

# Attentional Distribution and Spatial Language

Thomas Kluth and Holger Schultheis

Cognitive Systems, Universität Bremen  
Enrique-Schmidt-Str. 5, 28359 Bremen, Germany  
{kluth,schulth}@cs.uni-bremen.de

**Abstract.** Whether visual spatial attention can be split to several discontinuous locations concurrently is still an open and intensely debated question. We address this question in the domain of spatial language use by comparing two existing and three newly proposed computational models. All models are assessed regarding their ability to account for human acceptability ratings for how well a given spatial term describes the spatial arrangement of two functionally related objects. One of the existing models assumes that taking the functional relations into account involves split attention. All new models incorporate functional relations without assuming split attention. Our simulations suggest that not assuming split attention is more appropriate for taking the functional relations into account than assuming split attention. At the same time, the simulations raise doubt as to whether any of the models appropriately captures the impact of functional relations on spatial language use.

## 1 Introduction

Visual spatial attention allows selectively focusing on certain regions of perceivable space such that information processing in the attended regions is enhanced. According to two influential theories, the selective focus of visual spatial attention can be likened to a spotlight [1] or a zoom lens [2]. These conceptualizations highlight a number of important properties associated with attention: (a) enhanced processing is restricted to circumscribed regions in space; (b) processing enhancements are highest at the focal point of attention and gradually decrease with distance from this focus point; and (c) the size of the attended region is not fixed, but can be adjusted based on task demands. Both theories also assume that the attentional focus is unitary in the sense that it comprises a single focal point (perhaps with a graded surrounding) that cannot be split to several discontinuous locations in space. Under this assumption, attentional distribution is uni-focal with the mode coinciding with the focal point of the attentional “spotlight”.

---

<sup>0</sup> The final publication is published in “Spatial Cognition IX, edited by Christian Freksa, Bernhard Nebel, Mary Hegarty, and Thomas Barkowsky, Springer International Publishing 2014“ and available at SpringerLink via [http://dx.doi.org/10.1007/978-3-319-11215-2\\_6](http://dx.doi.org/10.1007/978-3-319-11215-2_6)

Whether attentional distribution is in fact uni-focal or whether it may be multi-focal (i.e., divided attention to several discontinuous locations concurrently) remains a controversial issue. While earlier studies mainly yielded evidence supporting uni-focal attentional distribution [3] there are now many studies that claim to show the existence of multi-focal attentional distribution (e.g., [4,5]). However, methodological complications render it difficult to interpret the results of these studies leading a recent review on divided attention to conclude that decisive evidence in favor of multi-focal attentional distribution still remains to be provided [6].

Producing and comprehending spatial utterances such as “The cups are on the shelf *above* the sideboard” is tightly related to the workings of visual spatial attention [7]: For example, attention is required to identify the objects (shelf and sideboard) that are related in the utterance [8] and also to apprehend the spatial relation between the related objects [9]. In accord with this importance, visual spatial attention plays a key role in a seminal computational cognitive model of processes involved in spatial language use. This model, called *attentional vector-sum* (AVS) model [10], computes acceptability ratings for how well a given spatial term such as *above* describes the spatial arrangement between two objects. In the light of evidence that functional relations between objects impact spatial language use [11,12], [13] proposed an extension of the AVS, henceforth called *functional AVS* (fAVS), that allows the computation of the acceptability ratings to be influenced by functional object relations.

Apart from its merit for our understanding of spatial language use, the fAVS is also of interest for our understanding of visual spatial attention more generally, because the mechanisms realized in the fAVS assume a multi-focal attentional distribution. Against the background of the existing debate in the visual spatial attention literature it is not immediately clear whether such an assumption is justified. On the other hand, if multi-focal attentional distribution is found to be a crucial component of a functional extension of the AVS, this would lend further support to the idea that visual spatial attention can be distributed multimodally. In this paper we investigate the role of multi-focal attentional distribution in the fAVS. We propose a number of alternative functional extensions of the AVS that employ uni-focal attentional distributions and compare all models’ ability to account for pertinent empirical data. In doing so, we not only examine the role of attentional distribution in the AVS, but also provide the first quantitative assessment of the fAVS.

