Consequences of an implicature-based account of evaluativity*

New York Philosophy of Language Workshop

Jessica Rett
rett@ucla.edu

October 16, 2017

Abstract

An adjectival construction is evaluative iff it entails that an individual hold a gradable property to a significantly high degree. Positive constructions like *John is short* are evaluative, and so are equatives like *John is as short as Mary*. But evaluativity cannot be encoded lexically because not all constructions containing e.g. short are evaluative (cf. *John is shorter than Mary*). And evaluativity cannot be assigned to particular constructions because not all e.g. equatives are evaluative (cf. *John is as tall as Mary*). I will review arguments from Rett 2015 that evaluativity arises as a conversational implicature, the flouting of the maxims of Quantity (in the case of positive constructions) and Manner (in other cases). I'll then argue that evaluativity as a conversational implicature places significant, novel constraints on recent attempts to use game-theoretic pragmatics to model the semantics of adjectival constructions (Lassiter and Goodman, 2013, 2015) and Manner implicature (Bergen et al., 2016).

1 Introduction

- A traditional problem (see Cresswell, 1976; Klein, 1980):

(1) a. Adam is tall. *positive construction*
b. Adam is 5ft tall. *measure phrase (MP) construction*

- MP constructions like (1-b) make gradable adjectives seem like transitive counterparts to non-gradable adjectives, cf. **Adam is 100-spot freckled.*

(2) a. $[\![\text{freckled}]\!] = \lambda x.\mathsf{freckled}(x)$
b. $[\![\text{tall}]\!] = \lambda d\lambda x.\mathsf{tall}(x, d)$

- This account of the "degree argument" in (1-b) results in two problems with the compositional analysis for (1-a): we predict (1-a) doesn't denote a proposition, and we fail to predict that it is evaluative.

(3) $[\![\text{tall(Adam)}]\!] = \lambda d.\mathsf{tall}(\mathsf{a}, d)$ *(with argument suppression)*

(4) A sentence is **evaluative** with respect to a gradable adjective $\mathcal{G}$ iff it entails that some individual instantiate $\mathcal{G}$ to a significant degree.

(5) a. Adam is tall. $\rightarrow$ Adam is not short.
b. Adam is 5ft tall. $\nrightarrow$ Adam is not short.

*Thanks to audiences at Michigan State University; PhLiP 2016; and the University of Arizona. Thanks especially to Chris Barker, Heather Burnett, Sam Cumming, Chris Kennedy, Michael Franke, and David Wolfsdorf for discussion.

- The response (beginning with Bartsch and Vennemann, 1972; Cresswell, 1976): a null operator that simultaneously 1) contributes evaluativity; and 2) binds (or values) the degree argument.

  (6) POS $\rightarrow \lambda\mathcal{G}_{\langle d,\langle e,t\rangle\rangle}\lambda x\exists d[\mathcal{G}(x,d) \wedge d > s]$, for some contextually valued standard s

- Which, in turn, presents a compositionality problem.

  > "As far as I can tell, there is no independent justification for introducing POS; it is merely a device for fixing up the semantics." (Klein, 1980, 3)

  > "The operator "positively," call it POS, is invisible, which made E. Klein think that it doesn't exist." (von Stechow, 1984, 59)

- However, unlike other null operators (e.g. *pro*), no language has an overt counterpart of POS

  > "The failure of the putative operator POS to correspond to overt material in language after language – in stark contrast to [comparative markers] – could be taken as evidence that POS does not actually exist and hence that any approach in which [evaluativity] requires a special operator is misguided." (Grano, 2012, 515)

- Worse, the POS account is wrong (Bierwisch, 1989; Rett, 2007, 2008): it incorrectly predicts evaluativity and overt degree-argument manipulation (valuation, binding) are in complementary distribution.
  - constructions that comply with this prediction:[1]

    (7) a. Adam is tall. *positive construction*
    b. Adam is 5ft tall. *(standard) MP construction*
    c. Adam is that tall (too). *positive-antonym degree demonstrative*
    d. How tall is Adam? *positive-antonym degree question*
    e. Adam is taller than Doug. *(synthetic) positive-antonym comparative*
    f. Adam is as tall as Doug. *positive-antonym equative*

  - constructions that do not comply with this prediction:

    (8) a. Adam is that short (too). *negative-antonym degree demonstrative*
    b. How short is Adam? *negative-antonym degree question*
    c. Adam is as short as Doug. *negative-antonym equative*

  - some other data worth noting (see Rett, 2015):

    (9) a. Adam is more tall than Doug. *analytic comparative*; Matushansky (2001)
    b. Tennis balls are as heavy as 60g. *MP equative*; Rett (2014)
    c. The soup cooled for/in an hour. *degree achievement*; Hay et al. (1999)
    d. He has considerable money. *evaluative DP*; Bolinger (1972); Morzycki (2009)
    e. He $[\text{cooks}]_F$, believe you me! *evaluative VP*; ibid.

