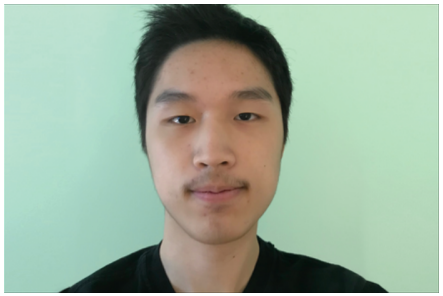
Granularity in the Semantics of Comparison

Helena Aparicio
Cornell University

haparicio@cornell.edu

joint work with



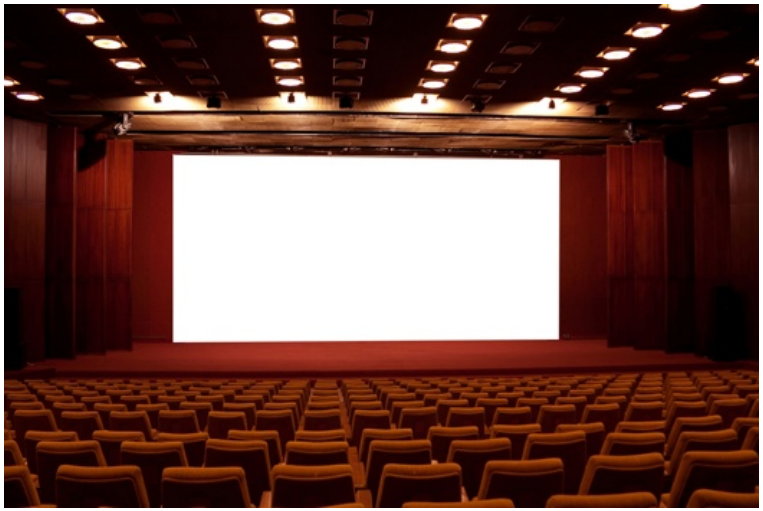
Curtis Chen (MIT)

Elizabeth Coppock (BU)

Roger Levy (MIT)

Imprecision & Gradable Adjectives

Positive Form
Maximum Standard Adjectives

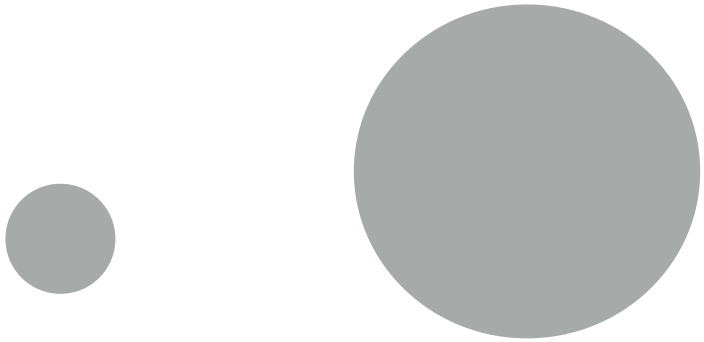


***Empty* movie theater**

Kennedy & McNally, 2005; Kennedy, 2007; Syrett et al., 2009;
Solt, 2015; Aparicio et al., 2015; Leffel et al., 2016; Qin, 2020, a.o.

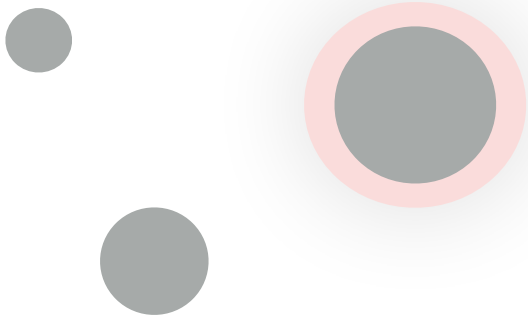
Definite Comparatives

The bigger circle.



Comparison Class

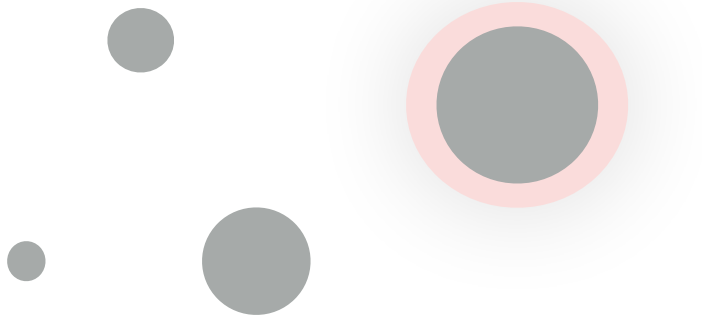
The biggest circle.



The bigger circle.

Comparison Class

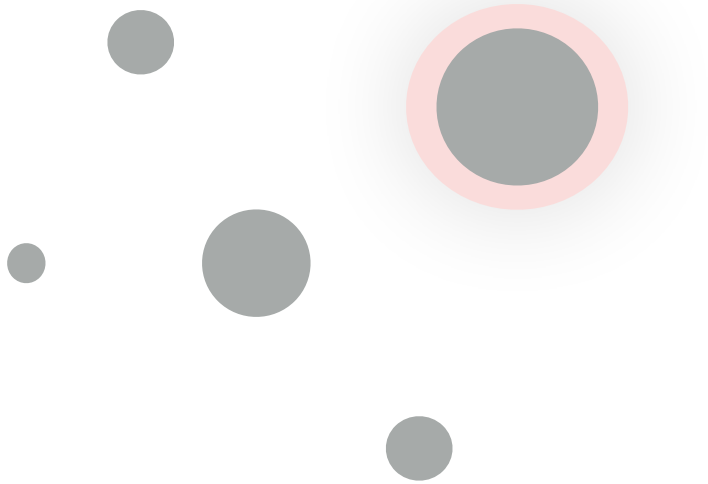
The biggest circle.



The bigger circle.

Comparison Class

The biggest circle.

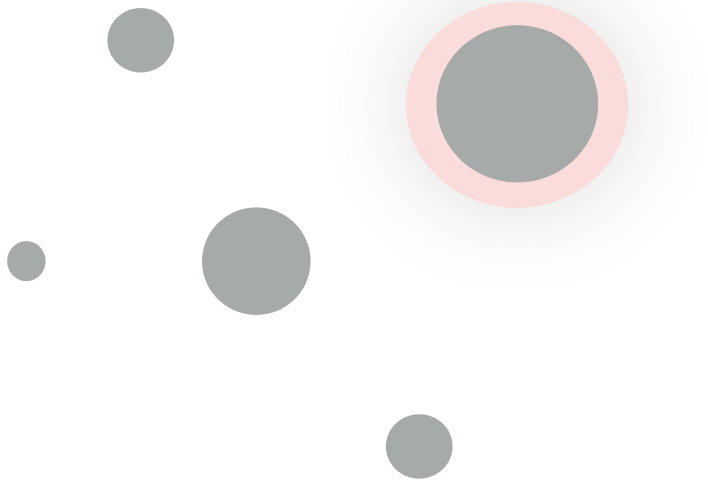


The bigger circle.

Comparison Class

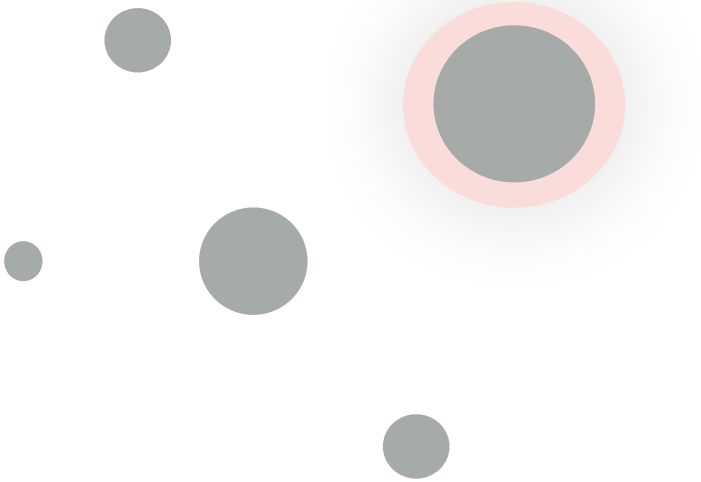
The biggest circle.

The bigger circle.

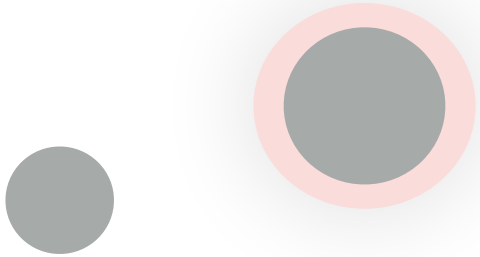


Comparison Class

The biggest circle.

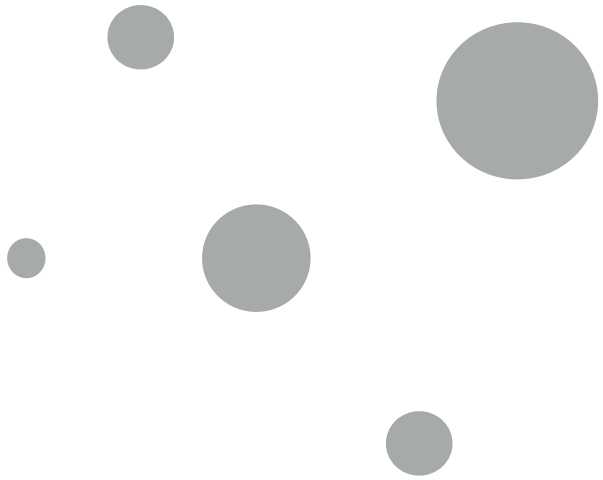


The bigger circle.

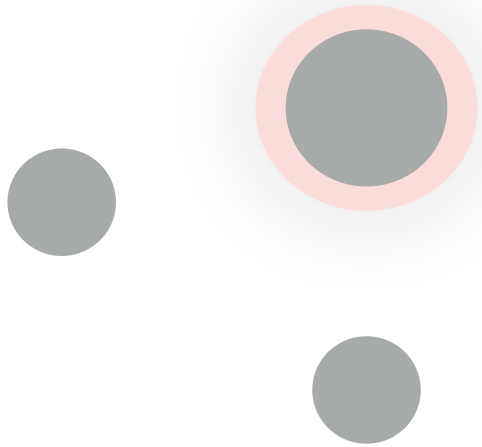


Comparison Class

The biggest circle.