## 2 Models

As in the utterance “The cups are on the shelf *above* the sideboard” spatial terms such as *above* are often used to indicate where to look for a certain object (called *trajector*) in relation to another object (called *landmark*). Determining how well a given spatial term matches the spatial relation between trajector and landmark is an important step in producing / comprehending spatial utterances [14] and all models considered in our simulations address this step. Put differently, the

models determine how acceptable a given spatial term is to describe the spatial relation between landmark and trajectory.

The considered models are the AVS, the fAVS, and three functional extensions of the AVS that employ a uni-focal distribution of attention. All five models are described in more detail in the remainder of this section.

## 2.1 AVS

The AVS-model takes a landmark-object, a trajectory position and a spatial preposition (e.g., *above*) as input. Its output consists of an acceptance rating on a given scale. If AVS returns a low rating, the spatial term *above* is not considered to adequately describe the spatial relation between landmark and trajectory. On the other hand, if the rating is high, the AVS predicts that most observers accept *above* as an appropriate description of the scene.

The rating is computed from two main sources: A height component and an angular component. The height component produces a value between 0 and 1 depending on the height relation between the trajectory and the top-side of the landmark. The higher the trajectory is compared to the top-side of the landmark, the higher the height component will be.

To compute the angular component, an attentional focus  $f$  is defined. Its center is the vertically aligned point on top of the landmark. In case the trajectory is not vertically aligned with the landmark, the attentional focus lies at the point that is closest to being so, i.e., the closest edge of the landmark. Fig. 1 shows an example landmark, trajectory and the corresponding attentional focus as assumed in the AVS.

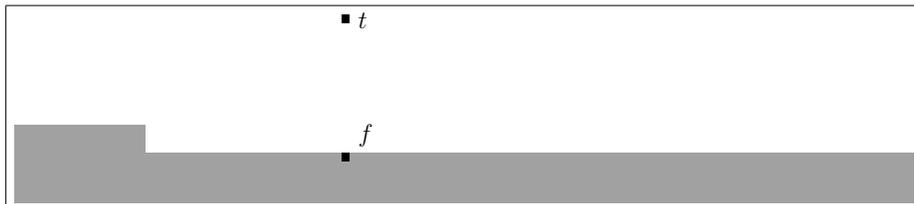


Fig. 1: An example landmark (toothbrush) with trajectory  $t$  and resulting focus  $f$ .

Every point  $i$  of the landmark gets an amount of attention  $a_i$ . The attention is highest in the attentional focus  $f$  and decreases exponentially with increasing distance from  $f$ . A vector  $v_i$  is rooted at every landmark point  $i$ , pointing to trajectory  $t$ . The length of each  $v_i$  is weighted by the amount of attention  $a_i$  at landmark point  $i$ . Fig. 2 visualizes this process with the same example trajectory  $t$  as in Fig. 1.

All these vectors are summed up to create a single vector that is compared to a reference vector. In case of *above* the reference vector is a vector aligned

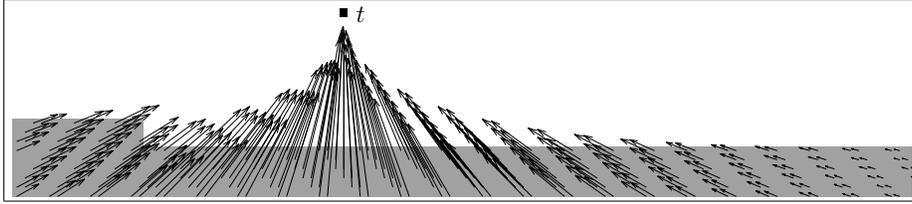


Fig. 2: An example trajectory  $t$  and resulting vectors  $v_i$  weighted by attention  $a_i$ .

with the upright vertical. The angular deviation between the vector sum and the reference vector determines the angular component: The higher the deviation, the lower the angular component will be. To arrive at a final value for the acceptance rating, the height value and the angular value are multiplied. For example, the trajectory used in Figs. 1 and 2 would result in an acceptance rating of 6.46 where 1 is the lowest rating and 7 is the highest rating.

