In D. Reitter & F. E. Ritter (Eds.), Proceedings of the 14th International Conference on Cognitive Modeling (ICCM 2016). University Park, PA: Penn State.

Distinguishing Cognitive Models of Spatial Language Understanding

Thomas Kluth (tkluth@cit-ec.uni-bielefeld.de)
Michele Burigo (mburigo@cit-ec.uni-bielefeld.de)
CITEC (Cognitive Interaction Technology Excellence Cluster), Bielefeld University, Inspiration 1, 33619 Bielefeld, Germany

Holger Schultheis (schulth@informatik.uni-bremen.de)
Cognitive Systems Group, Department of Computer Science, University of Bremen, Enrique-Schmidt-Str. 5, 28359 Bremen, Germany

Pia Knoeferle (pia.knoeferle@hu-berlin.de)
Department of German Language and Linguistics, Humboldt-University Berlin, Unter den Linden 6, 10099 Berlin

Keywords: spatial language; visual attention; model predictions; model flexibility.

Introduction

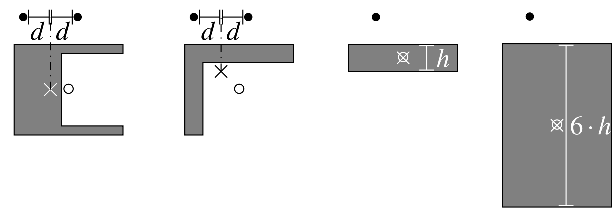
Consider the following sentence: "The picture (located object, LO) is above the desk (reference object, RO)." Given the locations of the picture and the desk – how acceptable is the use of "above"? Regier and Carlson (2001) proposed a cognitive model that computes an acceptability rating for the spatial preposition "above" in describing the spatial relation between a RO and a LO: the Attentional Vector Sum (AVS) model. In line with Logan and Sadler (1996), the AVS model assumes a shift of attention from the RO to the LO. However, in a study by Burigo and Knoeferle (2015) overt gaze shifts from the RO to the LO were infrequent during the comprehension of spatial relation utterances. By contrast, shifts in line with the mention of objects (from the LO to the RO) were highly frequent, suggesting they may be sufficient for understanding a spatial description (see also Roth & Franconeri, 2012).

Accordingly, Kluth, Burigo, and Knoeferle (2016) propose the reversed AVS (rAVS) model in which attention shifts from the LO to the RO (instead of shifting from the RO to the LO). The rAVS model accounts as well as the AVS model for the empirical data from Regier and Carlson (2001; see Kluth et al., 2016, for details). Thus, using these already existing data the two models cannot be distinguished, despite their different implementation of the attentional shift.

In order to assess whether one of these two models reflects human ratings of spatial language better than the other, we designed stimuli for which we hypothesized the models predict different acceptability ratings (see Fig. 1). With these stimuli, we conducted an empirical study to test the predictions.

Model Predictions

Based on the mechanisms of the models, we hypothesized two types of predictions for the stimuli in Fig. 1. The first type concerns the influence of asymmetrical ROs on the acceptability of spatial prepositions. Consider two LOs with equal horizontal distance d from the center-of-mass of an asymmetrical RO, as shown in Fig. 1a. The rAVS model predicts no difference in ratings for these LOs, because its computation is based on the center-of-mass of the RO. The AVS model, however, seems to predict higher ratings for the LO above the mass of the RO compared to the LO above the cavity of the RO. This is because the AVS model defines its population of vectors based on all points of the RO and thus gives more importance to the mass of the RO. The same reasoning applies for the LOs above the asymmetrical RO in Fig. 1b.



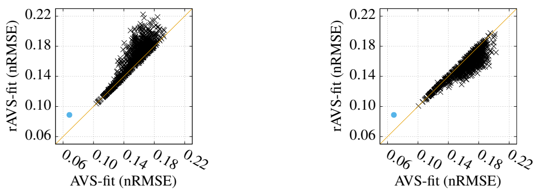
(a) Asym. "C" (b) Asym. "L" (c) "Thin" rect. (d) "Tall" rect.

Figure 1: Stimuli used for the computational and empirical studies. (• = LO; × = center-of-mass, ○ = center-of-object).

For the second type of predictions consider the two rectangular ROs in Figs. 1c and 1d. Here, the rAVS model predicts a lower rating for the LO above the "thin" rectangle compared to the LO above the "tall" rectangle. This is because the rAVS model explicity uses the *relative* distance of an LO from an RO. Here, relative distance is defined as absolute distance divided by the dimensions of the RO. Due to the free parameters of the AVS model, the prediction of the AVS model for this condition is unclear.

Empirical Study We were interested to see whether humans follow these hypothesized predictions. Thus, we conducted an empirical rating study with 28 LOs above each of the ROs in Fig. 1. Participants had to rate how well the German sentence "Der Punkt ist über dem Objekt." ("The dot is above the object.") describes a depicted RO-LO configuration. We also tested "unter" ("below") but do not report the results here. For the relative distance condition, we found that LOs above the "tall" rectangle were rated higher than LOs above the "thin" rectangle (mean difference: 0.078; 95% confidence intervals: 0.151, 0.007). This is in line with the prediction of the rAVS model.

