to set granularity to the finest contextually available level:

$$(36) \quad \llbracket \textit{exactly} A \rrbracket = \textit{gran}_{\textit{finest}}(A)$$

Applied to the example in Figure 1, the effect is to set the granularity level to the finest of the three levels, namely the 1-minute granularity level, such that *exactly 4 o'clock* is interpreted as the 1-minute interval around 4:00.

$$(37) \quad \llbracket \textit{exactly} 4 \textit{ o'clock} \rrbracket = \textit{gran}_{\textit{finest}}(4 \textit{ o'clock}) = 4 : 00 \pm 30 \textit{ sec}$$

Looked at slightly differently, scalar *exactly* can be viewed as operating on alternate interpretations of a scalar predicate that are ordered via an inclusion relationship. For example, the scalar interval denoted by *4 o'clock* at the finest granularity level is a proper subset of the

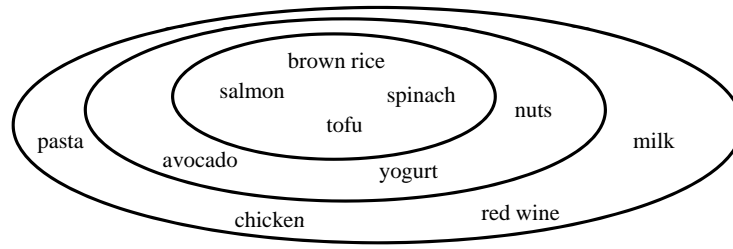


Figure 2: Alternate interpretations of *healthy*

interval denoted by the same expression at the medium granularity level, and so forth. *Scalar exactly* picks the maximum element in this ordering, i.e. the extension contained within all other contextually available extensions.

It is this that allows us to draw a connection to NPI *exactly*. Specifically, I propose that *exactly* in examples such as (26) operates over alternate interpretations of a vague predicate ordered in terms of strictness, which also may be construed as an inclusion relationship. To take a concrete example, the predicate *healthy* (applied to foods) might be interpreted in the various ways depicted in Figure 2. Here red wine, for example, is included in a looser extension of *healthy*, but not in stricter extensions.

To formalize this, I assume that vague predicates are interpreted relative to a contextual parameter that specifies, among other things, how vague predicates are made precise (this could be made compatible with various analyses of vagueness, e.g. supervaluationism per Kamp (1975)). As was the case with granularity, I assume that a given context may allow multiple possible interpretations of a predicate. For example, a particular context *C* might allow the following alternate interpretations of *healthy*:

- (38) a. $\llbracket \textit{healthy} \rrbracket^{c1} = \{\textit{tofu}, \textit{salmon}, \textit{spinach}\}$
 b. $\llbracket \textit{healthy} \rrbracket^{c2} = \{\textit{tofu}, \textit{salmon}, \textit{spinach}, \textit{brown rice}\}$
 c. $\llbracket \textit{healthy} \rrbracket^{c3} = \{\textit{tofu}, \textit{spinach}, \textit{nuts}\}$
 d. $\llbracket \textit{healthy} \rrbracket^{c4} = \{\textit{tofu}, \textit{salmon}, \textit{spinach}, \textit{brown rice}, \textit{avocado}, \textit{nuts}, \textit{yogurt}\}$; etc.

Here too, the alternate extensions may be ordered via an inclusion relation:

$$(39) \llbracket \textit{healthy} \rrbracket^c >_{\text{strict}} \llbracket \textit{healthy} \rrbracket^{c'} \text{ iff } \llbracket \textit{healthy} \rrbracket^c \subset \llbracket \textit{healthy} \rrbracket^{c'}$$

The ordering derived in this way is only a partial order, in that there may be elements *A* and *B* such that neither $A >_{\text{strict}} B$ nor $B >_{\text{strict}} A$; this is the case, for example, for (38a) and (38c) above.

As evidence that variation of this sort is relevant to the semantics of NPI *exactly*, note that *exactly* is infelicitous with predicates that are not vague, i.e. that do not allow stricter and looser interpretations. For example, (40) is quite odd, and while (41a) is acceptable (the point at which a computer might be considered dead being open to interpretation), (41b) is much less so.

- (40) ??Sue isn't exactly pregnant.
 (41) a. My laptop isn't exactly dead (but fixing it will cost more than buying a new one).
 b. ?Mr. Jones isn't exactly dead.

Building on this, I propose that the function of NPI *exactly* is to pick out the strictest interpretation of a predicate in the context, namely the one which is a proper subset of all other contextually available alternate interpretations.

$$(42) \quad \llbracket \textit{exactly healthy} \rrbracket^C = \llbracket \textit{healthy} \rrbracket^{c(\textit{strictest})} \\ = \iota P. P \in \{ \llbracket \textit{healthy} \rrbracket^c : c \in C \} \wedge \forall Q \in \{ \llbracket \textit{healthy} \rrbracket^c : c \in C \} [P \neq Q \rightarrow P >_{\textit{strict}} Q]$$

The parallel to scalar *exactly* should be apparent.

But why should the definition of *exactly* in (42) restrict its occurrence to negative contexts? The answer, I propose, is that there is no strictest interpretation for a predicate such as *healthy*. Recall that the ordering imposed on alternate extensions is only partial, and as such there is no guarantee that one possible extension will be stricter than all others. This can also be put in more conceptual terms. Whether something counts as *healthy* depends on a variety of potentially overlapping criteria (e.g. fat content, sugar content, etc.). The choice of two different sets of criteria could result in two distinct extensions, neither of which was a proper subset of the other. Determining a single strictest interpretation would require selecting some particular set of criteria, a choice that would be essentially arbitrary.

