Vagueness: insights from Martin, Jeroen, and Frank

Robert van Rooij

1 Introduction

My views on the semantic and pragmatic interpretation of natural language were influenced a lot by the works of Martin, Jeroen and Frank. What attracted me most is that although their work is always linguistically relevant, it still has a philosophical point to make. As a student (not in Amsterdam), I was already inspired a lot by their work on dynamics semantics, and by Veltman's work on data semantics. Afterwards, the earlier work on the interpretation of questions and aswers by Martin and Jeroen played a signifcant role in many of my papers. Moreover, during the years in Amsterdam I developped—like almost anyone else who ever teached logic and/or philosophy of language to philosophy students at the UvA—a love-hate relation with the work of Wittgenstein, for which I take Martin to be responsible, at least to a large extent.

In recent years, I worked a lot on vagueness, and originally I was again influenced by the work of one of the three; this time the (unpublished) notes on vagueness by Frank. But in recent joint work with Pablo Cobreros, Paul Egré and Dave Ripley, we developed an analysis of vagueness and transparant truth far away from anything ever published by Martin, Jeroen, or Frank. But only recently I realized that I have gone *too* far: I should have taken more seriously some insights due to Martin, Jeroen and/or Frank also for the analysis of vagueness and transparant truth.

- I should have listened more carefully to the wise words of Martin always to take Wittgenstein seriously. The tractatus does have an important point to make: not everything can be said. At it turns out, this holds even when we have a transparant truth predicate at our disposal, but also that some of the things that can only be shown, can be shown by implicature.
- Jeroen's recent move to inquisitive semantics is one that takes meanings to be more fine-grained than standard possible-world semantics allows for. As it turns out, such a more fine-grained view is important for the analysis of conversational implicatures in general, and for implicatures dealing with vagueness in complex sentences in particular.
- Perhaps the most important idea behind dynamic semantics is that the strict separation between semantics and pragmatics should be given up. Semantic notions like 'meaning' and 'consequence' should be 'pragmatized'. As it turns out, the pragmatized notions of meaning and logical consequence are very illuminating, also for the analysis of vagueness. The favoured pragmatic notion of logical consequence will be non-monotone, just like Veltman's notion of consequence in his default semantics

It is the aim of the rest of this paper to elucidate the above points.

2 3-valued semantics & non-transitive consequence

2.1 Non-transitive consequence

Let $\mathcal{M} = \langle \mathcal{D}, \mathcal{I} \rangle$, with $\mathcal{I}$ a total function from atomic sentences to $\{0, 1, \frac{1}{2}\}$. Now we can define the truth values of sentences as follows:[1]

- $\mathcal{V}_{\mathcal{M}}(\phi) \quad = \quad \mathcal{I}_{\mathcal{M}}(\phi)$, if ϕ is atomic
- $\mathcal{V}_{\mathcal{M}}(\neg\phi) \quad = \quad 1 - \mathcal{V}_{\mathcal{M}}(\phi)$
- $\mathcal{V}_{\mathcal{M}}(\phi \wedge \psi) \quad = \quad min\{\mathcal{V}_{\mathcal{M}}(\phi), \mathcal{V}_{\mathcal{M}}(\psi)\}$
- $\mathcal{V}_{\mathcal{M}}(\phi \vee \psi) \quad = \quad max\{\mathcal{V}_{\mathcal{M}}(\phi, \mathcal{V}_{\mathcal{M}}(\psi)\}$
- $\mathcal{V}_{\mathcal{M}}(\forall x\phi) \quad = \quad min\{\mathcal{V}_{\mathcal{M}}([^{d}/_{x}]\phi) : d \in D\}$

We say that ϕ is **strictly true** in $\mathcal{M}$ iff $\mathcal{V}_{\mathcal{M}}(\phi) = 1$, and that ϕ is **tolerantly true** iff $\mathcal{V}_{\mathcal{M}}(\phi) \geq \frac{1}{2}$. In terms of this semantics we can define some well-known logics: Kleene's $K3$ and Priest's LP. According to both logics, the consequence-relation is truth preserving. The only difference between the two is that while according to $K3$ only value 1 counts as true, according to LP, both 1 and $\frac{1}{2}$ do (while in $K3$ value $\frac{1}{2}$ stands for 'neither true nor false', in LP it denotes 'both true and false'). Thus, $\Gamma \models^{K3} \phi$ iff $\forall \mathcal{M}$: if $\forall \gamma \in \Gamma : \mathcal{V}_{\mathcal{M}}(\gamma) = 1$, then $\mathcal{V}_{\mathcal{M}}(\phi) = 1$, and $\Gamma \models^{LP} \phi$ iff $\forall \gamma \in \Gamma : \mathcal{V}_{\mathcal{M}}(\gamma) \geq \frac{1}{2}$, then $\mathcal{V}_{\mathcal{M}}(\phi) \geq \frac{1}{2}$. In some recent joint publications with Pablo Cobreros, Paul Egré, and David Ripley, we showed that a slight variant of $K3$ and LP can account for paradoxes of vagueness (Cobreros et al (2012) and transparant truth (Ripley 2012, Cobreros et al, to appear) in an, arguably, more satisfying way than either $K3$ or LP can. The crucial idea of the analysis of vagueness and transparant truth is that, although we don't give up the idea that entailment is truth-preserving, we allow the 'strenght' of truth of the conclusion to be weaker than the strength of truth of the premises. We say that a sentence ψ is ***st*-entailed** by a set of premises Γ, $\Gamma \models^{st} \phi$, iff $\forall \mathcal{M}$: if $\forall \gamma \in \Gamma : \mathcal{V}_{\mathcal{M}}(\gamma) = 1$, then $\mathcal{V}_{\mathcal{M}}(\phi) \geq \frac{1}{2}$. This analysis has two immediate consequences: (i) it interprets value $\frac{1}{2}$ as a notion of truth, just like LP does, and it thus allows for certain sentences to be both true *and* false, (ii) the notion of consequence is *non-transitive*. One appealing feature of the logic is that in contrast to either $K3$ and LP, it is a conservative extension of classical logic: it only differs from classical logic if we extend the language with (i) a similarity relation '$\sim$' (Cobreros et al, 2012), so that the tolerance principle $(\forall x, y((Px \wedge x \sim_P y) \rightarrow Py))$ becomes valid, or (ii) with a truth-predicate '$\mathcal{T}$' (Cobreros et al, to appear) that behaves fully transparant ($\mathcal{V}_{\mathcal{M}}(\mathcal{T}(\phi)) = \mathcal{I}_{\mathcal{M}}(\phi)$ for any ϕ). And even in these cases the resulting logical differences are minimal, it is only in very special cases (i.e., when it gives rise to paradox) that transitivity fails. For instance if we introduce a similarity relation '$\sim_P$' for each predicate P and we assume that $Px \wedge x \sim_P y \wedge y \sim_P z$ is strictly true, we can conclude via the validity of the tolerance-principle that Py is at least tolerantly true. And if $Py \wedge y \sim_P z$ is, or were, strictly true, we could conclude that Pz would be at least tolerantly true. However, the two inferences cannot be joined together to give rise to the Sorites-paradoxical conclusion: We cannot conclude from the strict truth of $Px \wedge x \sim_P y \wedge y \sim_P z$ to the tolerant truth of Pz.