The different functional extensions of the AVS differ only in how the angular component is computed. Computation of the height component is the same as in the AVS for all functional extensions discussed here.

## 2.2 fAVS

The AVS has been shown to successfully account for human acceptability ratings across a wide range of spatial arrangements and object shapes [10]. However, in its original formulation, the AVS does not take into account possible functional relationships between the related objects. Therefore [13] proposed the fAVS as a functional extension to the AVS. The idea behind the fAVS is that a functional part attracts more attention such that the amount of attention at every point lying in the functional part is increased:

$$A_i = \begin{cases} a_i \cdot (1 + \varphi) & \text{if } i \text{ lies in functional part} \\ a_i & \text{else} \end{cases} \quad (1)$$

Here,  $a_i$  denotes the amount of attention at point  $i$ , as defined in the AVS and  $\varphi$  is a free parameter. In the fAVS  $A_i$  is used instead of  $a_i$  to weight each vector  $v_i$ .

In [13] three values for  $\varphi$  were suggested:

- $\varphi = 2$  strong functional interaction between landmark and trajectory
- $\varphi = 1$  weak functional interaction between landmark and trajectory
- $\varphi = 0$  no functional interaction (i.e., AVS)

The fAVS has the flexibility to behave like the AVS with  $\varphi = 0$  – i.e., every point  $i$  of the landmark gets the amount of attention  $a_i$  as defined in the AVS model. However, if  $\varphi > 0$  the else condition of (1) is needed, because only the functional part of the landmark receives a higher amount of attention.

**Attentional Distribution.** Fig. 3 shows an exemplary attentional distribution as assumed by the fAVS. It can be seen that the fAVS leads to a bi-focal attentional distribution if  $f$  does not lie in the functional part, since the functional part receives more attention than its surroundings.

Accordingly, the mechanisms realized in the fAVS amount to the assumption that humans are able to divide their visual spatial attention.<sup>1</sup> As available empirical evidence does not unequivocally support whether or not this is possible (see [6]), we subsequently present three different functional extensions of the AVS that do not assume a multi-focal attentional distribution.

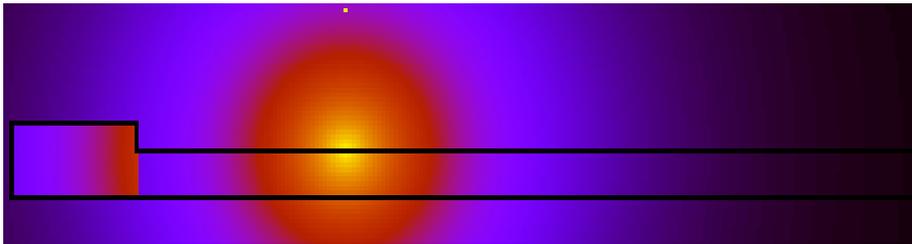


Fig. 3: Bi-focal attentional distribution when using the AVS extension proposed by [13]. Trajectory is assumed to be a toothpaste tube and is positioned at the yellow dot. Toothbrush bristles are defined as functional part. Brighter color means higher amount of attention. Borders of landmark are colored black.

### 2.3 Alternative functional extensions for AVS

In this section we present three alternative functional extensions to the AVS. In the first two extensions the location of the attentional focus is changed, in the last extension an attentional switch is proposed. Note that for all extensions the key idea – functional parts attract attention – stays the same, but that there is no assumption of divided visual spatial attention.