- a few unsatisfying responses:
  - lexically encoding evaluativity in relative adjectives... opposite, and harder, problem
  - a denial that the data in (7) and those in (8) exhibit the same phenomenon (various, p.c.)
  - a different null operator, EVAL (Rett, 2007, 2008), which is somehow attracted to marked antonyms but only in constructions in which antonymic pairs are otherwise mutually entailing
  - a distinct, optional, verbal POS (for degree achievements; Piñón, 2005; Kennedy and Levin, 2008)
  - a metaphysics-based analysis of evaluativity that focuses on antonymic pairs but does not speak towards the contrast between the data in (7) and those in (8) (Sassoon, 2011)

[1] I will only discuss relative adjectives (e.g. *tall, short*) here; other adjectives (e.g. *gorgeous, ugly*) display different, less complicated patterns of evaluativity (Rett, 2015).

2 Examples and diagnostics of conversational implicature

- Uninformativity-based Quantity implicatures (Grice, 1975; Geurts, 2011):

  (10) a. War is war.
  b. Boys will be boys.

  > "[A]t the level of what is said, [...tautologies] are totally noninformative and so...cannot but infringe the first maxim of Quantity in any conversational context. ...[T]he hearer's identification of their informative content at this level [of implicature] is dependent on his ability to explain the speakers selection of this particular patent tautology" (Grice 1975: 52).

- Manner implicatures: (Horn, 1989, 1991; Levinson, 2000)

  (11) a. not uncommon, not infrequent *litotes*
  b. He caused the sheriff to die. *periphrasis*

  > "[T]wo negatives...do not... cancel one another out...the longer expression is always weaker: *This is not unknown to me* or *I am not ignorant of this* means 'I am to some extent aware of it,' etc. The psychological reason for this is that the détour through the two mutually destructive negatives weakens the mental energy of the listener and implies...a hesitation which is absent from the blunt, outspoken *common* or *known*" (Jespersen, 1965, 332).

- Grice claimed that conversational implicatures are **cancellable** (e.g. (12)) and **detachable**...

  (12) Adam ate some of the pizza... in fact, he ate all of it.

- ...But, in fact, neither is a necessary property of all types of implicature (Hirschberg, 1991)
  - Manner implicatures don't seem to be cancellable or detachable ("insofar as the manner of expression plays no role in the calculation, it will not be possible to find another way of saying the same thing, which simply lacks the implicature in question"; Grice 1975:58):

    (13) a. I am not ignorant of this... #in fact, I am unaware of it.
    b. Adam caused the sheriff to die... #in fact, he killed him outright.

  - Nor are uninformativity-based Quantity implicatures (Huitink and Spenader, 2004):

    (14) a. War is war... #in fact, violence is avoidable in war.
    b. Boys will be boys... #in fact, they often behave unpredictably.

- How, then, can we identify implicatures?
  - Calculability: conversational implicatures have traditionally been distinguished from e.g. conventional implicatures in our ability to determine their meaning via pragmatic reasoning
  - Traditional differentiation between conversational implicature and entailed content (e.g. presupposition) wrt reinforceability with a conjunction (Horn, 1972) or something else (Sadock, 1978).
    * Horn's "Redundancy of Conjunction" differentiates presupposition and scalar implicature

      (15) a. #John is a bachelor and he is a man.
      b. Some people left early but not everybody did.

    * This seems to be a property of uninformativity-based Quantity and Manner implicatures, too

      (16) a. War is war... which is to say, violence is unavoidable in war.
      b. Adam caused the sheriff to die... which is to say, he killed him indirectly.

- The semantic status of implicatures varies with their discourse status:
  - At-issue content addresses the Question Under Discussion (QUD, Roberts, 1990), and doesn't project (Simons et al., 2011);
  - Not-at-issue content does not address the QUD, and does project.
  - If an implicature (even a scalar one) is at-issue, it isn't cancellable (van Kuppevelt, 1995).

    (17) A: Who passed some exams?
    B: John. In fact, he passed all of them.

    (18) A: How many exams did John pass?
    B: Some. #In fact, he passed all of them.

  - If an implicature (even a Manner one) has an 'anaphoric link', it is (van der Sandt, 1992).

    (19) A: Did John meet a woman at the bar last night?
    B: Yes (he met a woman), in fact he met his wife.

    (20) A: Adam looks quite shaken. Did he somehow cause the sheriff to die?
    B: Yes (Adam caused the sheriff to die), in fact he murdered him.

- And different types of implicatures seem to differ in their ablity to be interpreted locally
  - Scalar implicatures are interpreted locally (Chierchia, 2004)

    (21) John believes [$_{CP}$ that some students are waiting for him.]
    a. John believes that not every student is waiting for him. *local*
    b. #It is not the case that John believes that every student is waiting for him. *global*

  - Uninformativity-based Quantity implicatures, as well, must be interpreted locally:

    (22) A: Adam seems unfazed by the fact that he is a victim of a pyramid scheme.
    B: Well, you know, Adam has always believed that bankers will be bankers.

