

## The Sorites, Content Fixing, and the Roots of Paradox\*

Mario Gómez-Torrente

Instituto de Investigaciones Filosóficas, Universidad Nacional Autónoma de México

In earlier work (Gómez-Torrente 2010), I have proposed a theory of reference fixing for typical degree adjectives and scalar nouns,<sup>1</sup> together with an accompanying solution of sorites paradoxes involving those adjectives and nouns. According to the theory, the default mechanism of reference fixing for typical degree adjectives and scalar nouns works via the implicit contextual postulation of “preconceptions”, designed to single out extensions for these words relative to occasions of use. Sometimes, and in particular in occasions of use where the intended universe of discourse contains a relevant sorites series, the mechanism of preconceptions doesn’t manage to single out an extension for the relevant sorites-susceptible word. In such cases, and in particular in cases where a sorites reasoning involving the relevant sorites-susceptible word is *prima facie* compelling, the word in fact lacks an extension, and (utterances of) sentences containing it fail to have truth conditions, including the corresponding sentences used in the relevant sorites reasoning. However, the theory is not a “nihilist” theory of vague language; according to it, in a vast number of typical occasions of use of degree adjectives and scalar nouns, and in particular in cases where the objects in the universe of discourse cluster around the positive and negative paradigms of the relevant adjective or noun, this does get an extension, and (utterances of) sentences containing it do have truth conditions. That typical degree adjectives and scalar nouns get extensions and sentences containing them get truth conditions in typical occasions of use is, in fact, an element in the theory’s explanation of the intelligibility of vague language and of how communication with (utterances of) sentences containing vague predicates is possible; an explanation of this kind is unavailable to nihilist theories. In the mentioned earlier work, I also proposed that the central aspects of the theory can provide the basis for an understanding and solution of sorites paradoxes concerning predicates other than those involving typical degree adjectives and scalar nouns (if these are real paradoxes). Because of its postulate that occasions of use of sorites-

---

\* I thank Eugenia Allier and an anonymous referee for very helpful comments. Research was supported by the Mexican CONACyT (CCB 2011 166502) and by the Spanish MICINN and MINECO (research projects FFI2008-04263 and FFI2011-25626).

<sup>1</sup> A scalar noun is one that has an analytically associated dimension of comparison, usually also analytically connected with a degree adjective—a youngster, for example, is a *young* person (or animal).

susceptible words divide into those where an extension gets fixed and those in which none does, I called the theory the *dual* picture of vagueness.

In this paper I will supplement the presentation of the dual theory in my earlier work with a number of additional considerations, sometimes proposed as refinements, sometimes as rejoinders to potential objections to the theory, sometimes as arguments that the theory compares favorably to other theories in certain respects. In section 1 I will recall the main aspects of the dual picture with a bit more detail, laying more emphasis than in my earlier work on how the picture lends itself naturally to treatments of the contribution of a typical degree adjective to propositional content and to truth conditions. In section 2, I will present a number of reasonable refinements of the picture, especially concerning occasions of use of a degree adjective in which a class containing a sorites series is somehow involved in content fixing, but in such a way that an extension ultimately gets fixed for the predicate. In section 3, I will offer a new consideration in favor of the dual theory as compared with standard theories of the sorites. Standard theories of the logical and semantic paradoxes arguably rely on the antecedently plausible idea that the content-fixing mechanisms for the paradoxical expressions work well in general, but their functioning meets certain unexpected scope restrictions. Standard theories of the sorites paradox are clearly not based on this idea at all, but I will note that the dual theory, by contrast, is fully based on the idea. In section 4, I will offer a few remarks on some recent theories of vagueness that are broadly in the spirit of the dual picture, emphasizing the aspects where my picture compares favorably against those recent theories. In section 5 I will consider, with more detail than in my earlier work, the question whether the theory is stable with truth according to the theory itself, or equivalently the question of whether in the relevant theoretical occasions of use of the vague predicates that the theory employs, these predicates manage to get extensions.

## 1. Reviewing the dual picture of vagueness

In a typical formulation, we have an instance of the sorites when, in a particular occasion of use, we are confronted with a group of sentences of the following form, each of which seems highly compelling in that occasion of use:

- (Sorites Chain Schema)  $(\exists x_1) \dots (\exists x_n) ([Kx_1 \ \& \ Kx_2 \ \& \ \dots \ \& \ Kx_n \ \&] \ Rax_1 \ \& \ Rx_1x_2 \ \& \ \dots \ \& \ Rx_{n-1}x_n \ \& \ Rx_nz)$ ;  
 (Positive Paradigms Schema)  $[Ka \ \&] \ Fa$ ;  
 (Tolerance Schema)  $(\forall x)(\forall y)([Kx \ \& \ Ky \ \supset] \ (Fx \ \& \ Rxy \ \supset \ Fy))$ ;  
 (Negative Paradigms Schema)  $[Kz \ \&] \ \sim Fz$

(cf. Fara 2000, pp. 49f.). “F” is to be replaced with the sorites-susceptible predicate, “a” with a name of a case of application of the predicate which is clear in the occasion of use, “z” with a name of a case of negative application which is clear in the occasion of use, and “R” with a name of some binary relation. The occasion of use provides a *universe of discourse* (and range for the quantifiers) that contains the things named by the names replacing “a” and “z”. In some cases, the occasion of use provides a *contrast class*, named by a sign replacing “K”, which is the class over which the sorites-susceptible predicate is intended to be defined in the occasion of use; the contrast class may coincide with the universe of discourse, but in some cases will be properly included in it.<sup>2</sup> (The brackets indicate that the bracketed parts will appear in instances of the paradox involving a contrast class.) The contrast class may also coincide or not coincide with the *comparison class* of the context; in some contexts, this will be a class *C* that helps determine membership in the extension of the sorites-susceptible predicate *P* via the intuitive condition “is *P* compared to the members of *C*”. (For more on the distinction between the contrast class and the comparison class, see section 2 below.)<sup>3</sup>

The paradox is, of course, that every set of fully interpreted utterances of instances of the Sorites Chain Schema, the Positive Paradigms Schema, the Tolerance Schema and the Negative Paradigms Schema (whether we include the bracketed parts or not) is inconsistent

---

<sup>2</sup> Thus, if in an occasion of use we are talking both about some persons and about natural numbers, we don’t intend the predicate “is young” to be defined over the numbers, but normally only over the persons (or a subset of the persons). However, while the contrast class is the class over which a predicate is intended to be defined, classical logic requires that all predicates be defined over the whole universe of discourse. We will later assume, not unreasonably but perhaps somewhat artificially, that atomic sorites-susceptible predicates negatively apply to things in the universe of discourse from outside the contrast class (see note 12 below).

<sup>3</sup> There are of course other formulations of the paradox, in some of which there appear no sentences containing quantifiers. For example, instead of an instance of the Sorites Chain Schema we may be faced with a compelling sentence of the form “Rab<sub>1</sub> & Rb<sub>1</sub>b<sub>2</sub> & ... & Rb<sub>n-1</sub>b<sub>n</sub> & Rb<sub>n</sub>z”, where “b<sub>1</sub>”, “b<sub>2</sub>”, etc. are replaced by names; and instead of an instance of the (Tolerance Schema) we may be faced with a series of compelling conditional sentences of the forms “Fa & Rab<sub>1</sub> ⊃ Fb<sub>1</sub>”, “Fb<sub>1</sub> & Rb<sub>1</sub>b<sub>2</sub> ⊃ Fb<sub>2</sub>”, ..., “Fb<sub>n-1</sub> & Rb<sub>n-1</sub>b<sub>n</sub> ⊃ Fb<sub>n</sub>”, “Fb<sub>n</sub> & Rb<sub>n</sub>z ⊃ Fz”. In this formulation the universe of discourse in the occasion of use must contain, besides the referents of the names replacing “a” and “z”, those of the names replacing “b<sub>1</sub>”, “b<sub>2</sub>”, etc. It may also be good to note that the universe of discourse need not consist exclusively of existing objects, and may contain (or even consist exclusively of) individuals that are merely possible, imagined, postulated for the sake of some argument, etc. The basic considerations of the dual picture apply to these other formulations without substantive changes.

according to classical semantics and logic, and yet many such sets are highly compelling in their corresponding occasions of use. For example, “F” may be replaced with “is young”, “R” with “was born one day later than”, “a” with “Alvar”, the name of a living person born in 2011, and “z” with “Zellig”, the name of a living person born in 1931; the relevant contrast class (the interpretation of the replacement for “K”) may be taken to be the set of living persons, and the universe of discourse may be any set that includes the set of living persons, e.g. the set of living persons itself. Then the corresponding natural instance of the Sorites Chain Schema is presumably true, and the corresponding instances of the Positive Paradigms Schema, the Tolerance Schema and the Negative Paradigms Schema will all appear highly compelling to speakers in many normal occasions of use involving that contrast class,<sup>4</sup> even though all these instances are jointly inconsistent.

The core of the dual picture is a mechanism of reference fixing conjectured to underlie uses of typical degree adjectives and scalar nouns,<sup>5</sup> a mechanism that delivers referents for these words in some occasions of use—called *regular* occasions of use—and

---

<sup>4</sup> It is worth noting that the relevant instance of the Tolerance Schema (“ $(\forall x)(\forall y)(x \text{ is a living person} \ \& \ y \text{ is a living person} \ \supset \ (x \text{ is young} \ \& \ x \text{ was born one day later than } y \ \supset \ y \text{ is young}))$ ”) will presumably not appear compelling after the speaker’s exposure to the existence of a sorites series in the occasion of use and to the paradoxical aspects of the occasion in general. If the speaker has not been exposed to the relevant instance of the Sorites Chain Schema, or he has been exposed to it but has not been exposed to the paradoxical consequences of its truth, then he will presumably typically be strongly disposed to assert “ $(\forall x)(\forall y)(x \text{ is a living person} \ \& \ y \text{ is a living person} \ \supset \ (x \text{ is young} \ \& \ x \text{ was born one day later than } y \ \supset \ y \text{ is young}))$ ”. And he will similarly be strongly disposed to assert particular cases of this, such as *If Jones and Smith are living persons, Jones is young and Jones was born one day later than Smith, Smith is young*. These two kinds of dispositions are to be contrasted with the dispositions of the speaker after he has been exposed to the relevant instance of the Sorites Chain Schema and to its paradoxical consequences; in this case, he will not be (strongly) disposed to assert “ $(\forall x)(\forall y)(x \text{ is a living person} \ \& \ y \text{ is a living person} \ \supset \ (x \text{ is young} \ \& \ x \text{ was born one day later than } y \ \supset \ y \text{ is young}))$ ”, and he will even be on the lookout when presented with particular cases of this generalization. These presumable facts lend support to the dual theory we are about to review, for the theory will claim that occasions where a relevant instance of the Sorites Chain Schema is true make the use of the relevant sorites-susceptible predicate defective. This is one explanation of the fact that the relevant instance of the Tolerance Schema stops being compelling after exposure to paradoxicality. By contrast, the dual theory will claim that, in occasions where no relevant instance of the Sorites Chain Schema is true, the use of the relevant sorites-susceptible predicate is semantically in order. This is in syntony with the presumable fact that, in such occasions, speakers will view the relevant instance of the Tolerance Schema as compelling even after exposure to all the relevant facts.

