# Scalar implicature and negative strengthening: *What it takes to be happy*

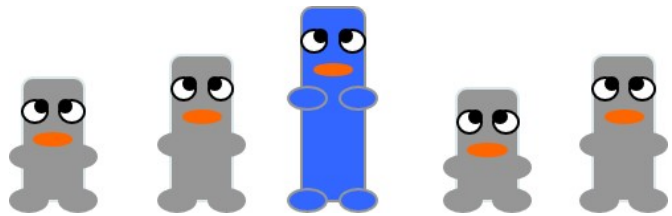
Nicole Gotzner, Stephanie Solt & Anton Benz

Leibniz Centre General Linguistics (ZAS)
Berlin

Annual Xprag.de Meeting
Graz

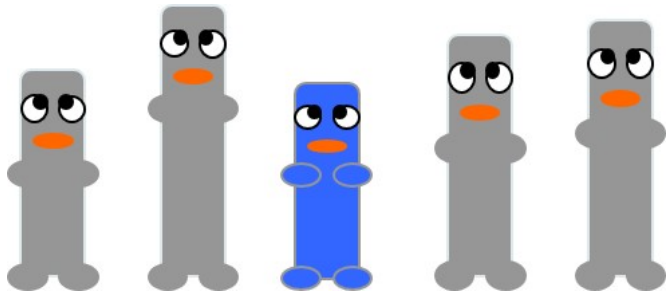
# Adjective semantics, standards and comparison classes

- Relative adjectives like *tall* are interpreted relative to a comparison class, their standard of comparison is determined contextually (Rotstein & Winter, 1994; Kennedy & McNally, 2005; Solt, 2009)
- Absolute adjectives like *empty* reference a fixed standard (experimental evidence: Frazier et al., 2008; Solt & Gotzner, 2012)



# Adjective semantics, standards and comparison classes

- Relative adjectives like *tall* are interpreted relative to a comparison class, their standard of comparison is determined contextually (Rotstein & Winter, 1994; Kennedy & McNally, 2005; Solt, 2009)
- Absolute adjectives like *empty* reference a fixed standard (experimental evidence: Frazier et al., 2008; Solt & Gotzner, 2012)



# The protagonist: Leo

- First word: absolute adjective *empty*
- 'Minimum standard' uses of relative adjectives: *hot* to refer to cold temperature (potential meaning: 'remarkable temperature')
- intermediate scale-mates *warm*, negative antonym *cold* produced later (*hot*: 16 months, *warm*: 19 months, *cold*: 21 months)

# Research questions

1. How does scale structure underlying the semantics of adjectives factor into implicature computation?
2. Which adjectives do children acquire first?

# Scalar implicature and the uniformity assumption

**Example (Scalar implicature)**

Leo ate some of the cookies.
Scalar implicature: Leo did not eat all of the cookies.
Scale: ⟨*some, all*⟩
Other scales: ⟨*or, and*⟩, ⟨*content, happy*⟩, ⟨*tall, huge*⟩, etc.

- **Uniformity assumption:**
- Scalar implicature is based on a single mechanism
- Behaviour of one scale generalises to the whole family of scales

- **Scalar diversity:**
- Experimental research demonstrates high variability in inference rates across triggers (Doran et al., 2009, 2011; van Tiel et al., 2016)
- Low rates for gradable adjectives

# Scalar implicature and negative strengthening

**Example (Scalar implicature, low endorsement)**

Leo is content.

Scalar implicature: Leo is not happy.

Scale: ⟨*content, happy*⟩

**Example (Negative Strengthening)**

Leo is not happy.

Negative Strengthening: Leo is not content.

Scale: ⟨*content, happy*⟩

# Negative Strengthening

- Negative strengthening: Pragmatic strengthening under negation, leading to stronger interpretation than semantic negation (Horn, 1989; Levinson, 2000)
- Q and R principle govern each other in conversation
- Q-principle: Say as much as you can (given R) $\rightsquigarrow$ SI
- R-principle: Say no more than you must (given Q) $\rightsquigarrow$ NegS

**Example (Politeness/Euphemism)**

'Leo is not very bright' $\rightsquigarrow$ Leo is rather stupid

# Scalar diversity: van Tiel et al. (2016)

- Inferential task with 32 adjectives, 6 main verbs, 2 auxiliary verbs, 2 quantifiers, 1 adverb

High variance in endorsement rates, for example:

- $\langle$*some, all*$\rangle$: 96%
- $\langle$*like, love*$\rangle$: 50%
- $\langle$*content, happy*$\rangle$: 4%

Factors explaining part of variability:

- Boundedness: Higher endorsement rate for upper-bounded scales
- Semantic Distance: Perceived difference in strength between weaker and stronger scale-mate

# van Tiel et al.: semantic interpretation, SI and NS

- Participants interpret semantically.

