

Shifts of Attention During Spatial Language Comprehension

A Computational Investigation

Thomas Kluth¹, Michele Burigo¹, and Pia Knoeferle²

- 1: Language & Cognition Group, CITEC (Cognitive Interaction Technology Excellence Cluster), Bielefeld University, Bielefeld, Germany
- 2: Department of German Language and Linguistics, Humboldt University, Berlin, Germany

February 24, 2016

Motivation // Remove the spider!



Motivation // Remove the spider!



image sources:

robot: by Mamirobothk (CC BY-SA 2.5, <https://commons.wikimedia.org/w/index.php?curid=25084931>)

spider: by L. Shyamal (CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=1309920>)

Robots comprehending human (spatial) language

- robot needs to know what you mean by “left”

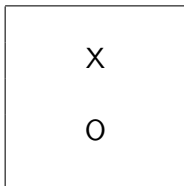
Robots comprehending human (spatial) language

- robot needs to know what you mean by “left”
- implement human-like processes

Robots comprehending human (spatial) language

- robot needs to know what you mean by “left”
 - implement human-like processes
- But:** How do humans comprehend spatial prepositions?

Previous Research // Logan and Sadler (1996, experiment 2)



The X is above the O.

Previous Research // Logan and Sadler (1996, experiment 2)

X	X	X	X	X	X	X
X	X	X	X	X	X	X
X	X	X	X	X	X	X
X	X	X	O	X	X	X
X	X	X	X	X	X	X
X	X	X	X	X	X	X
X	X	X	X	X	X	X



(image source: Logan & Sadler, 1996, p. 510)

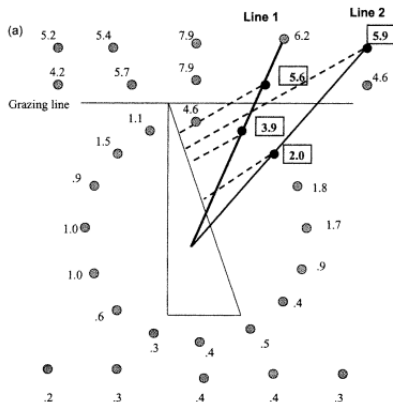
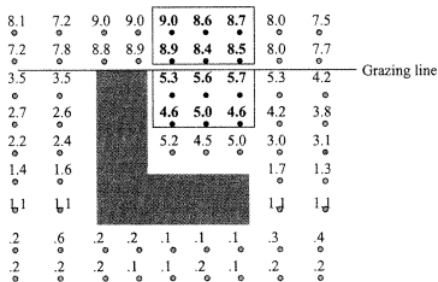
Previous Research // Logan and Sadler (1996, experiment 2)

X	X	X	X	X	X	X
X	X	X	X	X	X	X
X	X	X	X	X	X	X
X	X	X	O	X	X	X
X	X	X	X	X	X	X
X	X	X	X	X	X	X
X	X	X	X	X	X	X



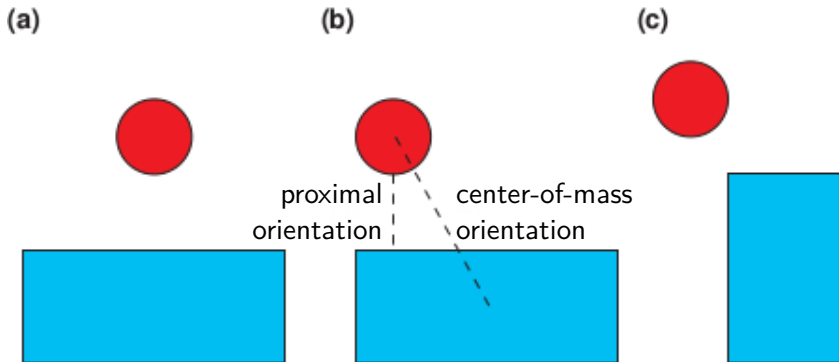
(image source: Logan & Sadler, 1996, p. 510)

Previous Research // Regier and Carlson (2001, exp. 5 & 6)



(image sources: Regier & Carlson, 2001, p. 287-288)

Proximal and center-of-mass orientation



(image adapted from Roy, 2005, p. 390)

AVS Model // Regier and Carlson (2001)

cognitive model:

Attentional **V**ector **S**um (AVS) model
(Regier & Carlson, 2001)