<sup>5</sup> Henceforth I will omit the parenthetical addition of “and/or scalar noun(s)” to “typical degree adjective(s)” in many cases in which it should be tacitly understood.

fails to deliver them in other—*irregular*—occasions of use, in particular in occasions of use where a relevant instance of the Sorites Chain Schema is true. The most popular standard theories of vagueness and the sorites propose that the logical expressions are governed by some non-classical semantics or logic, and do not trace the correct diagnosis of the sorites to an appropriate conception of reference fixing.<sup>6</sup> By contrast, the standard theories that don't abandon classical semantics or logic do crucially rely on some conception of reference fixing, but the conceptions of reference fixing they rely on are arguably implausible. Thus, epistemicism<sup>7</sup> and some forms of contextualism<sup>8</sup> must postulate that typical degree adjectives somehow get classical extension/anti-extension pairs over any universe of discourse, thus implying that there is always a counterexample to the relevant instance of the Tolerance Schema, even though what these extensions and counterexamples are is something not accessible to linguistic reflection. But while it is surely reasonable to think that the precise nature of the content of some words is not necessarily accessible to mere linguistic reflection, the suggestion that typical degree adjectives get unknowable classical extensions over universes of discourse like that of the preceding paragraph is bizarre. It is fair to say that these theories take too far the idea that the precise nature of content is not always accessible to mere linguistic reflection. On the other hand, weak nihilism,<sup>9</sup> which proposes that atomic vague predicates like “is young” always have an empty extension, has the bizarre consequence that there are never

---

<sup>6</sup> In particular, among these theories which propose that the logical expressions are governed by some non-classical semantics or logic I include the different varieties of supervaluationism, regardless of the extent to which the set of formal logical truths and logical consequences they determine coincides with the set of classical logical truths and consequences. Supervaluationist theories abandon the classical semantics for the logical expressions from the moment they give substantively non-classical truth conditions for sentences dominated by the logical expressions. Thus the supervaluationist says that an existential quantification is true simpliciter if it is classically true with respect to a certain set of classical interpretations (in particular, with respect to a certain set of classical interpretations of a relevant sorites-susceptible predicate). This has the well-known non-classical consequence that an existential quantification can be true in the supervaluationist sense without there being a true instance of it. Note that the supervaluationist rejects in an essential way the project of explaining how a privileged classical extension gets fixed for a sorites-susceptible predicate; instead, only a large set of equally acceptable interpretations play broadly semantic roles for such predicates.

<sup>7</sup> As championed e.g. by Sorensen (1988) and Williamson (1994).

<sup>8</sup> E.g. Fara's in (2000). The contextualist theory of Soames (1999) postulates that a typical degree adjective gets a precise extension/borderline extension/anti-extension triple over any universe of discourse, which is just as implausible as the view about to be described in the main text.

<sup>9</sup> A view suggested by Unger (1979).

semantically correct non-trivial distinctions to be made by means of sorites-susceptible degree adjectives. Finally, strong nihilism or incoherentism<sup>10</sup> relies on a conception of reference fixing opposite to epistemicism in the spectrum of possible theories: the relevant instances of the Positive Paradigms Schema, the Tolerance Schema and the Negative Paradigms Schema are taken to be constitutive of the meaning of the relevant sorites-susceptible predicate, so their inconsistency implies that the predicate is incoherent and lacks a referential content in all occasions of use. In view of these results, it is perhaps no wonder that theories of vagueness and the sorites that retain a classical semantics and logic are unpopular and theories that hypothesize the operation of some non-classical semantics or logic are generally preferred. But for those who find no independent reason to doubt that the logical expressions are in all essential respects governed by classical semantics and logic, like myself, the natural idea to pursue is that reference fixing for sorites-susceptible predicates proceeds in a subtler way than standard theories of vagueness allow.

The dual picture as presented in my earlier work focuses on the fixing of reference, or referential content, for vague predicates, and is not wedded to a particular conception of the (contribution to) propositional content of those predicates. However, the picture easily lends itself to natural ways of conceiving the relevant (contributions to) propositional contents. One basic idea of the picture is that a typical degree adjective, like “young”, has associated with it a quite abstract reference-fixing “instruction” that is meaning-constituting. This instruction is analogous to a Kaplanian character, in that when it is supplemented with some aspects specific to an occasion of use, it yields a more concrete instruction that one might see as the (contribution to) propositional content expressed by the relevant predicate in that occasion of use. These aspects include what I call the “preconceptions” specific to the occasion of use, about which more below. The more concrete instruction, in turn, determines (or fails to determine), in a way that is in principle fully accessible to linguistic reflection, the reference of the predicate in the occasion of use. When it exists, this reference, as assumed classically, is an extension. This extension is determined in a coordinated fashion together with the predicate’s anti-extension, i.e. the predicate’s complement over the universe of discourse. Let me illustrate this reference-fixing structure with examples.

The abstract, meaning-constituting instruction analogous to a Kaplanian character for “young” would look somewhat like this:

---

<sup>10</sup> As defended e.g. by Ludwig and Ray (2002) and Braun and Sider (2007).

In an occasion of use  $O$ ,

- where the universe of discourse is  $U$ ;
- where the relevant contrast class included in  $U$  is  $K$ ;
- where some members  $r_1, r_2$ , etc. of  $K$  are taken as cases of young things<sup>11</sup> and/or some members  $s_1, s_2$ , etc. of  $K$  are taken as negative cases of young things (in such a way that the relevant ages of  $r_1, r_2$ , etc. are lesser than the relevant ages of  $s_1, s_2$ , etc.);
- where some principles of the form “ $(\forall x)(\forall y)(Kx \ \& \ Ky \supset (x \text{ is young} \ \& \ x \text{ has an age inferior by } 1 \ u_i \text{ to } y \supset y \text{ is young}))$ ” are taken as holding (in such a way that the difference between the  $r_j$  with the greatest age and the  $s_k$  with the smallest age is always greater than  $1 \ u_i$ ):

if there is a unique pair  $\langle E, A \rangle$  of subsets of  $U$  which are mutually exclusive and jointly exhaustive over  $U$ , and are such that

- everything in  $K$  with the ages of  $r_1, r_2$ , etc. is in  $E$ ,
- everything in  $K$  that has an age inferior to something in  $E$  is in  $E$ ,
- everything in  $K$  that has an age superior by  $1 \ u_i$  to something in  $E$  is in  $E$ ,
- everything in  $K$  with the ages of  $s_1, s_2$ , etc. is in  $A$ ,
- everything in  $K$  that has an age superior to something in  $A$  is in  $A$ ,
- everything in  $K$  that has an age inferior by  $1 \ u_i$  to something in  $A$  is in  $A$ ,
- $U-K$  is included in  $A$ <sup>12</sup>,

then  $\langle E, A \rangle$  is the extension/anti-extension pair of “is young” relative to  $O$ .<sup>13</sup>

In essence, the abstract instruction for “young” says that, in an occasion of use where the things taken as positive paradigms plus the things that are taken to be tolerably close to

---

<sup>11</sup> When speaking of some things as being “taken as cases of young things” in an occasion of use, or (below) of some instances of the Tolerance Schema as being “taken as holding” in an occasion of use, I mean to stay neutral as to the ultimate sources of these attitudes. In some cases, that the participants in a conversation take a certain thing to be a paradigm of a young person may be caused by the fact that doing so is typical or has been typical in the past; and similarly for taking a certain tolerance level as holding. But in other cases, that the participants in a conversation take a certain thing to be a paradigm of a young person will be caused by idiosyncrasies of the occasion of use having nothing to do with typicality; and similarly for the tolerance levels taken as holding in an occasion of use. (See Gómez-Torrente 2010, p. 238 for some comments on the extreme variability of paradigms and tolerance levels across contexts of use of the same predicate.)

<sup>12</sup> I assume, perhaps somewhat artificially, that everything that is not in the contrast class is not young in the sense relevant to the occasion of use, and thus that it is in the anti-extension of “is young”.

<sup>13</sup> Here, as in all cases where a convention is merely implicit, an explicitation of the convention may not be fully understandable by a typical speaker. (Compare the case of Kaplanian characters and their use of relatively sophisticated concepts.) Also, different explicitations of the convention at work may be recognized as appropriate by different speakers (or even by single speaker) when they consider the question reflectively. The thesis put forward here is that the instruction in the text is a good representative explicitation of a linguistic convention about “young”.

them in age form a set, and the things taken as negative paradigms plus the things that are taken to be tolerably close to them in age form the complement set over the contrast class, the first set is the reference of “is young”.<sup>14</sup>

Now suppose that in a certain occasion of use  $P$  we are talking about Alvar, Zellig and Yannick, where Yannick is a living person born less than one year after Zellig. In  $P$  the set  $U_p = \{\text{Alvar, Zellig, Yannick}\}$  is the universe of discourse, and suppose that the contrast class is also  $U_p = K_p$ . Suppose further that  $P$  is such that Alvar is taken as young and Zellig is taken as not young (as allowed by the meaning-constituting instruction for “young”). Let  $a$  be the age of Alvar and  $z$  be the age of Zellig. Finally, suppose that  $P$  is such that it is accepted at least that one day, one year, and all the time units in between don’t make a difference as to whether someone is young or not. Then the more concrete instruction that one might see as the (contribution to) propositional content expressed by “is young” in the particular occasion of use  $P$  would be as follows:

If there is a unique pair  $\langle E, A \rangle$  of subsets of  $U_p$  which are mutually exclusive and jointly exhaustive over  $U_p$ , and are such that

- everything in  $K_p$  with age  $a$  is in  $E$ ,
- everything in  $K_p$  that has an age inferior to something in  $E$  is in  $E$ ,
- everything in  $K_p$  that has an age superior by one year or less to something in  $E$  is in  $E$ ,
- everything in  $K_p$  with age  $z$  is in  $A$ ,
- everything in  $K_p$  that has an age superior to something in  $A$  is in  $A$ ,
- everything in  $K_p$  that has an age inferior by one year or less to something in  $A$  is in  $A$ ,
- $U_p - K_p$  is included in  $A$ ,

then  $\langle E, A \rangle$  is the extension/anti-extension pair of “is young” relative to  $P$ .

Finally, given that its antecedent holds for  $\langle \{\text{Alvar}\}, \{\text{Zellig, Yannick}\} \rangle$ , this more concrete instruction determines that  $\langle \{\text{Alvar}\}, \{\text{Zellig, Yannick}\} \rangle$  is the extension/anti-extension pair for “is young” in the occasion of use. So  $\{\text{Alvar}\}$  is the reference of “is young” fixed by the (contribution to) propositional content of “is young” in the occasion of use.  $P$  is thus a regular occasion of use of “young”.

---

<sup>14</sup> If the meaning-constituting instruction for “young” is part of the meaning of “young” and the essential component of the way its reference in an occasion of use is fixed, this accords well with the desideratum that the semantics and logic of the logical expressions be classical. Arguably, the classical semantics and logic for the logical expressions makes sense only under the supposition that the predicates with which they interact have extension/anti-extension pairs which are mutually exclusive and jointly exhaustive over the universe of discourse.



By contrast, suppose that we are in a certain occasion of use  $Q$  where the relevant contrast class  $K_Q$  is the set of living persons, and the universe of discourse  $U_Q$  is that same set.  $Q$  is again such that Alvar is taken as young and Zellig is taken as not young; and it is such that it is accepted that one year or less doesn't make a difference as to whether someone is young or not young. Then the more concrete instruction that one might see as the (contribution to) propositional content expressed by "is young" in  $Q$  would be:

If there is a unique pair  $\langle E, A \rangle$  of subsets of  $U_Q$  which are mutually exclusive and jointly exhaustive over  $U_Q$ , and are such that

- everything in  $K_Q$  with age  $a$  is in  $E$ ,
- everything in  $K_Q$  that has an age inferior to something in  $E$  is in  $E$ ,
- everything in  $K_Q$  that has an age superior by one year or less to something in  $E$  is in  $E$ ,
- everything in  $K_Q$  with age  $z$  is in  $A$ ,
- everything in  $K_Q$  that has an age superior to something in  $A$  is in  $A$ ,
- everything in  $K_Q$  that has an age inferior by one year or less to something in  $A$  is in  $A$ ,
- $U_Q - K_Q$  is included in  $A$ ,

then  $\langle E, A \rangle$  is the extension/anti-extension pair of "is young" relative to  $Q$ .