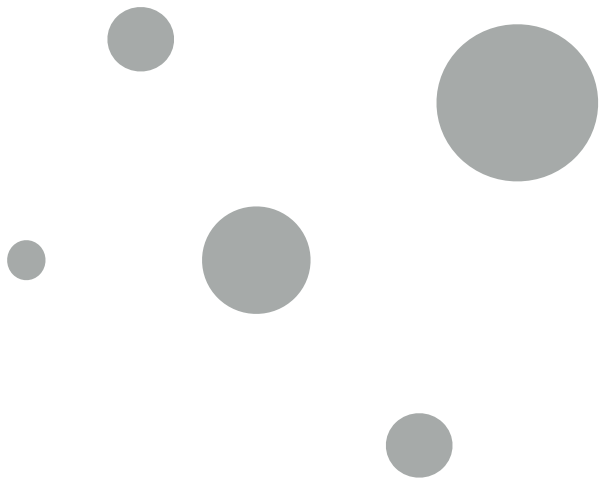


The bigger circle.

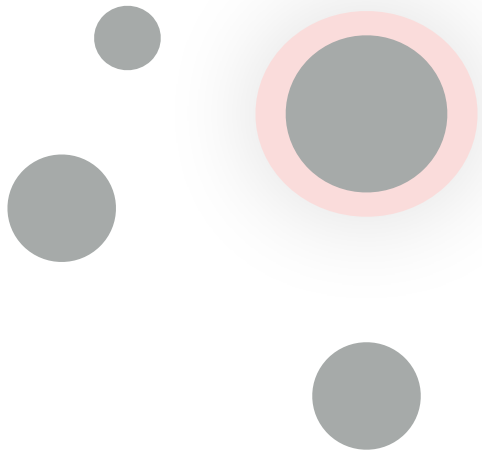


Comparison Class

The biggest circle.

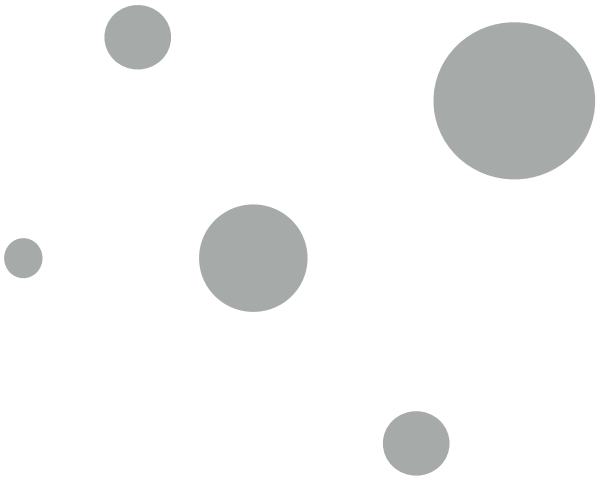


The bigger circle.

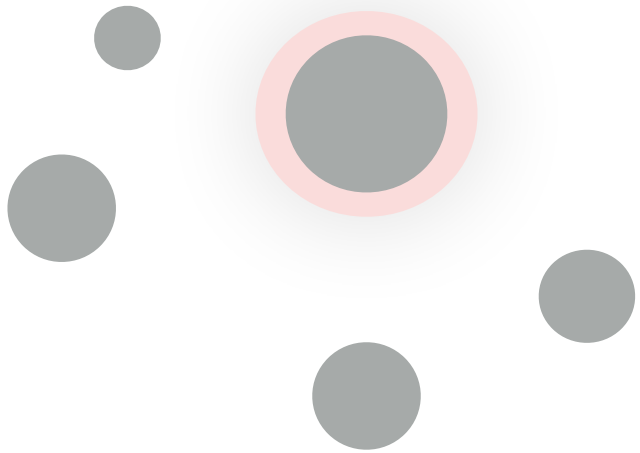


Comparison Class

The biggest circle.

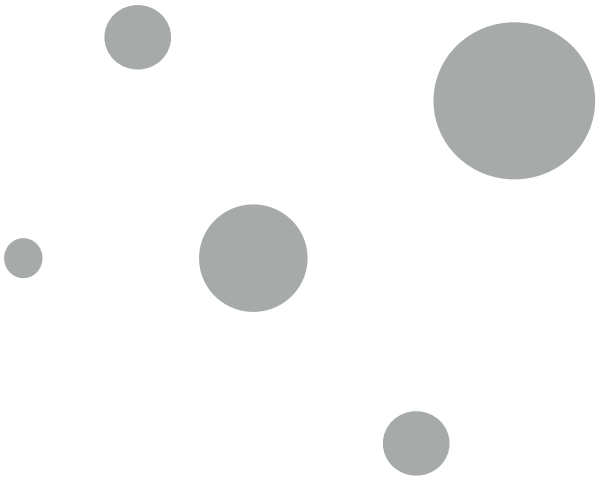


The bigger circle.

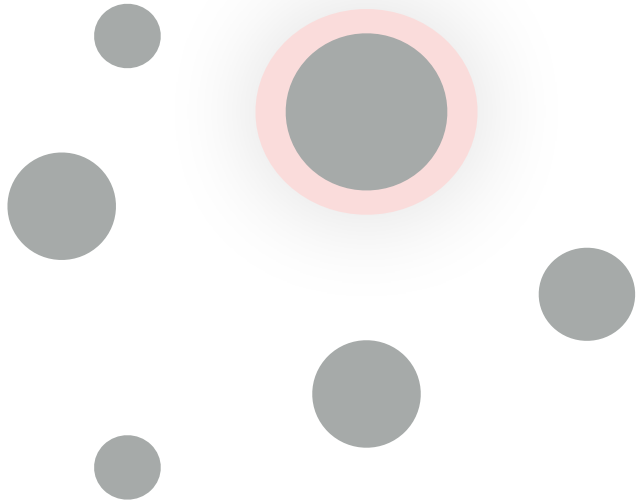


Comparison Class

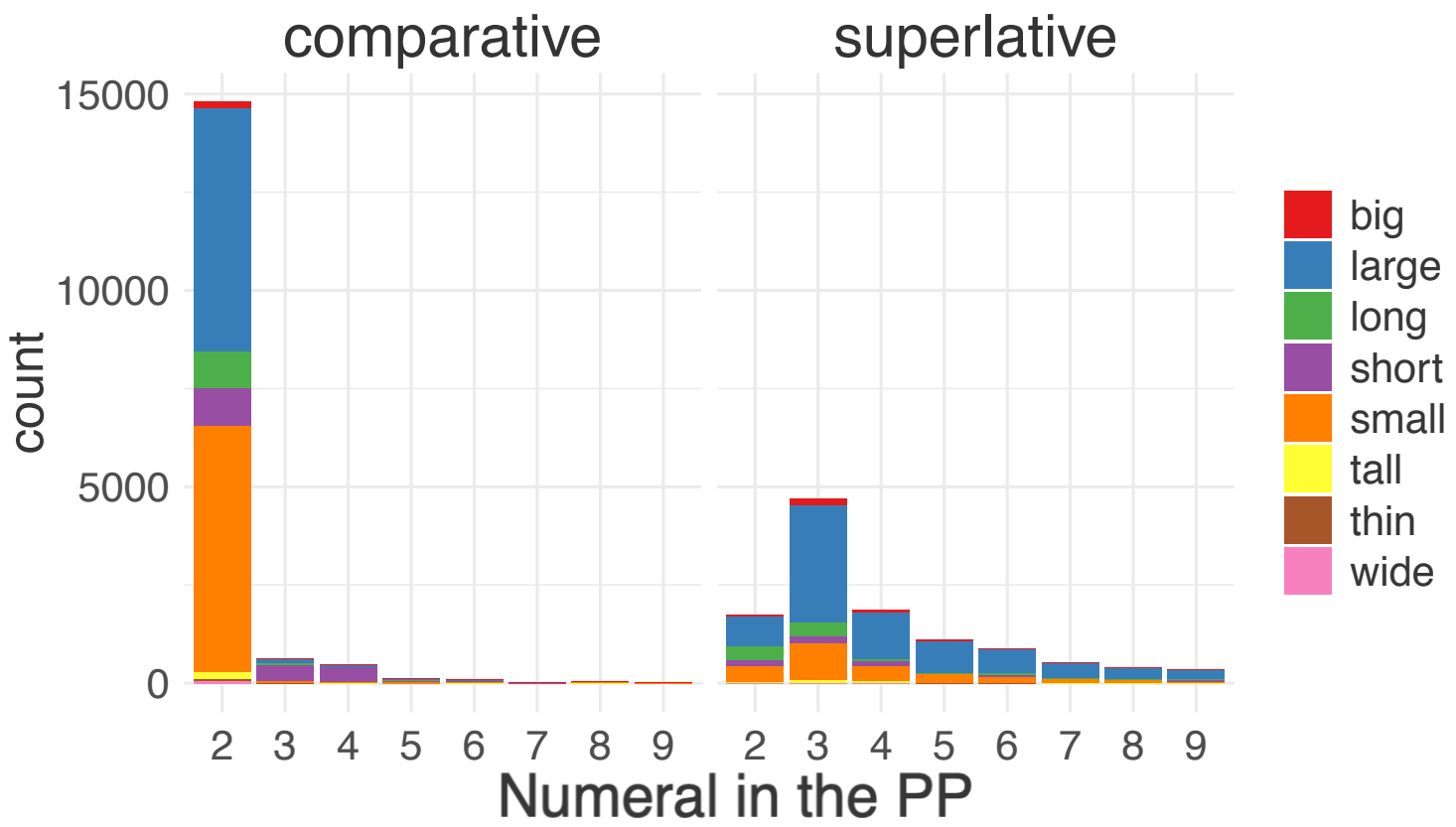
The biggest circle.



The bigger circle.



