Adjectives, force variability and pragmatic reasoning

SynSem Seminar at LLING, Nantes University

Cécile Meier

Institut für Linguistik
Johann Wolfgang Goethe-Universität
Frankfurt am Main

March, 27 2026, with some corrections in April 2026

The talk in a nutshell

- Question: How does pragmatic reasoning work with adjectives?
- Background on ⟨**some**, **all**⟩: Horn Scales.
- The Problem: Scalar Diversity. Adjectives sometimes allow for pragmatic enrichment, but not always.
- Observation (not mine): Horn Scales require a certain "conceptual" pattern in the meaning of the alternatives listed: $\langle \exists \ldots, \forall \ldots \rangle$.
- Hypothesis: Adjectives (in the Positive) may pattern with universal and existential quantifiers: the 'weaker' adjective is force variable.
- 'Quantifier-Flipping': **smart** may be existentially quantified in the vicinity of **brilliant** and universally quantified if not. Implicatures only arise if existentially interpreted. Force variability is costly. This explains infrequency.

Outline

Horn Scales, usually

- Horn Scales are lists of lexical items: expression alternatives (Ausdrucksalternativen)
- Horn Scales are used in order to derive scalar implicatures.
- Expressions higher on the scale entail expressions lower on the scale: ⟨**some**, **all**⟩, quantificational determiners

ex. **All penguins are in the pond ⇒ Some penguins are in the pond.**

- Are there Adjectival Horn Scales?

  ⟨**smart**, **brilliant**⟩

  ⟨**ugly**, **hideous**⟩

Adjectives are special

Doran et al. (2009, p. 239)

(1) a. [Irene asks:] What size is Jeremy?
b. [Sam says:] He's big.

...

(2) a. [Situation: Jeremy can't fit in an airplane seat.]
b. [Informant is asked:] "Is Sam telling the truth?"

- Informant's answer: "YES!" No pragmatic enrichment. "Sam is informative enough." (low rejection rate)

Adjectives are special

Doran et al. (2009, p. 239)

(3) a. [Irene asks:] Is Jeremy average, big or huge?
b. [Sam says:] He's big.

...

(4) a. [Situation: Jeremy can't fit in an airplane seat.]
b. [Informant is asked:] "Is Sam telling the truth?"

- Informant's answer: "NO!" There is pragmatic enrichment. "Sam is not informative enough." (higher rejection rate)
- Mentioning the alternatives in the question does play a role.

Within-Scale Variation

- Adjectives are special in that they show a special case of within-scale variation.
- The variability has nothing to do with a difference in comparison classes (Aparicio and Ronai, 2025). Those are identical in both examples.
- The Question under Discussion differs.
- The standards of comparison under discussion is different.

Quantificational Determiners

(5) a. [Irene asks:] What did Jeremy eat?
b. [Sam says:] He ate some of the cookies.

...

(6) a. [Situation: Jeremy ate all of the cookies.]
b. [Informant is asked:] Is Sam telling the truth?

- Answer "NO!" Pragmatic enrichment. "Sam is not informative enough."
- No mentioning of alternatives, scalar implicature is calculated anyway.
- That determiners and adjectives are different, is called across-scale variation.

Upshot

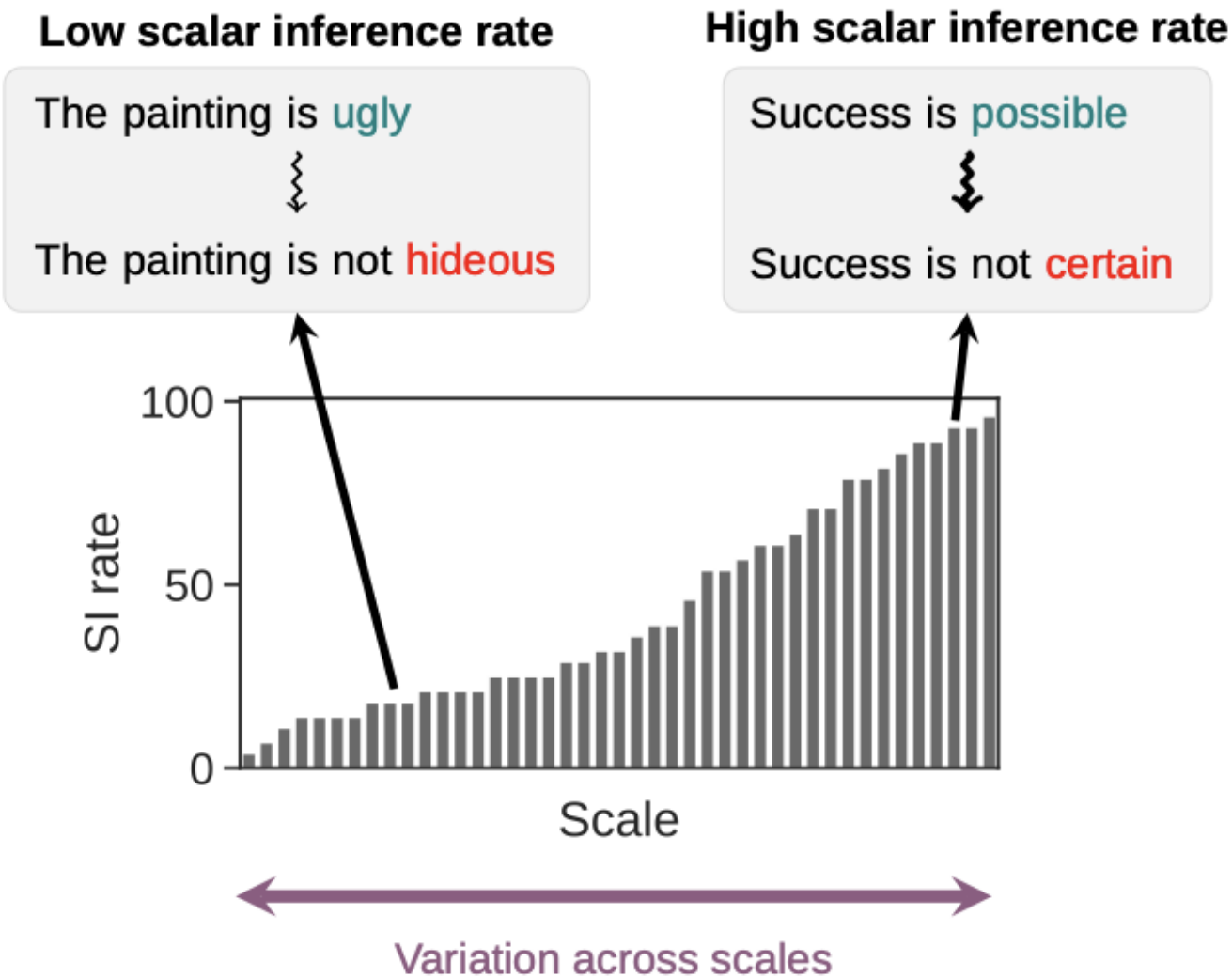
- Adjectives sometimes allow for deriving a scalar implicature.
- Here mentioning the stronger alternative played a role.
- Quantificational Determiner scales are not SO special. They trigger the scalar implicature almost independent on the previous linguistic context. The form of the restriction plays a role though (Degen and Tanenhaus, 2015).

Literature

- The psycholinguistic literature discusses what is called the Uniformity Hypothesis: It states roughly that all Horn scales are equal.
- Adjectival Horn Scales are the reason for the Uniformity Hypothesis to be rejected.
- Doran et al.'s work opened up a whole new field of investigation.
- Scalar implicature rates are well attested for different types of expressions in the meantime in the psycholinguistic literature. Context plays a role.
- Rates are low for adjectives, and high for quantificational items: Bob van Tiel et al. (2016), Gotzner, Solt, et al. (2018), Pankratz and B. van Tiel (2021), and Ronai and Xiang (2022), etc.

Scalar diversity: Visualization by Hu et al. (2023)

Bob van Tiel et al. (2016)



Explanations from empirical research

- Semantic distinctness between the scale mates (semantic distance / bounded - unbounded) (Bob van Tiel et al., 2016)
- Extremeness (Beltrama and Xiang, 2013),
- Polarity: Relation to other pragmatic inferences like negative strengthening (Gotzner, Solt, et al., 2018; Gotzner, Benz, et al., 2018),
- Context sensitivity (Ronai and Xiang, 2022; Aparicio and Ronai, 2025) and
- How relevant the scalar implicature is in the context (Pankratz and B. van Tiel, 2021).