[1]Notice that the semantics is just like that of Fuzzy Logic, but now limited to three truth values.

2.2 Logic and Expressibility

It is well-known that 2-valued logic gives rise to problems when one wants to treat vagueness or extend the language with a truth-predicate that behaves transparantly. In Cobreros et al (2012, to appear) we have argued that these problems can be solved when we make use of 3-valued semantics in combination with the new consequence relation $\models^{st}$. However, also this combination is not enough to solve all problems.

Consider, first, the extension of the language with a transparant truth predicate (and the possibility of self-reference). With such a transpararent truth predicate available, one can *express* much more than without. In particular, one can express within the language semantic properties of that language itself: i.e., the language becomes *semantically closed*. For instance, one can express within the language the truth-conditions of sentences of that language, and one can also easily define satisfaction in terms of truth (if one allows for enough proper names). Standardly, such semantically closed languages immediately give rise to paradox, in particular the liar paradox (for the transparant truth-predicate) and Grelling's heterological paradox (which can be expressed once one has a satisfaction-relation within the language of that same language). In Cobreros et al (to appear) it is shown, however, that these problems can be solved, making use of the non-transitive consequence-relation $\models^{st}$. But this doesn't mean that all problems of semantically closed languages disappear. It is sometimes claimed (e.g. Priest, 1979) that extending the standard logical language with a transparant truth predicate, and allowing for sentences to be both true and false, enables us to express everything without getting into problems. But this claim should be taken with a grain of salt.

The reason is that if a language would really be very expressive, we would also be able to say that a sentence is *only true*, or *only false*. But it is easy to see that this cannot be done.[2] We can try out two options: say it by means of a truth-predicate, or by introducing an explicit connective. Let us first consider the first option to express that a sentence ϕ is *only true*. We might try to express this by saying explicitly that ϕ is true, but not false: $\mathcal{T}(\phi) \wedge \neg\mathcal{F}(\phi)$. We know already that $\mathcal{V}_{\mathcal{M}}(\mathcal{T}(\phi)) = \mathcal{V}_{\mathcal{M}}(\phi)$. It seems reasonable, moreover, to assume that $\mathcal{F}(\phi) \equiv \neg\mathcal{T}(\phi)$ (or that $\mathcal{F}(\phi) \equiv \mathcal{T}(\neg\phi)$). The question now is whether $\mathcal{T}(\phi) \wedge \neg\mathcal{F}(\phi)$ can indeed only have value 1. Unfortunately, that is not the case. For suppose $\mathcal{V}_{\mathcal{M}}(\phi) = \frac{1}{2}$. Then $\mathcal{T}(\phi)$ has value $\frac{1}{2}$, just as $\neg\mathcal{T}(\phi)$ and $\mathcal{T}(\neg\phi)$. But that means that $\mathcal{F}(\phi)$ has value $\frac{1}{2}$ as well, just as $\neg\mathcal{F}(\phi)$ and $\mathcal{F}(\neg\phi)$, and thus also $\mathcal{T}(\phi) \wedge \neg\mathcal{F}(\phi)$. In other worlds, $\mathcal{T}(\phi) \wedge \neg\mathcal{F}(\phi)$ can be both true and false, in contrast to what we wanted.

Consider now the second option, introducing a new connective: $\mathcal{V}_{\mathcal{M}}(\sharp\phi) = 1$ if $\mathcal{V}_{\mathcal{M}}(\phi) = 1$, 0 otherwise. Unfortunately, as is well-known, this new connective gives rise to trouble: the extended liar sentence λ' that says of itself that it is not only true: $\lambda' = \neg\sharp\mathcal{T}(\lambda')$. The trouble is that there is no consistent way in which λ' can have a valuation: if it had value 1, it should have value $\frac{1}{2}$ or 0. If it had value 0, it should have value 1, and no sentence of the form $\neg\sharp\phi$ can have value $\frac{1}{2}$.

A similar problem arises when we try to express that a sentence is *only false*. Trying to express this by saying $\mathcal{F}(\phi) \wedge \neg\mathcal{T}(\phi)$ won't do, because *this* sentence can still have value $\frac{1}{2}$, i.e. be both true and false. Introducing a new connective $\sim$ with meaning $\mathcal{V}_{\mathcal{M}}(\sim \phi) = 1$ if $\mathcal{V}_{\mathcal{M}}(\phi) = 0$, 0 otherwise, allows one to express a new kind of extended liar giving rise to similar problems as the connective '$\sharp$' above. In conclusion, there just seems to be no way in which we can express in the object language that

[2] At least, if we don't allow the metalanguage to be paraconsistent. We won't consider that option here.

the semantic value of a sentence is 1 or 0.[3]

Wittgenstein always knew that there are limitations to what we can *express* in a language. Wittgenstein stated, however, that what cannot be expressed, can still be *shown*. But how to make sense of showing? I propose that to show that a sentence is 'only true' or 'only false', a speaker relies on a conversational implicature.