  - In contrast, Manner implicatures are only optionally embeddable (Bergen et al., 2016):

    (23) The judge believes that Adam caused the sheriff to die.
    a. Judge believes Adam was indirectly responsible for the sheriff's death *local*
    b. The speaker is being indirect for reasons of politeness or delicacy *global*

IMPLICATURE TYPE	**cancellable?**	**detachable?**	**reinforceable?**	**QUD-sensitive?**	**embeddable?**
scalar	✓	✓	✓	✓	✓
uninform. Quantity	x	x	✓	✓	✓
Manner	x	x	✓	✓	(✓)

Table 1: Properties of implicature

3 Evaluativity as conversational implicature

- Broad claim: evaluativity arises where it does, across languages, due to conversational implicature
  - In "degree tautologies" like the positive construction, it is an uninformativity-based Q implicature.
  - When it is tied to marked forms – e.g. negative antonyms – it arises as a Manner implicature.
  - Evaluativity seems presuppositional in some cases and assertoric in others because of the way we interpret various constructions out-of-the-blue... but all instances of evaluativity are at-issue or not-at-issue depending on the QUD (which is what we would expect if it were an implicature).
  - Evaluativity varies across adjective classes due to the relationship between antonyms in a given antonymic pair in a given language.

3.1 Some formalism

- Horn (1984) recasts Grice's maxims into a Q Principle (Make your contribution sufficient, say as much as you can) and an R Principle (Make your contribution necessary, say no more than you must). This "division of pragmatic labor" results in an equilibrium, namely:

(24) **Quantity implicatures**
   a. involve Q scales, which hold fixed markedness and order elements wrt informativity
   b. are calculated relative to the Q principle

(25) **Manner implicatures**
   a. involve M scales, which hold fixed informativity and order elements wrt markedness
   b. are calculated relative to the R principle (when a marked form is used) or the Q Principle (when an unmarked form is used)

(26) **Horn's Principle of Least Effort** (Horn, 1984, 22)
The use of a marked (relatively complex [...]) expression when a corresponding unmarked (simpler, less "effortful") alternate expression is available tends to be interpreted as conveying a marked message (one which the unmarked alternative would not or could not have conveyed).

- Katzir (2007) defines Q-alternatives based on 'parse trees': effectively trees with explicit morphology; he redefines the Q Principle accordingly, with $<, \sim$ denoting relations of structural complexity

(27) Q-ALTERNATIVES (Katzir, 2007, 679, modified slightly): Let ϕ be a parse tree. The set of **Q-alternatives** for ϕ, written as $A_Q(\phi)$, is defined as $A_Q(\phi) := \{\phi' : \phi' \sim \phi\}$.

(28) THE Q PRINCIPLE: Don't use ϕ if there is another $\phi' \in A_Q(\phi)$ st. ϕ' is assertable and $[\![\phi']\!]$ asymmetrically entails $[\![\phi]\!]$.

- We can define M-alternatives and the M Principle as effective duals of these, à la Horn:

(29) M-ALTERNATIVES: Let ϕ be a parse tree. The set of **M-alternatives** for ϕ, written as $A_M(\phi)$, is defined as $A_M(\phi) := \{\phi' : [\![\phi']\!] \leftrightarrow [\![\phi]\!]\}$.

(30) THE M PRINCIPLE: Don't use ϕ if there is another $\phi' \in A_M(\phi)$ st. ϕ' is assertable and $\phi' < \phi$.

- And formalize Horn's Principle of Least Effort ("Marked forms are associated with marked meaning"):

(31) **The Marked Meaning Principle**: For parse trees ϕ, ϕ' such that $\phi' \in A_M(\phi)$ and $\phi' < \phi$: ϕ carries the Manner implicature: "$[\![\phi]\!]$ is atypical in a relevant sense".

3.2 Evaluativity as implicature

- Three types of degree constructions with respect to evaluativity:

(32) **antonym-sensitive evaluativity**
   a. How short is Adam? *degree question*
   b. Adam is as short as Doug. *equative*
   c. Adam is that short too. *degree demonstrative*

(33) **antonym-insensitive evaluativity**
   a. Adam is tall/short. *positive construction*
   b. Is Adam tall/short? *polar degree questions*

(34) **non-evaluative**
   a. Adam is 5ft tall. *MP construction*
   b. Adam is taller/shorter than Doug. *comparative*

3.2.1 Antonym-sensitive evaluativity is a Manner implicature

- Lots of evidence negative antonyms are marked relative to positive ones (Lehrer, 1985; Heim, 2007)
- Degree demonstratives formed with negative (but not positive) antonyms are evaluative.

(35) A: Adam's really tall.
B: Doug's about that tall too.
B′:#Doug's about that short too.

(36) A: Adam's really short.
B: Doug's about that tall too.
B′: Doug's about that short too.

  - Holding fixed the context of evaluation (in which *that* picks out Adam's height), and ignoring evaluativity, the two responses in (35) are synonymous.
  - Because they qualify as M-alternatives, the marked construction (with the negative antonym) carries a Manner implicature... that Doug's height (in particular, his shortness) is atypical.

