Scalar implicatures vary within and across adjectival scales

Helena Aparicio (Cornell University) & Eszter Ronai (Northwestern University)

SALT 33

Contact: haparicio@cornell.edu
ronai@northwestern.edu

Scalar Diversity

Lexical scales differ in how likely they are to lead to scalar implicature (SI), e.g., (1) more likely than (2), see van Tiel et al. (2016):

(1) The museum is **old** →
The museum is **not ancient**

(2) The employee is **smart** →
The employee is **not brilliant**

Role of carrier sentences remains understudied (van Tiel et al. (2016) found no difference, cf. Degen (2015) for <some, all>).

Research Question

What is the role of sentential context in scalar diversity?

Does likelihood of Comparison Class (CC) having adjectival property modulate SI rates?

Manipulate CC: whether the noun (e.g., *scientist* vs. *employee*) is likely to have adjectival property (e.g., brilliance).

Gathering CCs (Norm. Study 1)

- Shown stronger scalemates (e.g., brilliant)
- Elicit nouns likely to have the property
- Two nouns selected per scale: one high frequency ("biased") and one very infrequent (≈ 1 count; "neutral").

Gathering Scales (Norming Study 2)

- 77 adjectival scales from previous work normed for cancellability and asymmetric entailment.
- Criterion: above 60% expected response
- Result: 45 scales

Task 1 (de Marneffe & Tonhauser, 2019)

X is brilliant... and even smart — "Odd"
X is smart... and even brilliant — "Not odd"

Task 2

X was brilliant... but not smart — "Contradictory"
X was smart... but not brilliant — "Not contradictory"

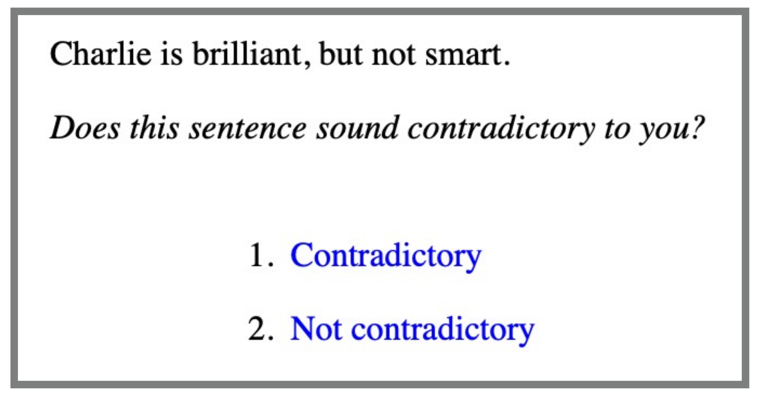


Fig. 1: Norm2 Trial Ex.

Hypothesis 1: Likelihood

- SI: reasoning about what was left unsaid (Grice, 1967; Horn, 1972)
- Biased nouns: the stronger adjective very likely to be true → its non-utterance is especially meaningful

H1 predicts higher SI rates for biased compared to neutral CCs.

How likely are {scientists, employees} to be brilliant?

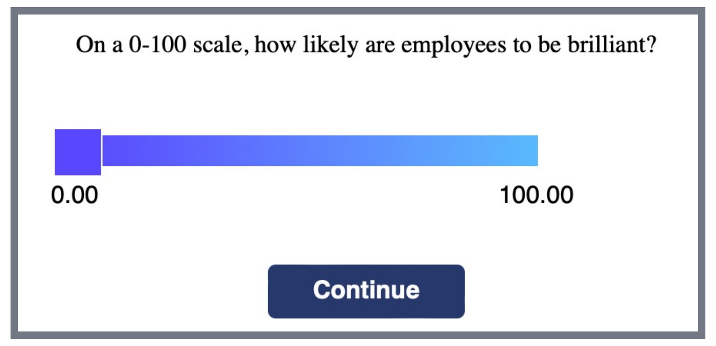


Fig. 2: Exp1 Trial Ex.

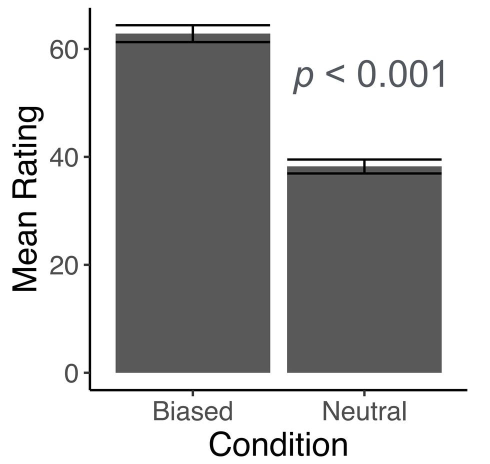


Fig. 3: Exp1 Results

Hypothesis 2: Threshold distance

- Semantic distance: close proximity between adjectival thresholds discourages SI calculation
- Elicit threshold (θ) distributions