Given that its antecedent doesn't hold for any pair of subsets of  $U_Q$ , this concrete instruction does not determine any pair of sets to be the extension/anti-extension pair for "is young" in the occasion of use  $Q$ . Assuming that the abstract instruction for "young" and its concrete applications are in essence all there is to reference fixing for "is young" in normal occasions of use, it follows that no reference for "is young" is fixed in  $Q$ , and thus also, always under classical assumptions, that no truth conditions are determined for sentences containing "is young" relative to  $Q$ .  $Q$  is an irregular occasion of use of "young".

According to the dual picture, the joint fulfillment of the relevant instances of the Positive Paradigms Schema, the Tolerance Schema and the Negative Paradigms Schema over the contextually intended contrast class is conventionally taken to imply that the relevant adjective applies to the relevant positive paradigms and to those things implied to fall under the adjective by the instances of the Positive Paradigms Schema and the Tolerance Schema; and it is also conventionally taken to imply that the adjective negatively applies to the relevant negative paradigms and to those things implied to fall under the negation of the adjective by the instances of the Negative Paradigms Schema and the Tolerance Schema. But a key idea of the dual picture is that the relevant instances of the Positive Paradigms Schema, the Tolerance Schema and the Negative Paradigms Schema taken as holding in an occasion of use are not constitutive of the meaning of the relevant

sorites-susceptible predicate. And yet, as the incoherentist presumably intuits, they are in a way the essence of the mechanism of reference fixing. That they are not meaning-constituting is independently plausible on account of the fact that epistemicism and typical contextualist theories, as well as weakly nihilist theories, are not intuitively *analytically* false. The relevant instances of the Positive Paradigms Schema, the Tolerance Schema and the Negative Paradigms Schema plausibly have a role in content fixing for typical degree adjectives, but it is not as meaning-constituting or even as reference-fixing via some direct a priori connection with the adjectives. On the dual picture, their role is as *preconceptions*, non-meaning-constituting, non-a priori connected, aspects of the occasion of use that work as auxiliary hypotheses in the mechanism of reference fixing.<sup>15</sup>

Compare the role of what are arguably preconceptions playing a role in reference fixing for natural kind terms. In the case of “gold”, for example, on a plausible Kripkean theory of reference fixing there is a preconception that certain things are paradigms of gold, and a preconception that the paradigms of gold are instances of a certain natural kind. These are not meaning-constituting for “gold”, or even a priori true of gold. But they do play a role in reference fixing, as hypotheses whose fulfillment is conventionally taken to imply that “gold” refers to the natural kind that has as instances the paradigms in question. A key suggestion of the dual picture is that something analogous happens with instances of the Positive Paradigms Schema, the Tolerance Schema and the Negative Paradigms Schema when it comes to reference fixing for a typical degree adjective. Note that the properly conventional or meaning-constituting content of such an adjective according to the dual picture is, as in the natural kind terms case, of hypothetical or conditional form. Note further that, also as in the natural kind terms case, the non-analyticity and non-apriority of the relevant hypotheses need not have been obvious prior to (what I take to be) the proper account of the semantic facts of the matter, and in fact that illusions of analyticity and apriority may have been widespread. If the dual picture’s account of the semantic facts is correct, it is not analytic or *a priori* that (someone with the age of) Alvar is young, just as it is not analytic or *a priori* that a given paradigm of gold is gold (even if there

---

<sup>15</sup> It may be worth insisting that, according to the dual picture, although many instances of the Positive Paradigms Schema, the Tolerance Schema and the Negative Paradigms Schema appear compelling, or are *prima facie* compelling, this is of course compatible with a belief, even a *prima facie* belief, that many instances are false and that many *prima facie* compelling instances may turn out false or truthvalueless. As noted in the text, the relevant instances are not intuitively like instances of a schema such as “If X is a bachelor, then X is unmarried”, which intuitively simply cannot have meaningful but untrue instances.

may have been a widespread illusion that “gold” applies analytically to things with the superficial characteristics of that paradigm).

There are analogies with natural kind terms also in cases of failure of reference. Take the case of “neurasthenia”, for example.<sup>16</sup> There used to be a preconception that certain persons were paradigms of neurasthenics, and a preconception that the paradigms of neurasthenia were instances of a certain natural kind, a certain nervous disease. These preconceptions turned out not to be jointly satisfied, for it has turned out that there are many unrelated diseases behind the syndrome afflicting traditional paradigms. On what seems to be the most reasonable view, there is no fact of the matter as to whether traditional paradigms are after all instances of neurasthenia (conceived perhaps as a name of a purely descriptive syndrome), or there is no such thing as an instance of neurasthenia; thus no extension is fixed for “is an instance of neurasthenia”, as there isn’t even a fact of the matter whether its extension should contain traditional paradigms or not. “Neurasthenia” is in some way defective, and the reason is presumably that its preexisting meaning together with its preexisting associated preconceptions fail to fix an extension for it in the presence of a conflict they were not designed to cope with. The dual picture proposes, analogously, that typical sorites-susceptible degree adjectives fail to have a reference in occasions of use such as  $Q$  above, in which the relevant contextual preconceptions are in conflict with a truth of the form of the Sorites Chain Schema, stating that a sorites series can be drawn from the contrast class. Consequently, utterances of usual sentences containing typical sorites-susceptible degree adjectives in such occasions of use will lack truth conditions.<sup>17</sup> We can ask again: does preexisting usage determine that some of the relevant paradigm preconceptions are false, or that the relevant instances of the tolerance schema are false? (Or is it the case that we have been under some illusion that classical semantics and logic govern our use of the logical expressions when they interact with “young”?) The most reasonable view seems again to be that “young”, as used in paradoxical occasions of use such as  $Q$ , is defective, as its meaning and associated

---

<sup>16</sup> I assume, following the American Psychiatric Association, that there is no single disease behind the syndrome traditionally attributed to neurasthenia.

<sup>17</sup> Perhaps some utterances of sentences containing sorites-susceptible predicates in some irregular occasions of use have truth conditions, e.g. some where the predicates are in the scope of locutions of propositional attitude.

preconceptions are not enough to get an extension for it in the presence of the unexpected sorites conflict.<sup>18</sup>

Unlike what happens with “gold” and natural kind terms in general, the reference of typical degree adjectives is intuitively a function of many contextual elements. Contextually intended paradigms and acceptable tolerance levels seem clearly to play a role in the determination of a contextual reference for a typical degree adjective, and presumably the contextually intended contrast class does as well; in fact, the intended contrast class presumably plays a role in constraining the intended paradigms and the acceptable tolerance levels. The contextual (contribution to) propositional content of a typical degree adjective, therefore, must probably incorporate in some way the values of a set of paradigms along a relevant dimension of variation, a set of acceptable tolerance levels in that dimension, and a contrast class. The picture just sketched satisfies this desideratum by postulating that the instances of the Positive Paradigms Schema and the Negative Paradigms Schema accepted in a context codify (at least some of) the relevant paradigms (and corresponding values in the relevant dimension of variation), and that the instances of the Tolerance Schema accepted in a context codify the set of relevant tolerance levels.

By contrast with the theories of vagueness with which it shares classical semantic and logical assumptions, the dual picture has a reasonable classical explanation of the intelligibility of vague language and of the possibility of communication using sorites-susceptible predicates. Epistemicism and typical contextualist theories must make a mystery of what would have to happen in order for a vague predicate to get a classical extension; they must thus make a mystery of the predicate’s (contribution to) propositional content, and specifically of how this content would determine a reference if this existed. Nihilist theories are in an indistinguishable position in this respect. They postulate either that a vague predicate must always get a trivial referent, and thus that sentences containing it must get trivial truth conditions, or that a vague predicate must never get referents, and thus sentences containing it can never get truth conditions. Thus nihilists can provide no explanation in terms of propositional content and its determination of truth conditions of

---

<sup>18</sup> The mechanism of reference fixing proposed by the dual picture predicts that there are also reference failures in some occasions of use in which no instance of the Sorites Chain Schema is true. This will happen in cases where, despite there being no relevant sorites chain, some objects are neither tolerably close to the positive paradigms nor tolerably close to the negative paradigms along the relevant dimension. While the occasions of use where an instance of the Sorites Chain Schema is true are irregular *paradoxical* occasions of use, these other irregular occasions of use are *non-paradoxical*. (See Gómez-Torrente 2010, p. 248, for discussion.)

how speakers can elicit from hearers desired, non-arbitrary responses by using vague predicates.<sup>19</sup> According to the dual picture, on the other hand, a speaker is perfectly familiar with occasions of use in which the content-fixing mechanism she implicitly masters delivers a (contribution to) propositional content which in turn delivers a reference for a vague predicate, and thus also plain truth conditions for the sentences involving it. (In fact, these occasions of use are probably the norm, as we will argue in the next section.) So a speaker knows the (contribution to) propositional content made by a vague predicate, and knows in particular how this determines its reference when it has one. In this way the dual picture can reasonably explain in terms of propositional content and its determination of truth conditions how speakers elicit from hearers desired, non-arbitrary responses by using vague predicates. If in occasion of use *P* above someone says to someone else *I need a photograph of the young person*, the (cooperative) hearer will (attempt to) get a photograph of Alvar and give it to the speaker. The explanation is simply that hearers grasp the (contributions to) propositional contents expressed by speakers (or some closely related ones) and what referents and truth conditions these determine.

If this explanation of communication with vague predicates in regular occasions of use is correct, then the dual picture can provide an explanation even of how a vague predicate can be successfully used in communication in irregular occasions of use. Knowing a predicate's (contribution to) propositional content, and specifically how this determines a reference when it does, suffices for successful communication even in irregular occasions of use, as speakers can elicit from hearers all sorts of desired responses under the implicit common understanding of how the referents and thus the truth conditions of their utterances would be determined *if they in fact existed*. Thus, even in the irregular occasion of use *Q* above someone can say to someone else *I need photographs of some young people* and get the hearer to look for a photograph of Alvar, but not of Zellig, as their

---

<sup>19</sup> Ludwig and Ray (2002, p. 448) perceptively acknowledge the similarity of the predicaments of the nihilist and the epistemicist when it comes to explaining the intelligibility of vague language and facts of apparent communication involving sorites-susceptible predicates. In contrast, Braun and Sider (2007) seem to me unduly optimistic when they repeatedly claim that these phenomena are to be explained by language users' "ignoring" vagueness (where "ignoring" is understood as a fully "negative" attitude of typical language users, not as a kind of "positive" knowing followed by a kind of "forgetting"). I don't think that any purely "negative" attitude can explain competence and communication. Some "positive" abilities must be involved, which Braun and Sider fail to account for. In fact, it seems clear that they could not consistently account for such abilities, since these must presumably involve some kinds of "positive" representation of vague concepts.

implicit knowledge of the concrete instruction for “young” in operation in  $Q$  makes it clear to them that provided “young” applies to some people in the occasion of use, it will apply to Alvar but not to Zellig. This particular instance of communication is not prevented by the lack of a reference for “is young” or by the lack of truth conditions for the speaker’s utterance. Similarly, someone can scare a bunch of people by uttering *They are coming to get us* even if he has no one in mind with his utterance of “they”, which thus presumably lacks a reference. It suffices that the bunch of people know in some implicit way how the reference of “they” gets determined *when it has one*. (Note also that someone can utter *Remember that neurasthenic?* and get her hearer to think of a particular person with a neurasthenic syndrome, even if both hearer and utterer are contemporary psychiatrists who know full well that “neurasthenic” is somehow defective. For this instance of communication to work out, it suffices that the hearer knows how the extension of “neurasthenic” would have been fixed if it had had one, and in particular to know what the paradigms were.)