AVS Model // Regier and Carlson (2001)

spatial preposition: *above*

located object: LO



reference object: RO

→ AVS model

AVS Model // Regier and Carlson (2001)

spatial preposition: *above*

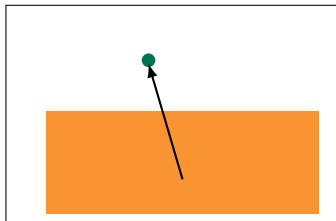
located object: LO



reference object: RO

→ AVS model → acceptability rating

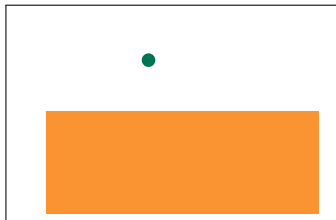
AVS Model // Regier and Carlson (2001)



→ AVS model → acceptability rating

AVS model assumes shift of attention from RO to LO

AVS Model // Regier and Carlson (2001)

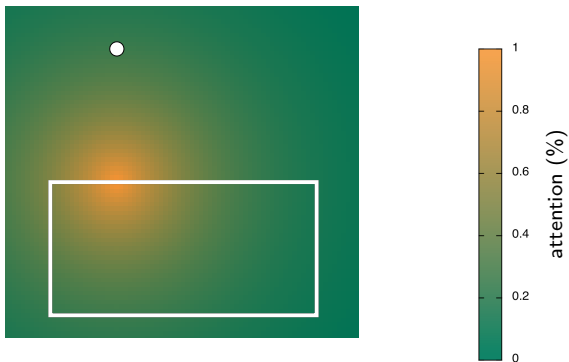


→ AVS model → acceptability rating

AVS model consists of

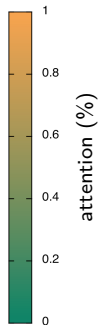
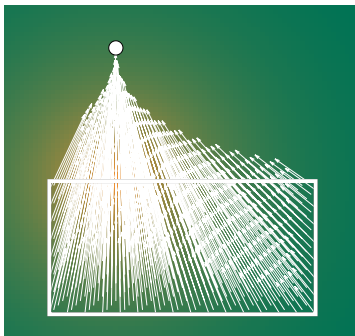
1. angular component
2. height component

AVS Model // Angular Component



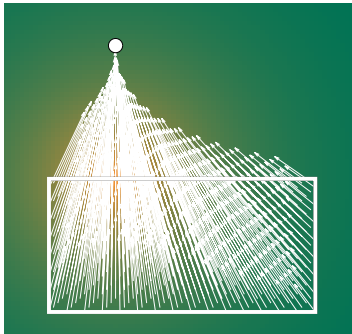
$$a_i = \exp\left(\frac{-d_i}{\lambda \cdot \sigma}\right)$$

AVS Model // Angular Component



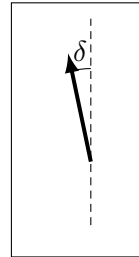
$$\overrightarrow{direction} = \sum_{i \in RO} a_i \cdot \vec{v}_i$$

AVS Model // Angular Component



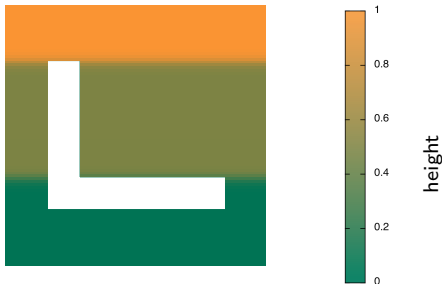
$$\sum a_i \cdot \vec{v}_i$$

A wavy arrow points from the equation to the right.



$$g(\delta) = \text{slope} \cdot \delta + \text{y-intercept}$$

AVS Model // Height Component



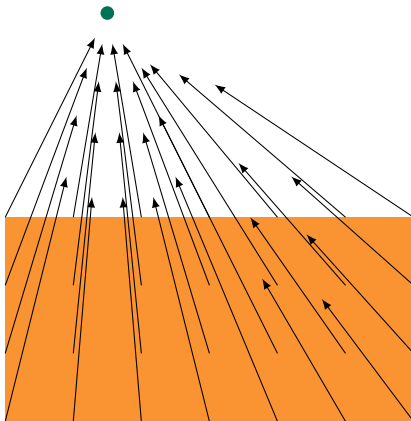
$$\text{height}(y_{LO}) =$$

$$\frac{\text{sig}(y_{LO} - \text{hightop}, \text{highgain}) + \text{sig}(y_{LO} - \text{lowtop}, 1)}{2}$$

$$\text{above}(LO, RO) = g(\delta) \cdot \text{height}(y_{LO})$$

rAVS Model // Motivation

- AVS assumes shift of attention from RO to LO

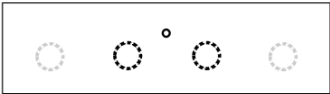


A

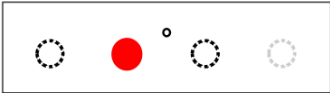
g L r ?	r L g ?	g R r ?	r R g ?	w L ?	w R ?
Y or N	Y or N	Y or N	Y or N	R or G	R or G