It is standardly assumed that from Gricean principles one can derive scalar implicatures: that from $\phi \vee \psi$ one can conclude that it is not the case that the stronger $\phi \wedge \psi$ is true, because otherwise the speaker would have said so. But once we think of value $\frac{1}{2}$ as 'true and false', the pragmatic reasoning from the assertion that ϕ is true, to the conclusion that ϕ is *only* true (and not also false) can be seen in the same light as well. Given that the speaker didn't say in addition that ϕ is also false, we pragmatically conclude that ϕ is only true, because othwerwise the knowledgeable speaker would have said so. Similarly, if the speaker says $p \vee \neg p$ we can conclude that not both p and $\neg p$ are true, because otherwise the speaker should have been explicit about this.

But before we show how this suggestion can be made precise, let us first consider a problem of the semantic/pragmatic account of vagueness in Cobreros et al (2012).

2.3 Taking borderline contradictions seriously

In Cobreros et al (2012) we argued that if Adam is a borderline case of a tall man, the sentence 'Adam is tall' is both true and false. We motivated this by a number of recent experiments (e.g., Alxatib and Pelletier , 2011; Ripley, 2011) that show that naive speakers find a logical contradiction like 'Adam is tall and Adam is not tall' acceptable exactly in case Adam is a bordeline tall man. In Cobreros et al (2012) we proposed that the explanation is that we always interpret a sentence pragmatically in the strongest possible way. This pragmatic interpretation accounts, on the one hand, for the intuitions that if one *says* that Adam is tall, what is *meant* is that Adam is only tall, but, on the other for the experimentally observed acceptability of contradictions at the border, because contradictions like '$Pa \wedge \neg Pa$' can only be interpreted as true when tolerant truth is at stake. In Cobreros et al (2012) we show that such a pragmatic interpretation also accounts for the observed unacceptability (cf. Serkuck et al., 2011) of classical tautologies like '$Pa \vee \neg Pa$' if Adam is borderline tall. Unfortunately, the interpretation rule gives rise to trouble for more complex sentences. Alxatib, Pagin, and Sauerland (2013) show that we wrongly predict that a sentence like 'Adam is tall and not tall, *or* John is rich' means that John is strictly rich, although it intuitively should mean that *either* Adam is borderline tall *or* John is strictly rich. Similarly, our analysis mispredict that a sentence like 'Adam is tall and Adam is not tall, *and* John is rich' can be appropriately asserted if John is not striclty rich.

2.4 Pragmatic interpretation

How should we account for pragmatic interpretation such that we can show what we cannot express, i.e. that a sentence is *only true*, and can solve the above problems with complex sentences involving borderline contradictions? Here is a proposal: the pragmatic interpretation of ϕ makes (exactly) one minimal truth-maker of ϕ as true

[3]Of course, it is possible to express in the object-language that ϕ has value $\frac{1}{2}$ by saying $\mathcal{T}(\phi) \wedge \mathcal{F}(\phi)$. The trouble is that negating this sentence does not rule out that the resulting sentence has value $\frac{1}{2}$.

as possible. What are the minimal truth-makers of ϕ? and how to think of 'as true as possible'? To start with the latter question, we can define an ordering $<_S$ between models, with S a set of sentences (a minimal truth maker), defined as follows: $\mathcal{M} <_S \mathcal{N}$ iff$_{def}$ $\{\phi \in S : \mathcal{V}_{\mathcal{M}}(\phi) = 1\} \subset \{\phi \in S : \mathcal{V}_{\mathcal{N}}(\phi) = 1\}$. As for the first question, we will make use of an idea also argued for in recent work by Groenendijk: that with a more fine-grained semantic interpretation we can do more than we could do standardly. In our case, a more fine-grained representation of meaning allows us to define the minimal truth-maker of ϕ, which will be thought of as a set of literals. The set of minimal truth-makers of ϕ, $T(\phi)$, can easily be defined recursively as follows:[4]

- $T(\phi) \quad = \quad \{\{\phi\}\}$, if ϕ is a literal.
- $T(\phi \vee \psi) \quad = \quad T(\phi) \cup T(\phi)$.
- $T(\phi \wedge \phi) \quad = \quad \{A \cup B | A \in T(\phi), B \in T(\psi)\}$.

Notice that according to these rules, $T(p) = \{\{p\}\}$, $T(\neg p) = \{\{\neg p\}\}$, $T(p \vee q) = \{\{p\}, \{q\}\}$, $T(p \vee \neg p) = \{\{p\}, \{\neg p\}\}, T(p \wedge \neg p) = \{\{p, \neg p\}\}$, $T((p \wedge \neg p) \vee q) = \{\{p, \neg p\}, \{q\}\}, T((p \wedge \neg p) \wedge q) = T(p \wedge (\neg p \wedge q)) = \{\{p, \neg p, q\}\}$, and $T((p \vee q) \wedge (r \vee s)) = \{\{p, r\}, \{p, s\}, \{q, r\}, \{q, s\}\}$.

Of course, the above definition does not deal with all sentences translated in a propositional language. For one thing, such a language also contains (bi)conditionals, for another, it might be that a negation has scope over a complex sentence. However, both problems can be solved easily when we stipulate that the minimal truth-makers should be determined only after a sentence is put in its so-called 'Negative Normal Form'.[5]

Notice that $T(\phi)$ can be thought of as a fine-grained *semantic* interpretation of ϕ, if we assume that with each literal there corresponds a (positive or negative) *fact*, or as *propositions*. Ciardelli et al, 2013) used meanings of a similar complexity to account for *hybrid* sentences containing declarative and inquisitive content. Indeed, it is worth observing that if we would think of the elements of $T(\phi)$ as propositions, $T(\phi)$ itself is of the same type as a question: a set of propositions. And indeed, it is natural to represent the yes-no-question p? simply as $p \vee \neg p$ as Ciardelli et al (2013) proposed. For this, 3-valued logic doesn't play any role. But let us now briefly turn to conditional questions. Consider the hybrid conditional question $p \rightarrow q?$. This tranlates first into $p \rightarrow (q \vee \neg q)$, and then into the following Negative Normal Form: $\neg p \vee (q \vee \neg q)$. But this means that $T(p \rightarrow q?) = \{\{\neg p\}, \{q\}, \{\neg q\}\}$. I will suggest soon that in terms of this, we can account for the meaning of such conditional questions in a very straightforward way. But for this we need to know how to interpret a sentence pragmatically, to which we turn now.