For the asymmetrical ROs, however, we found an effect that falsifies both models: LOs above the mass of an RO were rated lower than LOs above the cavity of an RO (mean difference: 0.518; 95% confidence intervals: 0.619, 0.428). This effect contradicts the influence of the center-of-mass orientation as suggested by Regier and Carlson (2001). Neither the AVS model nor the rAVS model can account for this empiri-



(a) AVS generated data (b) rAVS generated data

Figure 2: Results of the landscaping method (×) and fits of the models to empirical data (●).

cal finding, although both models account successfully for the data from Regier and Carlson (2001; see Kluth et al., 2016). In order to better understand the performance of both models on our stimuli, we analyzed the outcomes of the models for these stimuli using three different methods.

Parameter Space Partitioning Analysis To verify whether the two models actually generate our hypothesized predictions, we applied the PSP algorithm proposed by Pitt, Kim, Navarro, and Myung (2006; using their MATLAB implementation) with the ROs shown in Fig. 1 and up to 28 LOs above each RO. This analysis confirmed the hypothesized predictions for the rAVS model but disconfirmed the hypothesized predictions for the AVS model: The AVS model is able to generate the same patterns as the rAVS model but interestingly none of the patterns hypothesized above. Arguably then, the mechanisms of the AVS model are harder to translate into testable predictions. Moreover, the AVS model generates a greater range of possible outcomes, i.e., it is more flexible than the rAVS model. However, neither model generates the empirical pattern for the asymmetrical ROs.

Model Mimicry The PSP analysis revealed that both models are able to generate the same qualitative patterns for the stimuli in Fig. 1. To investigate the relative performance of the models on these stimuli, we used the landscaping analysis as proposed by Navarro, Pitt, and Myung (2004). Apart from assessing the ability of each model to mimic the other, this method also gives us another measure of model flexibility. We generated 1000 data sets from each model and fitted both models on these artificial data by minimizing the normalized Root Mean Square Error (nRMSE). The results are shown in Fig. 2. Model fits to the empirical data (nRMSE) are plotted as filled circles. There is a slight trend that the rAVS model fits the data generated by the AVS model worse than the AVS model fits the data generated by the rAVS model (maximal rAVS fit in Fig. 2a is greater than maximal AVS fit in Fig. 2b). However, overall, both models fit their own data better than the other data and thus, no model mimics the other model.

Model Flexibility Analysis The AVS model showed greater flexibility in the PSP analysis but not in the landscaping method. The MFA proposed by Veksler, Myers, and Gluck (2015) provides a quantitative measure of model flexibility (see Veksler et al., 2015, for the relation of the MFA to PSP and landscaping). We computed the MFA for the stimuli

Table 1: ϕ values of the Model Flexibility Analysis (MFA). The lower the ϕ value, the less flexible the model.

	stimuli from Fig. 1	stimuli from Regier and Carlson (2001)
AVS	$\phi = 0.000899$	$\phi = 0.000420$
rAVS	$\phi = 0.000544$	$\phi = 0.000292$

in Fig. 1 as well as for the stimuli used by Regier and Carlson (2001). We split the range of each of the four free parameters in 50 intervals and followed the procedure outlined by Veksler et al. (2015). As indicated by the lower ϕ values in Table 1, the rAVS model is less flexible than the AVS model.

Future Work In contrast to the landscaping results, the PSP and the MFA suggest that the AVS model is more flexible than the rAVS model. At the moment, we are investigating the cause of these differences in the model results.

More importantly, some of our empirical findings corroborate the rAVS model (effect of relative distance) while other findings falsify both models (effect of asymmetrical ROs). Currently, we are developing slightly modified models incorporating our suggestion that people base their acceptability ratings on the center-of-object (see ○ in Fig. 1) instead of on the center-of-mass (as suggested by Regier & Carlson, 2001). Preliminary simulation results support the use of the center-of-object over the center-of-mass.

References

Burigo, M., & Knoeferle, P. (2015). Visual attention during spatial language comprehension. *PLoS ONE*, *10*(1), e0115758.

Kluth, T., Burigo, M., & Knoeferle, P. (2016). Shifts of attention during spatial language comprehension: A computational investigation. In *Proceedings of the 8th International Conference on Agents and Artificial Intelligence – Volume 2* (pp. 213–222). SCITEPRESS.

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.

Navarro, D. J., Pitt, M. A., & Myung, I. J. (2004). Assessing the distinguishability of models and the informativeness of data. *Cognitive Psychology*, *49*(1), 47–84.

Pitt, M. A., Kim, W., Navarro, D. J., & Myung, J. I. (2006). Global model analysis by parameter space partitioning. *Psychological Review*, *113*(1), 57–83.

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.

Roth, J. C., & Franconeri, S. L. (2012). Asymmetric coding of categorical spatial relations in both language and vision. *Frontiers in Psychology*, *3*(464).

Veksler, V. D., Myers, C. W., & Gluck, K. A. (2015). Model flexibility analysis. *Psychological Review*, *122*(4), 755–769.