Corpus Data



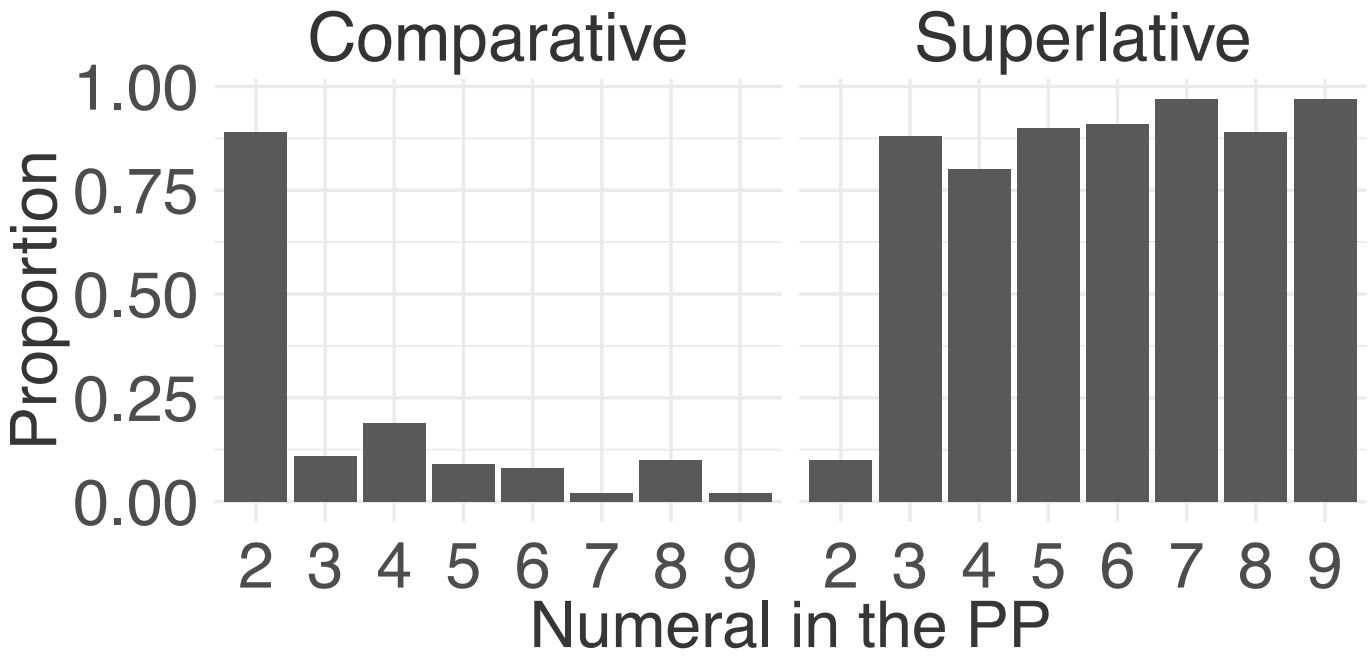
Query:

```
the [A] of [Num]
```

e.g., ‘the smaller of two’

Internet Archive Corpus (English)

Corpus Data


Comparative
Superlative
Proportion
1.00
0.75
0.50
0.25
0.00
2 3 4 5 6 7 8 9
2 3 4 5 6 7 8 9
Numeral in the PP

Question

What constraints Comparison Class calculation of a definite comparative?

2 Individuals (2I) Theory

(1) $[\![\text{-}er]\!] = \lambda C_{\langle e,t\rangle} \lambda A_{\langle d,\langle e,t\rangle\rangle} \lambda x_e : x \in C \wedge |C| = 2\,.$

$$\exists x' \in C : max\{d|A(d)(x)\} > max\{d|A(d)(x')\}$$

2 Degrees (2D) Theory

(2) $[\![\text{-}er]\!]^{\gamma} = \lambda C_{\langle e,t\rangle} \lambda A_{\langle d,\langle e,t\rangle\rangle} \lambda x_e : x \in C \wedge |\gamma_{A,C}| = 2\,.$

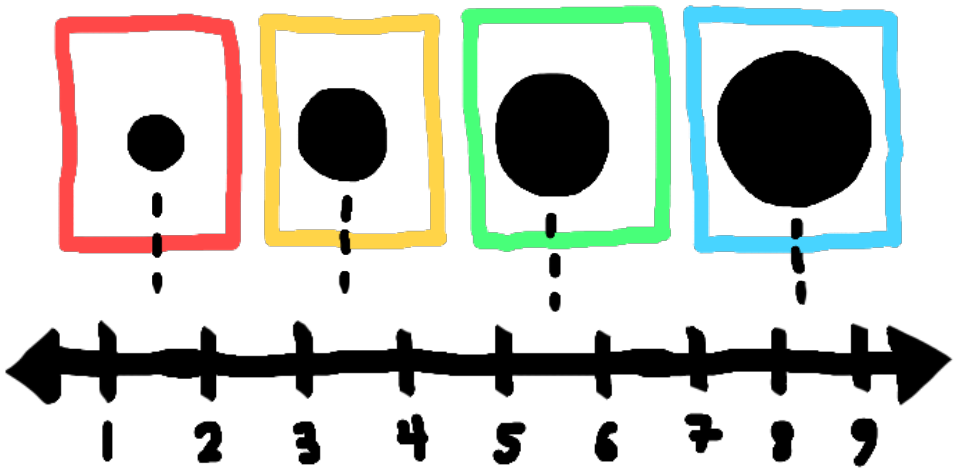
$$\exists d' \in \gamma_{A,C} : max\{d \in \gamma_{A,C} | A(d)(x)\} > d'$$

γ_A is a set of degrees of A-ness (a granularity of the A-scale)

$$\gamma_{A,C} = \{d \in \gamma_A : \exists x \in C : \max\{d' \in \gamma_A | A(d')(x)\} = d\}$$

Krifka (2007)

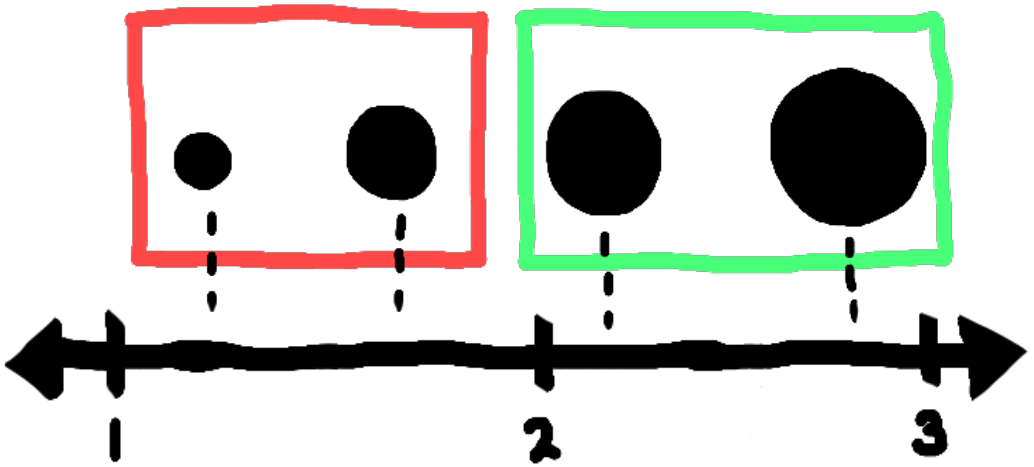
Granularity



$|\gamma_{A,C}| = 2$

✗

Granularity



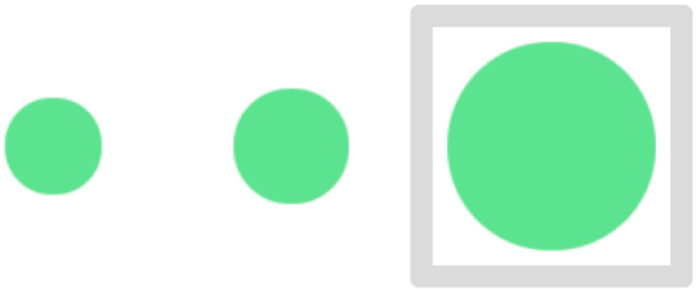
$|\gamma_{A,C}| = 2$

✓

The Experiment

Experimental Task

This is the biggest circle.

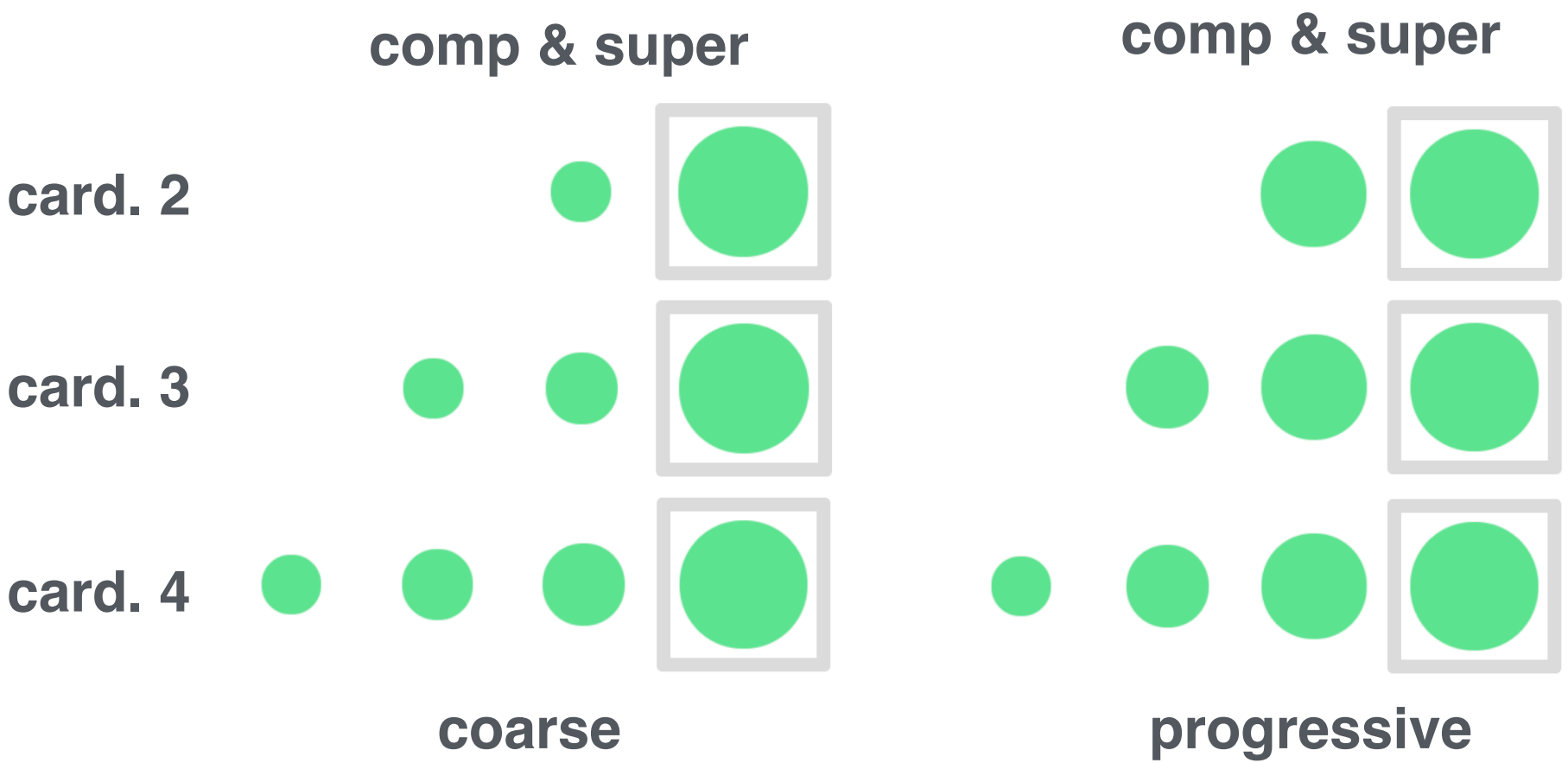


How acceptable is the above description of the scene?

