Reevaluating Evaluativity

XXX
YYY

1 Introduction

An **evaluative** degree construction is one that allows the inference to the positive (Bierwisch 1989). For example, the positive predication of a gradable predicate like *tall* in *Peter is tall* is - trivially - evaluative. A more interesting case, and the topic of the following discussion, is evaluativity triggered by negative antonymity.

(1) a. Peter is as short as Mary.
b. How short is Peter?
c. However short Peter is, he is tall enough to ride the roller coaster.
d. Peter is that short.

All of the examples under (1) allow the inference to the positive, namely to *Peter is short*. Here negative antonyms contrast with their positive counterparts which do not allow such an inference (2a). However negative antonymity is not in itself a sufficient condition for evaluativity. In other construction types, such as the comparative, we don't observe the inference to the positive use case (2b).

(2) a. Peter is as tall as Mary.
$\not\rightsquigarrow$ Peter is tall.
b. Peter is shorter than Mary.
$\not\rightsquigarrow$ Peter is short.

Thus we have to take account of negative antonym evaluativity as a product both of the construction type as well as the polarity of the adjective. The most explicit account proposed in the literature is the **statistical model** outlined in Rett (2014), Bumford & Rett (2021). Here evaluativity is construed as a **manner implicature**. The following proposal shall reevaluate this derivation. It is shown that the statistical model, following the principle of the **division of pragmatic labor** (Horn 1984, Levinson 2000), makes wrong predictions about positive antonym constructions. A countermodel is proposed that derives evaluativity by considerations about the ontology of degrees and the method of measurement.

2 The statistical model

In this section I first quickly review the notion of manner implicature or M-implicature. I then show how the statistical model assumes that negative antonym evaluativity can be calculated as an M-implicature. I conclude this section by pointing out what goes wrong with the statistical model.

2.1 Manner implicatures

Informally M-implicatures can be charaterized as the correlation of marked structure with marked meaning. A standard example is the contrast between the lexicalized causative *to kill* and the periphrastic causative *to cause x to die*. Although the two expressions are arguably truth conditionally equivalent, they lead to different implicatures. Whereas the first implicates that the act was done intentionally, the second implicates that the act was done unintentionally.

(3) a. Peter killed Paul.
$\rightsquigarrow$ He did it intentionally.
b. Peter caused Paul to die.
$\rightsquigarrow$ He did it unintentionally.

The derivation of M-implicatures can be cast as a form of (neo-)Gricean reasoning (Grice 1975, Horn 1984, Levinson 2000). By general considerations about effective communication and cooperativity between speech participants, the literal content of what has been said is enriched by reflection on what has *not* been said. For the well-known case of quantity implicatures, the content of *Some students smoke* gets enriched to *Some but not all students smoke* by counterfactual reasoning on the part of the listener about the speakers choice not to say *All students smoke*. Where for Q-implicatures the set of alternatives is determined by holding constant structural complexity (so to exclude *Some but not all students* from the set of alternatives) but varying over truth conditions, for M-implicature the reverse is the case. If the speaker chose a more verbose expression over one less so, without changing literal content, she must have had a reason for doing so. Rett (2020) gives the following encapsulation of reasoning over markedness.

(4) THE MARKED MEANING PRINCIPLE
For structures ϕ, ϕ' such that ϕ and ϕ' are truth conditionally equivalent ϕ' is a less complex structural alternative to ϕ, ϕ carries the manner implicature "& atypcial(ϕ)"

For a formal implementation of the notion of structural complexity, refer to Katzir (2007). One further remark is in place. The notion of typicality evoked here is not a content constituting notion, but rather a content regulating or licensing notion. What dimension of meaning the predicate "typical" targets has to be calculated case or construction specific (Levinson 2000).

2.2 Outline of the statistical model

Under the statistical model the derivation of negative antonym evaluativity as an M-implicature proceeds by the establishment of three conditions. First, for all the constructions exemplified under (1) positive and negative antonyms have equivalent truth conditional import. I will demonstrate this with the example of the equative. The basic outline of a compositional semantics for constructions of gradability (von Stechow 1984, Rullmann 1995, Beck 2011) assumes gradable predicates to be relations between individuals and degrees (measurement values belonging to a certain scale, like the scale of height, age, beauty, etc.). Furthermore, antonyms are only differentiated by the direction of comparison ($\geq$ *vs.* $\leq$). Lastly comparison constructions, such as the equative, are analyzed as quantifiers over degrees.