We define the pragmatic interpretation of ϕ, $PRAG(\phi)$ (where $[\![S]\!]^t_c$ abbreviates $\bigcap_{\psi \in S} [\![\psi]\!]^t_c$). We assume that $PRAG$ only takes sentences in Negative Normal Form.

- $PRAG(\phi) \quad = \quad \{\mathcal{M} |\ \exists S \in T(\phi) : \mathcal{M} \in [\![S]\!]^t \ \&\ \neg\exists \mathcal{N} \in [\![S]\!]^t : \mathcal{M} <_S \mathcal{N}\}$.

Notice that for literals and conjunctive sentences, this pragmatic interpretation rule simply tries to make its minimal truth-maker as true as possible, i.e., strictly

[4] The definition of $T(\phi)$ is inspired by the analysis of minimal truth-makers proposed by van Fraassen (1969). Like van Fraassen we could equivalently also have used both minimal truth *and* minimal false-makers

[5] ϕ' is a sentence in Negative Normal Form of ϕ if (i) it is logically equivalent with ϕ, and (ii) negations only stand in front of atomic formulas.

true. Thus, even if one cannot really *express* in language that a sentence like 'Adam is tall' is only true, it follows from the above pragmatic interpreation. But in general a sentence might have more than one minimal truth-maker, i.e., when the sentence is disjunctive: so the general pragmatic interpretation rule says that it is enough if one of its minimal truth-makers is as true as possible. As a result, $PRAG((p \wedge \neg p) \wedge q)$ singles out those models where p is tolerantly but not strictly true, *and* where q is strictly true, while $PRAG((p \wedge \neg p) \vee q)$ singles out those models where either p is tolerantly but not strictly true, *or* where q is strictly true.

Let us see what this pragmatic analysis predicts for some examples involving vagueness: (i) 'p' is interpret as being only true; (ii) '$p \vee \neg p$' is pragmatically interpreted as saying that either p is strictly true, or $\neg p$ is; (iii) '$p \wedge \neg p$' is predicted to be interpreted as saying that p is only tolerantly true; (iv) '$(p \wedge \neg p) \vee q$' is meant as saying that either p is only tolerantly true, or q is strictly true, while (v) '$(p \wedge \neg p) \wedge q$' is predicted to be interpreted as saying that p is only tolerantly true, and that q is strictly true. All these predictions seems to be in accordance with the experimental results and intuition.

Let us turn now to questions. First, the simple yes-no-question p?. We translate this as $p \vee \neg p$. This sentence is a classical tautology, and it is thus also always tolerantly true. However, this question is pragmatically interpreted as strongly as possible, meaning that either p has value 1 or value 0. This is as desired. Let us now consider the conditional question $p \rightarrow q$? It denotes via its disjunctive translation a classical tautology as well. We have seen above that $T(p \rightarrow q?) = \{\{\neg p\}, \{q\}, \{\neg q\}\}$. Let us now see what its conversational implicatures are. It is easy to determine the (quantity) conversational implicatures of a sentence ϕ simply by changing '$\exists S \in T(\phi)$' in the above definition of $PRAG(\phi)$ into '$\exists! S \in T(\phi)$': *Exactly one* of the minimal truth-makers of the sentence is as true as possible. The pragmatic interpretation of $p \rightarrow q$? can be expressed by $p \leftrightarrow (q \vee \neg q)$. And this has a natural 3-valued interpretation: if p is true, there is a real issue whether q or $\neg q$ is true, and if p is not true, neither q nor $\neg q$ will be (striclty) true.

2.5 Implicatures and Hurford's constraint

We mentioned already that our analysis easily accounts for conversational implicature. Notice that the suggested pragmatic interpretation rule—there is exactly one minimal truth-maker—can be reformulated as follows, when we would limit ourselves to 2-valued interpretations:

- $$\begin{aligned} PRAG_2(\phi) &= \{\mathcal{M} |\ \exists! S \in T(\phi) : \mathcal{M} \in [\![S]\!]\}. \\ &= \{\mathcal{M} |\ \exists S \in T(\phi) : \mathcal{M} \in [\![S]\!]\ \&\ \forall V \in T(\phi) : \mathcal{M} \in [\![V]\!] \rightarrow V = S\}. \end{aligned}$$

From our pragmatic interpretation rule (from both the two- and the three-valued ones) it follows that from '$p \vee q$' we conclude that '$p \wedge q$' is false. The conditional perfection inference from '$p \rightarrow q$' to $p \leftrightarrow q$ follows as a conversational implicature too. Perhaps (even?) more remarkable, this analysis account for 'embedded implicatures' as well. From the assertion of '$p \vee (q \vee r)$' it folows that only one of p, q and r is true, and from '$(p \vee q) \wedge (r \vee s)$' it immediately follows that exactly one of the following four alternatives is true: (i) $p \wedge q$; (ii) $p \wedge s$; (iii) $q \wedge r$, or (iv) $q \wedge s$. It is easy to see that it also accounts for scalar implicatures under universal (modal) qunatifiers: if it is asserted that 'Every boy kissed Mary or Sue', it is predicted by our pragmatic interpretation rule that every boy kissed only Mary, or only Sue. Similarity, 'John believes that $p \vee q$' is predicted to mean that in all of John's doxastic alternatives, exactly one of p or q is true.

Hurford's constraint bans disjunctions in which one of the disjuncts entails the other (Hurford 1974, Gazdar 1979, Singh 2008). This condition helps to explain the infelicity of

(1) *Jan is from (somewhere in) the Netherlands or from Amsterdam.