	1	2	3	4	5	**6**	7	
(Bad)	○	○	○	○	○	●	○	(Good)

Back *Next*

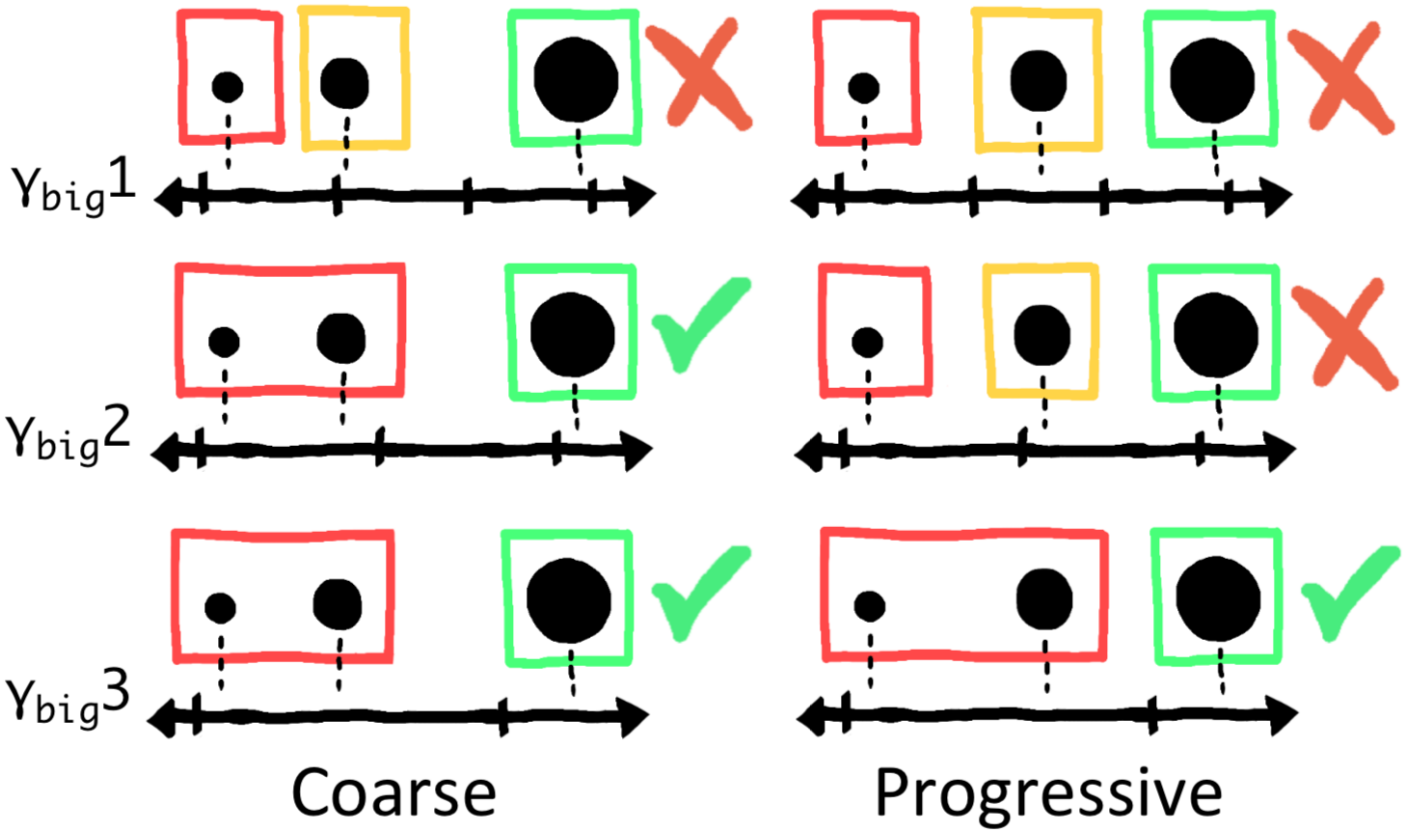
Study Design



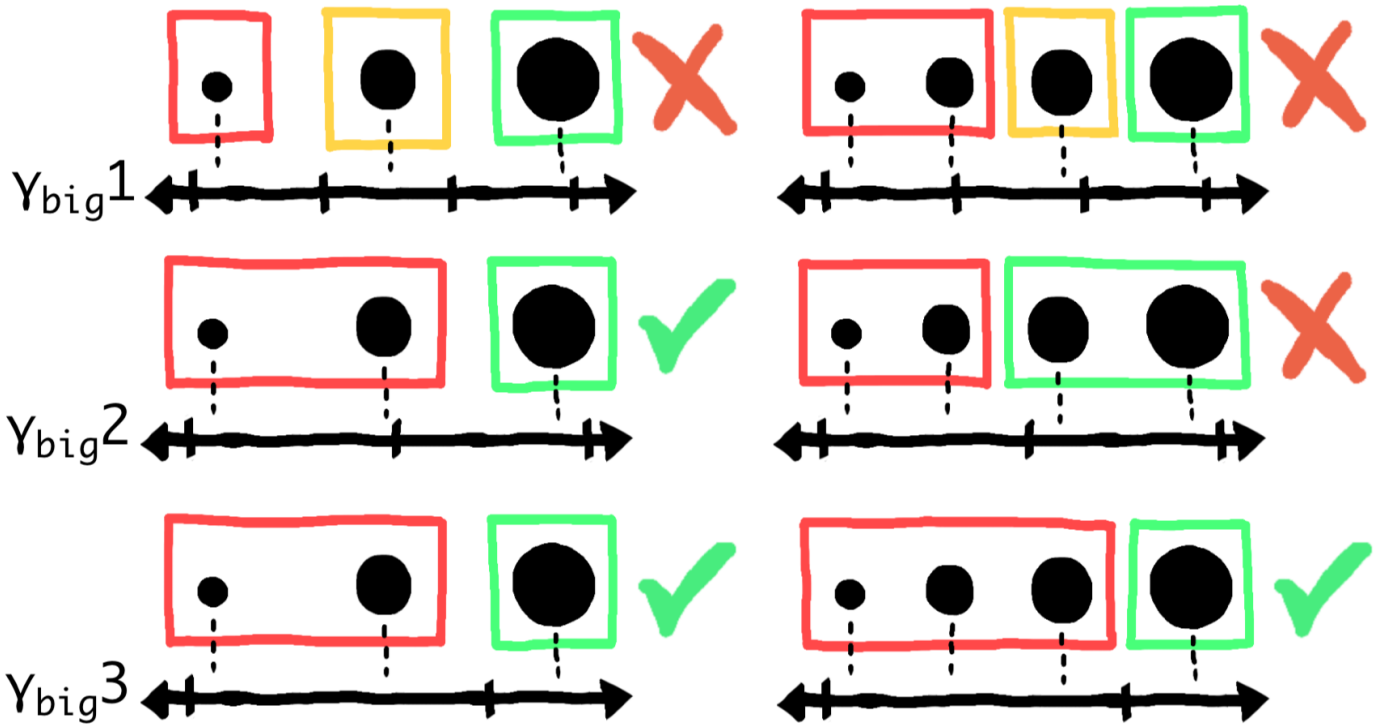
Linking Function

Acceptability of a description is modulated by its reference failure potential

Linking Function



Linking Function

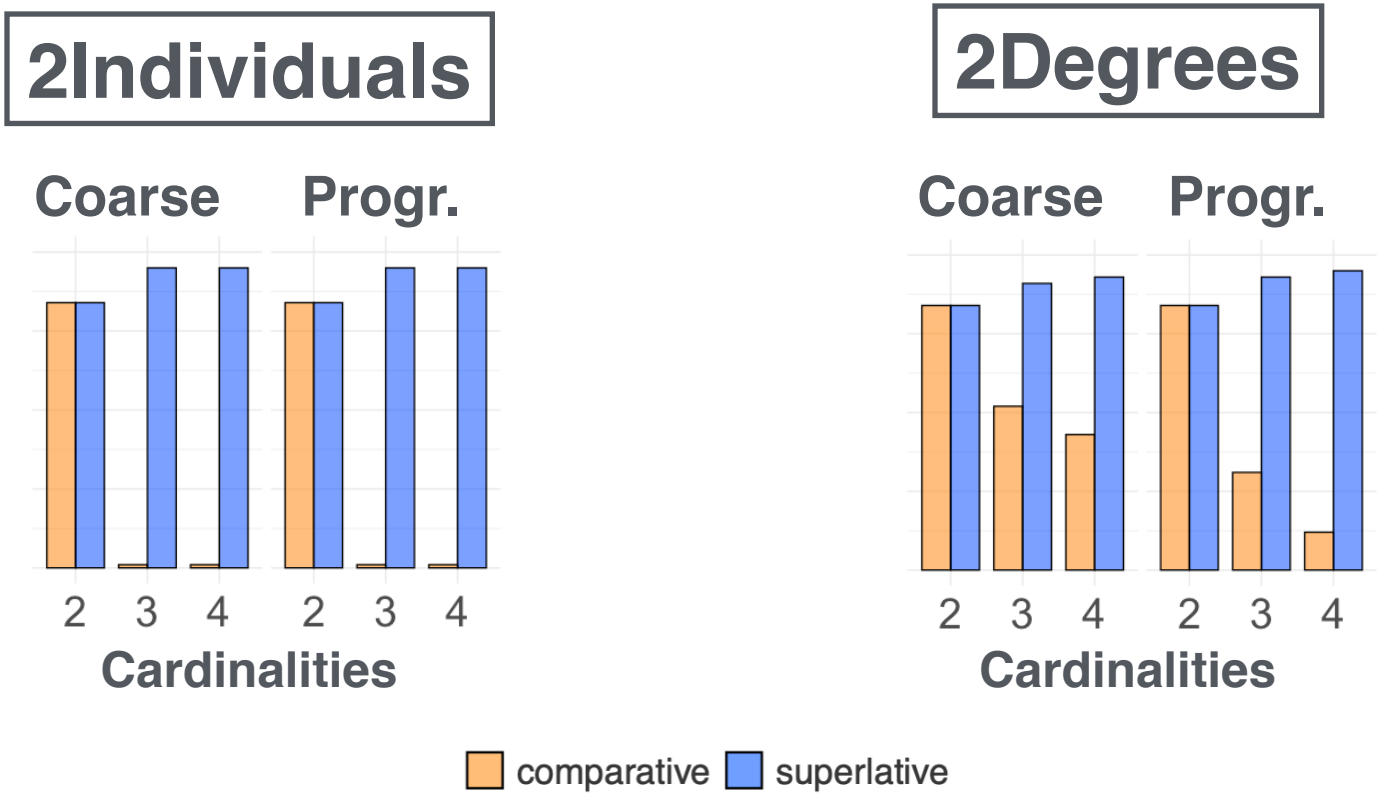


Progressive Card. 3

Progressive Card. 4

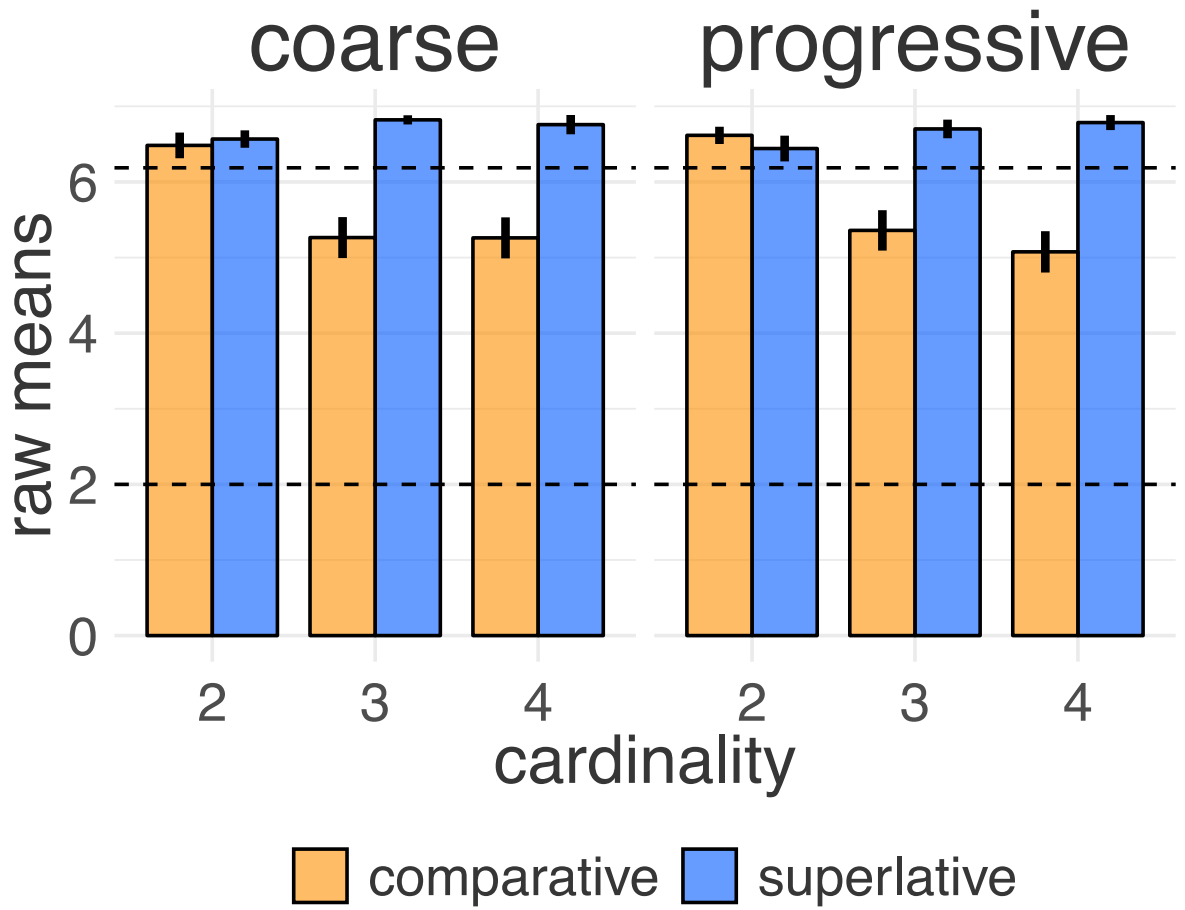
Qualitative Predictions