- Same for degree questions: aside from evaluativity, *How tall/short is Adam?* are synonymous.
  - Bonus prediction (see also Barker, 2002): the typically presuppositional evaluativity associated with degree questions becomes assertoric when we manipulate the QUD.

(37) A: I'm a little worried about the actress playing me in the movie. Is she tall or short?
B: *(to the casting agent)* How short is Susan again?
A: That's fine, as long as she's short/#tall.

(38) A: I'm a little worried about the actress playing me in the movie. Is she tall or short?
B: *(to the casting agent)* How tall is Susan again?
A: #That's fine, as long as she's short/tall.

- Equatives formed with negative antonyms are evaluative:

(39) a. Adam is as tall as Doug. $\not\rightarrow$ Adam is tall.
b. Adam is as short as Doug. $\rightarrow$ Adam is short.

  - Equatives are thought to semantically encode the weak $\geqslant$ ('at least') relation, but generally get strengthened to mean 'exactly' via scalar implicature, in competition with comparatives ($>$).
  - This seems to happen locally, even when equatives are explicitly modified by *at least*, just as the scalar implicature for *some* is calculated before modification in *At least some students passed the test* (Levinson, 2000); it could also be the result of something like conventionalization.
  - A bonus prediction: coupled with a discussion of at-issueness, this account correctly predicts that only the internal argument of a negative-antonym equative is evaluative (Rett, 2015)

- Like the Manner implicature in (23), the evaluativity here can but need not be embedded.

(40) a. (Robin's really confused about everyone's heights.) Robin thinks Doug is short, and she believes Adam is as short as Doug. *local*
b. (Robin thinks Adam and Doug are 5′0″, but she doesn't know that counts as short in this context.) Robin believes Adam is as short as Doug. *global*

- And also like other Manner implicatures, evaluativity in these constructions is reinforceable:

(41) a. Adam is as short as Doug, which is to say they're both short.
b. Adam knows how short Doug is, which is to say he knows he's short.

3.2.2 Antonym-insensitive evaluativity is an uninformativity-based Quantity implicature

- Positive constructions (and e.g. *Is Adam tall/short?*) qualify as tautologies ('degree tautologies')
  - $[\![\text{Adam is tall}]\!] = \exists d[\mathsf{tall}(\mathsf{adam}, d)]$
  - Gradable adjectives carry a 'positive extension presupposition' (Kennedy, 2007): they can only be predicated of individuals who instantiate the property to some degree, e.g. **This sauce is curvy*
  - (cf. the common assumption that there exists no null individual, i.e. one whose cardinality is 0)
- Like other Quantity implicatures (but unlike Manner implicatures), evaluativity in the positive construction is necessarily calculated locally.

(42) a. It's not always the case that war is war. Drones have complicated things immensely.
b. Adam isn't tall.

- Although it's hard to diagnose, the evaluativity of positive constructions seems to vary with the QUD.

(43) *In a lab full of typically smelly and colorful soap, two new varieties: one colorless and the other odorless; although the colorless one is significantly less smelly than the other soaps, and the odorless one is significantly more subdued in color.*
A: I can't tell these new soaps apart. Is this the odorless one?
B: No, that one's smelly.

- And like the implicature associated with other tautologies, the evaluativity associated with positive constructions is reinforceable, provided that the reinforcement helps disambiguate the intended reading.

(44) a. #War is war, which is to say that war is equivalent to war.
b. War is war, which is to say that violence is unavoidable in war. (16-a)

(45) a. #Adam is tall, which is to say he is tall.
b. Adam is tall, which is to say that he is taller than the average eighth-grader.

IMPLICATURE TYPE	**cancellable?**	**detachable?**	**reinforceable?**	**QUD-sensitive?**	**embeddable?**
scalar	✓	✓	✓	✓	✓
uninfo Quantity	x	x	✓	✓	✓
a-insensitive eval	x	x	✓	✓	✓
Manner	x	x	✓	✓	(✓)
a-sensitive eval	x	x	✓	✓	(✓)

Table 2: Properties of implicature, revised

3.2.3 Non-evaluative constructions

- MP constructions are not evaluative because a) they are not tautologous; and b) in English, they can only be formed with positive antonyms.
  - However, MP constructions can be formed with negative antonyms in Dutch (Doetjes, 2012), in which case they are evaluative. This would put Dutch MP constructions in the 'antonym-sensitive evaluativity' category, similar to degree demonstratives in English.