**Focus only at Functional Part.** This extension changes the location of the focus point: The focus always lies on the functional part. The way the focus point is chosen is very similar to the AVS. The only change is that the top of the functional part is used instead of the whole top of the landmark. Note

<sup>1</sup> One may speculate whether the workings of the fAVS could also be interpreted as a formalization of sequential attention shifts instead of instantiating a concurrent multi-focal attentional distribution. However, since nothing in the original formulation of the (f)AVS alludes to such attention shifts, an interpretation in terms of multi-focal attention seems appropriate. Furthermore, as our simulations show, explicitly including attention shifts into the AVS leads to significantly different performance of the resulting model.

that, although the focus always lies on the functional part, the whole landmark gets considered to compute the angular component. Fig. 4 shows an exemplary resulting focus point. This extension is not able to gradually account for the strength of the functional relationship between landmark and trajectory (cf. the  $\varphi$  parameter in the fAVS).

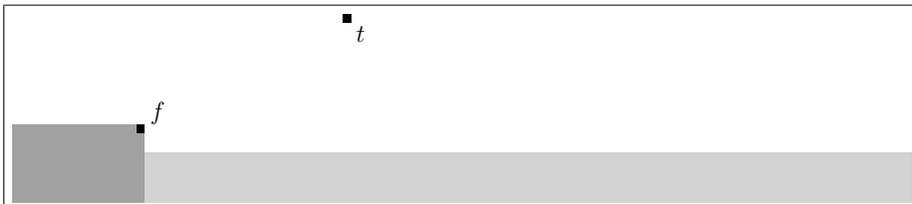


Fig. 4: An example trajectory  $t$  and its resulting focus  $f$  using the *focus only at function* extension. The functional part is colored in dark gray. Trajectory is assumed to be a toothpaste tube.

**Move Focus.** In this extension the location of the attentional focus  $f$ , that is, the point with the highest amount of attention, is changed. In the original model, the attentional focus  $f$  is the point, which is vertically aligned with the landmark (or closest to being so). This extension also starts with this focus point, but adds another step. The focus is moved into the direction of the functional part. The farther away the focus is from the functional part, the more it moves toward the functional part.

To compute the new focus the leftmost and rightmost point on the functional top are used, where the functional top is defined as all points in the functional part where one cannot find a point with the same  $x$ -value but a higher  $y$ -value. More precisely, the new focus  $mf$  is derived from the original focus  $f$ , the leftmost point on top of the functional part  $ltf$  or the rightmost point on top of the functional part  $rtf$  as follows:

$$mf = \begin{cases} f + w \cdot \overrightarrow{f, ltf} & \text{if } f \text{ is to the left of the functional part} \\ f + w \cdot \overrightarrow{f, rtf} & \text{if } f \text{ is to the right of the functional part} \\ f & \text{if } f \text{ lies in the functional part} \end{cases} \quad (2)$$

The parameter  $w$  controls the functional strength and is defined as  $w = \frac{\varphi}{2}$  to obtain values for  $\varphi$  that are comparable with the values proposed in [13], see Section 2.2. If  $\varphi = 0$ , this extension behaves like the AVS (no functional interaction) and if  $\varphi = 2$ , this extension behaves like the *focus only at function* extension.

Fig. 5 shows the focus point  $mf$  of the same example trajectory  $t$  used to illustrate the mechanisms of the AVS, using  $w = 0.5$  (i.e.,  $\varphi = 1$ ). Note that with

this extension the location of the focus point may not lie inside the landmark, as can be seen in Fig. 5.

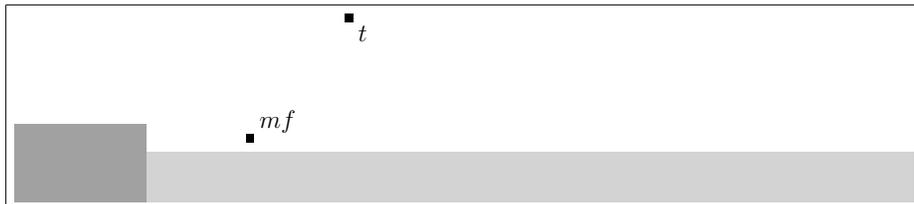


Fig. 5: An example trajectory  $t$  and its resulting focus  $mf$  using the *move focus* extension. The functional part is colored in dark gray. Trajectory is assumed to be a toothpaste tube.