But there are some examples where the constraint appears to be violated, even though the sentence is appropriate:

(2) a. Either John didn't do all of the readings or he didn't do any of them.
b. Either John came, or Mary, or both.
c. John had 3 cookies, or 4.
d. John had at least 3 cookies.

To account for these data,[6] Gazdar (1979) claims that Hurford's constraint should be *weakened* to: $\phi \vee \psi$ is infelicitous if ψ entails ϕ, unless ψ contradicts ϕ together with the implicatures of ϕ. Chierchia et al (in press) argue, instead, that these sentence do not violate Hurford's constraint, because one of the disjuncts gives rise to an embedded scalar implicature, and that because of that, there actually is no entailment relation between the disjuncts.

(3) a. Either John did some but not all of the readings or he didn't do any of them.
b. Either only John came, or only Mary, or both
c. John had only 3 cookies, or only 4.

I agree with Chierchia et al (in press) that Gazdar's weakening of Hurford's constraint is rather *ad hoc*. Unfortunately, however, Chierchia et al. (in press) make crucial use of a *local* notion of implicature: They would represent a sentence like (2-b) making use of two silent 'only's: $only(p) \vee only(q) \vee (p \wedge q)$. Of course, this analysis accounts for the data mentioned above. However, only at a huge price: the idea is given up that implicatures are a pragmatic affair, and that they do not influence the semantic meaning, let alone the syntactic representation of the sentence. So the question arises whether we cannot explain in a truly pragmatic way why, on the one hand, (1) is inappropriate, without being forced to say that, on the other hand, (2-a)-(2-d) are inappropriate as well.

We agree that (1) is inappropriate. However, we would like to claim that (1) is inappropriate, not so much because the first disjunct is already entailed by the second (as Hurford's constraint has it), but rather because the second disjunct is *redundant*, because already mentioned as a possibility in the first ($\phi \vee \psi$ is inappropriate if $T(\psi) \subseteq T(\phi)$). That is, we would like to represent (1) by a formula of the form $(p_1 \vee \cdots p_i \cdots \vee p_n) \vee p_i$. Of course p_i entails $p_1 \vee \cdots p_i \cdots \vee p_n$, but more relevantly, we feel, is that $T(p_1 \vee \cdots p_i \cdots \vee p_n) = T((p_1 \vee \cdots p_i \cdots \vee p_n) \vee p_i)$: adding the new disjunction doesn't add a new way to make the sentence (minimally) true. Notice that the constraint is really different from Hurford's constraint which is too strong: (2-a)-(2-d) are falsely ruled out as inappropriate by Hurford's constraint, but not by the constraint we proposed: $\phi \vee \psi$ is inappropriate if $T(\psi) \not\subseteq T(\phi)$. The reason is

[6] To account for some related data, Hurford (1974) argued that 'or' is ambiguous between an inclusive and an exclusive reading. See Gazdar (1979) for a thorough criticism of this view.

that $T(p \wedge q)$ is not a subset of $T(p)$ (and neither is it the case that $T(p) \subseteq T(p \wedge q)$).

Notice that our pragmatic interpretation rule $PRAG_2$ mispredicts for (2-a)-(2-d). We predict that (2-b) is not only semantically equivalent to 'Either John came or Mary came', but pragmatically as well. And this is true for almost any global analysis of pragmatic interpretation, in particular when exhaustification is used.[7] But in terms of our pragmatic interpretation rule, we can say that it *almost* predicts correctly: we just have to *weaken* the force of the rule $PRAG_2$ slightly into $PRAG_2^*$ as follows (by changing $V = S$ into $V \subseteq S$):

- $PRAG_2^*(\phi) = \{\mathcal{M} | \ \exists S \in T(\phi) : \mathcal{M} \in [\![S]\!] \ \& \ \forall V \in T(\phi) : \mathcal{M} \in [\![V]\!] \to V \subseteq S\}$.

Suppose we talk about two atomic propositions, p and q, and thus four relevant models $\mathcal{M}, \mathcal{N}, \mathcal{K}, \mathcal{L}$ such that $\mathcal{V}_{\mathcal{M}}(p) = \mathcal{V}_{\mathcal{M}}(q) = 1$, $\mathcal{V}_{\mathcal{N}}(p) = 1, \mathcal{V}_{\mathcal{N}}(q) = 0$, $\mathcal{V}_{\mathcal{K}}(p) = 0, \mathcal{V}_{\mathcal{K}}(q) = 1$ and $\mathcal{V}_{\mathcal{L}}(p) = \mathcal{V}_{\mathcal{L}}(q) = 0$. Now, the sentences $p \vee q$ and $p \vee q \vee (p \wedge q)$ give rise to the same semantic meaning: $\{\mathcal{M}, \mathcal{N}, \mathcal{K}\}$, but to different sets of minimal truthmakers: $T(p \vee q) = \{\{p\}, \{q\}\}$, and $T(p \vee q \vee (p \wedge q)) = \{\{p\}, \{q\}, \{p, q\}\}$. Although according to our old pragmatic interpretation rule it holds that $PRAG_2(p \vee q) = \{\mathcal{N}, \mathcal{K}\} = PRAG_2(p \vee q \vee (p \wedge q))$, our new pragmatic interpretation rule predics that they also give rise to different pragmatic meanings: $PRAG_2^*(p \vee q) = \{\mathcal{N}, \mathcal{K}\}$, while $PRAG_2^*(p \vee q \vee (p \wedge q)) = \{\mathcal{M}, \mathcal{N}, \mathcal{K}\}$. This is exactly as desired.

The analysis even accounts for so-called 'intermediate implicatures': 'Every student read some of the books or every student read all books', showing that Sauerland's (2012) reason to adopt a grammatical approach to implicatures is not well-motivated.[8]

3 Reasoning by taking what was meant seriously

In Cobreros et al (2012, to appear) we defined the consequence relation $\models^{st}$ as going from strict to tolerant. It was shown that when restricted to the classical vocabulary, our logic is identical with classical logic, and is more permissive if we add (i) a distinguished similarity relation, or (ii) a truth-predicate to the language. In that case, (i) the tolerance principle $\forall x, y((Px \wedge x \sim_P y) \to Py)$ and (ii) Tarski's $\mathcal{T}$-equivalence $\mathcal{T}(\phi) \leftrightarrow \phi$ became valid, but the consequence relation became *non-transitive*.