B

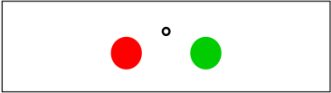
1200-1600ms
(Random)



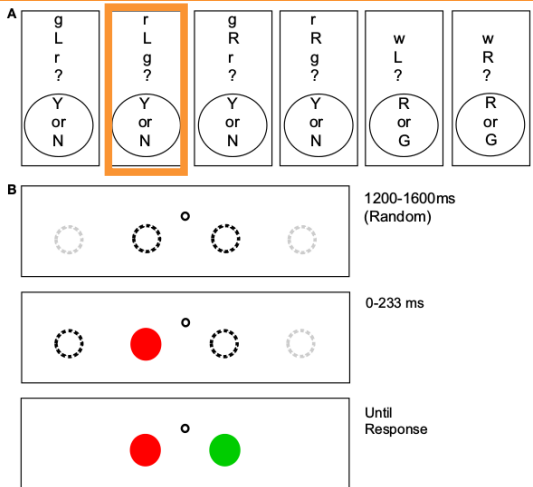
0-233 ms



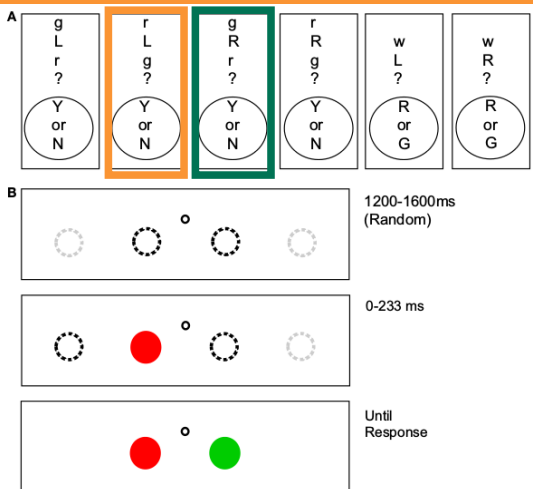
Until
Response



(image source: Roth & Franconeri, 2012, p. 5)



(image source: Roth & Franconeri, 2012, p. 5)



(image source: Roth & Franconeri, 2012, p. 5)

Visual World Paradigm // Burigo and Knoeferle (2015)



“Die Box ist *über* der Wurst”
‘The box is above the sausage’

(image source: Burigo & Knoeferle, 2015, p. 6)

Visual World Paradigm // Burigo and Knoeferle (2015)



“Die Box ist *über* der Wurst”
‘The box is *above* the sausage’

(image source: Burigo & Knoeferle, 2015, p. 6)

Visual World Paradigm // Burigo and Knoeferle (2015)



“Die Box ist *über* der Wurst”
‘The box is above the sausage’

(image source: Burigo & Knoeferle, 2015, p. 6)

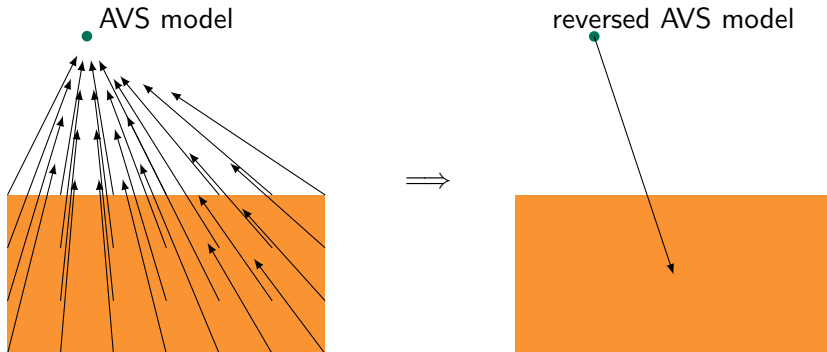
Visual World Paradigm // Burigo and Knoeferle (2015)