## 2. Refining the dual picture

The dual picture seems to me especially appealing as an account of content fixing for typical degree adjectives in occasions of use where only a small and clearly divided contrast class appears to be involved in content fixing. As argued in my earlier work, these occasions of use appear to be quite frequent. If I show Figure 1 to my little son,

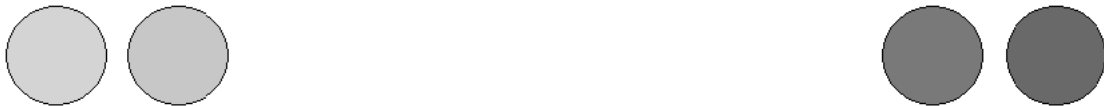


Figure 1

and I say to him, *Oh, look, there are two dark circles and two that are not dark*, the extension of “is dark” is intuitively the set consisting of the two circles to the right, and its anti-extension intuitively contains the two circles to the left. Essentially similar pictures and situations are often used when teaching degree adjectives to small children (e.g. in pedagogical TV programs), which further suggests that they are (among) the paradigmatic occasions of use for such adjectives. The reference-fixing mechanism of section 1 arguably delivers the right result, simply under the assumption that the two circles to the right are taken as paradigms of dark things in the context, the two circles to the left are taken as paradigms of non-dark things in the context, and only the difference in lightness between the circles in either side

is taken as tolerable for the application of “dark”; or, alternatively, under the assumption that the rightmost circle is taken as a paradigm of a dark thing, the leftmost circle is taken as a paradigm of a non-dark thing, and only the difference in lightness between the circles in either side is taken as tolerable for the application of “dark”. Note that in a case like this no plausible appeal, however implicit, is made to a large comparison class, such as the class of all things or the class of all circles.

In contrast, cases where a large or not clearly divided contrast class is involved in content fixing seem comparatively rare and often have an air of artificiality. An example complementary to the one involving Figure 1 would be an occasion of use of “dark” involving Figure 2 (from which Figure 1 comes, by deletion of the five middle circles).

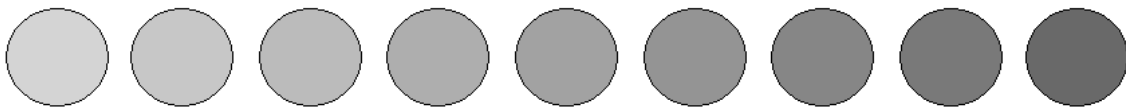


Figure 2

The reference-fixing mechanism of the dual picture delivers the result that such an occasion of use is an irregular one in which “is dark” lacks an extension, provided that the rightmost circle is taken as a paradigm of a dark circle, the leftmost circle is taken as a negative paradigm of a dark circle, and the difference in lightness between two adjacent circles is taken as tolerable for the application of “dark”. This gives a reasonable explanation of the fact that some questions and requests do not seem sensible in such an occasion of use, such as *What are the dark circles?*, *How many dark circles are there?*, and *Point out to me the dark circles*. One would not normally use a figure of this sort when teaching the meaning of “dark” to a small child. (As noted above, a declaration of irregularity will not prevent the picture from explaining at least many instances of communication in such occasions of use. For example, if while looking at Figure 2 I say to my little son *Point out to me a dark circle*, he will probably point to the rightmost circle or the one next to it. On the dual picture, as we said, this instance of communication is explained by the existence of a convention to the effect that if “is dark” is to have an extension in such an occasion of use, this will include these circles, which are taken as paradigms or as tolerably close to paradigms.)

But there may be occasions of use in which a large or not clearly divided class is somehow involved in content fixing for some purpose of comparison but in such a way that the occasion of use is not really irregular, because the contrast class is ultimately small

and divided. If so, then the dual picture as presented in my earlier work does not offer a fully adequate description of these cases, and may need to rely on an unnecessarily indirect explanation of instances of communication in them.

Here is the beginning of a newspaper article chosen at random, where degree adjectives have been italicized for future reference:

The film crew walked out on the *old* railroad trestle *high* above Georgia's Altamaha River, then placed a metal-frame bed on the tracks for actor William Hurt. The plan called for Hurt to lie on the bed in a dream sequence for the film "Midnight Rider", in which he plays rock singer Gregg Allman. Two trains had already crossed the bridge that day, and the crew was told no more were scheduled, hairstylist Joyce Gilliard recalled. Then a train came barreling toward them. "We all ran for our lives", Gilliard said. "All I could think of was my family getting that call..." Crew members scrambled to get themselves and their gear to the side, but couldn't get the bed off the tracks. The locomotive smashed into it. Sarah Jones, a 27-year-old camera assistant, was killed, struck by metal shards from the bed and by the train itself, according to witnesses and a police report. Six other crew members on the bridge were injured. Film and TV sets have always been *dangerous* places. Last year, three people died during a shoot for a Discovery Channel show when a helicopter crashed in Acton. In 2012, a diver died preparing an underwater scene for the Disney movie "The Lone Ranger". But the Feb. 20 fatality in Georgia has triggered an *extraordinary* response, with Hollywood's "below-the-line" corps of working crew members seizing on Jones' death to rally for *safer* working conditions (*Los Angeles Times* online edition, March 21, 2014, retrieved on that same date from <http://www.latimes.com/entertainment/envelope/cotown/la-et-ct-film-safety-20140322,0,7173322.story#ixzz2wewrzJ3h>)

Standard semantic theories of degree adjectives postulate that the extension of, say, "is old" as used here, is calculated via some algorithm that operates on a contextually relevant comparison class, presumably the class of railroad trestles in this case, and that delivers a contextually relevant class norm, some kind of standard value of the ages or age-related appearances of railroad trestles. The contextual extension of "is old" is then determined to be the class of trestles that have an age or an appearance that is greater or older-looking in a significant way than that class norm.<sup>20</sup> The class of railroad trestles is presumably very large, and if it is the contrast class in this occasion of use, then the dual picture in its earlier version will probably declare the occasion of use as irregular. But taking at least some cues from standard semantic theories of degree adjectives, we can plausibly view at least many similar occasions of use as regular.

First let me make it clear that the details of standard semantic theories of degree adjectives seem implausible to me, for a number of reasons. In fact, some of these reasons played a motivating role in my coming up with the mechanism of reference fixing of the dual picture. One not very important reason is that there is no plausibility in the suggestion

---

<sup>20</sup> See e.g. the theories in Bartsch and Vennemann (1972), Kennedy (2007), and Bale (2011). These theories nevertheless differ substantively as to the nature of the algorithm that calculates the contextually relevant class norm.



that a precise algorithm is involved in the calculation of a class norm by speakers or hearers of utterances of degree adjectives; but this idealization in particular might be taken as slight and harmless. A more important reason is the related one that it's implausible that any such algorithm, precise or not, would have something reasonably accessible to operate on. The class of ages or age-related appearances of trestles surely exists, but what knowledge can the typical reader (or writer) have of this class? Or take the use of "high" that follows the use of "old" in the text: can the class of heights of trestles, the presumable argument for the relevant algorithm, really be involved in some implicit calculation of a class norm by writer or reader in this case? I very much doubt it. A third reason why standard semantic theories of degree adjectives are implausible is that, to the extent that they don't idealize in a harmful way, they lead to an infinite regress: they must use some vague degree adjective or functionally equivalent construction when stating something to the effect that the class norm must be *significantly* far away from the value of an object in the relevant dimension in order for the object to fall in the relevant extension; the extension of "is significant" must then be calculated in some implicit way following the same procedure, which will involve a new appeal to "significant"; and so on. To be sure, a theory of this kind may work for some purposes under the idealization that there is a numerical threshold that precisifies the relevant notion of being "significantly far away", an idealization that of course stops the mentioned regress. But this is not a harmless idealization when theorizing about vagueness, for the stipulation of such a threshold generates semantically unlikely cut-off points.

The role of the comparison class in cases such as these occasions of use of "old" and "high" is probably not that of providing a sufficiently detailed input to an algorithm (however imprecise) computing a class norm and indirectly an extension, but that of making the writer or reader think of some remembered or imagined paradigm of trestle, perhaps a taken-to-be normal paradigm of trestle as regards age or height. This remembered or imagined paradigm, however, is probably not determined in any way by the comparison class alone. While the classes of trestles or trestle ages or age-related appearances and the class of trestle heights are surely not known in any kind of sufficient detail by the writer or reader, surely their knowledge of the concept of trestle and the minimal factual information they have about trestles are enough to make them think of some remembered or imagined paradigm of trestle, and probably of a paradigm of trestle that is intended to be normal as regards age or age-appearance and height. If this is so, then the reference-fixing mechanism of section 1 can deliver referents, i.e. extension/anti-extension pairs, for "is old" and "is high" in these occasions of use in such a way that the

trestle in the news story falls under the extension of “is old” and “is high” in those occasions. This will happen, for example, if: no other trestles are being actively considered in the contrast class, nor are hence taken to be part of the universe of discourse; the trestle in the story is taken as a paradigm of an old and a high trestle in the context; the remembered or imagined trestle is taken as a negative paradigm of an old and a high trestle in the context; and the differences in age or age-related appearance and in height between the remembered or imagined trestle and the trestle in the story are not among those taken as tolerable for the application of “old” and “high”. Or alternatively, if: no other trestles are being actively considered in the contrast class; the remembered or imagined paradigm of trestle is taken as a paradigm of old trestle (as regards age or age-appearance) and as a paradigm of high trestle; and the differences in age or age-related appearance and in height between the remembered or imagined trestle and the trestle in the story are among those taken as tolerable for the application of “old” and “high”.

To be sure, the paradigms of trestle remembered or imagined by writer and reader will probably be somewhat different, so the propositional content expressed by the writer and the propositional content grasped by the reader will probably be different. Does this threaten the possibility of communication? No, provided those propositional contents are sufficiently similar as to determine the same extension for “old” and “high”, or at the very least extensions that contain the trestle in the story, as they plausibly will. In particular, if we take, say, the second alternative above, it will be enough if the propositional contents (i.e. the conditional reference-fixing instructions) entertained by writer and reader are similar in the following respects: in both propositional contents no other trestles are being actively considered in the contrast class; in both propositional contents the remembered or imagined paradigm of trestle is taken as a paradigm of old trestle (as regards age or age-appearance) and as a paradigm of high trestle; and in both propositional contents the differences in age or age-related appearance and in height between the remembered or imagined trestle and the trestle in the story are among those taken as tolerable for the application of “old” and “high”.<sup>21</sup> Similarly, on a standard Fregean account of proper names or natural kind terms, the propositional contents assigned to a sentence containing one such term will vary, however slightly, from speaker to speaker, even if successful

---

<sup>21</sup> In obvious ways, the propositional contents of writer and reader will also typically be sufficiently similar as to determine the same intension for “old” and “high”, thus making feasible instances of communication involving embeddings of sorites-susceptible predicates under intensional constructions. Compare the example involving the modal sentence *It could have been the case that three of these circles had been dark* in section 4 below.

communication using the sentence will not be prevented provided there is agreement in term reference and sentential truth conditions. (Note that analogously, on standard semantic theories of degree adjectives, communication cannot be explained by supposing that different speakers implicitly calculate the same class norm and indirectly the same extension. There is no plausibility in that. At best, standard semantic theories must be coupled with a psychological theory on which the implicit calculation of different standards by conversational participants nevertheless usually leads to the determination of the same or a closely similar extension for the relevant adjective.)