(46) het 20 cm diepe / ondiepe water
the 20 cm deep / shallow water
'the 20cm-deep water' / 'the 20cm-deep (which is shallow) water'

- Comparatives are not evaluative because a) they are not tautologous; and b) antonymic pairs of comparatives are not synonymous, and therefore do not count as M-alternatives

3.3 Why evaluativity?

- Prediction: evaluative constructions carry an implicature that the property is "atypical in a relevant sense". In the case of a gradable predicate, atypicality amounts to evaluativity.
- *non-gradable predicates:* There are several potential ways someone can be an atypical e.g. student...
- *gradable predicates:* But there are really only two ways someone's e.g. height can be atypical: she can be more tall than average (the evaluative meaning); or she can be less tall than average.
- There's lots of evidence that natural language privileges intensification over diminunization (Kirchner, 1955; Bolinger, 1972; Morzycki, 2012):

(47) a. He frightened me so!
b. What attempts they made!
c. It's a long, long way.
d. She cried her heart out.

- The evaluative meaning is the stronger of the two; maybe something like Kennedy's (2007) Interpretive Economy principle is at work ("Maximize the contribution of the conventional meanings of the elements of a sentence to the computation of its truth conditions," p.36).

4 Implicature in Game-Theoretic Pragmatics

- Rett (2015) doesn't present a formal account of evaluativity, but the driving force behind the analysis – the equilibrium of Horn's (1989) Principle of Least Effort – seems ripe for a game-theoretic treatment.
- GT models have the added bonus in that they contribute an elegant analysis of embedded implicature of all types (not just Q-implicatures); see in particular Potts et al. (2016).

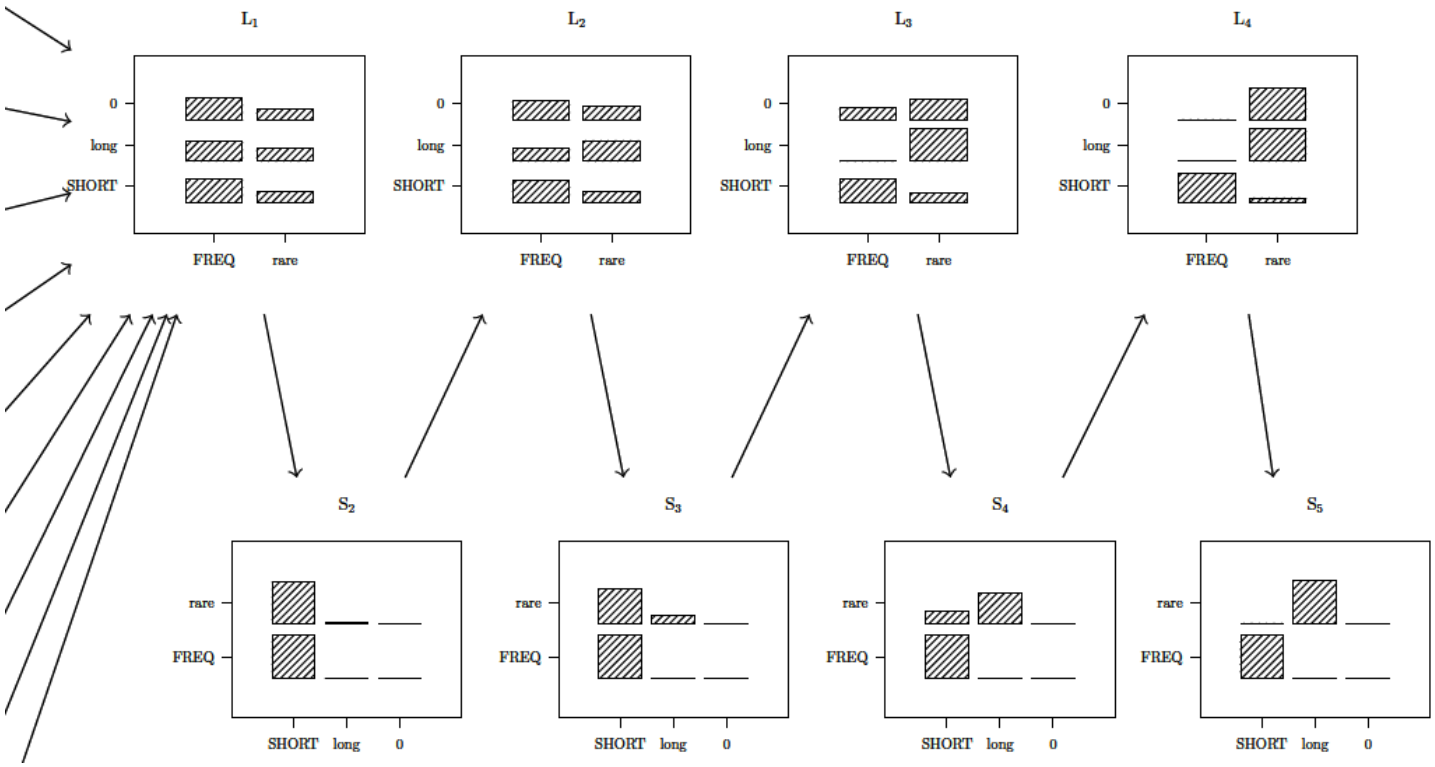
4.1 Goodman & Lassiter is a meta-semantic theory