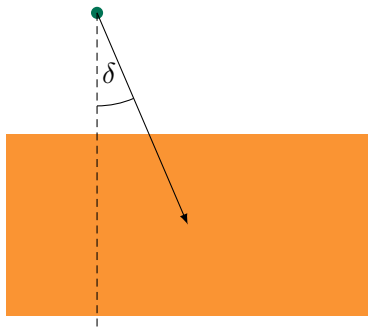
“Die Box ist über der Wurst”
‘The box is above the sausage’

(image source: Burigo & Knoeferle, 2015, p. 6)

rAVS Model // Main Idea

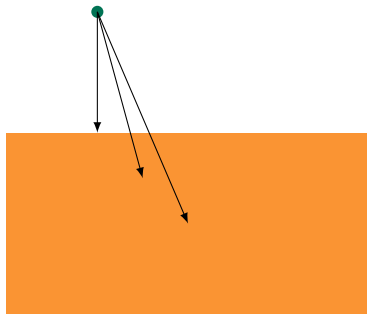


rAVS Model // Details

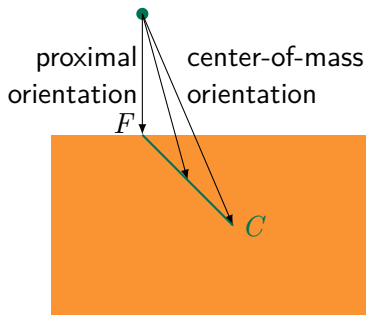


$$\text{above}(LO, RO) = g(\delta) \cdot \text{height}(y_{LO})$$

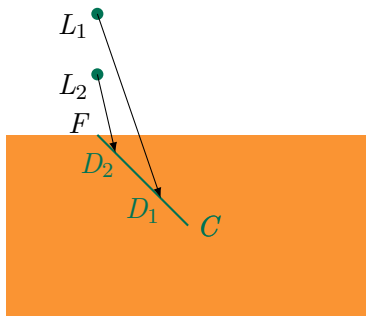
rAVS Model // Details



rAVS Model // Details



rAVS Model // Details



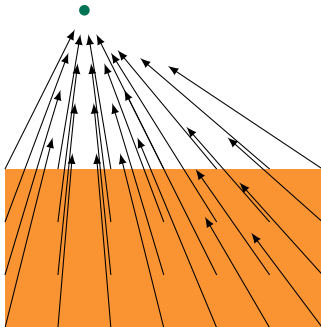
$$D = \begin{cases} \overrightarrow{LC} + (-\alpha \cdot dist_{rel.} + 1) \cdot \overrightarrow{CF} & \text{if } (-\alpha \cdot dist_{rel.} + 1) > 0 \\ C & \text{else} \end{cases}$$

rAVS Model // Relative Distance

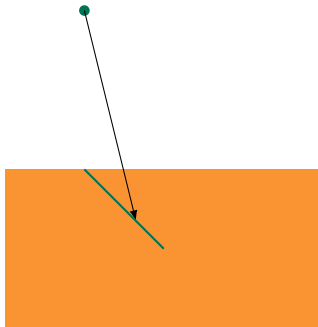


$$\text{relative distance} = \frac{|LO, P|_x}{RO_{width}} + \frac{|LO, P|_y}{RO_{height}}$$

Method // Model Comparison



free parameters:
slope, intercept, highgain, λ



free parameters:
slope, intercept, highgain, α

Method // Model Comparison

AVS & rAVS
4 free parameters

Method // Model Comparison

Regier and Carlson (2001):
7 experiments
→ 10 ROs, 337 LOs

input (ROs, LOs)

AVS & rAVS
4 free parameters

Method // Model Comparison

Regier and Carlson (2001):
7 experiments
→ 10 ROs, 337 LOs

input (ROs, LOs)

AVS & rAVS
4 free parameters

empirical ratings

model ratings

Method // Model Comparison

Regier and Carlson (2001):
7 experiments
→ 10 ROs, 337 LOs

input (ROs, LOs)