---

John says:

*She is content.*

Would you conclude from this that, according to John, she is not happy?

☐ Yes ☑ No

---

# van Tiel et al.: semantic interpretation, SI and NS

- Participants interpret semantically.
- Participants draw scalar implicature.

---

John says:

*She is content.*

Would you conclude from this that, according to John, she is not happy?

☑ Yes ☐ No

---

# van Tiel et al.: semantic interpretation, SI and NS

- Participants interpret semantically.
- Participants draw scalar implicature.
- Participants apply negative strengthening.

---

John says:

*She is content.*

Would you conclude from this that, according to John, she is not happy?

☐ Yes ☑ No

---

# Experiment 1: Research Goals

- Investigate diversity in a large set of adjectival scales varying in scale structure
- Assess the interplay between scalar implicature and negative strengthening
- Determine factors that explain scalar diversity and negative strengthening

# Experiment 1: Materials and Tasks

- 70 adjective pairs with weak and strong scale-mates
- Main tasks: Rating of statement involving scalar implicature (SI) and negative strengthening (NegS)
  - **Annotation of scale structure:**
  - **Boundedness**: open, upper bounded, lower bounded scale (Kennedy & McNally, 2005)
  - **Polarity**: positive form vs. negative (Cruse, 1986)
  - **Extremeness**: extreme vs. non-extreme adjectives (Morzycki, 2012)
    - **Additional measures:**
    - Cloze task: Mention 3 words associated with weaker term
    - Semantic Distance: Evaluate difference in strength on 7-point scale
    - Politeness rating: How nice is weaker, stronger and negated stronger statement on 7-point scale (based on Bonnefon et al., 2009)
    - Relative frequency (COCA) and LSA values

# NegS Task

Mary says:

*He is not brilliant.*

Would you conclude from this that, according to Mary, he is not intelligent?

Yes No

Figure: Sample item of the negative strengthening task

# Main results: Correlation of endorsements in SI and NegS task

Table: Example scales and their respective endorsement rates in the scalar implicature (SI), negative strengthening (NegS) task

| **Weak** | **Strong** | **Scale structure** | **SI** | **NegS** |
|---|---|---|---|---|
| attractive | stunning | unbounded extreme positive | 0.08 | 0.6 |
| wet | soaked | lower bounded extreme positive | 0.24 | 0.44 |
| cool | cold | unbounded non-extreme negative | 0.43 | 0.54 |
| difficult | impossible | upper bounded extreme negative | 0.76 | 0.35 |

**Anti–correlation:** -.66, $p < .0001$

(**Correlation with scale reversal:** .34, $p < .01$)

# Predictors of SI and NegS

- **Scalar implicature:**
- Extremeness: lower for extreme adjectives
- Semantic distance: higher for distant scale-mates
- Boundedness: higher for upper bounded scales
- Polarity: higher endorsement for negative antonyms
- **Negative strengthening:**
- Extremeness: higher for extreme adjectives
- Semantic distance: higher for distant scale-mates
- Cloze probability: higher endorsement for strongly associated scale-mates

Model for SI with NegS rates has improved fit (R-squared = 0.58 vs. 0.64) and other factors remain significant

# Discussion: Relevance of factors

- **Boundedness**: obligatory computation of standard with absolute adjectives (Frazier et al., 2008), constraints on implicature related to vagueness (Leffel, Cremers, Gotzner & Romoli, accepted)
- **Extremeness**: restricted attention to a particular salient portion of a scale (Morzycki, 2012); emotive component
- **Polarity**: for positive form the whole scale is relevant while negative items introduce presupposition (Cruse, 1986)
- Scale structure determines relevant alternatives in inference computation

# Discussion: Interaction between scalar implicature and negative strengthening

- Q and R principle govern each other
- Negative strengthening stands in competition with scale reversal/indirect scalar implicature
- Current work: Investigate clusters of implicature triggers, develop account of negative strengthening

# Corpus study: The acquisition of adjectives

- **Research question**: Do factors related to scale structure predict the order of acquisition of gradable adjectives?
- **Current Method**: Corpus search of production frequencies in Childes (Childfreq, Baath, 2010)
- **Annotation**: boundedness, polarity, negative morpheme, positioning on scale (basic level, intermediate scale mate, extreme term), subjectivity
- Additional predictors: word length, number of syllables, frequency (COCA)

# Previous studies on adjective acquisition

- 4-year olds shift the standard of comparison in relative adjectives like *tall* depending on context (Barner & Snedeker, 2008; Syrett et al., 2006)
- 3-year olds know distinction between relative and absolute adjectives (Syrett et al., 2009)
- comprehension of positive adjectives before negative ones (Clrk, 1973; Barner & Snedeker, 2008)