My Problem

- It is surprising that adjectival scales sometimes allow for Horn/Gricean reasoning (see above the examples from Doran et al. 2009)
- Uniformity hypothesis may still be right – but not at the level of the lexical items but at the level of representation. Some adjectives are ambiguous?
- Change the mode of calculating implicatures!

A Relation to the Theory of Alternatives

- Hu et al. (2023) discuss two possible reasons for the low scalar inference rates:
- Insecurity about what the right scale is: string-based ⟨**smart**,—?—⟩
- Insecurity about how the concept representing the strong scale-mate is expressed: ⟨**SMART**, **VERYSMART**⟩

A Relation to the Theory of Alternatives

Language Models

- Hu et al. (2023) calculate model surprisal using GPT-2 model with the SyntaxGym command-line interface.
- They ask how likely is it that a strong scale-mate appears in the construction [NP] is [WEAK], but not [STRONG]?
- And, they claim that these GPT-likelihoods correlate with the empirical research on Horn Scales (see above Tiel et al. Gotzner et al. Ronai/Xiang and Pankratz and van Tiel): Low scalar implicature rates for elements with high surprisal rates.

A Relation to the Theory of Alternatives

- Hu et al. (2023) use two types of likelihood calculation: raw surprisal vs. similarity-weighted surprisal
- Similariy-weighted surprisal captures the inference rates in the empirical data (see above Tiel et al. Gotzner et al. Ronai/Xiang and Pankratz and van Tiel) to a greater extent than raw surprisal.

$\Rightarrow$ Alternatives (as part of a suitable scale) are conceptually and not word related.

- ... Conceptual alternatives???

Conceptual Alternatives

Buccola et al. (2022)

- The Symmetry Problem: ⟨**some**, **all**⟩ is a suitable scale, ⟨**some**, **some but not all**⟩ is not (would be equally more informative, though more complex)
- Horn: Alternatives are constrained by the lexicon . . .
- Buccola et al. (2022): Alternatives are IN ADDITION related by what counts as a primitive concept: Concepts are quantificationally restricted.
- Alternatives look like: ⟨∃. . . , ∀ . . .⟩ with gradual shades . . .
- etc. pp.

Sometimes alternatives are inexpressible

Charlow (2018), pace Buccola's presentation

(7) **John repeated the rumor that Mary or Bill was expelled.**

(8) The 'naive' scalar alternative is not suitable for calculating this inference: **John repeated the rumor that Mary and Bill was expelled.**

(9) 'Conceptual representation': **Either Mary or Bill are such that John repeated the rumor that he or she was expelled.**

(10) Actual implicature: **It is not the case that Mary and Bill are each such that John repeated the rumor that he or she was expelled.**

Charlow (2018) argues (according to Buccola et al.)

- Sentence (4) (**rumor . . . or**) involves wide scope existential quantification ($\exists$) over choice functions.
- Alternative (5) involves universal quantification ($\forall$) over choice functions
- such universal alternatives "are more abstract than we might have thought", "do not seem to correspond to any expressible lexical items" (in English)

$\rightarrow$ Alternative is not expressible.

Question and Hypothesis

- **Question**: Why could pragmatic reasoning for adjectives be different dependent on the alternatives mentioned?
- **Hypothesis**: This has something to do with quantificational force (existential and universal).
- **Background**: Adjectives in the positive are universally quantified (von Stechow 2009).
- **WEAK-STRONG**: The stronger adjective comes with a wider domain than the weaker one and can induce domain widening on the weaker one. Widening of the domain triggers force variablity (only the prototypical ones?).

$\Rightarrow$ If the domain of the weaker one is widened it becomes existentially quantified. Weak and strong adjectives become conceptual scale mates $\exists \ldots \forall$ and pragmatic enrichment may take place.

Conceptual adjectival scales

- **Bigger Picture**: Scales of suitable alternatives are conceptually determined (weaker elements correspond to existentials and stronger ones to universals)
- My work seems related to Buccola et al. (2022): Conceptual Alternatives.

- . . . step by step . . .

Interpretation of prototypical (WEAK) adjectives

Following Stechow (2009)

- The interpretation of gradable adjectives is associated with points on a scale (Cresswell, 1976).
- This scale is constructed on the basis of an "ordering source" and restricted to the relevant items (comparison class).
- The points on the scale are called degrees and values of a measure function, i.e. numbers.
- **big** is interpreted with respect to a measure function μ: μ_{SIZE}.
- Adjectives express relations between an individual from a set of individuals A and a degree from the set of reels $\mathbb{R}$.

- $[\![\textbf{big}]\!]^s = \lambda d.\lambda x.\vdash \mu_{\text{SIZE}}(s)(x) \geq d \dashv$ (adjectival stem)

The Positive, Stechow (2009) a universal quantifier

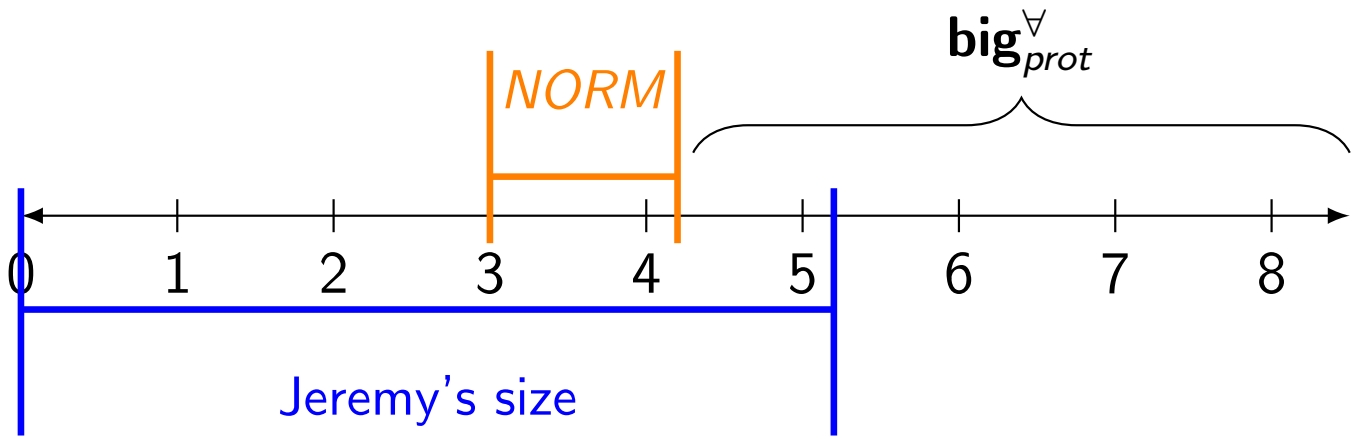
- The positive form of adjectives is more complex: **POS-big**. I adopt the view that there is an (invisible) positive morpheme that relates two sets of degrees. One set is a contextually determined interval somewhere in the middle of the scale: delineation interval. What the norm is, depends on the adjective. The other is a set that collects degrees below the actual measure of the individual measured by μ. That is: The positive morpheme corresponds to a universal quantifier.

Definition Positive

- $[\![\textbf{POS}^{\forall}]\!]^{s,c} = \lambda D_1.\vdash NORM_s^c \subseteq \downarrow D_1 \dashv$

- $[\![\textbf{Jeremy is POS-big}]\!]^{s,c} = 1$ iff
$\forall d[d \in NORM_s^c \rightarrow \mu_{\text{SIZE}}(s)(\text{J}) \geq d]$

Positively **big**



- $[\![\textbf{Jeremy is POS-big}]\!]^{s,c} = 1$ iff
  $\forall d[d \in \mathit{NORM}^c_s \rightarrow \mu_{\mathrm{SIZE}}(s)(\mathrm{J}) \geq d]$

Extreme (STRONG) Adjectives

Morzycki (2012)