An unfortunate consequence of using $\models^{st}$ as our consequence relation was that we were forced to make a distrinction between the Sorites reasoning with and without the tolerance principle as explicit premise. Without the principle as explicit premise we predicted in Cobreros et al (2012) that although each step in the argument is valid, the argument as a whole is *in*valid, because the arguments cannot be joined together. We felt, and still feel, that this is intuitively the correct diagnosis of of the Sorites paradox. However, in CERvR we had to claim that with the tolerance principle as

[7] Though see Schulz & van Rooij (2006) for an exception, making use of 'dymamic' exhaustification.

[8] That we can account for the phenomena discussed in the main text doesn't mean that we thereby have accounted for the pragmatics of disjunctive sentences in general. Singh (2008), for instance, observes that implicatures can obviate Hurford's constraint only in earlier disjuncts: **Either John and Mary came, or John, or Mary.* Perhaps Hurford's constraint is in play for left to right inferences between disjuncts. However, Singh also claims that disjunctive sentences are inappropriate not only if one disjunct entails another, but sometimes also when two disjucnts are mutually consistent, as in **John is from Russia or Asia.* A constraint for $\phi \vee \psi$ that $[\![\phi]\!] \cap [\![\psi]\!] = \emptyset$ seems much too strong (cf. (3-b)). However using $T(\cdot)$ we can state the much weaker demand that $T(\phi) \cap T(\psi) = \emptyset$. Sentences given by Singh of the form $(p \vee r) \vee (q \vee r)$ are now ruled out because the two disjuncts share a minimal truthmaker.

explicit premise, the argument is valid, but that one of the premises (i.e., the tolerance principle) is not true enough to be used as a premise in an $\models^{st}$ inference. Can we not come up with another consequence relation such that the Tolerance principle *can* be used as a substantial premise?

It turns out that we can, if we adopt the following notion of pragmatic inference (from *pr*agmatic to *t*olerant interpretation):

- $\Gamma \models^{prt} \psi \quad \text{iff}_{df} \quad \bigcap_{\phi \in \Gamma} PRAG(\phi) \subseteq [\![\psi]\!]^t$

Thus, for inference we take into account what is (pragmatically) meant by the premises. According to this notion of entailment it follows that $\phi \wedge \psi \models^{prt} \phi$ and also $\phi \models^{prt} \phi \vee \psi$, for any ϕ and ψ. The fact that we look at what was meant by the premises means that, even though $\phi \wedge \neg\phi \models^{prt} \phi$, it does not hold that $\phi \wedge \neg\phi \models^{prt} \psi$. Thus, explosion is not valid. In this sense, *prt*-entailment is a type of paraconsistent entailment relation. On the other hand, this notion coincides with *st*-entailment in case Γ is contradiction-free.[9]

Making use of the new consequence-relation, we can also diagnose the Sorites reasoning with the tolerance principle as explicit premise as invalid, even though all the steps are valid. The fact that the tolerance principle $\forall x_i, x_j(Px_i \wedge x_i \sim_P x_j) \rightarrow Px_j)$ (with $1 \leq i, j \leq n$) cannot be strictly true if both Px_1 and $\neg Px_n$ are taken as premises that are strictly true, does not rule out that it can be used appropriately in an inference where the premises are interpreted pragmatically.

Notice that, so far, we have not yet made full use of the pragmatic machinery developed in this paper. Our notion $\models^{prt}$ only looks for the pragmatic interpretation of the *premises*. What would result if we also interpreted the *conclusion* as strongly as possbile? Two new consequence relations immediately come to mind: if we abbreviate $\bigcap_{\phi \in \Gamma} PRAG(\phi, c)$ by $PRAG(\Gamma, c)$, the first one is:

- $\Gamma \models^{prpr_s} \psi \quad \text{iff}_{df} \quad \text{(for all } c\text{)}\ PRAG(\Gamma, c) \subseteq PRAG(\psi, c)$

Notice that in constrast to what is predicted with $\models^{prt}$, according to $\models^{prpr_s}$ the Tolerance-principle is not predicted to be valid anymore. It becomes a contingent propositions. Although this is certainly a defendable position, $\models^{prpr_s}$ is still a rather strange consequence relation. For one thing, because $\models^{prpr_s}$ does not satisfy conjunction-elimination, $p \wedge \neg p \not\models^{prpr_s} p$. Fortunately, we can do better by adopting a strategy Martin, Jeroen and Frank have used already for a long time: by going *dynamic*.

- $\Gamma \models^{prpr_d} \phi \quad \text{iff}_{df} \quad \text{(for all } c\text{)}\ PRAG(\Gamma, c) = PRAG(\psi, PRAG(\Gamma, c))$.

Interestingly, neither consequence-relation allows for explosion, i.e. $\phi \wedge \neg\phi \not\models^{prpr} \psi$, and also the law of excluded middle is not a validity anymore for either logic: $\not\models^{prpr} \phi \vee \neg\phi$. However, while $\models^{prpr_s}$ does not satisfy conjunction-elimination, $\models^{prpr_d}$ does: $p \wedge \neg p \not\models^{prpr_s} p$ but $p \wedge \neg p \models^{prpr_d} p$.[10]

When I started to think about vagueness, Frank pointed me to Wittgenstein's radical pragmatic solution to the problem posed by vagueness in his *Philosofische Untersuchungen*:[11] don't worry too much, because *in practice* contradiction is avoided.

[9] Observe that there exists a difference between ϕ, ψ as premises on the one hand, and the single premise $\phi \wedge \psi$ on the other: $p, \neg p \vee q \models^{prt} q$ but $p \wedge (\neg p \vee q) \not\models^{prt} q$. This is not unlike what happens in Veltman's (1996) Update semantics.