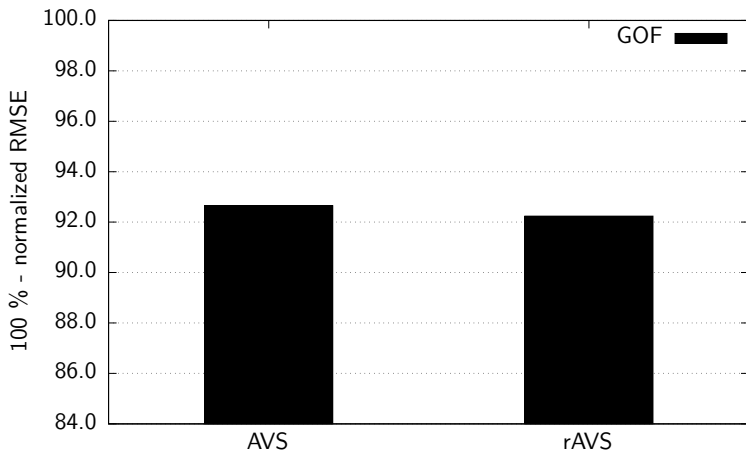
AVS & rAVS
4 free parameters

empirical ratings

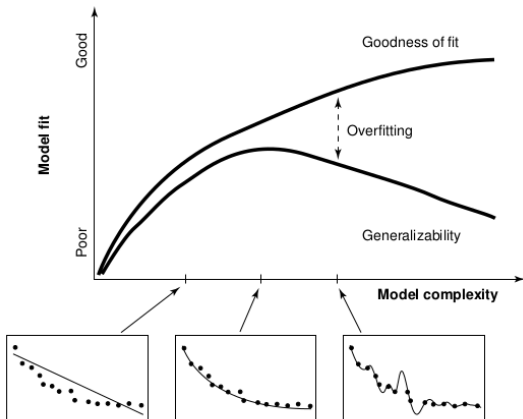
model ratings

$$RMSE = \sqrt{\frac{1}{n} \sum_i^n (data_i - modelOut_i)^2}$$

Results // Goodness of Fit, Regier and Carlson (2001, all experiments)

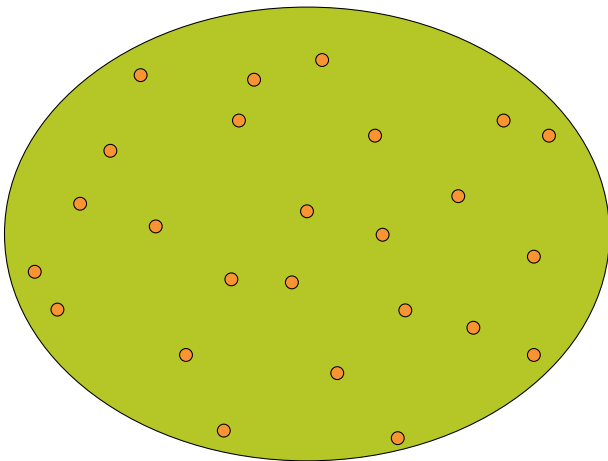


Method // Problems of GOF

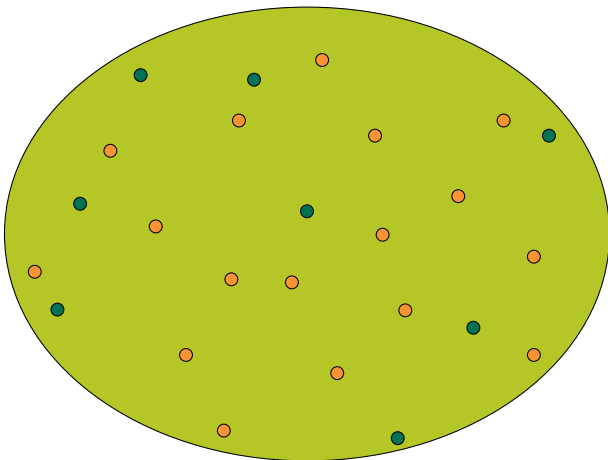
*TRENDS in Cognitive Sciences*

(image source: Pitt & Myung, 2002, p. 424)

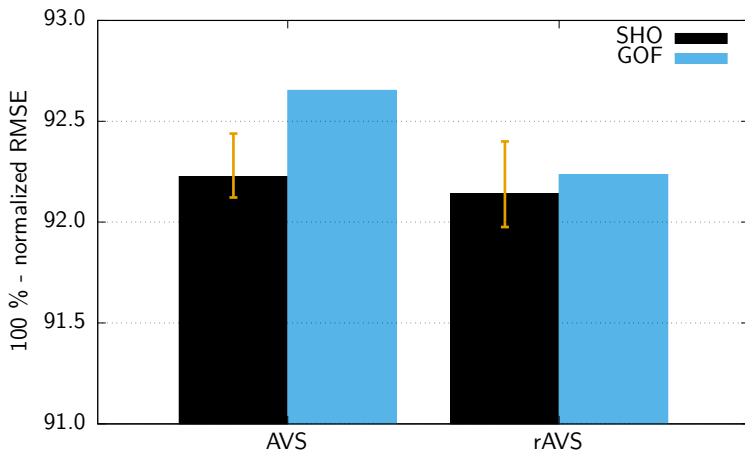
Method // Simple Hold-Out (Schultheis, Singhaniya, & Chaplot, 2013)