**Attentional Switch.** In this last alternative extension an attentional switch is assumed to operate when the landmark contains functional parts. First the attentional focus lies on the functional part of the landmark and then the landmark is attended as if it had no functional part. More precisely, the extension consists of the following steps:

1. One vector-sum is computed as in the AVS, but with the focus  $f_1$  chosen like in the *focus only at function* extension (i.e., the focus lies on the functional part).
2. A second vector-sum is computed, but this time the attentional focus  $f_2$  is the one from the AVS, that is, the landmark is handled as if it were not containing any functional part.
3. The deviation from upright vertical (in the case of *above*) is measured for both vectors.
4. Both deviations are combined to get a final value for the angular component. Here,  $w$  controls again for the functional strength between landmark and trajectory:

$$angle = \frac{w \cdot funcDev + geomDev}{w + 1} \quad (3)$$

Once more, the parameter  $w$  is defined as  $w = \frac{\varphi}{2}$  to obtain comparable values for  $\varphi$ . If  $\varphi = 0$  this extension is the same as AVS – i.e., equivalent to no functional interaction. If  $\varphi = 2$  function and geometry both play an equal role. For values between 0 and 2 the functional strength is weighted accordingly.

Fig. 6 shows the two vectors as defined in step 1 (left vector) and step 2 (right vector) in light gray. The vector drawn in solid black is the vector with the average deviation from upright vertical, as stated in step 4, using  $w = 1$  (i.e.,  $\varphi = 2$ ). However, note that, to avoid visual clutter, the starting points and lengths of the vectors have been modified.

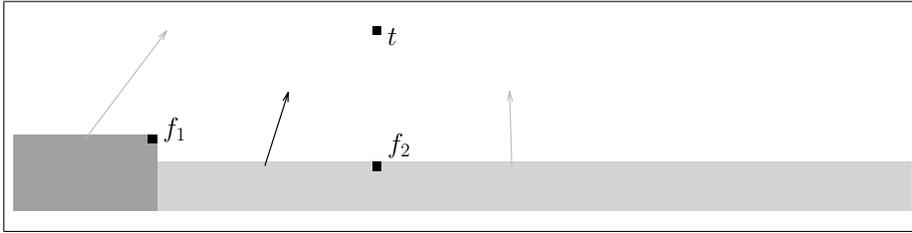


Fig. 6: An example trajector  $t$ , two foci  $f_1$  and  $f_2$  and the resulting vectors according to the *attentional switch* extension. The functional part is colored in dark gray. Trajector is assumed to be a toothpaste tube.

### 3 Model Comparison

Two questions seem of prime importance regarding our inquiry into attentional distribution and spatial language use. The first question is whether any of the functional extensions is able to better account for functional effects in human spatial language use than the other extensions. If, for example, the fAVS would outperform the other extensions this would lend support to the idea that visual spatial attention can be multimodally distributed. The second question is whether any of the four functional extensions is able to better account for functional effects in human spatial language use than the AVS. If, for example, the AVS would perform as well as or better than the functional extensions, this would raise doubt whether the extensions appropriately capture the impact of functional relationships between landmark and trajector.

To provide a first answer to these questions, we assessed the ability of the AVS and its four extensions to simulate human acceptability ratings from two empirical studies that employed functionally related landmark and trajector objects.