My earlier presentation of the dual picture identified the contrast class of an adjective  $A$  (the class over which  $A$  is intended to be defined) and what is usually called the comparison class for  $A$ , the class  $C$  that intuitively determines membership in the extension of  $A$  via the intuitive condition of “being  $A$  compared to the members of  $C$ ”. (This intuitive condition may in turn be theoretically explained in terms of  $C$ 's role in helping a language user fix a class norm for  $A$  that can then be used in contrastive content fixing in the occasion of use, or in terms of its role of helping the language user fix on a paradigm in the class, or quite possibly in terms of other apparatuses.) In the refined version of the dual picture postulated here, the comparison class need not be the ultimate relevant contrast class in the occasion of use. Therefore, the comparison class need not always be included in the universe of discourse of the occasion of use. In particular, if the comparison class is large, it nevertheless need not be the case that a large universe of discourse is being actively considered in the occasion of use. This sounds especially plausible given that the things in the comparison class need not be being actively considered in the occasion of use, as the class will often be considered via some descriptive representation, under a general concept that characterizes it. (Compare Fara's 2000 insistence that comparison class relativization is not an extensional phenomenon.) And when the large comparison class is not a part of the universe of discourse, the chances of the occasion of use being a regular one increase greatly.<sup>22</sup>

Returning to the text, consider the uses of “dangerous” and “extraordinary”. By contrast with the reviewed uses of “old” and “high”, these seem to be what in my earlier

---

<sup>22</sup> It may be worth stressing that, while comparison classes do not enter into a specification of the extension-fixing clauses for vague predicates of section 1, this doesn't make comparison classes an idle wheel of the dual theory. An account of the determination of elements of an occasion of use such as the contrast class, the paradigms, etc. (which appear mentioned in those clauses) must in turn be part of a full development of the theory, and according to the refinement presented here, for example, comparison classes play a role in the determination of paradigms.

work I have called *contrastive* uses: uses that do not seem to rely even indirectly on large, sorites-prone, or even not clearly divided comparison classes, and instead directly presuppose a contrast class consisting of two clearly separated sets of objects that need to be forcefully contrasted for the conversational purposes at stake. Thus, the use of “dangerous” as applied to film and TV sets could scarcely presuppose a large comparison class of “places”; presumably film and TV sets are not among the top dangerous places in the universe. The author just means to contrast film and TV sets, where as has been shown people sometimes die or get hurt, with places where people don’t die or get hurt. Similar remarks apply to the use of “extraordinary”. This is surely not meant to compare the response to the reported death with the class of all responses ever; it is just intended to highlight a contrast with the apparently lackadaisical responses to the earlier deaths in film and TV sets.

One of the criticisms voiced against the dual picture, made by Peter Pagin (2011, 306), is that the picture condemns instances of the Positive Paradigms Schema and the Negative Paradigms Schema to defectiveness in all cases where the comparison class contains a relevant sorites series, while according to Pagin those instances should come out at least nearly always as true. I disagree with Pagin that those instances should come out nearly always as true, and in my earlier work I have given several considerations intended to show that they need not (including the reasons for the analogy with the case of failed natural kind terms, where paradigm sentences can be seen on reflection to have an air of indeterminacy, regardless of how compelling they may have been previously). But given the present plausible refinement of the picture, the number of irregular occasions of use of typical degree adjectives must be considerably smaller than previously admitted, and consequently the number of occasions of use where instances of the Positive Paradigms Schema and the Negative Paradigms Schema lack truth conditions must be correspondingly smaller. In particular, occasions of use in which the comparison class involved in content fixing is large, but the contrast class is small and divided, will be occasions in which relevant instances of the Positive Paradigms Schema and the Negative Paradigms Schema will have truth conditions and in fact will be true according to the refined picture. Furthermore, it must also be remembered that, even though in irregular occasions of use relevant instances of the Positive Paradigms Schema and the Negative Paradigms Schema will lack truth conditions, the dual picture has a specific explanation of why they will be thought of as true even in these occasions. The explanation is again that, given the concrete instruction associated with the relevant predicate in the occasion of use,

it will be clear to speaker and hearer that provided the predicate applies to some objects in the occasion of use, it will apply to the things preconceived as positive paradigms and negatively apply to the things preconceived as negative paradigms.

### **3. The roots of paradox**

The most widely accredited, standard theories of the logical and semantic paradoxes do not invoke a non-classical semantics or logic, and instead arguably rely on a shared abstract diagnosis involving content fixing: they all exploit in some way the idea that the content-fixing mechanisms for the paradoxical expressions work well in general, even if they turn out to fail in certain relatively localized cases. The exact delimitation of the problematic cases is of course a matter of dispute among the different theories, but there can be little doubt that the theories follow this abstract pattern. Tarski's theory of truth begins by laying bare the central role of T-biconditionals in the specification of the truth conditions of truth ascriptions. It then proposes a hierarchy of truth predicates and restricts meaningful truth ascriptions to ascriptions to sentences not containing the ascribed truth predicate nor any truth predicate of a higher level in the hierarchy. Truth ascriptions not abiding by these restrictions are considered meaningless. Kripke's theory of truth relaxes somewhat the Tarskian restriction, noting for example that in some natural non-paradoxical cases two speakers can make truth attributions about each other's sets of claims. This cannot be straightforwardly accommodated within a hierarchy of Tarskian truth predicates, and Kripke in fact proposes a theory with one unique truth predicate. Nevertheless, contentful truth ascriptions, and thus truth ascriptions susceptible of getting their truth conditions fixed by T-biconditionals, are delimited in another way: they are restricted to those where the sentence to which truth is ascribed is "grounded"—roughly, those whose truth conditions can ultimately be traced to statements not containing semantic vocabulary.

Logical or set-theoretic paradoxes are generated by the unrestricted application of comprehension intuitions: against intuition, applications of comprehension do not always give rise to a bona fide set. Russell's theory of types proposes a hierarchy of types of sets: the sets of individuals, the sets of sets of individuals, the sets of sets of sets of individuals, etc.; and it restricts comprehension to conditions defined over sets (or individuals) in types inferior to a given type. Instances of comprehension not abiding by these restrictions are considered meaningless. Zermelo's set theory relaxes somewhat the Russellian restriction, noting that in some natural non-paradoxical cases sets can contain elements from infinitely many types. This cannot be straightforwardly accommodated within a hierarchy of

Russellian types, and Zermelo in fact proposes a theory with one unique comprehension schema—and a single untyped notion of membership. Nevertheless, contentful instances of comprehension are delimited in another way: they are restricted to those where the condition at stake is defined over an existing set—one whose existence is yielded by a succession of applications of the operations of set formation legislated by the set-theoretic axioms, starting with individuals and/or the empty set.

It is not unreasonable to see in these abstract coincidences a correspondingly abstract diagnosis of the roots of paradox, at least for the main semantic and logical paradoxes. Certain procedures—the reliance on unrestricted T-biconditionals, or on unrestricted comprehension—are implicitly taken to determine truth conditions or referents for appropriate expressions—truth ascriptions, straightforward set-naming descriptions. But it turns out that these procedures, while working well in many, or even a vast majority of cases, give problematic results in some—at least for truth ascriptions where the sentence to which truth is ascribed is “ungrounded”, and at least for some straightforward set-naming descriptions where the condition at stake is not defined over a set formable by a succession of applications of the operations codified by Zermelo’s axioms starting with individuals and/or the empty set, e.g. when it is defined over “all sets”. The intelligibility and usefulness of the paradoxical concepts is arguably to be explained by the non-problematic character of the procedures in a vast number of cases, and presumably in the central ones—the liar and truth-teller sentences, or the descriptions for Russell’s class or for the class of all things in the universe are presumably not central cases of the use of the concepts of truth and of set. And while the failure of the procedures in some cases reveals that the scope under which the procedures function appropriately is or was not well understood, it would not seem to show that the procedures are thoroughly worthless.

If this diagnosis is correct as a diagnosis of the roots of the main semantic and logical paradoxes, as I think it is, one might expect it to provide at least a *prima facie* sensible blueprint for a diagnosis of the sorites paradox. However, it is clear that standard theories of vagueness and the sorites are incompatible with any diagnosis of this abstract kind. To begin with, the most popular theories (the varieties of supervaluationism and degree theories) postulate that sorites-susceptible concepts work only in tandem with a non-classical semantics for the logical expressions. Their diagnosis of what goes wrong in the sorites reasoning does not appeal to unforeseen necessary restrictions in the implicitly used mechanisms of content fixing, but to somewhat mysterious adaptations of the logic of natural language when vagueness is concerned. On the other hand, when one focuses on

standard theories that don't abandon the classical semantics and logic for the logical expressions, one never gets a diagnosis of the mentioned kind either. Epistemicists and most contextualists are led to postulate that the content-fixing mechanisms for sorites-susceptible concepts work well without restrictions, but in some unknowable way. And weak and strong nihilists postulate that those mechanisms just turn out not to work well at all.

It may be worth noting that epistemicist and nihilist views about truth or sethood are likely to strike the reader as theories of last resort, and certainly as theories that provide as little illumination on the intelligibility and usefulness of those notions as epistemicist and nihilist views of sorites-susceptible concepts provide of their intelligibility and usefulness. An epistemicist about truth would say that a liar sentence is not only completely meaningful, but has determinate classical truth conditions provided by whatever it is that fixes the content of truth ascriptions, only we don't or even can't know which conditions those are; an epistemicist about sethood would say that "the set of all sets that do not belong to themselves" has a determinate classical set as referent provided by whatever it is that fixes the content of straightforward set-naming descriptions, only we don't or even can't know which set that is. Weak nihilists about truth would say that nothing is true, weak nihilists about sethood that nothing is a set. Strong nihilists about truth would say that no statement involving truth is meaningful, strong nihilists about sethood that every statement involving sethood is meaningless. While I doubt that more than a few of these positions have actual defenders, pointing out this very fact may help convince us that the analogous theories of vagueness and the sorites, tempting though they may have appeared to some theorists of vagueness in moments of despair, ought to have few defenders as well.

But more importantly, calling attention to these analogies suggests an abstract argument for the desirability of a theory of vagueness and the sorites that aligns itself with the most widely accredited, standard theories of the logical and semantic paradoxes: given that the main logical and semantic paradoxes have their root in the localized failure of a content-fixing procedure, a failure that takes place only in a fairly restricted number of cases, a theory of vagueness and the sorites that shares this feature is desirable. To the extent that this argument is compelling, so will be the dual picture, for the picture certainly shares that feature with standard theories of the logical and semantic paradoxes. To begin with, it is based on a particular description of a mechanism of reference fixing for typical sorites-susceptible degree adjectives, postulated to underlie uses of these adjectives. Furthermore, the picture describes how some occasions of use of these adjectives will be

non-paradoxical and other such occasions of use will be paradoxical—depending on whether the universe of discourse in the occasion of use provides a relevant sorites series or not. And we have already seen how the non-paradoxical occasions of use of typical degree adjectives will be occasions in which the proposed mechanism of reference fixing does deliver referents—classical extensions—for those adjectives. In order for the dual picture to be in perfect alignment with the standard diagnoses of the main logical and semantic paradoxes, it only remains to be argued that these occasions of use provide a vast number of cases, and in fact the central ones. But both in section 2 and in my earlier work I have argued that, while only a large empirical study can fully confirm this hypothesis, the hypothesis appears more than likely in view of the fact that a great number of uses of degree adjectives one comes across are essentially contrastive, and thus regular. Furthermore, given the refinement of the earlier picture proposed in section 2, the number of occasions of use of typical degree adjectives in which the proposed mechanism of reference fixing does deliver referents for the adjectives must be even much greater than previously envisioned. In this way, there is an evidently strong case that the number of regular occasions of use of typical degree adjectives is vast, and includes the central cases. In conclusion, then, if we accept the reasonable abstract argument at the beginning of this paragraph, we will have also a strong consideration in favor of the dual theory.