The result is that *exactly healthy* is undefined. But in spite of this, *not exactly healthy* is still assertable. For something (say, red wine) to be not exactly healthy, it must simply be the case that there is some extension of *healthy* that is narrow enough to exclude it. There is no need to identify the single strictest interpretation. That is, *not exactly healthy* can be interpreted as ‘outside of the strictest interpretation of healthy (whatever that may be)’.

The example that I have been considering up to this point, *healthy*, is not associated with a unidimensional scale. The situation is slightly different for a predicate such as *tall*, which has such a scale. In this case, the ordering of alternate extensions is presumably at least a weak order, in that for any *A* and *B* it must be the case that either $A \geq_{\textit{strict}} B$ or $B \geq_{\textit{strict}} A$. The reason is that as the standard for *tall* is set higher, the resulting extension will be a (possibly proper) subset of preceding ones; the situation exemplified by (38a,c) will never arise. But here too, the maximum element in the ordering is undefined. However strict we set the standard for what counts as *tall*, we can always imagine a stricter standard. Thus the predicate *exactly tall* is undefined, but an individual may nonetheless be *not exactly tall*. To relate this to the discussion in the preceding section, it was noted there that in any context, the choice of a standard for a predicate such as *tall* is essentially arbitrary. To this we can add that in any context, the choice of the strictest possible standard is likewise arbitrary.

Finally, consider again scalar *exactly*. Here, too, it is true that no matter how fine the granularity level, we can typically imagine yet a finer level (for example, *4 o'clock* could be interpreted not just at the 1-minute level, but also more precisely at the 10-second level, and so forth). But in this case there is, at least in principle, a non-arbitrary maximally strict interpretation, namely the point itself; the same does not hold for predicates such as *tall* and *healthy*. Thus just as in the case of *slightly*, the felicitous and infelicitous uses of *exactly* in positive contexts are distinguished by the existence of a non-arbitrary single point on which the interpretation of a predicate can be based.

4 Concluding observations

In this paper, I have discussed two types of arbitrariness in interpretation, both of which have implications for how predicative expressions can compose with degree modifiers. Large distances,

but not small distances, can be measured from a scalar standard whose location is arbitrary (*slightly*). Somewhat similarly, it is possible to say what is not contained within an essentially arbitrary maximally strict predicate interpretation (*not exactly*) but not what falls within it (*exactly*).

One wonders if this pattern might be broader. A further possibility involves the lexical item *much*, which like *exactly* has both NPI and non-NPI uses. What is interesting is that arbitrariness seems to play a role in its NPI use as well. For example, the partitive *Sue lost much of her money in the stock market* is better than the corresponding non-partitive *??Sue lost much money in the stock market*. Correspondingly, the standard for what counts *much of her money* is more readily fixed than what counts as *much money*. I leave this as a topic for future investigation.

References

- Bylinina, Elizaveta. 2011. Functional standards. In *ESSLLI 2011 Student Session*, 46–57.
- Davies, Mark. 2008-. The Corpus of Contemporary American English (COCA): 400+ million words, 1990-present. available online at <http://americancorpus.org>.
- Horn, Laurence R. 1989. *A natural history of negation*. Chicago: University of Chicago Press.
- Israel, Michael. 1996. Polarity sensitivity as lexical semantics. *Linguistics and Philosophy* 19:619–666.
- Kagan, Olga, and Sascha Alexeyenko. 2010. Degree modification in Russian morphology: The case of the suffix *-ovat*. In *Proceedings of Israel Association for Theoretical Linguistics 26*, ed. Yehuda N. Falk. Bar Ilan University.
- Kamp, Hans. 1975. Two theories about adjectives. In *Formal semantics of natural language*, ed. Edward L. Keenan, 121–155. Cambridge: Cambridge University Press.
- Kennedy, Christopher. 2007. Vagueness and grammar: The semantics of relative and absolute gradable predicates. *Linguistics and Philosophy* 30:1–45.
- Kennedy, Christopher, and Louise McNally. 2005. Scale structure and the semantic typology of gradable predicates. *Language* 81:1–37.
- Krifka, Manfred. 1995. The semantics and pragmatics of polarity items. *Linguistic Analysis* 25:1–49.
- Raffman, Diana. 1996. Vagueness and context relativity. *Philosophical Studies* 81:175–192.
- Rotstein, Carmen, and Yoad Winter. 2004. Total adjectives vs. partial adjectives: Scale structure and higher order modifiers. *Natural Language Semantics* 12:259–288.
- Sassoon, Galit. 2011. A slightly modified economy principle: Stable properties have non-stable standards. Talk at Workshop on Degree Semantics and its Interfaces, Utrecht.
- Sauerland, Uli, and Penka Stateva. 2007. Scalar vs. epistemic vagueness. In *Proceedings of Semantics and Linguistic Theory (SALT) 17*, ed. Masayuki Gibson and Tova Friedman. Ithaca, NY: CLC Publications.
- Solt, Stephanie. 2011. Notes on the comparison class. In *Vagueness in communication (ViC2009), revised selected papers (LNAI 6517)*, ed. Rick Nouwen, Robert van Rooij, Uli Sauerland, and Hans-Christian Schmitz, 189–206. Berlin, Heidelberg: Springer.
- Toledo, Assaf, and Galit W. Sassoon. 2011. Absolute vs. relative adjectives - variance within vs. between individuals. In *Proceedings of Semantics and Linguistic Theory 21*, ed. Neil Ashton, Anca Chereches, and David Lutz, 135–154. New Brunswick, NJ: Rutgers University.