Predictions: Superlative



Results

Results

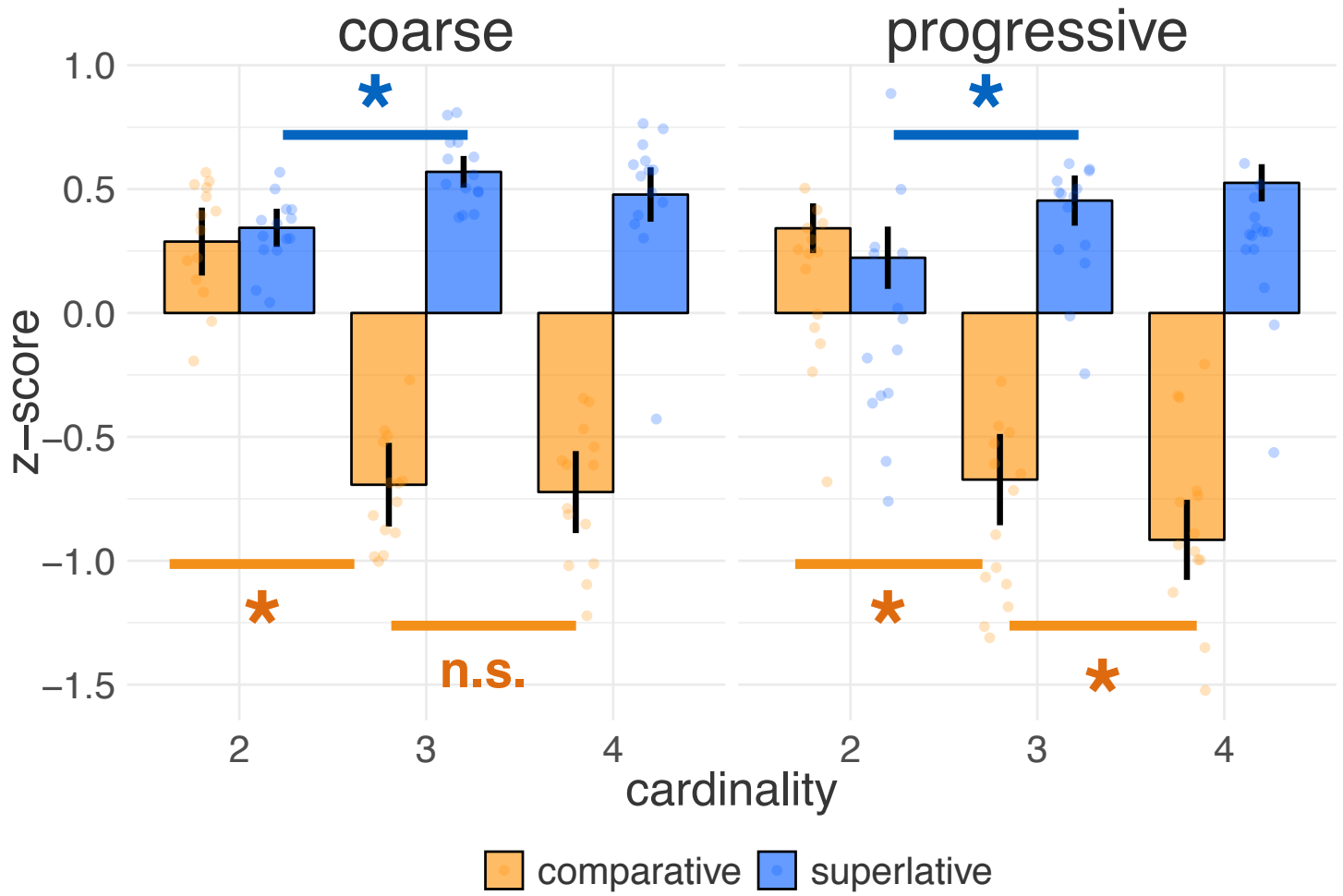


Grammatical Fillers
No Reference Failure

Global
Reference Failure
Fillers

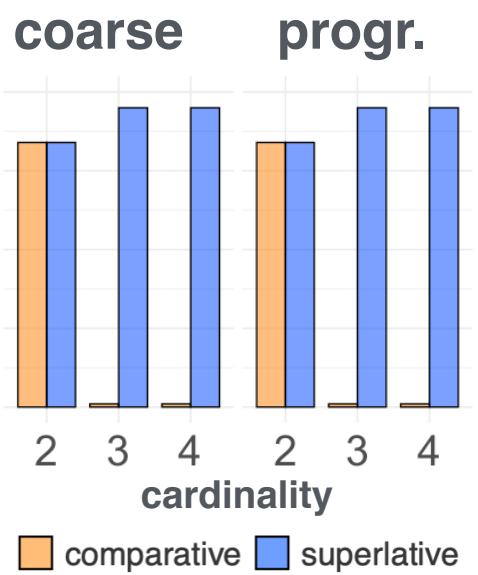
n=140

z-scored acceptability judgements

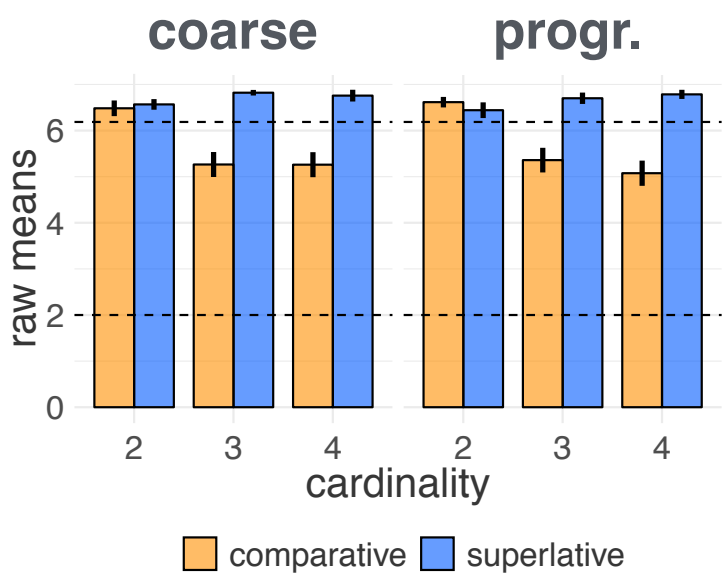


Discussion: 2Individuals

2Individuals Predictions



Experimental Results



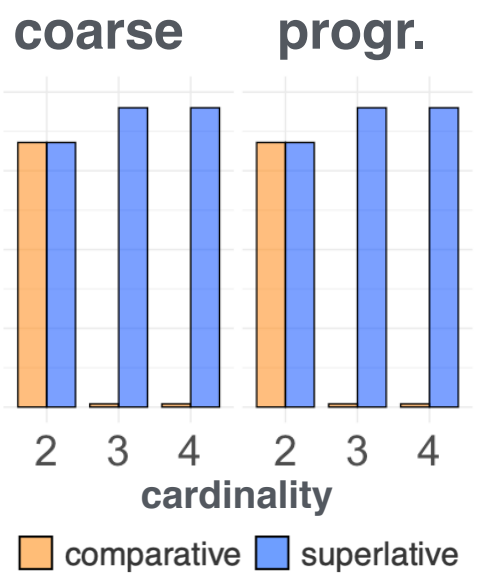
Grammatical Fillers
No Reference Failure