#### **4. Remarks on recent related theories**

A number of theories of vagueness and the sorites proposed in recent years are fairly closely related to the dual picture, in that they all exploit in some way distinctions between occasions of use in which a relevant universe of objects does not contain a sorites chain and occasions of use where the relevant universe does. These include the theories in Manor (2006), Rayo (2008, 2013), Gaifman (2010), Pagin (2010, 2011), and van Rooij (2011a, 2011b).<sup>23</sup> As will be clear, several of the proposals in these theories are correct from the point of view of the dual picture, and some could probably be used to supplement it. However, in this section I will call attention to some of the major differences between these theories and the present theory, seeking to emphasize some of the present theory's comparative advantages.

---

<sup>23</sup> Kamp (1981) already exploits a similar distinction, but to so very different effects that it cannot be counted as being in the same class as these recent theories. On Kamp's theory, for example, every particularization of an instance of the Tolerance Schema is true in all contexts, and yet the instance in question can be false (in so-called "incoherent" contexts). This requires the postulation of a non-classical semantics and logic for the logical expressions.



Rayo's (2008, 2013) theory is different from the dual picture and the other theories in that it basically rejects the conceptual apparatus of reference and truth conditions for vague language. In particular, Rayo (2013) sketches a "grab bag" model of linguistic understanding and communication, according to which with a basic expression a subject does not associate a linguistic meaning or rule but only "a 'grab bag' of mental items: memories, mental images, pieces of encyclopedic information, pieces of anecdotal information, mental maps, and so forth" that "will typically not be enough to determine a range of application for the relevant lexical item independently of the subject's general-purpose abilities" (Rayo 2013, 648) but that will usually help him come to sensible decisions about what things the expression is to count as applying to in particular occasions. This is to be contrasted with a model on which a speaker's use of a basic expression is guided by semantic rules "which determine the range of application of a given expression (relative to a context of utterance) independently of the subject's general-purpose abilities, such as common sense" (647). Though independent of each other, the "rules" model goes hand in hand with "semantic globalism", the view that "in order for an assertion to be in good order, it must succeed in determining a partition of the entire space of possibilities" (650). The "grab bag" model, on the other hand, goes hand in hand with "semantic localism", the view that "all that is required for an assertion to be in good order is for it to succeed in dividing the possibilities that are relevant for the purposes of the assertion into verifiers and falsifiers" (650).

The "grab bag" model and semantic localism inspire Rayo's approach to vague predicates. When relatively few items are being actively considered, a "grab bag" can help a subject decide which items to count as falling under the range of application of a vague predicate and which not. The classification chosen by the subject will be one that stands out "as significantly more salient than any rival, in light of the grab bag corresponding to [the vague predicate] and general information about the situation at hand" (660). But when the items being actively considered give rise to a relevant sorites chain, the "grab bag" will be insufficient to provide any such classification, as "any potential location for the cutoff-point will seem roughly as salient as its immediate neighbors" (660). According to the "grab bag" localist, in such an occasion both a relevant instance of the Tolerance Schema and its negation "should be rejected. For they both presuppose that [the vague predicate] has a range of application in the context at hand, and the Grab Bag Localist thinks that there is no such thing" (663). On the other hand, if one is under the grip of a model of understanding and communication based on rules or meanings (which determine

propositions and truth conditions), one is almost bound to think incorrectly, with “semantic globalism”, that a classification must be in place even in sorites contexts, while at the same time making the mistake of accepting problematic instances of the Tolerance Schema.

I think the dual picture shows by example that the prospects for an account of vagueness and the sorites that relies on the apparatus of rules, propositions and truth conditions are probably not so bleak as Rayo suggests, and certainly that the apparatus is compatible with the sensible “semantic localism” he advocates. The dual picture sees the abstract instruction for a typical degree adjective as a meaning-constituting rule, suggests that the adjective’s (contribution to) propositional content is a concrete filling out of that instruction, and claims that the adjective’s reference (or contribution to truth conditions), in occasions of use where it has one, is determined by that propositional content. But this is compatible with the rule being meager in content and modest in its aim, and in particular with its inability to partition just any universe of discourse into a range of application and a range of negative application of the adjective. Also, to the extent that the picture coincides with the “grab bag localist” diagnosis of what goes wrong in sorites contexts, the successes of the “grab bag localist” picture in dealing with these contexts will be presumably reproducible by the dual picture.

It is also clear that Rayo’s “grab bag localism” cannot deliver all the time-honored benefits of the familiar apparatus of propositional and referential content unless it’s supplemented with a whole new apparatus of suitable theories. That the dual picture is at least in principle susceptible of delivering these benefits by relying on that familiar apparatus is in itself an important consideration in its favor. This is not the place to go in detail over the virtues of semantic theories of propositional and referential content or the advantages often claimed for them over other theories of understanding and communication. But some brief remarks are in order. First, semantic theories of referential content provide reasonably accurate models of speaker intuitions about the notions of a word standing for an object and of a word substitution making a difference to the truth or falsehood of a sentence of statement. Second, semantic theories of propositional content provide reasonably accurate models of speaker intuitions about the notions of what a speaker says, believes, knows, and so on, as well as of the notions of what is necessarily true, possibly true, sometimes true, compatible with what one knows, and so on. Third, while some or all of these notions can in principle be dismissed on the basis of philosophical argumentation, the fact that they feed back the process of construction of

linguistic conventions itself is no mean reason to take them seriously when one devises accounts of those conventions: for example, if we define the material conditional in terms of its truth conditions, or a new modality in terms of the truth of a proposition in a restricted subset of the set of possible worlds, and the device seems to work, we have reason to believe that referential and propositional notions get some aspects of linguistic reality right. The dual picture provides a way of reaping the presumable benefits of a “localist” picture of vagueness and the sorites without giving up the benefits of semantic theories of propositional and referential content.

As an example specializing the second consideration above, think of the embedding of sentences containing vague adjectives under modal operators. Go back to the occasion of use involving Figure 1 above—call it  $R$ —and imagine me saying *It could have been the case that three of these circles had been dark*. This is intuitively a correct thing to say (I could have given a much darker gray to one of the circles to the left after drawing it). But an explanation for this cannot merely rely on a class of mental associations with “dark” that allow the typical subject to make decisions as to how to apply the term in concrete situations. It has to rely on some kind of contribution of “dark” to the interpretation of utterances containing it (possibly a “grab bag”), on some kind of contribution of “it could have been the case” to the interpretation of utterances containing it, and on some kind of interaction between those contributions. One natural explanation is provided by the hypotheses that there is such a thing as the proposition that three of those circles are dark, and that this proposition is true in some modal circumstances of evaluation (and false in others). A natural part of the first hypothesis is then that “dark” makes a distinctive contribution to the constitution of the proposition in question, as a thing that can be evaluated in different modal circumstances. The dual picture suggests in an obvious way a candidate for such a (contribution to) propositional content (and in fact other candidates are compatible with the picture as well). This (contribution to) propositional content might be a concrete instruction looking somewhat like this. Let  $U_R = K_R$  be the set of the circles in Figure 1, let “i”, “ii”, “iii” and “iv” be names of these circles from left to right, let  $l_i$  and  $l_{iv}$  be the lightness values of i and iv, respectively, and let one  $\#$  be the difference in lightness between i and ii or iii and iv.

If there is a unique pair  $\langle E, A \rangle$  of subsets of  $U_R$  which are mutually exclusive and jointly exhaustive over  $U_R$ , and are such that

- everything with  $l_{iv}$  is in  $E$ ,
- everything in  $K_R$  that has a lightness inferior to something in  $E$  is in  $E$ ,

- everything in  $K_R$  that has a lightness superior by one  $u$  or less to something in  $E$  is in  $E$ ,
  - everything with  $l_i$  is in  $A$ ,
  - everything in  $K_R$  that has a lightness superior to something in  $A$  is in  $A$ ,
  - everything in  $K_R$  that has a lightness inferior by one  $u$  or less to something in  $A$  is in  $A$ ,
  - $U_R-K_R$  is included in  $A$ ,
- then  $\langle E, A \rangle$  is the extension/anti-extension pair of “is dark” relative to  $R$ .

Then this concrete instruction determines {iii, iv} as the extension of “is dark” relative to  $R$  in the actual world, but determines {ii, iii, iv} as that extension relative to  $R$  in a world where, after drawing ii, I give it the gray tone that iii has in the actual world. With respect to that other world, an appropriate utterance of *Three of these circles are dark* is true, which given the usual apparatus of possible worlds semantics makes true with respect to any world an utterance of *It could have been the case that three of these circles had been dark* relative to  $R$ . Rayo’s “grab bag localism” does not in and of itself come with a comparable explanation of the intuitive truth of some modal claims involving vague predicates, and needs a supplement if it is to have one.

Rayo’s picture would also appear to have the weakness that it does not allow for certain instances of apparently feasible communication involving vague adjectives. According to Rayo, in sorites contexts the “grab bag” is not enough to provide a range of application for the relevant vague predicate, and hence presumably not enough for the subject to classify intuitive positive (negative) paradigms as positive (negative) paradigms. But surely in such contexts one can manage to communicate something to someone else by means of instances of the Positive Paradigms Schema, of instances of the Negative Paradigms Schema, and by means of other sentences, such as *I need photographs of some young people* in the occasion of use  $Q$  above. By contrast, as already pointed out, the dual theory has a ready explanation of these instances of communication, based partly on its postulated (contribution to) propositional content for vague predicates.

Manor (2006) calls attention to the intuitive difference between occasions of use such as  $Q$  or the one involving Figure 2—sorites occasions of use—and occasions of use such as  $P$  above, or the one involving Figure 1. In occasions of the first kind, according to Manor, there is no non-arbitrary extension/anti-extension pair for the relevant vague predicate, while in occasions of the second kind there is a non-arbitrary extension/anti-extension pair, afforded by the gap between the objects in the context along the relevant dimension of comparison. Manor suggests that in contexts of the first kind there is a failure of a conventional presupposition that the use of the relevant predicate must be based on a non-

arbitrary demarcation of its extension, while in contexts of the second kind the presupposition is fulfilled. Manor identifies this presupposition with the presupposition that the relevant vague predicate has a reference, and sees the predicate as composed in part of a description subject to a Russellian analysis in sentences in which it appears; in fact, Manor appears to assume that on the Russellian analysis the description will always get wide scope. Thus, for example, the predicate “is young” is identified with the predicate “is among the young (people)”, and *Alvar is young*, say, is analyzed as *Alvar is among the young* and in turn as *There is a non-arbitrary demarcation of the extension of “young” such that Alvar belongs to it*. This has the consequence that in contexts of the first kind instances of the Positive Paradigms Schema, the Tolerance Schema and the Negative Paradigms Schema will all be false, while in contexts of the second kind instances of the Positive Paradigms Schema and the Negative Paradigms Schema will be true and (according to Manor) instances of the Tolerance Schema will be false. (This last claim of Manor appears to be false given her apparent assumptions. “ $(\forall x)(\forall y)(x \text{ is young} \ \& \ x \text{ was born one day later than } y \supset y \text{ is young})$ ” gets analyzed as “ $(\forall x)(\forall y)(x \text{ is among the young} \ \& \ x \text{ was born one day later than } y \supset y \text{ is among the young})$ ” and in turn as “There is a non-arbitrary demarcation of the extension of ‘young’ such that  $(\forall x)(\forall y)(x \text{ belongs to it} \ \& \ x \text{ was born one day later than } y \supset y \text{ belongs to it})$ ”, which is vacuously true in the occasion *P* above.)