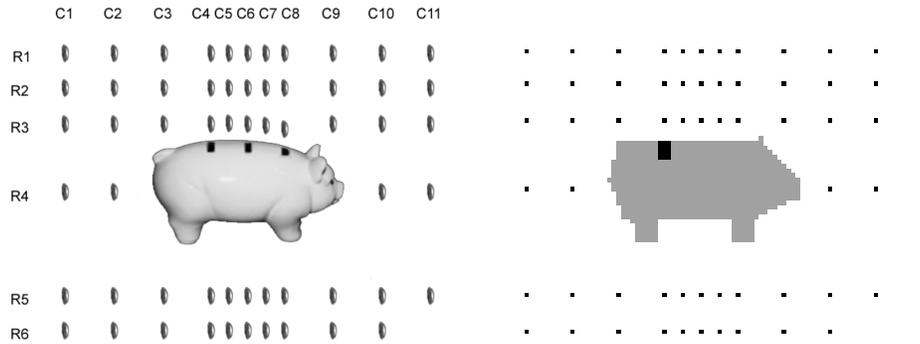
In the following we will first describe in more detail the data and method used for model assessment. Then we will present and discuss the results of the model assessment.

#### 3.1 Data and Method

To compare the AVS and its four extensions we used data from [12, experiment 2] and [15]<sup>2</sup>. For both data sets the experimental setup was the following: Participants were shown a landmark and a trajector at different positions around the landmark. For each trajector (shown at a specific position), participants had to rate the appropriateness of the sentence “The [trajector] is [spatial-preposition] the [landmark]” on a scale from 1 to 7, with 1 being lowest acceptance and 7 highest acceptance.

<sup>2</sup> We thank Thomas Hörberg for sharing his data.

**Carlson-Radvansky et al.** Fig. 7(a) shows the landmark and trajectory positions used in [12, experiment 2], Fig. 7(b) shows the schematized landmark used in our simulations. As can be seen in Fig. 7(a), experimental conditions manipulated the location of the functional part of the landmark (i.e., the coin slot) to investigate how the location impacts acceptability ratings. For each of the 3 slot positions ratings were obtained for all 58 trajectory positions from each participant. The considered spatial term was *above*.



(a) Landmark, trajectory positions and functional parts used in [12]. In the experiment only one coin and slot were shown at the same time. (reproduced from [12, p. 519]. Copyright © 1999 American Psychological Society. Reprinted by permission of SAGE Publications.)

(b) Schematized piggy bank with slot in the back and trajectory positions used in the simulations.

Fig. 7: Landmarks and trajectory positions for [12] data

**Hörberg.** In [15] experiments are presented that investigated the swedish spatial prepositions *ovanför*, *över*, *nevanför* and *under* (corresponding to the english prepositions *above*, *over*, *below* and *under*) with respect to their acceptability when influenced by a functional relationship between landmark and trajectory. As the AVS simulations in [10] and Experiment 2 in [12] only consider *above*, we restricted our simulations to the corresponding swedish preposition *ovanför*.

The functional interaction between landmark and trajectory in [15] is divided in two types: *center-of-mass aligned* and *center-of-mass deviant* interactions. For functional interactions of the first type the center of mass of the trajectory needs to be above the landmark in a strict geometric way (e.g., coin over piggy bank, see Fig. 8(a)). For *center-of-mass deviant* interactions the center of mass of the trajectory is either to the left or to the right of the landmark (e.g. ketchup bottle over hot dog or Fig. 8(c)).

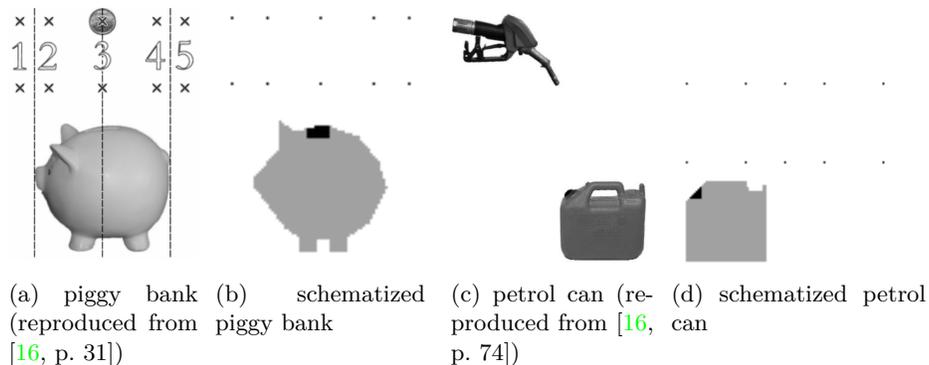


Fig. 8: (a) and (c): Landmarks and trajectory positions used in [15]. (Reprinted by permission of Thomas Hörberg.) (b) and (d): Schematized landmarks and trajectory positions used for simulations. Functional parts are colored in black.