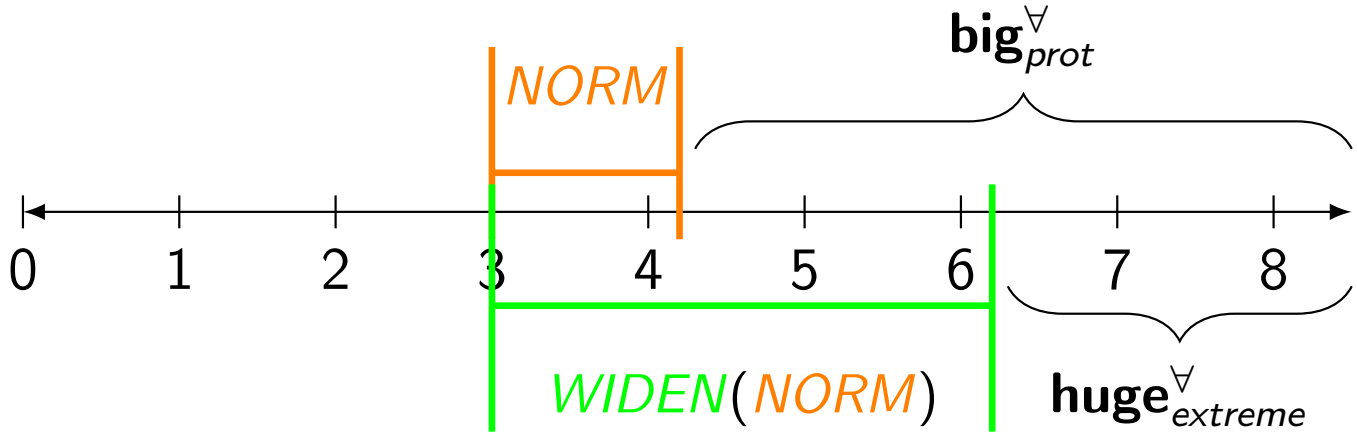
- Extreme adjectives are different in semantic type, a property of individuals. No combination with intensifiers or a comparative operator.

Definition Positive (not Morzycki)

- $[\![\textbf{huge}]\!]^{s,c} =$
  $\lambda x.\vdash \mathit{WIDEN}(\mathit{NORM}_s^c) \subseteq \ \{d : \mu_{\mathit{SIZE}}(s)(\mathrm{J}) \geq d\} \dashv$

- $[\![\textbf{Jeremy is huge}]\!]^{s,c} = 1$ iff
  $\forall d[d \in \mathit{WIDEN}(\mathit{NORM}_s^c) \rightarrow \mu_{\mathit{SIZE}}(s)(\mathrm{J}) \geq d]$

Direct comparison: **big** and **huge**



- $[\![\textbf{Jeremy is POS-big}]\!]^{s,c} = 1$ iff
  $\forall d[d \in NORM^c_s \rightarrow \mu_{\text{SIZE}}(s)(\text{J}) \geq d]$
- $[\![\textbf{Jeremy is huge}]\!]^{s,c} = 1$ iff
  $\forall d[d \in WIDEN(NORM^c_s) \rightarrow \mu_{SIZE}(s)(\text{J}) \geq d]$

$\Rightarrow$ No implicatures calculated. Two universal quantifications.

Scale mates should share the restriction

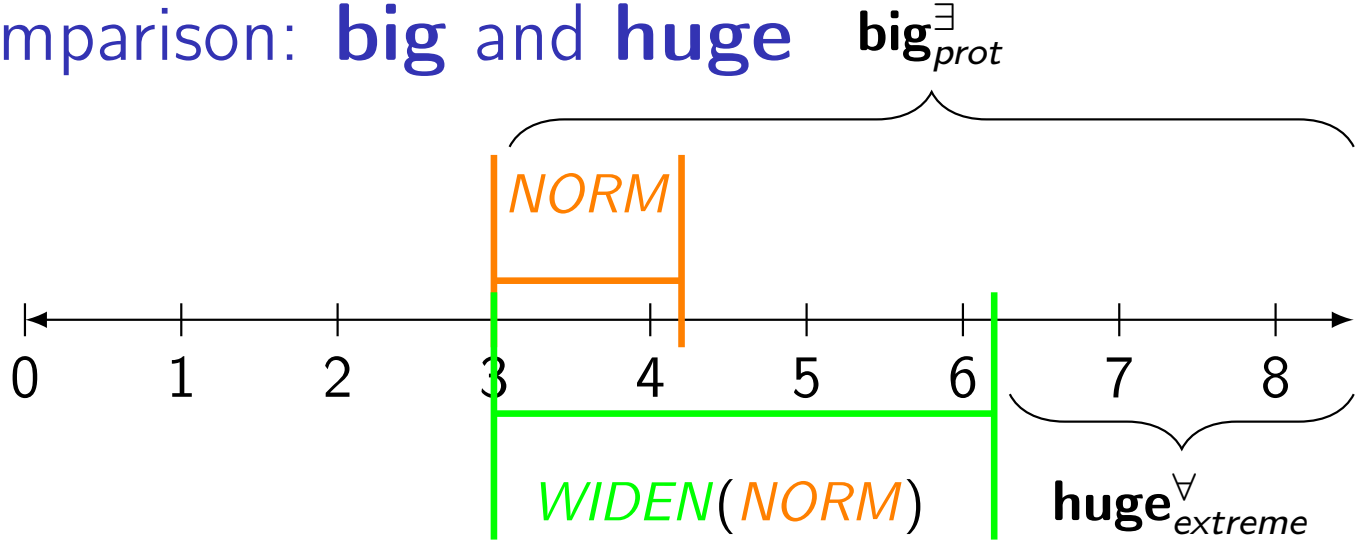
Otherwise no 'entailment' in the nominal domain!

- $[\![\textbf{Jeremy is huge}]\!]^{s,c} = 1$ iff
  $\forall d[d \in \mathit{WIDEN}(\mathit{NORM}_s^c) \rightarrow \mu_{\mathit{SIZE}}(s)(\mathrm{J}) \geq d]$

- $[\![\textbf{Jeremy is POS}^{\forall}\textbf{-big}]\!]^{s,c} = 1$ iff
  $\forall d[d \in \mathit{NORM}_s^c \rightarrow \mu_{\mathrm{SIZE}}(s)(\mathrm{J}) \geq d]$

- Reformulation as an existential quantification (on the basis of widened delineation interval):
- $[\![\textbf{Jeremy is POS}^{\exists}\textbf{-big}]\!]^{s,c} = 1$ iff
  $\exists d[d \in \mathit{WIDEN}(\mathit{NORM}_s^c)\ \&\ \mu_{\mathrm{SIZE}}(s)(\mathrm{J}) \geq d]$

Direct comparison: **big** and **huge**



- $[\![\textbf{Jeremy is POS}^{\exists}\textbf{-big}]\!]^{s,c} = 1$ iff
  $\exists d[d \in \mathit{WIDEN}(\mathit{NORM}^c_s)\ \&\ \mu_{\mathrm{SIZE}}(s)(\mathrm{J}) \geq d]$
- $[\![\textbf{Jeremy is huge}^{\forall}]\!]^{s,c} = 1$ iff
  $\forall d[d \in \mathit{WIDEN}(\mathit{NORM}^c_s) \rightarrow \mu_{\mathit{SIZE}}(s)(\mathrm{J}) \geq d]$

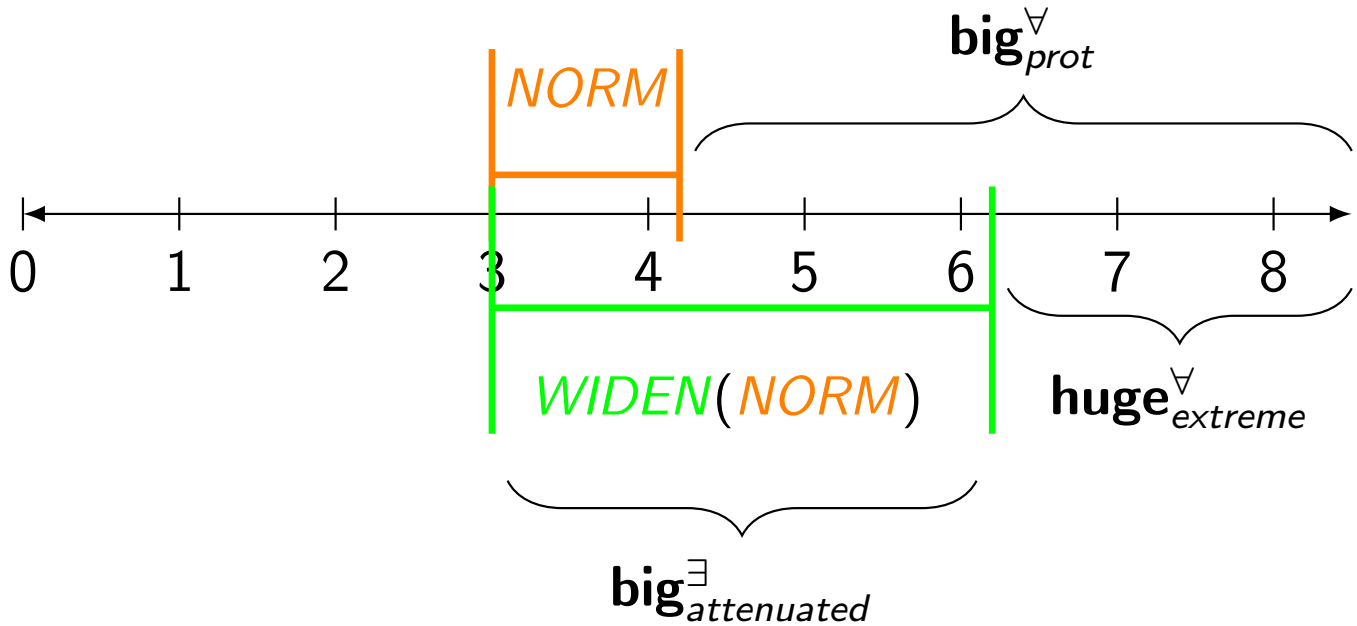
$\Rightarrow$ Conceptual pattern $\langle \exists \ldots, \forall \ldots \rangle$ is met, implicatures calculated.