Global
Reference Failure
Fillers

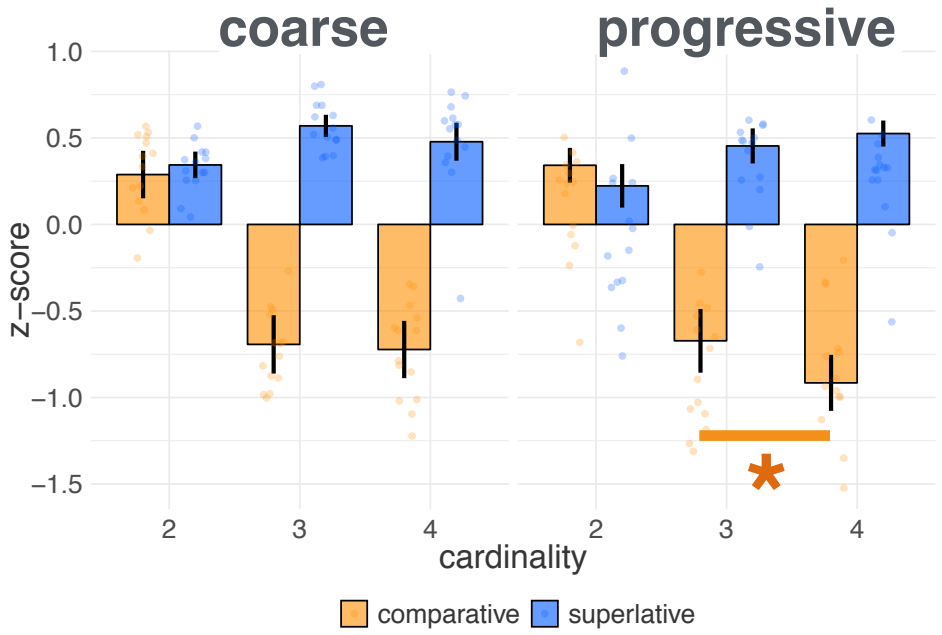
No global reference failure for the comparative at higher cardinalities

Discussion: 2Individuals

2Individuals Predictions



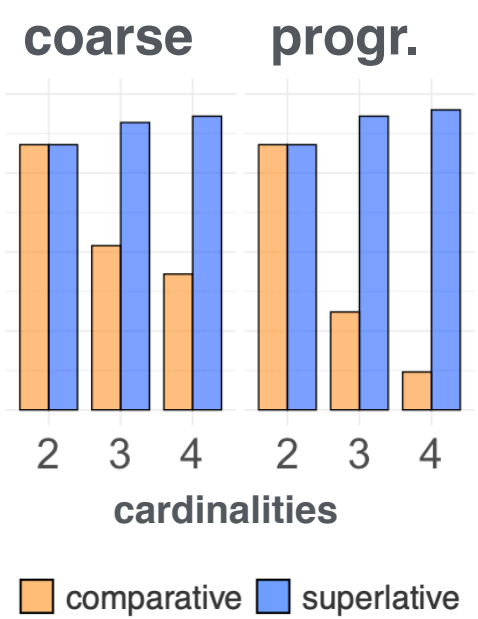
Experimental Results



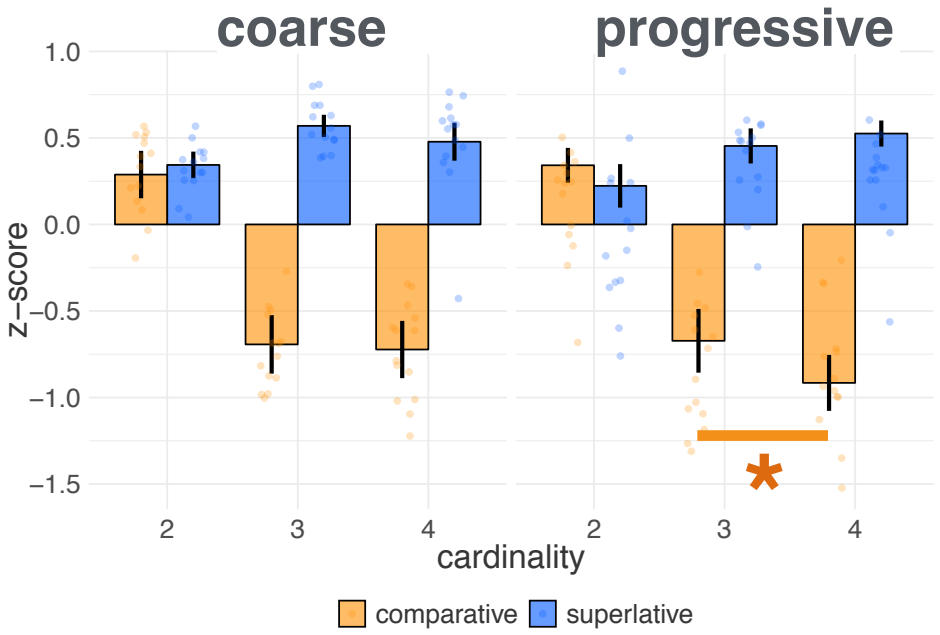
2I Theory does not predict gradient acceptability effects for comparative