Now, one clear drawback of Manor’s theory is its implication that in some contexts “young”, say, is false of a baby and “not young” is similarly false of an octogenarian (regardless of the speaker’s intention that “young” should apply to the baby and “not young” to the octogenarian). It seems clear that there should be no contexts where the truth conditions of, say, *Alvar is young* are such that it is false regardless of the speaker’s contextual intentions that “young” is to apply to Alvar). The negation of *Alvar is young*, whatever that is, should express the proposition that Alvar is not young, and that proposition is certainly not true, though it is implied to be true by Manor’s theory. The root of the problem is Manor’s strange suggestion that vague predicates are (Russellian) descriptions (taking always wide scope). This implies in fact that *Alvar is young* and *Alvar is not young* are not contradictory, despite clear intuitions to the contrary. Note, by contrast, that on the dual picture *Alvar is young* and *Alvar is not young* are contradictory, and that the worst that can happen to *Alvar is young* is that it lacks truth conditions in some presumably strained occasions of use where the principles guiding the normal application of “young” lead to an unexpected conflict.

Gaifman (2010) divides contexts into unfeasible and feasible, according as to whether the objects that have to be classified by a vague predicate in them form a sorites chain or not. In fact, in Gaifman's paper a context is identified with a set of objects. In feasible contexts there is a sufficiently large gap between the extension and the anti-extension of the predicate along the relevant sorites dimension, and instances of the Tolerance Schema, which are taken to be meaning-constitutive, are satisfied. In the case of the practically unrealized unfeasible contexts, some mechanism determines that the relevant vague predicate "loses its tolerance" and gets a sharp cut-off point in the sorites chain. According to Gaifman, only feasible contexts arise in practice and are needed for actual usage and actual reasoning. The bulk of Gaifman's theory is the development of a model theory and proof theory for a language that represents contexts as thus understood. Perhaps the main weakness of Gaifman's theory is that it's not based on a discernible account of content fixing, and in particular on an account of how vague predicates get their reference, either in feasible or in unfeasible contexts. This makes it quite mysterious what makes it the case that a vague predicate "loses its tolerance" and gets a sharp cut-off point in unfeasible contexts. This appears to mean that in the weird contexts where sorites chains are considered, e.g. philosophical contexts, vague predicates do get sharp cut-off points; but this is of course implausible.

Van Rooij (2011a, 2011b) explicitly takes his inspiration from Gaifman, and adds the proposal that in Gaifman's unfeasible contexts the use of the relevant vague predicate is "pragmatically inappropriate". Van Rooij's presentation is brief and does not provide details of the kind of pragmatic inappropriateness at stake, but certainly suggests that in unfeasible contexts the predicate is still semantically kosher and gets a classical extension/anti-extension pair (as in Gaifman) but its use is somehow merely pragmatically infelicitous. If so, then the obvious complaint (aside from the one about the inspecificity of the infelicitousness at stake) is, as with Gaifman's proposal, that van Rooij's suggestion leaves it unclear what content-fixing mechanism is responsible for the kosher semantics of vague predicates in unfeasible contexts.

Pagin's (2010, 2011) theory is quite different from the preceding theories and from the dual picture, in that he seems to labor under the assumption that most vague predicates are used by default with respect to large universes of discourse, or domains of quantification, containing relevant sorites sequences. The theory's main proposal is then that there is a mechanism that takes the context as input and delivers as output a class of possible restrictions of the domain of quantification. Each of these restrictions generates a

“central gap” in the sorites series, which in turn determines in the obvious fashion an extension/anti-extension pair, in such a way that the extension can play the role of referent for the predicate in the context. Each of the possible restrictions thus determines a possible “corrected” context by means of which the utterances of a speaker could be interpreted, and the idea is that those utterances, and in particular utterances of instances of the Tolerance Schema and of the Positive Paradigms Schema and the Negative Paradigms Schema will normally have their intuitive truth value in a “corrected” context. Nevertheless, Pagin concedes that there will be recalcitrant cases where not even a “corrected” context can provide an interpretation of a speaker’s utterances that makes all of them true, or even consistent.

I think an implausible aspect of Pagin’s theory is that it requires the mechanisms of quantifier domain restriction and the mechanisms of predicate interpretation to work in an unlikely fine-tuned coordination. If the two mechanisms interact, the model case would appear to occur when predicate content fixing relies on a comparison class, possibly identical with the quantifier domain; and here predicate interpretation presupposes a previously given domain, not the other way around. Further, quantifier domain restriction would seem to arise from motivations that have little to do with the avoidance of sorites paradoxes. Typically, the motivation is for the speaker to be able to make a claim that holds for the restricted domain but not for bigger domains that might be taken as the relevant universes of discourse. In such cases, the explicit quantifier domain restriction will often be in conflict with the postulated existence of a restriction creating a gap in the relevant sorites chain(s). If I say *Few living persons residing in small towns are young*, I seem to be explicitly saying that few people in the whole set of living persons residing in small towns are young; the restriction effected by “residing in small towns” is intended to make true a relatively weak claim, while the stronger unrestricted claim *Few living persons are young* will normally be considered false. But on Pagin’s theory, when I say *Few living persons residing in small towns are young*, I am to be interpreted as saying that few people in a more restricted set, which doesn’t contain all the living persons residing in small towns, are young; for this restricted set must contain a central gap in the sorites series that goes from the clearly young to the clearly not young people within the set of living persons residing in small towns.<sup>24</sup>

Pagin’s theory is also afflicted by some higher-order vagueness problems evaded by the dual picture and the other theories in this section. How is the class of possible restrictions of the domain of quantification to be determined by Pagin’s hypothetical

---

<sup>24</sup> Pagin (2011) mentions examples of this kind as a potential worry for his theory.

mechanism? Pagin limits himself to postulating that there is a function assigning to a context its class of acceptable restrictions of the domain of quantification. But the idea that such a function exists is implausible. In the occasion of use  $Q$  above, for example, it is simply implausible that there is a youngest living person that will not appear in the extension of any of the extension/anti-extension pairs that might all equally well play the role of interpreting “young” in  $Q$ . What could possibly determine the appropriate cut-off point for the concept ‘person that appears in the extension of some of the extension/anti-extension pairs that might all equally well play the role of interpreting “young” in  $Q$ ? Note that by contrast, the dual picture simply relegates  $Q$  to irregularity. And when the picture declares an occasion of use as regular, as in the case of  $P$ , it only relies on the fact that in such a case the universe of discourse is already clearly divided into two subsets clustering around positive and negative paradigms of a predicate. If such a clear division exists, then the meaning-constituting instruction for “young” and its induced (contribution to) propositional content suffice to generate an extension/anti-extension pair for “is young” in the occasion of use. The dual picture postulates no other mechanism as necessary for the interpretation of “young”, and in particular no mechanism required to deliver an unlikely cut-off point in the occasion of use, whether for the predicate or for higher-order concepts involving the predicate.

A different question is whether the truth of the *general* statement of the dual theory requires the existence of implausible divisions in its presupposed universe of discourse. This question will be taken up in the next section.

## 5. Does the dual picture have truth conditions?

Although typical degree adjectives and scalar nouns are the paradigms of vague and sorites-susceptible expressions, many, in fact most, other expressions have been, not implausibly, claimed to be vague and sorites-susceptible as well. In particular, the theoretical expressions used in the statement of the dual picture itself are perhaps sorites-susceptible. There is thus a risk that the statement of the picture lacks truth conditions. Does it?

When predicates other than those whose predicative element is a degree adjective or a scalar noun are claimed to be sorites-susceptible, the reason is normally intuitively quite different from the reason why a degree adjective or a scalar noun are sorites-susceptible. For those other predicates there is typically no analytically connected dimension of variation, and instead it is postulated that there is some dimension that, under reasonable but perhaps not water-tight assumptions, gives rise to a sorites conflict. If a cat is identical



with a certain aggregate of molecules (something that could be questioned), the set of all aggregates of molecules included in this aggregate contains, it is claimed, a sorites chain: start with the cat aggregate, continue with an aggregate in which a molecule has been removed, and so on, until you get a single molecule of the initial aggregate, which is surely not identical with any cat. Suppose that the implicitly used instance of the Tolerance Schema is somehow compelling (something that could be questioned), and we have that “is a cat” is sorites-susceptible. Assuming that this kind of argument, assumptions included, is correct, can the dual picture of degree adjectives be extended to a reasonable picture of what goes wrong in sorites occasions of use of vague predicates in general? As suggested in my earlier work, I think it can. The idea is simply that whatever instructions for reference fixing are implicitly in play for a predicate such as “is a cat”, these instructions will normally presuppose a universe of discourse in which there is no relevant sorites chain. In the case of “is a cat” and the mentioned sorites chain, this is especially plausible, for presumably the reference-fixing instructions for it presuppose that they are to be applied in occasions of use involving a universe of discourse containing only ordinary things, which turn out to be either clear cats or clear non-cats, but not more than a few of the enormous amount of weird aggregates of molecules in that chain. In such occasions of use, “is a cat” will get a reference, even if it will lack one in the “metaphysical” contexts where enormous amounts of weird aggregates of molecules are considered.

Even if some appropriate generalization of these considerations turns out to be the generically correct account of sorites occasions of use of vague predicates, it is conceivable that it might not be the correct account of the predicates appearing in the statement of the dual picture. Is it the correct account? Is the universe of discourse that has to be presupposed in the theoretical occasions of use in which the picture is stated such that the predicates employed in the statement divide its objects into clear positive cases and clear negative cases?

For some of the predicates employed in the statement of the dual picture, the answer seems clearly affirmative. The dual picture can be stated briefly as follows: *All sorites-susceptible predicates, in paradoxical occasions of use, lack an extension, and all sorites-susceptible predicates, in occasions of use in which the mechanism of preconceptions works, have an extension.* Think of “is sorites-susceptible”. Theorists of vagueness typically assume that the words of English (or at least those we find in a current dictionary) are sharply divided into sorites-susceptible and non-sorites-susceptible. A common view is that except for mathematical words in general and a few logical expressions, every other word is sorites-susceptible. And

presumably the universe of things a theorist of vagueness needs to quantify over can be restricted to include only clear cases and clear non-cases of words. The proponent of the dual picture may accept this and consequently assume that “is sorites-susceptible” divides clearly the set of words of his universe of discourse into two mutually exclusive and jointly exhaustive subclasses. It does seem hard to see what could be a word of an English dictionary that is neither clearly sorites-susceptible nor clearly not sorites-susceptible. By assuming that “is sorites-susceptible” divides clearly the set of words of his universe of discourse, the dual theorist is pretty clearly not endangering the explanatory power of his theory at all.