(5) a. $[\![\text{tall}]\!] = \lambda d \lambda x.HEIGHT(x) \geq d$
b. $[\![\text{short}]\!] = \lambda d \lambda x.HEIGHT(x) \leq d$

(6) a. Peter is as tall as Mary.
b. $[\![[[\text{as } [\lambda_1[\text{Mary is } d_1 \text{ tall}]]][\lambda_2[\text{Peter is } d_2 \text{ tall}]]]]\!] = 1$ iff
$\{d : HEIGHT(M) \geq d\} = \{d : HEIGHT(P) \geq d\}$

(7) a. Peter is as short as Mary.
b. $[\![[[\text{as } [\lambda_1 \text{ [Mary is } d_1 \text{ short}]][\lambda_2 \text{ [Peter is } d_2 \text{ short}]]]]\!] = 1$ iff
$\{d : HEIGHT(M) \geq d\} = \{d : HEIGHT(P) \geq d\}$

The two truth conditions in (6) and (7) are, assuming linear order between degrees, provably equivalent. If Mary and Peter share all heights they reach, they necessarily share all heights they don't reach.[1]

The second condition in the derivation of evaluativity is the establishment of an order of structural markedness or complexity between positive and negative antonyms. I will not go into the details of this discussion here and will take the point for granted. For a discussion, compare Rett (2014), Moracchini (2018).

The last step is deriving evaluativity as the meaning contribution of atypicality. In the statistical model, this is done by considerations about the positive use form of relative adjectives. Following Kennedy (2007) we assume that relative adjectives in their positive use case compare the measurement values of individuals against a contextually specified standard value.

(8) a. Peter is tall.
b. $HEIGHT(P) \geq d_{s(tall)}$

The determination of $d_{s(tall)}$ can now be construed as based on the statistical distribution of heights (Solt & Gotzner 2012). If we regard the adult population of some country as the relevant context against which to evaluate whether Peter counts as tall or not, then $d_{s(tall)}$ is determined by the statistical normal distribution of heights over this population. Roughly speaking, someone counts as short if her height is somewhere left of the bell curve peak of the normal distribution, she counts as neither short nor tall if her height is somewhere around the peak and as tall if on the right side of it.

(9) HEIGHT: 0 ——short—— $|_{d_{s(short)}}$ ——neither short nor tall—— $|_{d_{s(tall)}}$ ——tall—— $> \infty$

Since what defines these regions is precisely the typicality of the measurement values they delimit, evaluativity can be cast as an inference about which region of the scale the adjective refers to.[2]

(10) a. Peter is as short as Mary.
b. $\{d : HEIGHT(M) \geq d\} = \{d : HEIGHT(P) \geq d\} \wedge \underline{HEIGHT(M) \leq d_{s(short)}}$

2.3 Problems with the statistical model

What the statistical model misses is the **division of pragmatic labor**. This is a term coined by Horn (1984), indicating that M-implicatures are parasitic on **I-implicatures**. An I-implicature - or informativity implicature - is the inference from unmarked structures to unmarked meanings. The reasoning behind I-implicatures builds on the assumption that, if something were special or abnormal about the situation reported, the speaker would have indicated so by explicit mention. If the speaker didn't, then the hearer is licensed to infer from the lack of indication to the lack of abnormalness. The MARKED MEANING PRINCIPLE of (11) is complemented by an UNMARKED MEANING PRINCIPLE.

(11) THE UNMARKED MEANING PRINCIPLE
For structures ϕ, ϕ' such that ϕ and ϕ' are truth conditionally equivalent and ϕ' is a more complex structural alternative to ϕ, ϕ carries the informativity implicature "& typcial(ϕ)"

For illustration of an I-implicature and the relationship to M-implicatures, look at (3). If this principle generally holds true, then we should expect positive antonym equatives to come with a typicality inference. The problem is that, under the statistical model, typicality is encoded as a constraint on

[1] For ease of exposition, I make the simplifying assumption that equatives are analyzed as *exactly*-relations, not as *at least*-relations. Rett (2014) herself makes this assumption. In consequence however she faces the problem that truth conditional equivalence is not validated by an *at least*-construal. This problem has been solved in Bumford & Rett (2021). The proposal nonetheless faces the same challenge that will be pointed out in section 2.3.