Method // Simple Hold-Out (Schultheis, Singhaniya, & Chaplot, 2013)



Results // GOF and SHO, Regier and Carlson (2001, all experiments)



Conclusion

- rAVS model: a modification of the AVS model that integrates recent findings (Burigo & Knoeferle, 2015; Roth & Franconeri, 2012)

Conclusion

- rAVS model: a modification of the AVS model that integrates recent findings (Burigo & Knoeferle, 2015; Roth & Franconeri, 2012)
- rAVS is less complex than AVS

Conclusion

- rAVS model: a modification of the AVS model that integrates recent findings (Burigo & Knoeferle, 2015; Roth & Franconeri, 2012)
- rAVS is less complex than AVS
- both models perform equally well on the data from Regier and Carlson (2001)

Conclusion

- rAVS model: a modification of the AVS model that integrates recent findings (Burigo & Knoeferle, 2015; Roth & Franconeri, 2012)
- rAVS is less complex than AVS
- both models perform equally well on the data from Regier and Carlson (2001)
- simulations do not favor any of the two models

Conclusion

- rAVS model: a modification of the AVS model that integrates recent findings (Burigo & Knoeferle, 2015; Roth & Franconeri, 2012)
- rAVS is less complex than AVS
- both models perform equally well on the data from Regier and Carlson (2001)
- simulations do not favor any of the two models

both directionalities of the
attentional shift are equally well supported

Future Work

- experiment to distinguish the models

Future Work

- experiment to distinguish the models
- extend model with
 - the LO

Future Work

- experiment to distinguish the models
- extend model with
 - the LO
 - timing



(image source: Hustvedt, CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=5743799>)

Future Work

- experiment to distinguish the models
- extend model with
 - the LO
 - timing
 - functionality of objects



(images adapted from: Hörberg, 2008, p. 200)

Future Work

- experiment to distinguish the models
- extend model with
 - the LO
 - timing
 - functionality of objects
- implement into technical systems
 - C++ source code available under an open source license at Kluth (2016)



(image source: Mamirobothk, CC BY-SA 2.5, <https://commons.wikimedia.org/w/index.php?curid=25084931>)

Thank you for your attention!

References

- Burigo, M., & Knoeferle, P. (2015). Visual attention during spatial language comprehension. *PLoS ONE*, *10*(1), e0115758. doi: 10.1371/journal.pone.0115758
- Hörberg, T. (2008). Influences of form and function on the acceptability of projective prepositions in Swedish. *Spatial Cognition & Computation*, *8*(3), 193–218. doi: 10.1080/13875860801993652
- Kluth, T. (2016). *A C++ Implementation of the reversed Attentional Vector Sum (rAVS) model*. Bielefeld University. doi: 10.4119/unibi/2900103
- Logan, G. D., & Sadler, D. D. (1996). A computational analysis of the apprehension of spatial relations. In P. Bloom, M. A. Peterson, L. Nadel, & M. F. Garrett (Eds.), *Language and Space* (pp. 493–530). The MIT Press.
- Pitt, M. A., & Myung, I. J. (2002). When a good fit can be bad. *Trends in Cognitive Sciences*, *6*(10), 421–425.
- Regier, T., & Carlson, L. A. (2001). Grounding spatial language in perception: An empirical and computational investigation. *Journal of Experimental Psychology: General*, *130*(2), 273–298. doi: 10.1037//0096-3445.130.2.273
- Roth, J. C., & Franconeri, S. L. (2012). Asymmetric coding of categorical spatial relations in both language and vision. *Frontiers in Psychology*, *3*(464). doi: 10.3389/fpsyg.2012.00464
- Roy, D. (2005). Grounding words in perception and action: computational insights. *Trends in Cognitive Sciences*, *9*(8), 389–396.
- Schultheis, H., Singhaniya, A., & Chaplot, D. S. (2013). Comparing model comparison methods. In *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 1294 – 1299). Austin, TX: Cognitive Science Society.

RMSE // GOF and SHO, Regier and Carlson (2001, all experiments)