# Corpus study: Results

- minimum standard adjectives produced more frequently than relative ones
- positive before negative adjectives (see also Katsos et al., 2016 for similar finding in quantifier domain)
- basic level terms before intermediate and extreme scale-mates

# Example adjectives produced at age 1

- minimum standard adjectives: dirty (9), wet (12)
- max.: clean (12), full (18), dry (12)
- relative: big (9), good (9), hot (12), warm (18)
- later onset:
  - modals: allowed (30), likely (66), possible (78)
  - extreme: huge (21), enormous (60)

# Discussion

- categorical uses of relative adjectives (Smith et al., 1986)
- potential reason minimum standard adjectives are easier: simple learning by ostension strategy
- acquiring compositional semantics of relative adjectives is more demanding
- what is the role of input frequency and how are adjectives used in child directed speech?

# Conclusions

- Numerical correlation between different types of implicature:
  - Scalar implicature (Q implicature)
  - Negative strengthening (R/M implicature)
- Scale structure underlying the semantics of gradable adjectives predicts variation in inference rates and order of adjective acquisition
- Scale structure should be integrated in theories of pragmatic strengthening (see Leffel et al., accepted for concrete proposal)

Thank you for your attention! And thanks to our student assistants Franziska Conradts and Henry Salfner!

# Leffel, Cremers, Gotzner & Romoli (accepted): Vagueness and Implicature

- Effect of scale structure: Contrast between minimum standard and relative adjectives:
- 'John is not very tall' $\rightsquigarrow$ John is not tall
- 'John was not very late' $\rightsquigarrow$ John was late
- Constraint on vague implicatures: Do not derive implicature if it leads to borderline contradiction
- Definition borderline contradiction: If F and G are gradable predicates on the same scale and with the same polarity (Fa $\wedge\neg$ Ga) is a borderline contradiction iff a is a borderline case for both F and G.

# Predictors: Endorsements in SI task

Table: Predictors of endorsements in (a) the scalar implicature and (b) negative strengthening task.

| **(a) Scalar implicature** | Estimate | SE | t-value | p-value | |
|---|---|---|---|---|---|
| (Intercept) | -0.450962 | 0.200752 | -2.246 | 0.0285 | |
| lower bounded | -0.025124 | 0.052719 | -0.477 | 0.6355 | |
| upper bounded | 0.091486 | 0.045143 | 2.027 | 0.0473 | * |
| extremeness | -0.273715 | 0.052290 | -5.235 | 2.39e-06 | *** |
| polarity | 0.114126 | 0.048065 | 2.374 | 0.0209 | * |
| semantic distance | 0.155524 | 0.028026 | 5.549 | 0.000 | *** |
| politeness weak | 0.008772 | 0.036372 | 0.241 | 0.8103 | |
| politeness strong | 0.020949 | 0.021826 | 0.960 | 0.3411 | |
| relative frequency | 0.009177 | 0.022509 | 0.408 | 0.6850 | |
| LSA | 0.012314 | 0.105014 | 0.117 | 0.9071 | |

# Predictors: Endorsements NegS task

Table: Predictors of endorsements in (a) the scalar implicature and (b) negative strengthening task.

| **(b) Negative strengthening** | Estimate | SE | t-value | p-value | |
|---|---|---|---|---|---|
| (Intercept) | 1.472420 | 0.307221 | 4.793 | 0.000 | |
| lower bounded | -0.039224 | 0.043193 | -0.908 | 0.36764 | |
| upper bounded | -0.046990 | 0.037755 | -1.245 | 0.21837 | |
| extremeness | 0.146204 | 0.045073 | 3.244 | 0.00197 | ** |
| polarity | -0.008700 | 0.039276 | -0.222 | 0.82548 | |
| semantic distance | -0.109197 | 0.025060 | -4.357 | 0.000 | *** |
| politeness weak | -0.043342 | 0.025452 | -1.703 | 0.09403 | . |
| politeness 'not' strong | -0.049070 | 0.045032 | -1.090 | 0.28044 | |
| relative frequency | 0.005728 | 0.018199 | 0.315 | 0.75409 | |
| LSA | -0.131505 | 0.091573 | -1.436 | 0.15645 | |
| cloze probability | 0.494745 | 0.222365 | 2.225 | 0.03006 | * |

# Overview of Tasks

| Label | Task | Intended measure |
|---|---|---|
| Main task SI | Inference judgment (yes\no) | scalar implicature |
| Main task NegS | Inference judgment (yes\no) | negative strengthening |
| Semantic distance | strength rating (1-7 scale) | scale distinctness |
| Cloze task | free word production | association strength |
| Politeness weak | kindness rating (1–7 scale) | weak statement |
| Politeness strong | kindness rating (1–7 scale) | strong statement |
| Politeness 'not' strong | kindness rating (1–7 scale) | negated strong statement |

Table: Overview of tasks