- Lassiter and Goodman (2013, 2015) model how the value of the contextual variable in positive constructions is determined
  - They use a Bayesian-pragmatic theory to model how the free degree variable in positive constructions comes to receive a 'higher than average' sort of reading
  - The analysis assumes a POS operator or equivalent to introduce a contextual variable ('the threshold') into the semantics to begin with, which is what the present account pertains to
- Why this isn't sufficient:
  - Because it builds on standard POS accounts, it fails where they fail with respect to predicting the distribution of evaluativity
  - Their account thus requires rejection of a unified account of the semantics of evaluativity
- However:
  - Their model is compatible with an alternate source of evaluativity for the positive construction and elsewhere, like the present account
  - Their model is best interpreted as a (better!) answer to the question posed in §3.3: it uses Bayes to explain why markedness in adjectival constructions amounts to evaluativity in particular
  - It has the added bonus of being intuitive and useful in its connections to the Sorites Paradox

4.2 M-implicatures in the Rational Speech Act Framework

4.2.1 RSA and LUM

- Bergen et al. (2016) use the Rational Speech Act (RSA) framework (Frank and Goodman, 2012; Goodman and Stuhlmüller, 2013) to model Manner implicatures à la Horn (1989)[2]
  - RSA models communication as iterated speaker-listener inference from an utterance to the observed reality (or vice-versa)
  - It must be modified to deal with M-implicatures; the baseline RSA merely encodes markedness as cost, leading to an across-the-board dispreference for the marked form
  - Bergen et al. develop the 'Lexical Uncertainty Model' (LUM) to explain 'one-shot' M-implicatures:
    * Instead of a fixed lexicon, LUM incorporates a probability distribution over a set Λ of lexica
    * Each semantically equivalent utterance is associated with a (markedness-based) cost;
    * Each utterance must receive an interpretation consistent with its semantics; it can be associated with FREQ or *rare* (corresponding to the typical and atypical properties above, respectively); FREQ having higher prior probability
  - Via iterated speaker-listener inferences, Bergen et al. model an equilibrium betwee interlocutors in which any marked utterance is associated with a *rare*/atypical meaning, and any unmarked utterance is associated with a FREQ/typical meaning.



[2]There are a few earlier attempts at using GT to model M-implicature: van Rooij (2004) and De Jaeger (2008) employ an EGT analysis that is more appropriate for characterizing diachronic facts (or, possibly, Generalized Implicature); Franke (2009) and Jäger (2012) derive M-implicatures within the Iterated Best Response (IBR) model which (unlike the RSA) does not assume agents are merely approximately rational. They do make this assumption in Franke and Jäger (2014) in the Iterated Quantal Response (IQR) model; it's unclear if the worries expressed here about M-implicatures in RSA extend to that account.

4.2.2 Problems posed by evaluativity

- Recall from §3.2 that there are three patterns of evaluativity:
  1. antonym-sensitive evaluativity (e.g. in *how* questions, equatives)
  2. antonym-insensitive evaluativity (e.g. in positive constructions, polar degree questions)
  3. non-evaluativity (e.g. MP constructions, comparatives)
- The LUM predicts antonym-sensitive evaluativity (Type 1), the straightforward M-implicatures
- And it also predicts the absence of evaluativity in Type 3 constructions:
  - in which the marked (*long*) utterance is morphologically blocked (i.e. MP constructions)
  - or in which two utterances containing antonyms are not in fact synonymous (i.e. comparatives)
- But it fails to predict the existence of Type 2 constructions.
  - It predicts that *John is tall* will be associated with FREQ and *John is short* with *rare*
  - That is, while it does a great job of deriving markedness-based M-implicatures, it does so **across the board**, without accounting for the observed construction-specificity we see with evaluativity
- We need a way of modeling uninformativity-based Q-implicatures as primary to M-implicatures
  - There's evidence of this primacy (Levinson, 2000; Rett, 2015): equatives arguably involve both, but a markedness competition is only possible after strengthening
- LUM, in its current form, can't do this because it only deals in cost and content.
  - "[T]he basic lexical uncertainty model can only derive distinct pragmatic interpretations for a pair of utterances by leveraging either differences in semantic content or complexity" (p58).
  - (Example (p16): in the scalar competition between *some* and *all*, the option *some but not all* is not entertained (which would lead to false predictions) because it's more costly than the others.)
  - The positive constructions *John is tall* and *John is short* are in competition because they're synonymous... but because it's relatively less costly, LUM predicts *John is tall* is never evaluative.
  - Worse, LUM also predicts that *John is tall* is itself in competition with other, even more unmarked tautologies (with consistent presuppositions...)
- In light of evaluativity, LUM must be adapted to incorporate non-stipulative Horn scales (cf. Fox and Katzir, 2011) to restrict M-implicatures for reasons other than the scalar-based ones usually discussed
- Other areas ripe for extension:
  - Bergen et al. refine LUM to perform on sub-sentential expressions, useful for explaining why equatives modified by *at least* pattern with equatives (not comparatives) re: evaluativity
  - Bergen et al. (2016) and Potts et al. (2016) show how the RSA naturally accounts for implicature embedding... can it account for the difference between M- and Q-implicatures with respect to optional embeddability?
  - An extension of the theory to relative adjectives with more than one antonym (e.g. the quadruple *hot; warm; cool; cold*), which behave differently with respect to evaluativity than relative adjectives with one clear antonym (e.g. *tall; short*).