It is worth stressing that the dual theory does not deny that “is sorites-susceptible” and other predicates that appear in the statement of the theory are sorites-susceptible. On the contrary, the dual theorist accepts that they are. For him, the key question is whether this creates a problem for the possibility of stating the theory with truth. The sought answer is that the theory needs only to quantify over a universe of discourse that the theory’s predicates divide neatly into clear positive instances and clear negative instances, that the theory can do this without suffering from any loss in explanatory power. We have just made it plausible to think that the dual theorist can be satisfied with quantifying over things which are clear positive cases or clear negative cases of “is sorites-susceptible”. But the same can be said fairly clearly of “is a predicate”, “is an occasion of use”, “is the mechanism of preconceptions” and “is an extension”. Arguably, say, the linguist does not theorize about things which are neither clear positive cases of predicates nor clear negative cases of predicates, but this does not endanger the explanatory power of her theoretical considerations about the notion of a predicate. The dual theorist cannot be held to a higher standard than the linguist, at least as far his use of “is sorites-susceptible”, “is a predicate”, “is an occasion of use”, “is the mechanism of preconceptions” and “is an extension” is concerned.

The case of “is paradoxical”, as applied to occasions of use, is not so clear. A paradoxical occasion of use is one where the operating preconceptions don’t single out an extension/anti-extension pair over the universe of discourse. And it is not so clear that the dual theorist can adequately restrict the universe of discourse so that the predicates “singles out an extension/anti-extension pair” (as applied to a set of speaker’s preconceptions relative to an occasion of use) and “is the universe of discourse of” (as applied to universes of discourse relative to occasions of use) divide the things in the universe into clear positive cases and clear negative cases.

Think first of “is the universe of discourse of” (as applied to universes of discourse relative to occasions of use). Fix a universe of discourse  $U$ . Is there a sorites series of occasions of use that begins with an occasion of use where  $U$  is clearly the universe of discourse and ends with an occasion of use where  $U$  clearly is not the universe of discourse? The universe of discourse is determined by speaker’s intentions, so it would seem that what we are looking for is a sorites series of sets of speaker’s intentions, that begins with a set of intentions that determine  $U$  and ends with a set of intentions that does not determine  $U$ . But what would be the nature of the minute switch from one set of intentions to the next? The only possibility that occurs to me is that there could be some kind of slight deterioration of the mental representations that single out  $U$ . For example, suppose I think of  $U$  by having mental pictures of two individuals, Alvar and Zellig. But suppose that somehow a little part of my mental picture of Zellig is taken away from my mind or brain, then that another little part is taken, and so on, until I’m left with something that clearly does not depict Zellig; at the end of this series of representations it would seem I don’t have a representation of  $U$  nor hence an intention singling out  $U$  as the universe of discourse.

Assume something like this is what makes “is the universe of discourse of” sorites-susceptible. If the statement of the dual picture must quantify over all the possible occasions of use in sorites series like this, then the statement will lack truth conditions. But it doesn’t seem to me that the statement of the dual theory must ultimately quantify over the immeasurable amount of weird possible occasions of use that would appear in such sorites series. It will be enough if the theory quantifies over occasions of use one comes across in customary linguistic interactions, among which there is likely to be a clear division between occasions of use in which a determinate universe of discourse is born in mind by a speaker and occasions of use in which no such determinate universe is. This case is thus not so different from the case of “is sorites-susceptible”, “is a predicate” and the like. The theorist of vagueness is only marginally interested in theorizing about things one does not come across for practical purposes, such as things that are not clear predicates nor clear non-predicates, predicates that are not clearly sorites-susceptible nor clearly not sorites-susceptible, or occasions of use in which no universe of discourse is clearly determined and no universe of discourse is clearly not determined.

That there are occasions of use one comes across in which no universe of discourse is plausibly determined is a different, and very realistic possibility. But the existence of such occasions of use doesn’t present even a *prima facie* problem for the dual theory. If, as is

reasonable to assume, a vast majority of occasions of use are such that a determinate universe of discourse is fixed in them, occasions of use where this does not happen will often be susceptible of a derivative treatment without this hurting the explanatory power of the theory. The theory does not presuppose that a universe of discourse is fixed or determined in all occasions of use of vague predicates, and in fact it can explain communicative successes even in occasions of use of these predicates where no universe of discourse is actually fixed. Presumably the following is one such occasion. Out of the blue, I say to my little son *Draw me a dark circle and a circle that is not dark*. I am presumably not asking him to draw for me two circles I have in mind or two among some circles I have in mind. However, it's reasonable to think that my utterance has certain satisfaction conditions statable with the help of the reference-fixing mechanism of the dual picture: my utterance will be satisfied just in case there are two possible circles such that my son draws them and paints them in such a way that the difference in lightness between them is not among those I take as tolerable for the application of "dark" and "not dark".

Now think of "singles out an extension/anti-extension pair". Suppose we are talking about a certain set of speaker preconceptions, e.g. those in the occasion of use  $Q$  above, and about occasions of use of "young" by a speaker with these preconceptions. Consider then a series of occasions of use with corresponding universes of discourse, beginning with the set of all living persons as in occasion of use  $Q$  above, continuing with a universe of discourse in which persons born on January 1<sup>st</sup>, 1971 and December 31<sup>st</sup>, 1970 have been removed, then with a universe of discourse in which persons born on January 2<sup>nd</sup>, 1971 and December 30<sup>th</sup>, 1970 have been removed, and so on, until we are left with a universe of discourse clearly divided between positive and negative paradigms for "is young". If all these occasions of use with their respective universes of discourse are part of the universe of discourse of the statement of the dual picture, it would seem that the dual theorist is in trouble, for it would seem that the predicate "singles out an extension/anti-extension pair" does not then divide the universe of discourse of the statement of the picture into clear positive cases (suitable pairs of the form <set of speaker's preconceptions, occasion of use>) and clear negative cases.

But I think this impression is deceptive. Even if all these occasions of use need to be in the universe of discourse of the statement of the dual picture if its explanatory power is not to suffer, a determinate set of speaker's preconceptions will presumably single out an extension/anti-extension pair precisely in those occasions of use in the sorites series where there is a sufficiently big gap of middle cases. Specifically, a gap of at least the greatest

tolerance level appearing in the preconceptions which are instances of the Tolerance Schema. In the case of the preconceptions borrowed from  $Q$  above, this is one year. So this determinate set of speaker's preconceptions will single out an extension/anti-extension pair precisely in those occasions of use in the sorites series where there is a gap of at least one year between cases at the extremes. One might insist that the predicate "is an accepted tolerance level" (as applied to a speaker relative to an occasion of use) must be vague as well, and thus that it's not so clear that there will always be an appropriate gap.<sup>25</sup> For example, there will be sorites series in which one starts with a clear case of an instance of the Tolerance Schema being a preconception in an occasion of use, and ends with a clear case of that instance not being a preconception in another occasion of use, with many intermediate occasions of use along the way. But first, as in our discussion of mental deterioration above, it is again unclear that more than a few among the relevant occasions of use at the extremes of such sorites series will be among the customary occasions of use that the dual theory needs to quantify over. And second, it is unclear that "is an accepted tolerance level (of speaker  $S$  in occasion of use  $O$ )", being a psychological predicate about whose self-ascription a speaker has a special authority, gives rise to suitable sorites series. For it would seem that in order to ascertain whether, say, an instance of the Tolerance Schema is a preconception of a given speaker relative to an occasion of use, one could in principle simply ask him: if she has the disposition to answer "yes", it is; otherwise, it's not.<sup>26</sup> It may not make sense to ask what is the cut-off point for "is young" in an occasion of use such as  $Q$ , but it does make sense to ask what is the cut-off point for "is an accepted tolerance level (of speaker  $S$  in occasion of use  $O$ )", at least with respect to the customary occasions of use that the theorist of vagueness is primarily interested in.

In this section I have admittedly been talking about large universes containing sorites series for the concepts involved. Presumably many of my utterances here thus lack truth conditions according to the dual picture. But I think this is acceptable, provided the statement of the dual picture itself, as given in earlier sections, does not quantify over those large universes. I would be happy to accept that in this section I have only attempted to "show" things that cannot be "said", to use the Tractarian metaphor. We must not forget

---

<sup>25</sup> Similar remarks, here and in what follows, will hold for "is taken as a positive (negative) paradigm" and for the more general "is a preconception".

<sup>26</sup> Compare Rayo's related reply (in Rayo 2013, p. 661) to a similar potential objection to his account, as well as the related remarks about psychological predicates about whose self-ascription a speaker has a special authority, in Gómez-Torrente (2010, pp. 347-8, n. 28). (Against what he seems to think, I don't think that Rayo's kind of response is peculiarly facilitated by his "grab bag localism".)

that communication is possible even in irregular occasions of use according to the dual theory. I hope to have communicated some truths in this section, even if by the dual theory's lights, I have not said those truths.

## References

- Bale, A. C. (2011), "Scales and Comparison Classes", *Natural Language Semantics*, vol. 19, 169-190.
- Bartsch, R., and T. Vennemann (1972), *Semantic Structures. A Study in the Relation between Semantics and Syntax*, Athenäum, Frankfurt.
- Braun, D. and T. Sider (2007), "Vague, so Untrue", *Noûs*, vol. 41, 133-156.
- Fara, D. G. (2000), "Shifting Sands: An Interest-Relative Theory of Vagueness", *Philosophical Topics*, vol. 28, 45-81. (Originally published under the name Delia Graff.)
- Gaifman, H. (2010), "Vagueness, Tolerance and Contextual Logic", *Synthese*, vol. 174, 5-46.
- Gómez-Torrente, M. (2010), "The Sorites, Linguistic Preconceptions, and the Dual Picture of Vagueness", in R. Dietz and S. Moruzzi (eds.), *Cuts and Clouds. Vagueness, its Nature and its Logic*, Oxford University Press, Oxford, 228-253.
- Kamp, H. (1981), "The Paradox of the Heap", in U. Mönnich (ed.), *Aspects of Philosophical Logic. Some Logical Forays into Central Notions of Linguistics and Philosophy*, Reidel, Dordrecht, 225-277.
- Kennedy, C. (2007), "Vagueness and Grammar: The Semantics of Relative and Absolute Gradable Predicates", *Linguistics and Philosophy*, vol. 30, 1-45.
- Ludwig, K. and G. Ray (2002), "Vagueness and the Sorites Paradox", *Philosophical Perspectives*, vol. 16, 419-461.
- Manor, R. (2006), "Solving the Heap", *Synthese*, vol. 153, 171-186.
- Pagin, P. (2010), "Vagueness and Central Gaps", in R. Dietz and S. Moruzzi (eds.), *Cuts and Clouds. Vagueness, its Nature and its Logic*, Oxford University Press, Oxford, 254-272.
- Pagin, P. (2011), "Vagueness and Domain Restriction", in P. Égré and N. Klinedinst (eds.), *Vagueness and Language Use*, Palgrave Macmillan, New York, 283-307.
- Rayo, A. (2008), "Vague Representation", *Mind*, vol. 117, 329-373.
- Rayo, A. (2013), "A Plea for Semantic Localism", *Noûs*, vol. 47, 647-679.
- Soames, S. (1999), *Understanding Truth*, Oxford University Press, New York.
- Sorensen, R. (1988), *Blindspots*, Oxford University Press, Oxford.
- Unger, P. (1979), "There Are No Ordinary Things", *Synthese*, vol. 41, 117-154.
- van Rooij, R. (2011a), "Implicit versus Explicit Comparatives", in P. Égré and N. Klinedinst (eds.), *Vagueness and Language Use*, Palgrave Macmillan, New York, 51-72.
- van Rooij, R. (2011b), "Vagueness and Linguistics", in G. Ronzitti (ed.), *Vagueness: A Guide*, Springer, Dordrecht, 123-170.
- Williamson, T. (1994), *Vagueness*, Routledge, London.