Existential Degree Quantification

Definition Positive as an existential operator

- $[\![\mathbf{POS}^{\exists}]\!]^{s,c} = \lambda D_1.\vdash NORM_s^c \cap \downarrow D_1 \neq \emptyset \dashv$

Domain Widening triggers force variability



- Attenuated meaning of a weak adjective is only derived if the more informative adjective **huge** widens the gap of big and widening triggers the existential reading.
- The implicature calculation explains the blocking effect of the more informative adjective.

Interim conclusion

- The uniformity hypothesis may be valid at a deeper level of logical representation.
- **big** patterns with **possible** and **some** in one reading: if interpreted existentially.
- Force variablity "that the quantifier flips in force from universal to existential"
- $\mathbf{POS}^{\exists}$ is the dual of $\mathbf{POS}^{\forall}$
- Quantifier flipping/domain widening is costly.
- The psycholinguistic literature discusses contextual salience in detail. I only relate to one of the factors connected to salience: semantic distance "How wide are the gaps?" The wider the contextually relevant gap the more likely it is that quantifier flipping occurs.

Measurement Scale and Square of Opposition

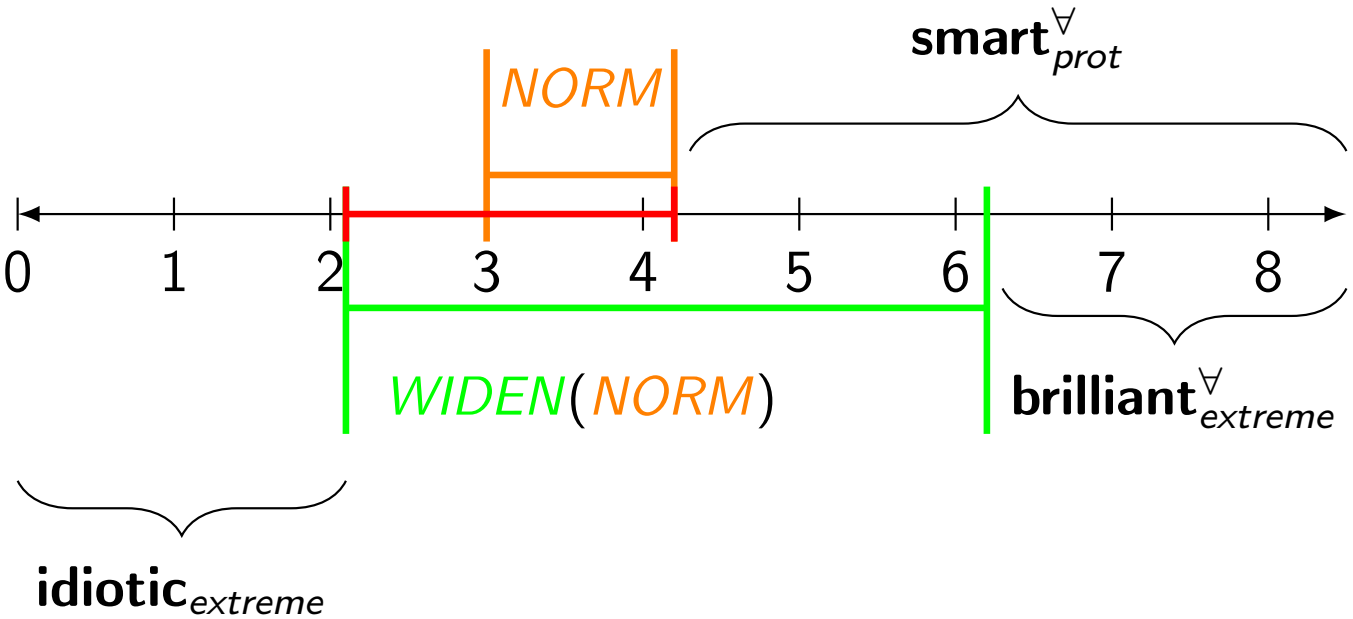
Gotzner, Benz, et al. (2018, p. 412): Uniformity again

"It might be tempting to take Aristotle's square of opposition as a template to be applied to all kinds of Horn scales. However, it is particularly important in the context of adjectival scales that the meaning relations of the square of opposition do not generalise."

Quadruples:
⟨**brilliant**, **smart**, **idiotic**, **not brilliant**⟩
⟨**all**, **some**, **no**, **not all**⟩

The argument: **smart** and **idiotic** raise no contradiction but **some** and **no** do.

Gotzner et al.'s problem: Mind the gap



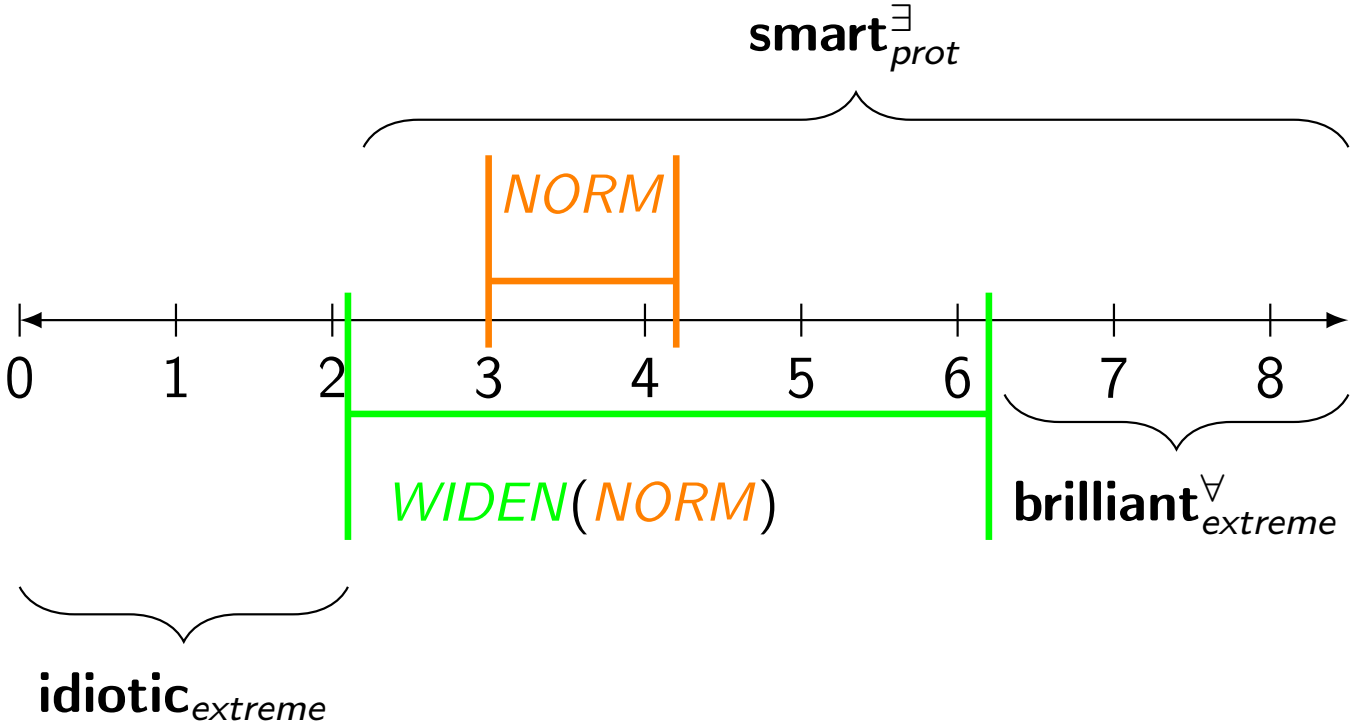
The Problem: **smart** and **idiotic** are not contradictions but contraries. Adjectives seem not to fit the Square of Opposition. But why should they? They cannot ... different restrictions . . . gaps.