[10] Notice that $p \wedge \neg p \models^{prpr_d} p$ and $p \models^{prpr_d} p \vee \neg p$, and also $p \wedge \neg p \models^{prpr_d} p \vee \neg p$. On the other hand $p \models^{prpr_s} p \vee \neg p$, but $p \wedge \neg p \not\models^{prpr_s} p \vee \neg p$, because $p \wedge \neg p \not\models^{prpr_s} p$.

[11] See in particular section 85-87: 'A rule stands like a signpost ... The signpost in order in in normal circumstances it fulfils its purpose. See also Waismannn's (1968) notion of 'open texture'.

In normal discourse, we talk about relatively few objects, all of which are easily discernible from the others. In those circumstances, the tolerance principle will not give rise to inconsistency, but serves its purpose quite well. Only in exceptional situations i.e., when we are confronted with long sequences of pairwise indistinguishable objects — do things go wrong. But in such situations, we should not be using vague predicates like 'tall' but precisely measurable predicates involving, in this case, millimeters. I defended such an approach in my first paper on vagueness that I wrote in 2006, but that was only published in 2010 (van Rooij, 2010), and soon afterwards discovered that others came up with a very similar pragmatic solution (e.g. Rayo 2010, Pagin 2011).

I am still very much attracted to this view because of its naturalness. The hypothesis (the gap hypothesis) that we can appropriately use a predicate P in a context if and only if it helps to clearly demarcate the set of objects that have property P from those that do not seems to make good pragmatic sense: the division of the set of all relevant objects into those that do have property P and those that do not is (i) easy to make and (ii) worth making. Still, almost from the very beginning I also had my doubts: the gap-principle doesn't seem to be a necessary condition for the appropriate use of relative adjectives. Even if there is no clear demarcation between the bigger and the smalller entities of the domain, certainly the tallest object can be called 'tall', can't it? And similarly for the smallest object. This was the reason why I developped an alternative view as defended in van Rooij (2011) and Cobreros et al (2012a,b).

But still, isn't there something true about the gap-hypothesis? Isn't it still the view that is correct for normal cases? Can't we have an approach according to which we use and interpret sentences containing adjectives like 'tall' according to the gap-hypothesis in normal situations, but that we still allow for the use of such adjectives in abnormal situations? By making use of our dynamic consequence relation $\models^{prpr_d}$ we can, at least if we also make a difference between tolerant and strict truth of similarity statements as follows:

- $\mathcal{V}_{\mathcal{M}}(x \sim_P y) \quad =_{df} \quad 1 - |\mathcal{V}_{\mathcal{M}}(Px) - \mathcal{V}_{\mathcal{M}}(Py)|$.

We can immediately observe that with this meaning given to the similarity relation the following inferences are predicted to be valid:

1 Tolerance, $Px_1, \neg Px_n \models^{prpr} \neg(x_1 \sim_P x_2 \wedge \cdots \wedge x_{n-1} \sim_P x_n)$

2 (Tolerance,) $Px_1, x_1 \sim_P x_2 \cdots x_{n-1} \sim_P x_n \models^{prpr} Px_n$

(though not if one knows that $x_1, \cdots x_n$ are all the individuals). The following, however are *not* valid:

3 Tolerance, $Px_1, \neg Px_n, x_1 \sim_P x_2 \wedge \cdots \wedge x_{n-1} \sim_P x_n \not\models^{prpr} Px_n$

or even

4 Tolerance, $Px_1, \neg Px_n, x_1 \sim_P x_2 \wedge \cdots \wedge x_{n-1} \sim_P x_n \not\models^{prpr} Px_2$

Prediction 4 is of course really different from what CERvR predicted using $\models^{st}$. But the differences in 1 and 2 are also interesting:

- Prediction 1 basically says that if we have a sequence going from truth value 1 to 0, you expect this to be due to a gap (a pair x_i, x_j such that $x_i \not\sim_P x_j$). I take this to be in agreement with the gap-hypothesis suggested by Wittgenstein in the Philosophische Untersuchungen as noted by Veltman (1987).

- Prediction 2 means that if you explicitly say that $x \sim_P y$ (and do not say much more) then you expect that Px and Py have the same truth value: in this case, both value 1. The truth of the tolerance principle is not even needed for this prediction. I also think that this is natural, and is actually just the contrapositive side of prediction 1.

- Prediction 4 (and prediction 3) shows that this expectation can be cancelled if it is explicitly said that another individual in the transitive closure of the similarity relation (of course, the similarity relation is only transitively closed with respect to strict truth) doesn't have property P. But this shows that $\models^{prpr}$ is non-monotonic.

Notice that the Deduction Theorem is not valid for $\models^{prpr}$: $Px, x \sim y \models^{prpr} Py$ for any x and y. But $\not\models^{prpr} (Px \wedge x \sim y) \rightarrow Py$ for all x and y: if there is one model where $|\mathcal{V}_{\mathcal{M}}(Px) - \mathcal{V}_{\mathcal{M}}(Py)| = 0$ and another where $\mathcal{V}_{\mathcal{N}}(Px) = 1$ and $\mathcal{V}_{\mathcal{N}}(Py) = \frac{1}{2}$, then the latter model is not in $Prag((Px \wedge x \sim y) \rightarrow Py)$. This also shows, again, that the tolerance principle is not valid.

References

[1] Alxatib, S. and F.R. Pelletier (2011), 'The Psychology of Vagueness: Borderline Cases and Contradictions'. *Mind and Language*. **26**: 287-326.

[2] Alxatib, S., P. Pagin and U. Sauerland (2013), 'Acceptable contradictions: pragmatics or semantics? A reply to Cobreros et al., *Journal of Philosophical Logic*.

[3] Chierchia, G., D. Fox & B. Spector (ms), 'Hurford's constraint and the theory of scalar implicature: evidence for embedded implicatures'.

[4] Ciardelli, I., J. Groenendijk & F. Roelofsen (2013), 'Inquisitive semantics: a new notion of meaning', *Language and Linguistic Compass*, **7**, 9, 459-476.

[5] Cobreros, P., P. Egre, D. Ripley, and R. van Rooij, (2012a), 'Tolerant, classical, strict', *Journal of Philosophical Logic*, **41**: 347-385.