5 Summary

- Evaluativity is wide-spread! But it's not everywhere. So we don't want it lexically encoded in the meaning of adjectives.
- Some constructions are not evaluative (e.g. synthetic comparatives with relative adjectives): these are a) meaningful and b) not synonymous with their antonymic counterparts.
- Some constructions are evaluative regardless of their antonym (e.g. the positive construction): these are degree tautologies, giving rise to evaluativity as an uninformativity-based Quantity implicature.
- Some constructions are only evaluative when formed with a negative antonym or are otherwise marked: they give rise to evaluativity as a Manner implicature.
- Evaluativity behaves as predicted, given independent observations about the behavior of (non-scalar) conversational implicatures with respect to discourse-sensitivity, embeddability, and cancellability.
- We need a theory of the semantics/pragmatics interface that allows for the embeddability, etc., of (non-syntactically-encoded) implicature, like evaluativity, but also other Manner and Quantity implicatures.
- Recent RSA-based accounts are great candidates: they treat M-implicatures using cost (Bergen et al., 2016); and very naturally model embedded conversational implicature of all types (Potts et al., 2016).
- But, as it stands, RSA models are incapable of treating **both** M-implicature **and** uninformativity-based Q-implicature.
- The behavior of evaluativity suggests that we need to incorporate lexical or syntactic knowledge (i.e. Horn scales) into the models to restrict competition for uninformativity-based Q-implicatures.

References

Barker, C. (2002). The dynamics of vagueness. *Linguistics and Philosophy*, 25:1–36.

Bartsch, R. and Vennemann, T. (1972). The grammar of relative adjectives and comparison. *Linguistische Berichte*, 20:19–32.

Bergen, L., Levy, R., and Goodman, N. (2016). Pragmatic reasoning through semantic inference. *Semantics and Pragmatics*, 9.

Bierwisch, M. (1989). The semantics of gradation. In Bierwisch, M. and Lang, E., editors, *Dimensional Adjectives: Grammatical Structure and Conceptual Interpretation*, pages 71–237. Springer-Verlag.

Bolinger, D. (1972). *Degree words*. Mouton, Paris.

Breakstone, M. (2012). Inherent evaluativity. In Aguilar, A., Chernilovskaya, A., and Nouwen, R., editors, *Proceedings of Sinn und Bedeutung 16*. MITWPL.

Chierchia, G. (2004). Scalar implicatures, polarity phenomena and the syntax/pragmatic interface. In Beletti, A., editor, *Structures and Beyond*, pages 39–103. Oxford University Press.

Cresswell, M. (1976). The semantics of degree. In Partee, B., editor, *Montague Grammar*, pages 261–292. Academic Press.

De Jaeger, K. (2008). The evolution of Horn's rule. *Journal of Economic Methodology*, 15:275–284.

Doetjes, J. (2012). On the (in)compatibility of non neutral adjectives and measure phrases. In Guevara, A., Chernilovskaya, A., and Nouwen, R., editors, *Proceedings of SuB 16*. MITWPL.

Fox, D. and Katzir, R. (2011). On the characterization of alternatives. *Natural Language Semantics*, 19:87–107.

Frank, M. and Goodman, N. (2012). Predicting pragmatic reasoning in language games. *Science*, 336:998–998.

Franke, M. (2009). *Signal to act: Game Theory in pragmatics*. PhD Thesis, Universiteit van Amsterdam.

Franke, M. and Jäger, G. (2014). Pragmatic back-and-forth reasoning. In Pistoia Reda, S., editor, *Pragmatics, semantics, and the case of scalar implicatures*, pages 170–200. Palgrave Macmillan.

Geurts, B. (2011). *Quantity implicatures*. Cambridge University Press.

Goodman, N. and Stuhlmüller, A. (2013). Knowledge and implicature: modeling language understanding as social cognition. *Topics in Cognitive Science*, 5:173–184.

Grano, T. (2012). Mandarin *hen* and universal markedness in gradable adjectives. *Natural Language and Linguistic Theory*, 30:513–565.

Grice, P. (1975). Logic and conversation. In Cole, P. and Morgan, J., editors, *Syntax and Semantics*, volume 3. Academic Press.

Hay, J., Kennedy, C., and Levin, B. (1999). Scalar structure underlies telicity in "degree achievements". In Matthews, T. and Strolovitch, D., editors, *Proceedings of SALT IX*, pages 127–144. CLC Publications.

Heim, I. (2007). Little. In Gibson, M. and Howell, J., editors, *Proceedings of SALT* XVI. CLC Publications.