Hey wait a minute . . .

My talk given at Workshop on Negation, SALT 30, New Haven

- This conclusion seems only correct if there is NO existential interpretation for **smart** (and no alignment of all the gaps in the square)!
- Adjectival scales DO generalize if we take into account that prototypical adjectives are force variable: 'think conceptually'.
- No argument against uniformity.
- Gotzner et al.'s problem vanishes if we interpret **smart** existentially and keep the restrictions of all the degree quantifiers THE SAME.

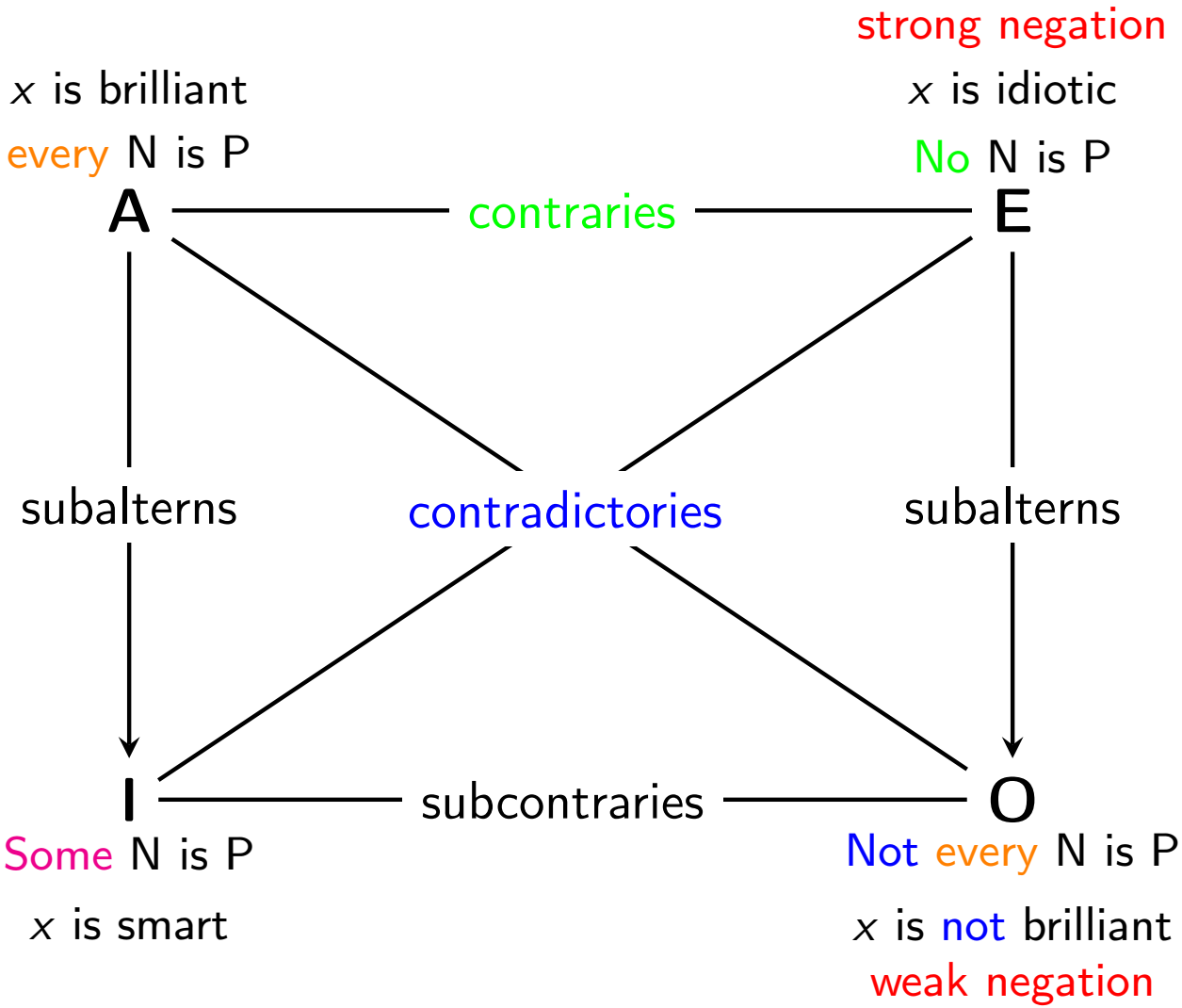
Gotzner's problem solved: No gap



The Observation: **smart** and **idiotic** may be contradictions but only if the weak adjective is interpreted existentially. Then adjectives seem to fit the Square of Opposition without any problem.

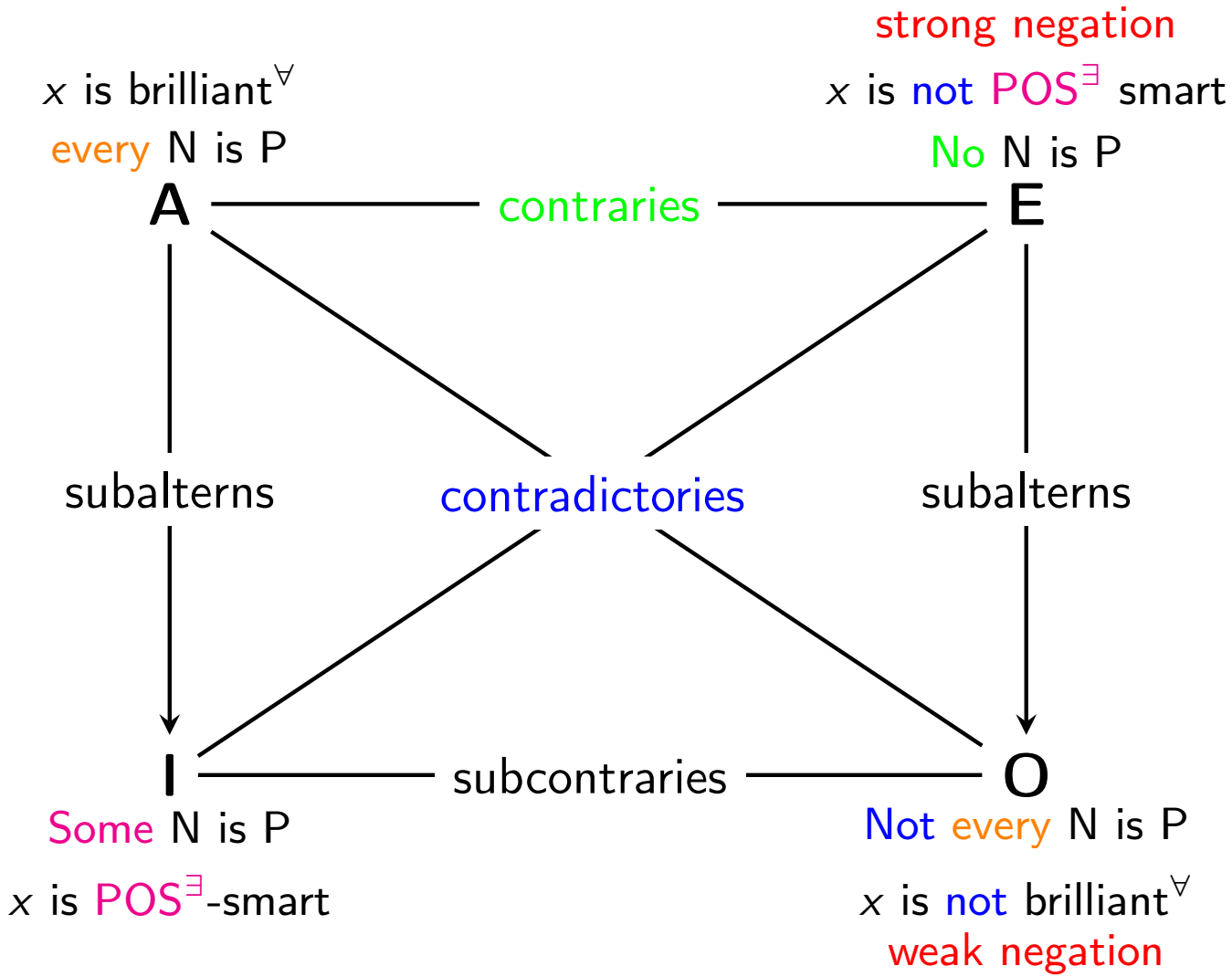
The Square of Opposition: Quantification in general