[2] Why precisely the strengthening does not happen in the direction of the atypical range above the standard for tallness is left unexplained by Rett (2014).

the range of measurement values the gradable predicate refers to. Thus if the domain of reference for *short* in *Peter is as short as Mary* is restricted to the atypical range of degrees below the standard for shortness, then - by I-implicature - *tall* in *Peter is as tall as Mary* should be restricted to the typical range of heights above the standard for shortness. But of course no such implicature arises. The sentence *Peter is as tall as Mary* just doesn't come with any generalized implicature about how tall or short Peter and Mary are.[3]

3 The deviation model

In this section I outline a model that encompasses the problem identified above. The perspective will shift from dividing up scale segments to dividing up ways of measuring. The central claim is that evaluativity is correlated with measuring deviations instead of absolute values. This proposal is inspired by Kennedy (2001).

3.1 Comparisons of deviation

Kennedy (2001) points out a type of degree construction he calls **comparison of deviation** (COD). Constructions of this type contain two antonymically related adjectives.

(12) a. The Red Sox are more legitimate than the Orioles are fraudulent.
b. This year Grace was as cold as he was hot in the 1989 playoffs.

Two things to note about COD constructions. Firstly, they are not accounted for under a standard degree approach as presented in the last section. (12b), for example, would be predicted to be true iff Grace performed equally well this year as he did in the 1989 playoffs. But clearly the sentence expresses some qualitative difference between Grace's performance then and now. Secondly, COD constructions are evaluative. Thus from (12a) I can both infer that the Red Sox are legitimate as well as that the Orioles are fraudulent.

Kennedy argues that COD constructions compare, not absolute measures, but rather **deviation measures**. He gives the following paraphrase for the meaning of (12a).

(13) The degree to which the Red Sox exceed a standard of legitimacy is greater than the degree to which the Orioles exceed a standard of fraudulence.

With this paraphrase we see that evaluativity is an effect of the way of measuring. By measuring how much something deviates from a given standard, . I want to propose that we understand the evaluative degree constructions discussed in the previous sections in the same way. That is, a sentence like (1a) has the following paraphrase.

(14) The degree to which Peter exceeds a standard of shortness is greater than the degree to which Mary exceeds a standard of shortness.

The distinction between deviation measures and absolute measures is the dimension targeted by the

[3] To convince you of this, we can apply the *Are you saying...* diagnostic developed by Meyer (2013).

(i) a. A: Some students smoke.
B: Are you saying that not all students smoke? ✓
b. A: Peter is as short as Mary.
B: Are you saying that Mary is short? ✓
c. A: Peter is as tall as Mary.
B: Are you saying that Mary is normal in height? ×

M-implicature. Typically measures of height are construed as absolute values. Atypically however they can be construed as deviation measures.

3.2 Implementation

Kennedy (2001) proposes to formalize the deviation intuition by switching from a point based to an interval based degree semantics. This switch allows us to compare lengths. In the following I present a modified version of Kennedy's interval semantics.

Firstly, an interval is simply a convex set of degrees belonging to a certain scale S. On the basis of S we can define the set of positive and negative intervals as follows.

(15) a. $Pos(S, d_R) := \{I \subseteq S : \exists d_1 \in I \forall d_2 \in S[d_R \leq d_2 \leq d_1 \leftrightarrow d_2 \in I]\}$
b. $Neg(S, d_R) := \{I \subseteq S : \exists d_1 \in I \forall d_2 \in S[d_1 \leq d_2 \leq d_R \leftrightarrow d_2 \in I]\}$

d_R is a reference value which determines either the lower limit (in the case of positive antonyms) or the upper limit (in the case of negative antonyms). Antonym pairs are correlated with different interval domains. Positive adjectives relate individuals to positive intervals such that the relation expressed by *tall*, for example, holds between an individual and an interval iff the individual's height falls above the maximal element of the interval (16a). Negative adjectives relate individuals to negative intervals such that the individual's height falls below the minimal element of the interval (16b). Comparison constructions are analyzed as comparing interval lengths. The length of an interval I, with lower limit d_l and upper limit d_u, is defined as $d_u - d_l$. Furthermore we define an order relation $\geq$ between intervals such that for any two intervals I_1 and I_2, $I_1 \geq I_2$ iff $LENGTH(I_1) \geq LENGTH(I_2)$. Lastly I define a function $LONGEST$ which maps a set of intervals to its longest member.