Discussion: 2Degrees

2Degrees Predictions



Experimental Results

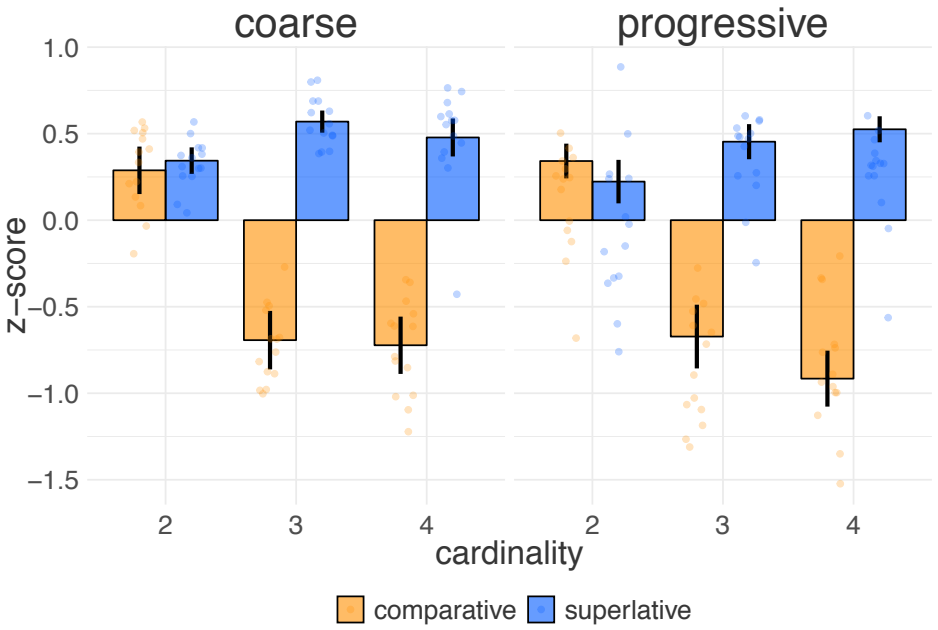


2D theory predicted gradient acceptability effects for comparative

Interim Conclusion

—No global reference failure for comparative at higher cardinalities

—Gradient effects of acceptability for the comparative



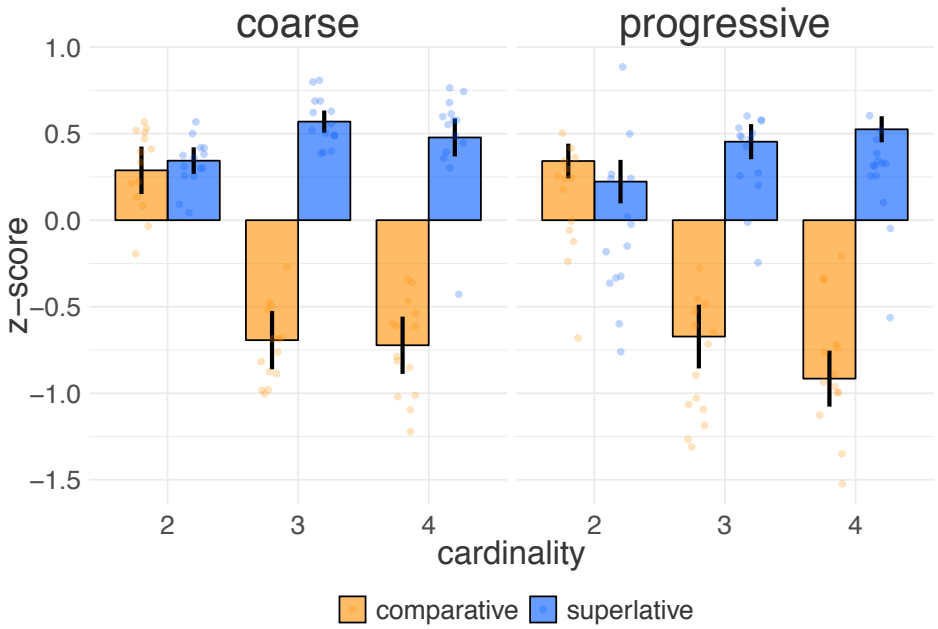
~~2Individuals Theory~~

2Degrees Theory

Interim Conclusion

—No global reference failure for comparative at higher cardinalities

—Gradient effects of acceptability for the comparative

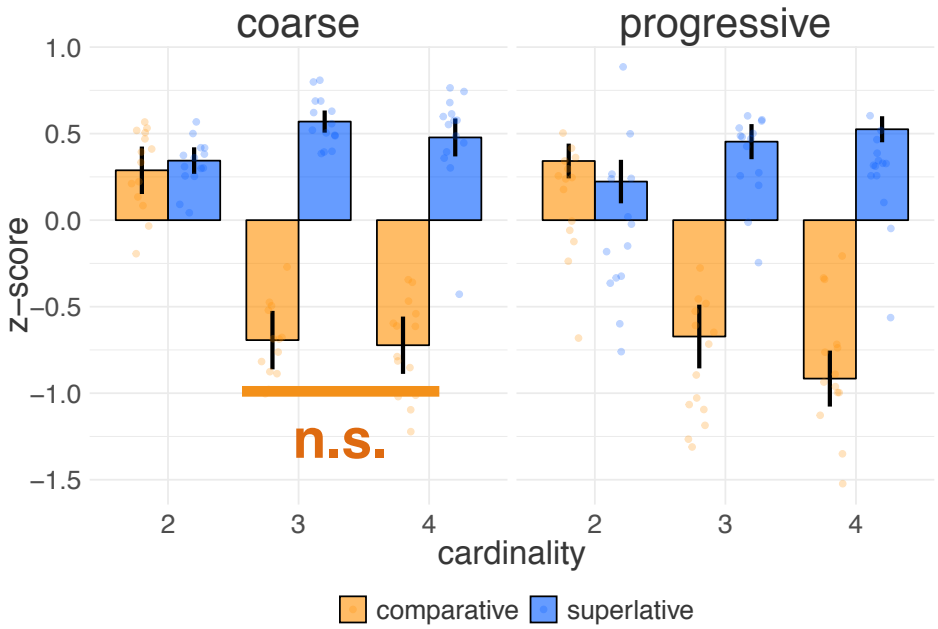


~~2Individuals Theory~~

2Degrees Theory

Open Questions

- Why no significant effect in coarse condition for comparative?



Proposal:

Penalty against pragmatic weakening

$|\gamma_{A,C}| = |C|$

$|\gamma_{A,C}| < |C|$

(Cf. Schwarzschild 1996; Champollion 2016)

Computational Modeling

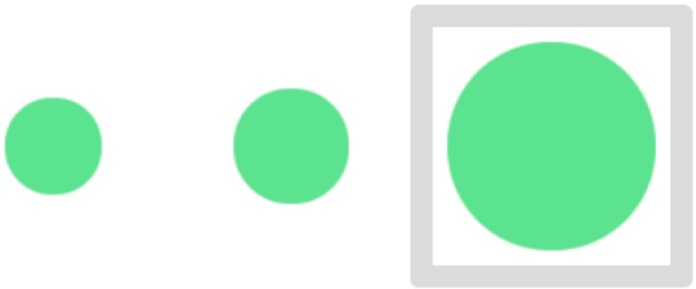
Rational Speech Act Models

- Probabilistic models of language interpretation
- Language understanding as social reasoning
- Speaker & Listener recursively reason about each other
- Formalization of Gricean pragmatics that uses Bayesian reasoning

Frank & Goodman (2012);
Goodman & Frank (2016)

Our Experiment in RSA

This is the biggest circle.

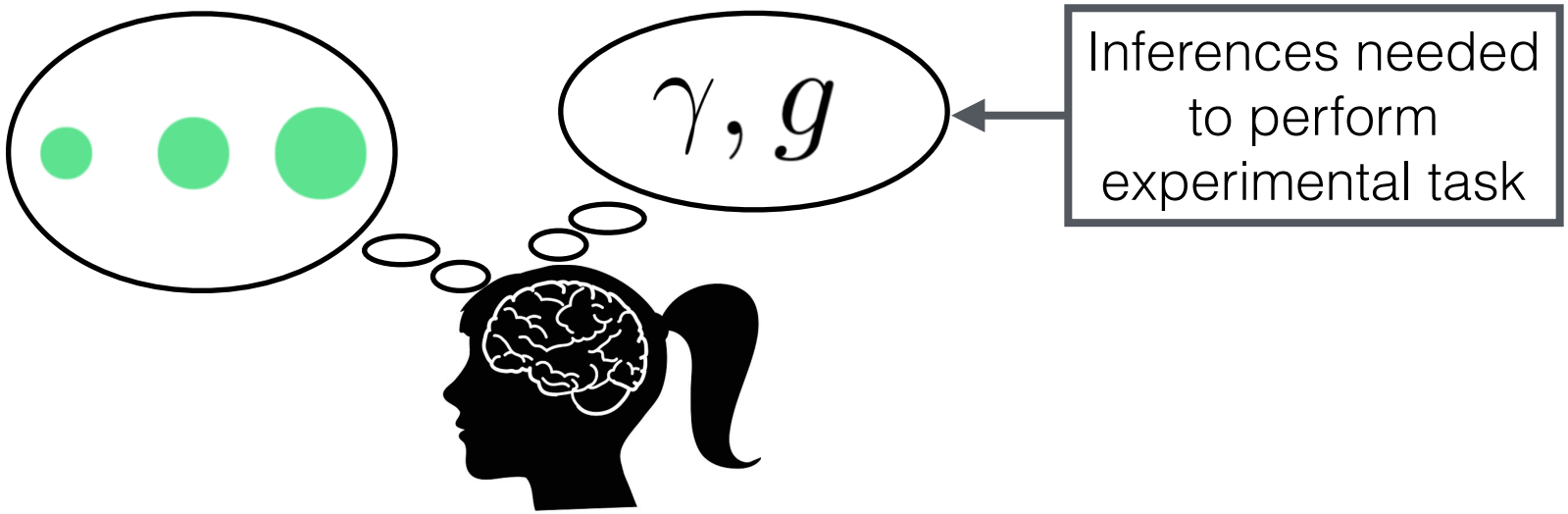


How acceptable is the above description of the scene?

	1	2	3	4	5	**6**	7	
(Bad)	○	○	○	○	○	●	○	(Good)

Back *Next*

Modeling our Experiment in RSA



Pragmatic Listener (L1)

The bigger circle.

$$L_1(r, \gamma \mid d = \text{The bigger circle})$$

Pragmatic Listener (L1)

$$L_1(r, \gamma \mid d = \text{The bigger circle})$$

Bayesian Update

Pragmatic Listener (L1)

$$L_1(r, \gamma \mid d = \text{The bigger circle}) \propto$$

$$S_1(d = \text{The bigger circle} \mid r, \gamma) \cdot P(r) \cdot P(\gamma \mid d)$$

Likelihood **Priors**

Speaker (S1)

- Utility defined as a trade-off between two communicative pressures:

$$S_1(d \mid r, \gamma) \propto \exp(\alpha \times \ln(L_0(r \mid d, \gamma)) - \text{cost}(d))$$

Maximize informativity of the utterance to the listener (efficiency)

Speaker (S1)

- Utility defined as a trade-off between two communicative pressures:

$$S_1(d \mid r, \gamma) \propto \exp(\alpha \times \ln(L_0(r \mid d, \gamma)) - \text{cost}(d))$$

Minimize production cost for the speaker (efficacy)

Speaker (S1)

- Utility defined as a trade-off between two communicative pressures:

$$S_1(d \mid r, \gamma) \propto \exp(\alpha \times \ln(L \ (r \mid d, \gamma)) - \mathsf{cost}(d))$$

Speaker and Listener models are mutually recursive

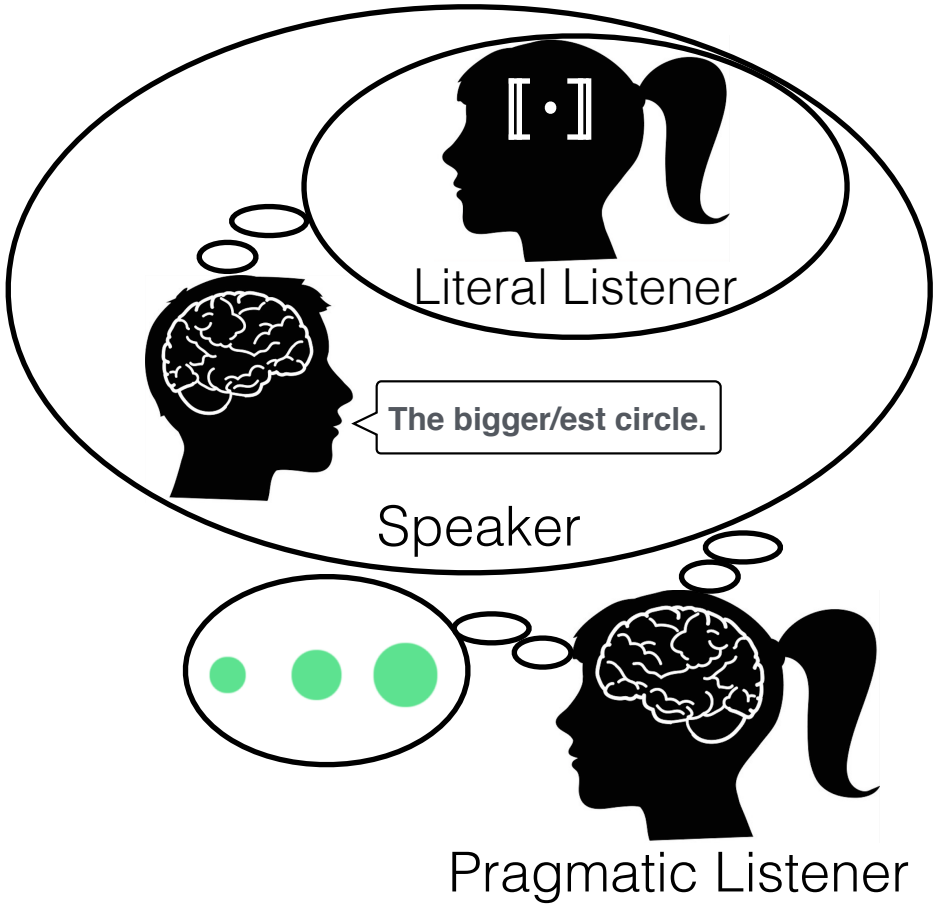
Literal Listener (L0)