Aristotle, Horn (1989): Relations between Quantifications

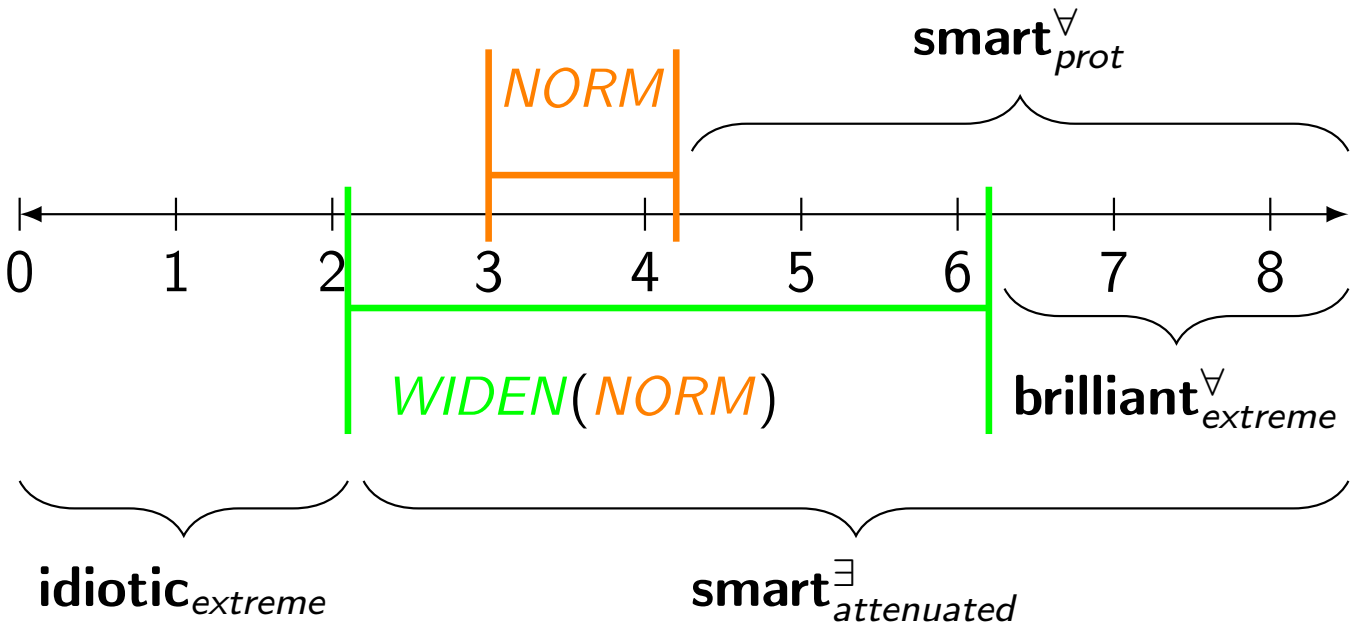


The Square of Opposition: **brilliant**

Restrictions adjusted to the widened domain



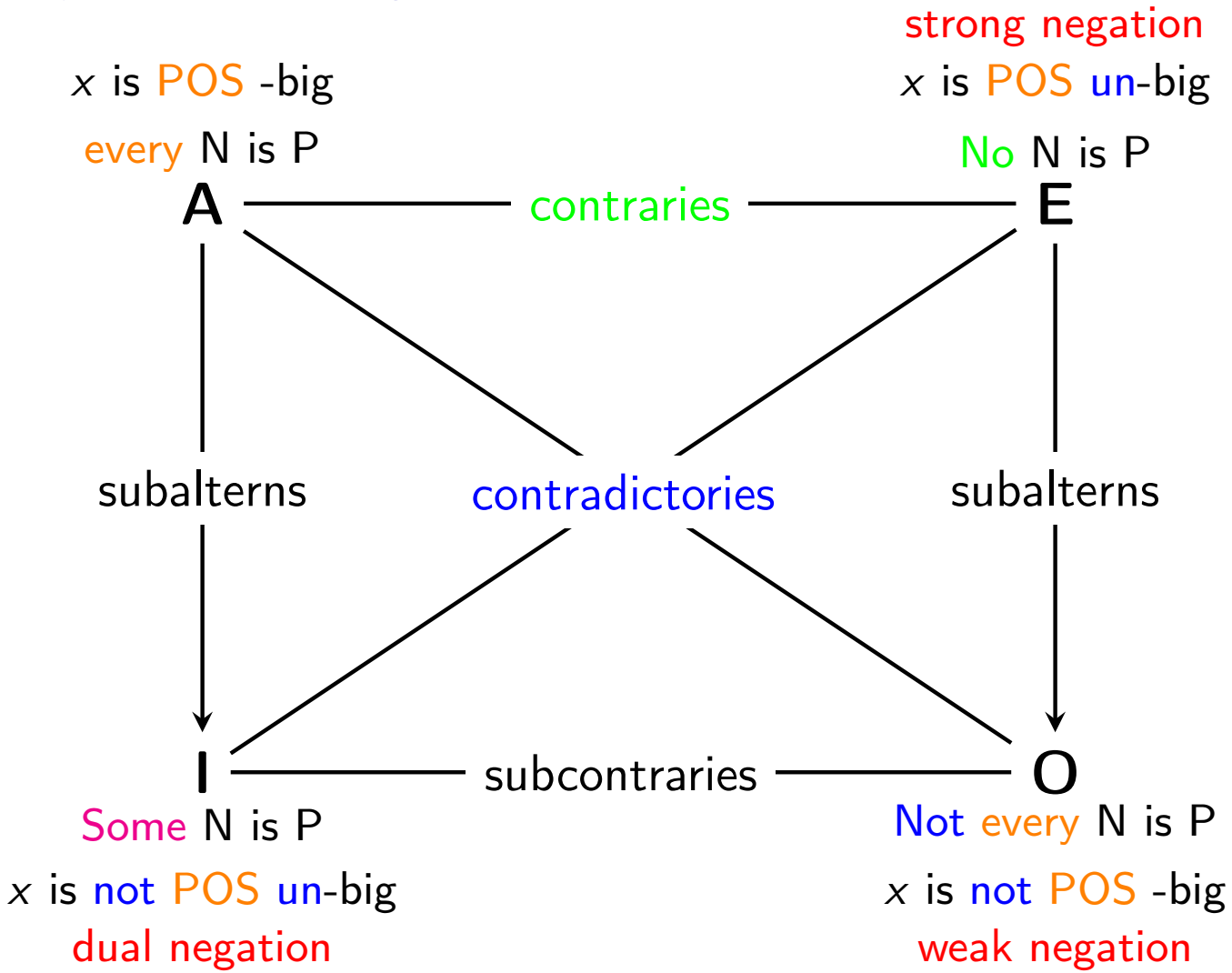
Gotzner's problem: Mind the gap



Problem solved: ∃**smart** and ∀**brilliant** turn out to be contradictions. Adjectives fit the Square of Opposition. Restrictions aligned.