[6] Cobreros, P., P. Egre, D. Ripley, and R. van Rooij, (2012b), 'Tolerance and mixed consequence in a S'valuationist setting', *Studia logica*, **100**: 853-876.

[7] Cobreros, P., P. Egre, D. Ripley, and R. van Rooij, (to appear), 'Reaching transparent truth', *Mind*.

[8] Cobreros, P., P. Egre, D. Ripley, and R. van Rooij, (2013),'Vagueness, truth and permissive entailment', in Achourioti et al. (eds.), *Unifying the philosophy of truth*, Springer.

[9] Deemter, K. van (1995) 'The Sorites fallacy and the context-dependence of vague predicates', in M. Kanazawa, et al., *Quantifiers, Deduction, and Context*, CSLI Publications, Stanford, Ca., pp. 59-86.

[10] Fine, K. (1975), 'Vagueness, truth, and logic', *Synthese*, **30**, 265300.

[11] Fraassen, B. van (1969), 'Facts and tautological consequence', *Journal of Philosophy*, **66**: 477-487.

[12] Gaifman, H. (1997), 'Vagueness, tolerance and contextual logic', manuscript, Columbia University.

[13] Gazdar, G. (1979), *Pragmatics: Implicature, Presupposition and Logical Form*, New York: Academic Press.

[14] Grice, H.P. (1967), 'Logic and conversation', *William James Lectures*, Harvard University, reprinted in *Studies in the Way of Words*, 1989, Harvard University Press, Cambridge, Massachusetts.

[15] Groenendijk, J. and M. Stokhof (1984), *Studies in the Semantics of Questions and the Pragmatics of Answers*, Ph.D. thesis, University of Amsterdam.

[16] Hurford, J. R. (1974), 'Exclusive or inclusive disjunction', *Foundations of Language*, **11**: 3, 409-411.

[17] Kamp, H. (1981), 'The paradox of the heap', in U. Mönnich (ed.), *Aspects of Philosophical Logic*, D. Reidel, Dordrecht, pp. 225-277.

[18] Kraus, S., D. Lehman and M. Magidor (1990), 'Nonmonotonic reasoning, preferential models and cumulative logics', *Artificial Intelligence*, **4**: 167-207.

[19] Lewis, C.I. (1918), *A Survey of Symbolic Logic*, Berkeley, Calif.

[20] McCarthy, J. (1980), 'Circumscription - a form of non-monotonic reasoning', *Artificial Intelligence*, **13**: 27-39.

[21] Pagin, P. (2011), 'Vagueness and domain restriction', in P. Egré and N. Klinedinst (eds.), *Vagueness and Language Use*, Palgrave MacMillan, Basingstroke, pp. 283-307.

[22] Pinkal, M. (1984), 'Consistency and context change: The Sorites paradox', in F. Landman & F. Veltman (eds.), *Varieties of Formal Semantics, Proceedings of the fourth Amsterdam Colloquium*, Groningen-Amsterdam Studies in Semantics, Foris Publications, Dordrecht.

[23] Priest, G. (1979), 'The logic of paradox', *Journal of Philosophical Logic*, **8**: 219-241.

[24] Priest, G. (1991), 'Minimally inconsistent LP', *Studia Logica*, **50**: 321-331.

[25] Rayo, A. (2010), 'A metasemantic account of vagueness', in R. Dietz and S. Moruzzi (eds.), *Cuts and Clouds: Vagueness, its Nature and its Logic*, Oxford University Press, Oxford, pp. 23-45.

[26] Ripley, D. (2011). 'Contradictions at the borders', In R. Nouwen, R. van Rooij, H.-C. Schmitz, & U. Sauerland (Eds.), (2012), *Vagueness in communication*, Berlin: LICS, Springer.

[27] Rooij, R. van (2010), 'Vagueness and linguistics', in Ronzitti (ed.), *Vagueness: A Guide*, 2010, Springer, pp. 122-170.

[28] Rooij, R. van (2012), 'Vagueness, tolerance and non-transitive entailment', in P. Cintula, C Fermüller, L. Godo & P. Hajek (eds.), *Reasoning Under Vagueness: Logical, Philosophical, and Linguistic Perspectives*, College Publications, 2012, pp. 205-223.

[29] Rooij, R. van and K. Schulz (2004), ‘Exhaustive interpretation of complex sentences’, *Journal of Logic, Language, and Information*, **13**: 491-519.

[30] Sauerland, U. (2012), ‘The computation of scalar implicatures:pragmatic, lexical or grammatical?’, *Language and Linguistic Compass*, **6**: pp. 36-49.

[31] Schulz, K. & R. van Rooij (2006), ‘Pragmatic Meaning and Non-monotonic Reasoning: The Case of Exhaustive Interpretation’, *Linguistics and Philosophy*, 2006, **29**: 205-250.

[32] Serchuk, P., Hargreaves, I., & Zach, R. (2011), ‘Vagueness, logic and use: Four experimental studies on vagueness’, *Mind and Language*, **26**: 540-573.

[33] Singh, R. (2008), ‘On the interpretation of disjunction: asymmetric, incremental, and eager for inconsistency’, *Linguistics and Philosophy*, **31**: 245-260.

[34] Stokhof, M. (2008), ‘The architecture of meaning: Wittgenstein’s *Tractatus* and formal semantics’, In: D. Levy & E. Zamurer (eds.), *Wittgenstein’s Enduring Arguments*, London, Routledge, pp. 201-244.

[35] Veltman, F. (1985), *Logics for Conditionals*, PhD thesis, Universiteit van Amsterdam.

[36] Veltman, F. (1996), ‘Defaults in update semantics’, *Journal of Philosophical Logic*, **25**: 221-261.

[37] Veltman, F. (1987), *Syllabus Logische Analyse 2: Vaagheid*, Universiteit van Amsterdam.

[38] Wittgenstein, L. (1921), ‘Tractatus Logico-Philosophicus, *Annalen der Naturphilosophische*, **19**.

[39] Wittgenstein, L. (1953), *Philosophische Untersuchungen/Philosophical Investigations*, translated by G.E.M. Anscombe, Blackwell, Oxford.‘,