(3) $P(\theta_w|\theta_{\neg s}) \propto P(\theta_{\neg s}|\theta_w)P(\theta_w)$

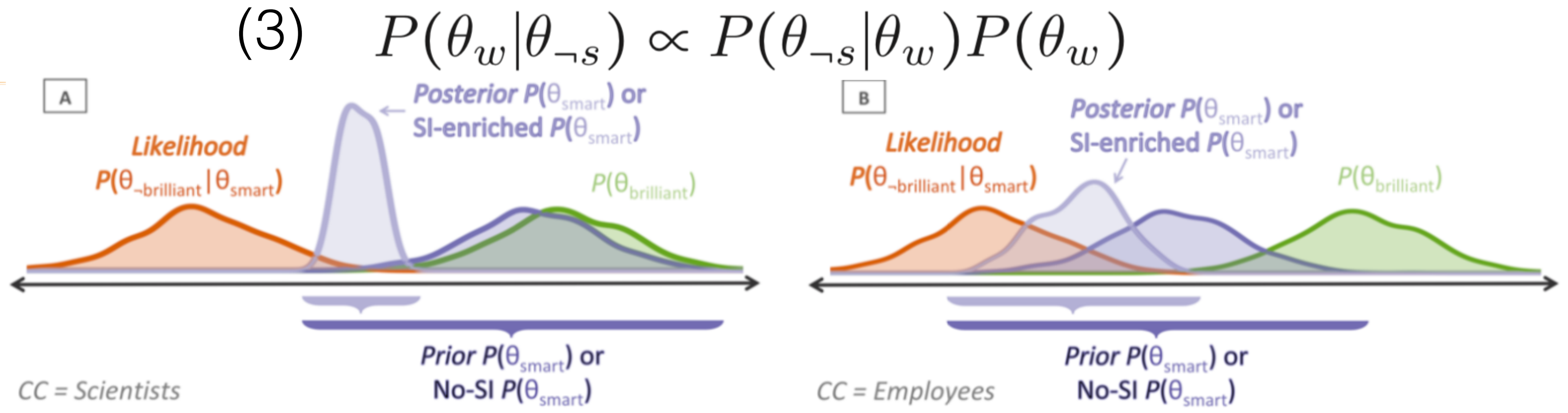


Fig. 4: Model Predictions

H2 predicts higher SI rates for neutral compared to biased CCs.

The {scientist/employee} is {smart, possibly brilliant/brilliant}

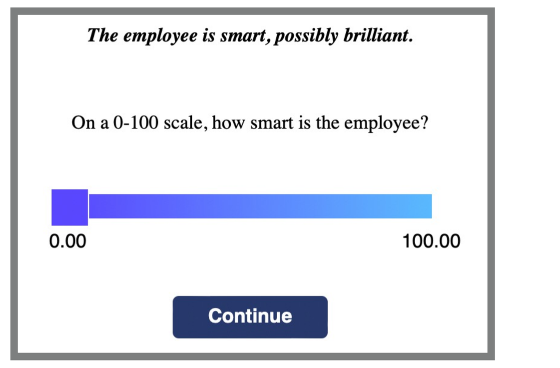
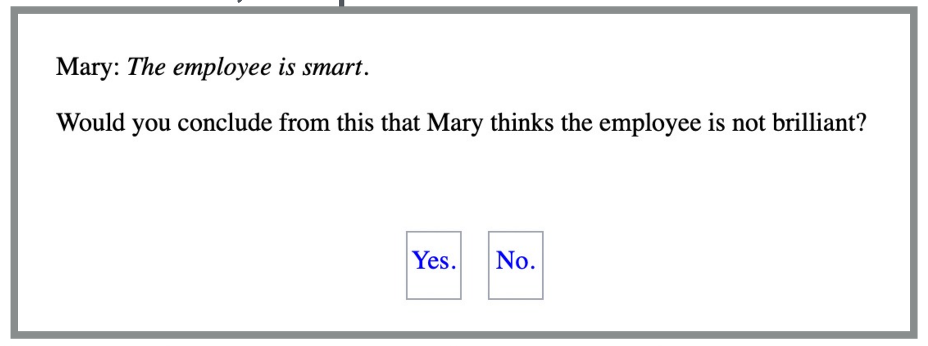


Fig. 5: Exp2 θ-elicitation Trial

SI Calculation

- SI task, Experiment 3:



"Yes" = SI calculation; "No" = no SI calculation

Fig. 6: Exp3 Trial Example

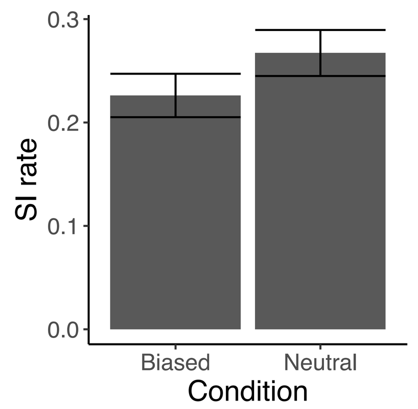


Fig. 7: Exp3 Results

Lower rates of SI calculation in biased condition ($p < 0.05$)

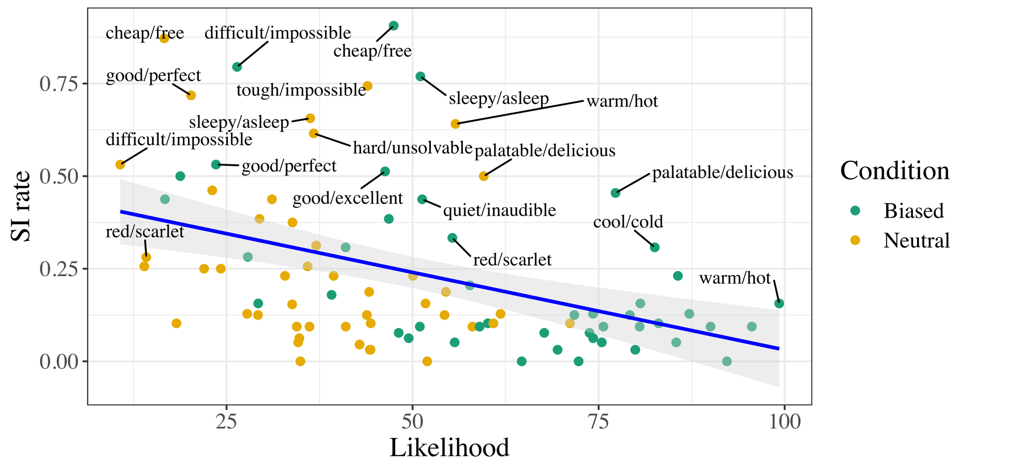


Fig. 8: SI rate ~ likelihood

Contra H1

Likelihood and SI negatively correlated

($r = -0.42$, $p < 0.001$)

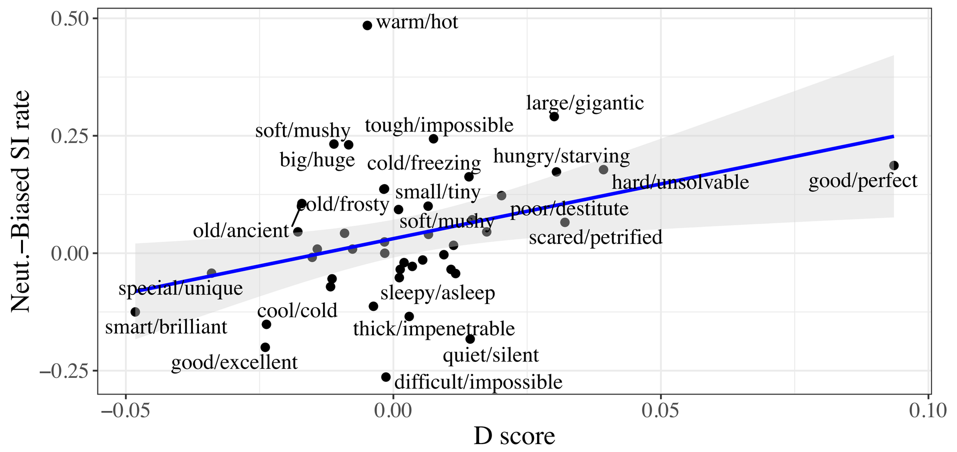


Fig. 9: Neut-Bias SI rate ~ *D*-score

In line with H2

D-score positively correlated with neutral−biased SI rate

($r = 0.36$, $p < 0.02$)

- Distance (*D*)-score: (3) $d_n = (\mu_{s_n} - \mu_{w_n})/\sigma_{s_n}\sigma_{w_n}$ (4) $D = d_{n_{neut.}} - d_{n_{bias.}}$

Conclusion

- **Contra H1, biased nouns lead to less SI.**
- **Semantic distance (e.g., van Tiel et al. 2016; Horn 1972) between adjective thresholds better predictor of SI, supporting H2.**
- **Results highlight the methodological importance of controlling for carrier sentences.**