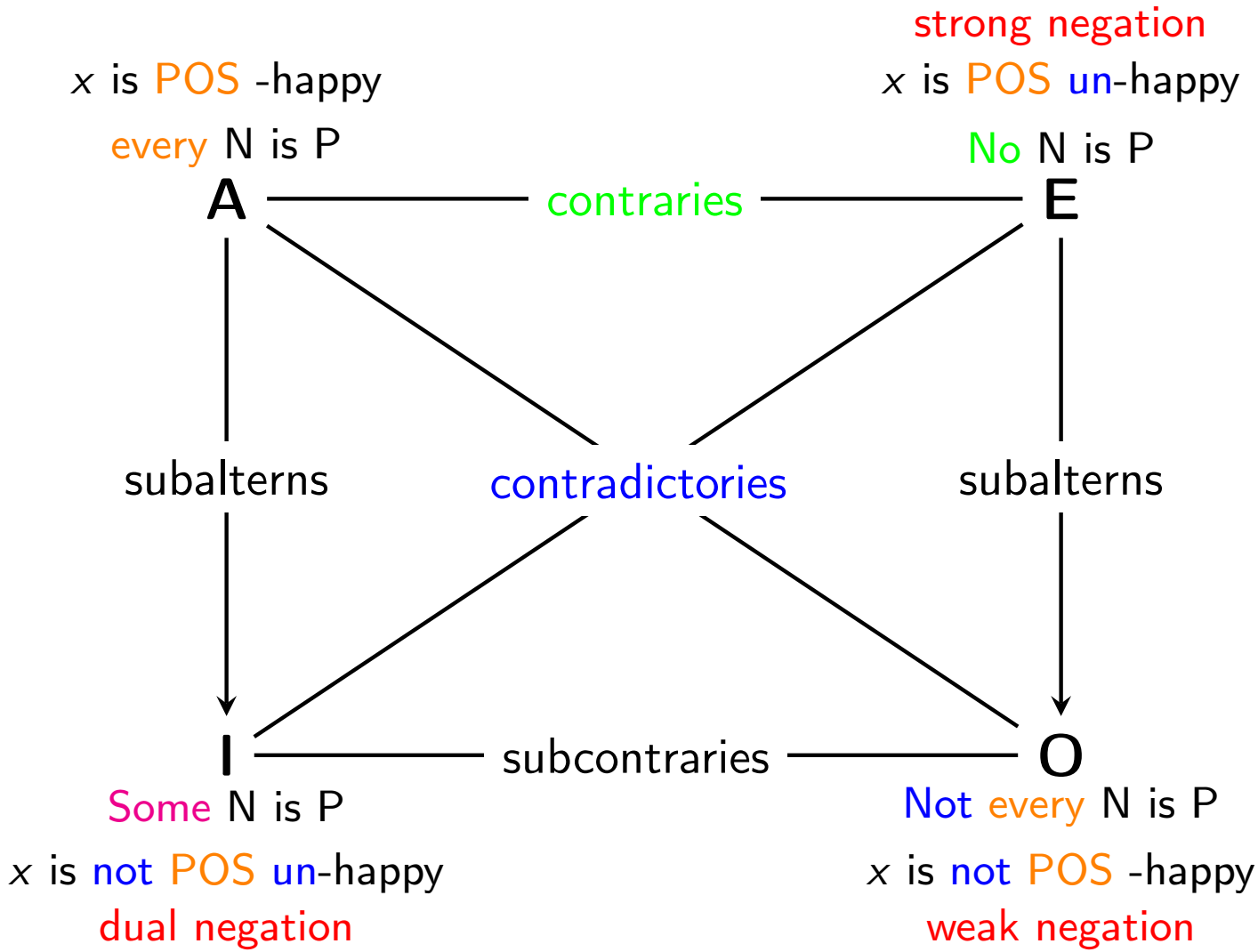
The Square of Opposition: prototypical **big**