- Grounds model in truth-conditional semantics

$$L_0(r \mid d, \gamma) \propto [\![d]\!]^{\gamma}(r) \cdot P(r)$$

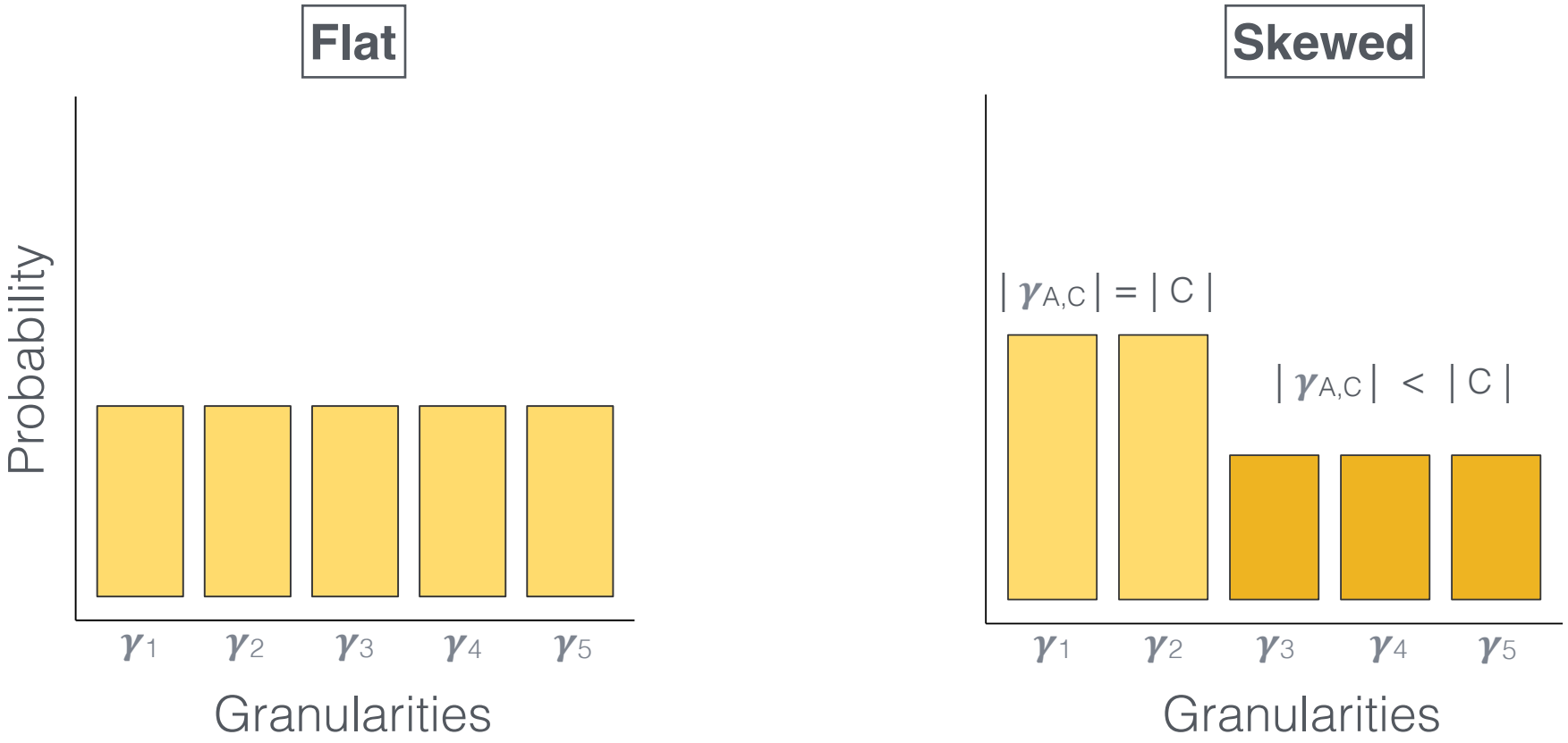
$$P(r) = \begin{cases} \epsilon & \text{if } r = \text{fail} \\ \text{uniform} & \text{otherwise} \end{cases}$$

RSA Model


⟦·⟧
Literal Listener
The bigger/est circle.
Speaker
Pragmatic Listener

Pragmatic Weakening

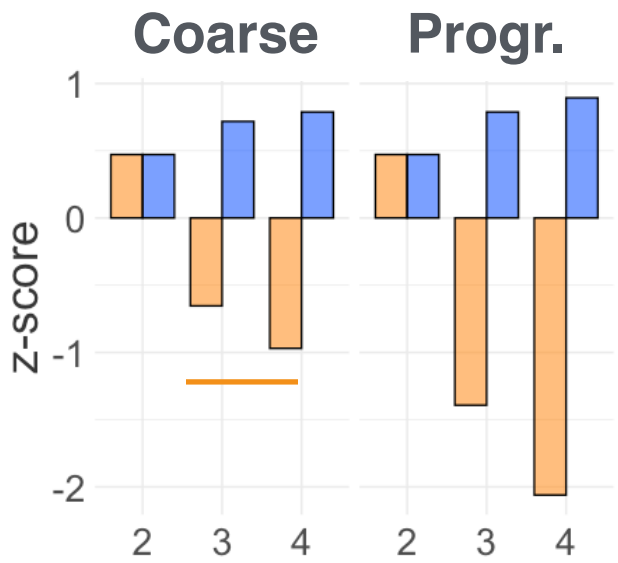
Granularity Priors



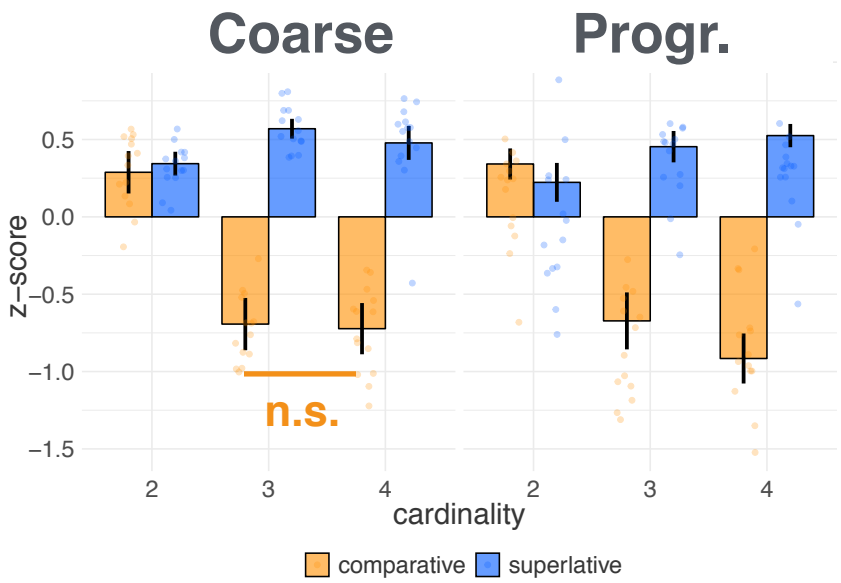
Simulations

Flat Granularity Prior

Model



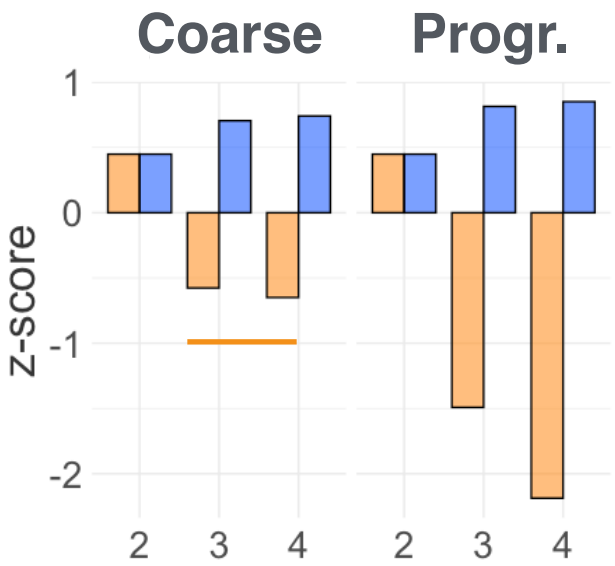
Experiment



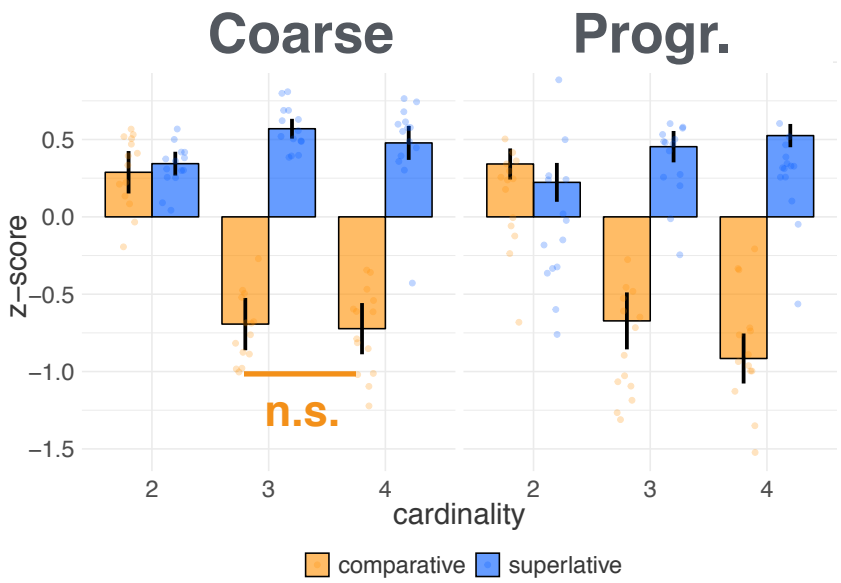
✚ No significant effect in coarse condition comparative ✗

Skewed Granularity Prior

Model



Experiment



✚ No significant effect in coarse condition comparative ✓

Conclusion

1. Our experimental results are incompatible with 2Individuals Theory

2. Our results can be explained by the **2Degree Theory**

→ Listeners flexibly infer a granularity that satisfies the semantic requirements of the definite comparative

3. 2Degrees Theory alone fails to account for lack of effect for the comparative in Coarse condition

→ Post hoc computational modeling: dispreference for pragmatic weakening

Thank you