(16) a. $[\![\text{tall}]\!] = \lambda I \lambda x : I \in Pos(S_{Height}, d_R).MAX(I) \geq HEIGHT(x)$
b. $[\![\text{short}]\!] = \lambda I \lambda x : I \in Neg(S_{Height}, d_R).MIN(I) \leq HEIGHT(x)$
c. $LONGEST = \lambda \mathscr{I}_{\langle dt,t \rangle}.\iota I[I \in \mathscr{I} \wedge \neg \exists I' \in \mathscr{I}[I' \geq I \wedge I' \neq I]]$

(17) a. Peter is as tall as Mary.
b. $[\![[[\text{as } [\lambda_1[\text{Mary is } I_1 \text{ tall}]]][\lambda_2[\text{Peter is } I_2 \text{ tall}]]]]\!] = 1$ iff
$LONGEST(\lambda I.\textbf{tall}_{d_R}(I)(M)) \geq LONGEST(\lambda I.\textbf{tall}_{d_R}(I)(P)) \wedge \underline{d_R = 0}$

(18) a. Peter is as short as Mary.
b. $[\![[[\text{as } [\lambda_1[\text{Mary is } I_1 \text{ short}]]][\lambda_2[\text{Peter is } I_2 \text{ short}]]]]\!] = 1$ iff
$LONGEST(\lambda I.\textbf{short}_{d_R}(I)(M)) \geq LONGEST(\lambda I.\textbf{short}_{d_R}(I)(P)) \wedge \underline{d_R = d_{s(short)}}$

The underlined sections point out the addition of pragmatic information. In (17) the I-implicature sets d_R to 0, whereas in (18) the M-implicature sets d_R to the standard for shortness $d_{s(short)}$

4 Conclusion

The preceding pages have outlined, first the case of negative antonym evaluativity; second, the statistical model introduced by Rett (2014), Bumford & Rett (2021); and third, a new model that gives a different derivation of evaluativity as M-implicature. The statistical model fails to account for the predicted I-implicature upon which the M-implicature is parasitic. If negative antonym constructions are restricted to atypical scale ranges, positive antonym constructions should be restricted to typical ranges. In the deviation model the missing restriction is explained by linking typicality to the way of measuring. Typical is a way of measuring from 0 upwards. Atypical the measuring from some standard value downwards. Together these considerations allow an account of evaluativity that accommodates the division of pragmatic labor.

References

Beck, Sigrid. 2011. Comparison constructions. In Klaus von Heusinger, Claudia Maienborn & Paul Portner (eds.), *Semantics: an international handbook of natural language meaning*, 2–1341. De Gruyter Mouton.

Bierwisch, Manfred. 1989. The semantics of gradation. *Dimensional adjectives* 71(261). 35.

Bumford, Dylan & Jessica Rett. 2021. Rationalizing evaluativity. *Proceedings of Sinn und Bedeutung* 25. 187–204.

Grice, Herbert P. 1975. Logic and conversation. In *Speech acts*, 41–58. Brill.

Horn, Laurence. 1984. Towards a new taxonomy for pragmatic inference: q-and r-based implicature. *Meaning, form and use in context*.

Katzir, Roni. 2007. Structurally-defined alternatives. *Linguistics and Philosophy* 30. 669–690.

Kennedy, Christopher. 2001. Polar opposition and the ontology of 'degrees'. *Linguistics and philosophy* 24. 33–70.

Kennedy, Christopher. 2007. Vagueness and grammar: the semantics of relative and absolute gradable adjectives. *Linguistics and philosophy* 30. 1–45.

Levinson, Stephen C. 2000. *Presumptive meanings: the theory of generalized conversational implicature*. MIT press.

Meyer, Marie-Christine. 2013. *Ignorance and grammar*. Massachusetts Institute of Technology dissertation.

Moracchini, Sophie. 2018. Evaluativity and structural competition. In *Semantics and linguistic theory*, 727–746.

Rett, Jessica. 2014. *The semantics of evaluativity* (Oxford Studies in Theoretical Linguistics). London, England: Oxford University Press.

Rett, Jessica. 2020. Manner implicatures and how to spot them. *International Review of Pragmatics* 12. 44–79.

Rullmann, Hotze. 1995. *Maximality in the semantics ofwh-constructions*. University of Massachusetts Amherst.

Solt, Stephanie & Nicole Gotzner. 2012. Who here is tall? comparison classes, standards and scales. *Sfb* 833. 79–83.

von Stechow, Arnim. 1984. Comparing semantic theories of comparison. *Journal of semantics* 3(1-2). 1–77.