Negative Polarity and Double Negation

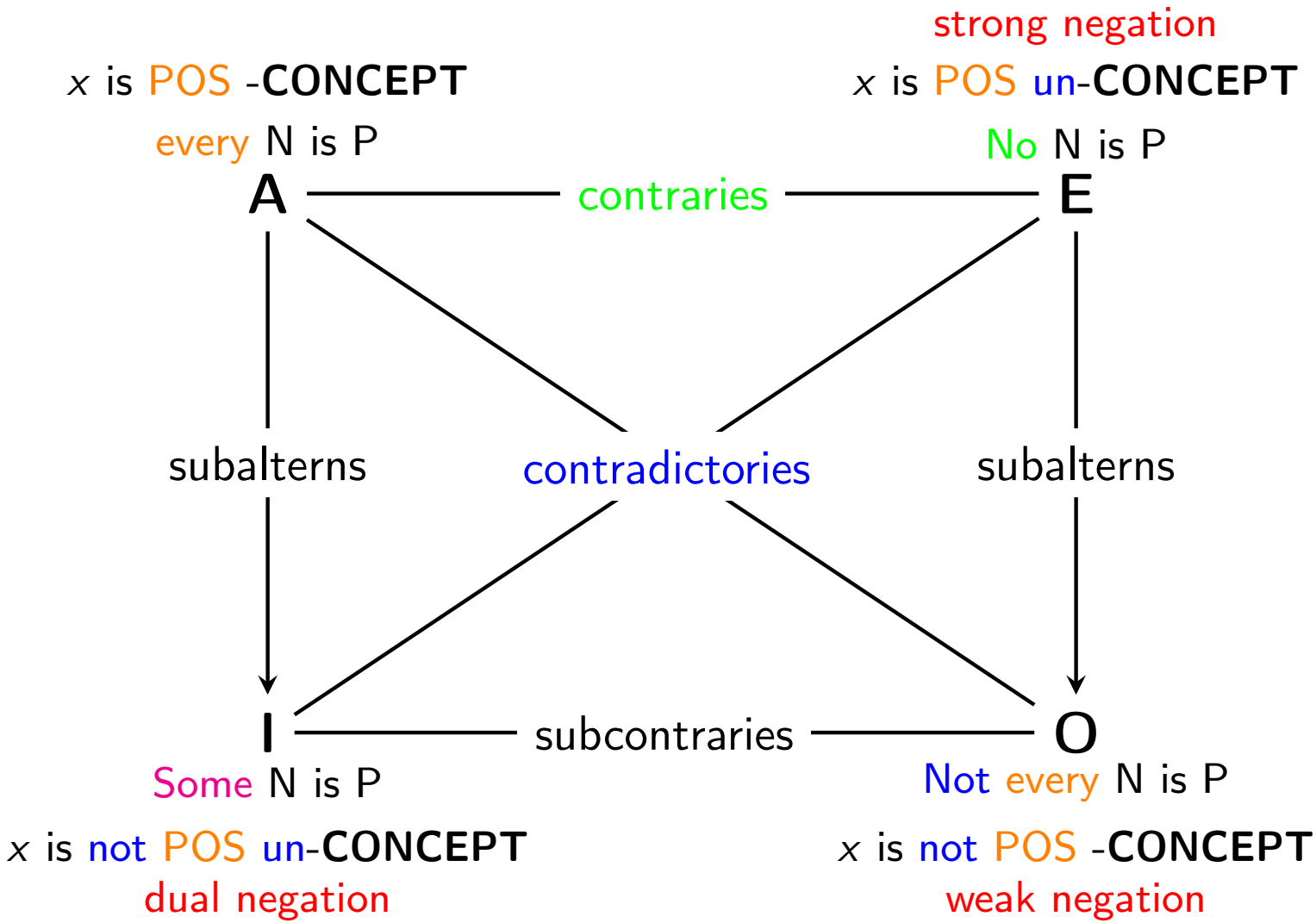


The Square of Opposition: **happy**

Evaluative Adjectives

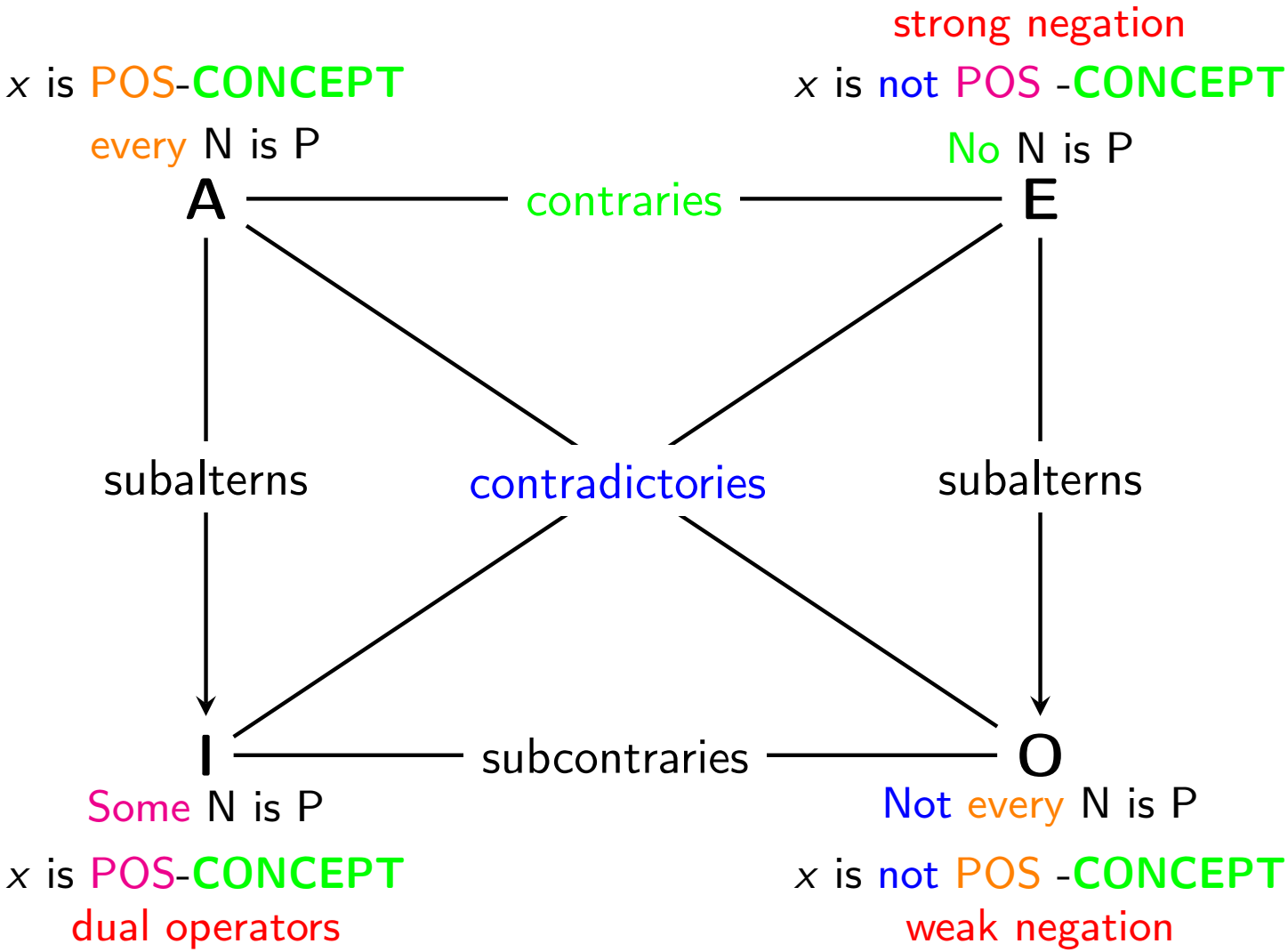


Prototypical **DEGREE-CONCEPT**



Adding extreme DEGREE-CONCEPTs

Domain widening triggers existential interpretation



Spell out: Quadruples

Mentioning the extreme adjective sets the widened gap (relative to the normal gap) and this triggers existential quantification.

- ⟨**happy**, **not unhappy**, **unhappy**, **not happy**⟩
- ⟨**all**, **some**, **no**, **not all**⟩
- ⟨**huge**, **big**, **small**, **not huge**⟩
- ⟨**big**, **not unbig**, **small**, **not big**⟩
- etc.

One and the same adjective can be part of different sets of alternatives. Mentioning the adjectives eliminates insecurity about the spell out.

Conclusions I

- Scalar implicatures with adjectives are more rare because they require a reinterpretation of the restriction. This is costly.
- Prototypical adjectives (those that are used in order to construct measurement scales) have an existential (basic?) and a universal meaning (stereotypical): polysemy? [This I changed because of Oana Lungu's remarks on how children develop quantification after the talk, see (Cournane et al., 2026)]
- Classical degree semantics Stechow (2009), Kennedy (1999), and Heim (2008) only capture the unversal (or stereotypical) meaning, not the basic one: Force variability.

Conclusions II

- The story fits well into work by Buccola et al. (2022): Alternatives are conceptual. and conceptual alternatives are more likely to be part of human reasoning Hu et al. (2023) [Language models and human reasoning?] The insecurity in building SCALES could also be about the force of the quantifier and scopal issues in the interaction with negation.
- The uniformity hypothesis might not be so far away anymore. If there is variation, it has to do with the context (Degen and Tanenhaus, 2015). The adjectival interpretation is context-sensitive in more ways than just the comparison class. Here, the variation is about the standard of comparison - the width of the gap.
- The square of opposition is compatible with measurement scales (contra Gotzner, Solt, et al. (2018))

Conclusions III

- In double-negation constructions, the degree quantifier intervenes between the two negations.
- Open questions? Yes, many!

Thank you! Comments are very welcome.

c.meier@lingua.uni-frankfurt.de

- I would like to thank my colleagues in Frankfurt, especially Jan Koepping, Caro Reinert and Helmut Weiß, and furthermore the audiences in New Haven, Saarbrücken especially Remus Gergel, in memoriam, and Nantes, especially Orin Percus, Oana Lungu, Mitcho Erlewine, Jérémy Pasquereau and Hamide Demirdache for the very fruitful discussion. Mistakes remain mine.

References I


Aparicio, Helena and Eszter Ronai (June 2025). "Scalar implicature rates vary within and across adjectival scales". In: Journal of Semantics 42.1-2, pp. 97–126. ISSN: 0167-5133. DOI: 10.1093/jos/ffaf002. eprint: https://academic.oup.com/jos/article-pdf/42/1-2/97/62787382/ffaf002.pdf. URL: https://doi.org/10.1093/jos/ffaf002.


Beltrama, Andrea and Ming Xiang (2013). "Is good better than excellent? An experimental investigation on scalar implicatures and gradable adjectives". In: Proceedings of Sinn und Bedeutung 17 held in Paris "Is excellent better than good?" Adjective scales and scalar implicatures.17, pp. 81–98.


Buccola, Brian et al. (2022). "Conceptual Alternatives. Competition in Language and Beyond.". In: Linguistics and Philosophy 45, pp. 265–291. DOI: https://doi.org/10.1007/s10988-021-09327-w.


Charlow S. (2019). . Ms., Rutgers University. (2018). Scalar implicature and exceptional scope. Tech. rep. Rutgers University. URL: https://ling.auf.net/%20lingbuzz/003181.

References II


Cournane, Aliís et al. (2026). "Word learning challenges explain nonadult possibility language comprehension in preschoolers". In: Developmental psychology 5, pp. 1048–1066.


Cresswell, Maxwell J. (1976). "The semantics of degree". In: Montague Grammar. Ed. by Barbara Partee. New York: Academic Press, pp. 261–292.


Degen, Judith and Michael K. Tanenhaus (2015). "Processing Scalar Implicature: A Constraint-Based Approach". In: Cognitive Science 39.4, pp. 667–710. DOI: https://doi.org/10.1111/cogs.12171. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/cogs.12171. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/cogs.12171.


Doran, Ryan B. et al. (2009). "On the Non-Unified Nature of Scalar Implicature: An Empirical Investigation". In: International Review of Pragmatics 1, pp. 211–248.

References III

- Gotzner, Nicole, Anton Benz, et al. (2018). "Adjectival scales and three types of implicature". In: Proceedings of SALT 28, pp. 409–432.
- Gotzner, Nicole, Stephanie Solt, et al. (2018). "Scalar Diversity, Negative Strengthening, and Adjectival Semantics". In: Frontiers in Psychology 9, p. 1659. ISSN: 1664-1078. DOI: 10.3389/fpsyg.2018.01659. URL: https://www.frontiersin.org/article/10.3389/fpsyg.2018.01659.
- Heim, Irene (2008). "Decomposing Antonyms". In: Proceedings of Sinn und Bedeutung 12.212-225.
- Horn, Laurence R. (1989). A natural history of negation. Reissue edition 2001. With a forew. by Laurence R. Horn. Stanford, CA: CSLI.
- Hu, Jennifer et al. (2023). "Expectations over unspoken alternatives predict pragmatic inferences". In: Transactions of the Association for Computational Linguistics. Vol. 11, pp. 885–901. DOI: https://doi.org/10.1162/tacl_a_00579.

References IV


Kennedy, Christopher (1999). Projecting the adjective: The syntax and semantics of gradability and comparison. New York: Garland.


Morzycki, Marcin (2012). "Adjectival extremeness: degree modification and contextually restricted scales". In: Natural Language and Linguistic Theory 30.2, pp. 567–609. DOI: 10.1007/s11049-011-9162-0. URL: http://dx.doi.org/10.1007/s11049-011-9162-0.


Pankratz, E. and B. van Tiel (2021). "The role of relevance for scalar diversity: A usage-based approach". In: Language and Cognition 13, pp. 562–94.


Ronai, Eszter and Ming Xiang (Dec. 2022). "Three factors in explaining scalar diversity". In: Proceedings of Sinn und Bedeutung 26, pp. 716–733.


Stechow, Arnim von (2009). "The temporal degree Adjectives *früher/später* and the semantics of the positive.". In: Quantification, Definiteness, and Nominalization. Ed. by Monika Rathert and Anastasia Giannakidou. Oxford: Oxford University Press.

References V

 Tiel, Bob van et al. (2016). "Scalar diversity". In: Journal of Semant 33, pp. 137–75.