Hirschberg, J. (1991). *A theory of scalar implicature*. Garland.

Horn, L. (1972). *On the semantic properties of the logical operators in English*. PhD Thesis, UCLA.

Horn, L. (1984). Toward a new taxonomy for pragmatic inference: Q- and R-based implicature. In Shiffrin, D., editor, *Meaning, form, and use in context*, pages 11–42. Georgetown University.

Horn, L. (1989). *A natural history of negation*. University of Chicago.

Horn, L. (1991). Duplex negatio affirmat... the economy of double negation. *Chicago Linguistics Society*, 27:80–106.

Huitink, J. and Spenader, J. (2004). Cancelation resistant PCs. In van der Sandt, R. and Geurts, B., editors, *Proceedings of the ESSLLI 2004 Workshop on Implicature and Conversational Meaning*.

Jäger, G. (2012). Game theory in semantics and pragmatics. In Maienborn, C., Portner, P., and von Heusinger, K., editors, *Semantics: An international handbook of natural language meaning*, pages 2487–2425. De Gruyter Mouton.

Jespersen, O. (1965). *The Philosophy of Grammar*. W. W. Norton & Co. Originally published 1924, Allen & Unwin, London.

Katzir, R. (2007). Structurally-defined alternatives. *Linguistics and Philosophy*, 30:669–690.

Kennedy, C. (2007). Vagueness and grammar: The semantics of relative and absolute gradable predicates. *Linguistics and Philosophy*, 30:1–45.

Kennedy, C. and Levin, B. (2008). Measure of change: the adjectival core of degree achievements. In McNally, L. and Kennedy, C., editors, *Adjectives and adverbs: syntax, semantics and discourse*. Oxford University Press.

Kirchner, G. (1955). *Gradadverbien: Restrikiva und Verwandtes im heutigen Englisch*. Halle.

Klein, E. (1980). A semantics for positive and comparative adjectives. *Linguistics and Philosophy*, 4:1–45.

van Kuppevelt, J. (1995). Discourse structure, topicality and questioning. *Journal of Linguistics*, 31:109–147.

Lassiter, D. and Goodman, N. (2013). Context, scale structure, and statistics in the interpretation of positive-form adjectives? In *Proceedings of SALT 23*, page 587610. Cornell Linguistics Society.

Lassiter, D. and Goodman, N. (2015). Adjectival vagueness in a bayesian model of interpretation. *Synthese*.

Lehrer, A. (1985). Markedness and antonymy. *Journal of Linguistics*, 21:397–429.

Levinson, S. (2000). *Presumptive meanings: the theory of generalized conversational implicature*. MIT Press.

Matushansky, O. (2001). More of a good thing: Russian synthetic and analytic comparatives. In *Proceedings of FASL 10*. University of Michigan Press.

Morzycki, M. (2009). Degree modification of gradable nouns: size adjectives and adnominal degree morphemes. *Natural Language Semantics*, 17:175–203.

Morzycki, M. (2012). Adjectival extremeness: degree modification and contextually restricted scales. *Natural Language and Linguistic Theory*, 30:567–609.

Piñón, C. (2005). Adverbs of completion in an event semantics. In Verkuyl, H., de Swart, H., and van Hout, A., editors, *Perspectives on Aspect*, pages 149–166. Springer.

Potts, C., Lassiter, D., Levy, R., and Frank, M. (2016). Embedded implicatures as pragmatic inferences under Compositional Lexical Uncertainty. *Journal of Semantics*, 33:755–802.

Rett, J. (2007). Antonymy and evaluativity. In Gibson, M. and Friedman, T., editors, *Proceedings of SALT XVII*, pages 210–227. CLC Publications.

Rett, J. (2008). *Degree Modification in Natural Language*. PhD thesis, Rutgers University.

Rett, J. (2014). Modified numerals and measure phrase equatives. *Journal of Semantics*, 32:425–475.

Rett, J. (2015). *The semantics of evaluativity*. Oxford University Press.

Roberts, C. (1990). *Modal subordination, anaphora, and distributivity*. Garland Press.

van Rooij, R. (2004). Signalling games select Horn strategies. *Linguistics and Philosophy*, 27:493–527.

Sadock, J. M. (1978). On testing for conversational implicature. In Cole, P., editor, *Pragmatics*, volume 9 of *Syntax and Semantics*, pages 281–298. Academic Press.

van der Sandt, R. (1992). Presupposition projection as anaphora resolution. *Journal of Semantics*, 9:333–377.

Sassoon, G. (2011). Be positive! Norm-related implications and beyond. In Reich, I., editor, *Proceedings of Sinn und Bedeutung 15*, pages 531–546. Universaar – Saarland University Press.

Simons, M., Tonhauser, J., Beaver, D., and Roberts, C. (2011). What projects and why. In Li, N. and Lutz, D., editors, *Proceedings of SALT XX*, pages 309–327. CLC Publications.

von Stechow, A. (1984). Comparing semantic theories of comparison. *Journal of Semantics*, 3:1–77.