To evaluate the AVS and its extensions with both types of functional interaction, we chose one landmark-trajectory pair for each type, shown together with their corresponding polygons used for simulation in Fig. 8. The piggy bank with its trajectory (a coin) is an example of a *center-of-mass aligned* functional interaction, the petrol can with its trajectory (a gas pump handle) is an example of a *center-of-mass deviant* interaction. The functional part of most of the other landmarks used in [15] spanned the whole top of the landmark – we did not use these landmarks, because all AVS extensions would have behaved the same for such landmark objects.

Every landmark was tested with a prototypical trajectory (e.g., coffee mug and a sugar cube) and a non-prototypical trajectory (e.g., coffee mug and an ice cube). Our simulations focus on prototypical trajectories.

Since the AVS represents trajectories only as a point, but the trajectories were quite big objects in the experiments, we had to decide, which trajectory positions to use in the simulation. We used the center of mass of the coin trajectories for the data from both [15] and [12]. This seems reasonable, because a coin is a small trajectory and the functional interaction shown in Fig. 8(a) and Fig. 7 are *center-of-mass aligned* interactions. Fig. 8(c) shows an example of *center-of-mass deviant* functional interaction. We decided to use the positions of the functionally important parts of the trajectory (i.e., the bottom right of the gas pump handle in Fig. 8(c)) as trajectory positions in the simulation and not the center of mass positions.

**Method.** The AVS and its extensions were implemented in C++.<sup>3</sup> Each of the models was fit to three different data sets: (i) ratings for all three slot positions in [12, Experiment 2]; (ii) ratings for the piggy bank landmark in [15]; and (iii)

<sup>3</sup> The sourcecode can be found at <https://bitbucket.org/kluth/avs>.

ratings for the petrol can landmark in [15]. Using the RMSE of model ratings and empirical ratings as the criterion function, all models were fit using a variant of the Metropolis algorithm [17]. The RMSE that resulted from fitting the models was used as a measure for the *Goodness of Fit* (GOF) of the model to the data.

### 3.2 Results: Goodness of Fit.

All model variations are able to tightly fit the first data set as can be seen in Fig. 9(a). However, in comparison, the *focus only at function* extension is the worst. The extensions *attentional switch* and *move focus* result in neither the best nor the worst fit. The fAVS-model fits the data best, but interestingly, the AVS-model fits nearly equally well – without taking any functional parts of the landmark objects into account.

One may speculate that the reason for the nearly identical performance of the fAVS and the AVS is that the complete data set includes many trajectory positions that are far away from the functional parts. These positions may not be affected by object function as strongly as positions closer to the functional parts. Accordingly, the functional effect may be clouded when fitting the models to ratings from all 58 trajectory positions leading to nearly identical performance of the fAVS and the AVS. To investigate this possibility, we fit all models to two *functional subsets* of the data. The first functional subset (called *large functional subset*) consists of the trajectory positions in the rows R1-R3 and the columns C2-C10 (see Fig. 7(a)). The second functional subset (called *small functional subset*) comprised trajectory positions directly above the positions of the slots, i.e., rows R1-R3 in columns C4, C6 and C8 (see Fig. 7(a)).

All models except *focus only at function* achieve closer and nearly similar fits to the large functional subset (Fig. 9(b)). The GOF achieved by the models on the small functional subset is different from the GOFs on both the complete set of trajectory positions and the large functional subset, as can be seen from Fig. 9(c). All models are able to fit even closer to the small functional subset (cf. magnitude of the RMSE). Interestingly, the models with rather bad GOFs for all trajectory positions or the large functional subset (i.e., *attentional switch* and especially *focus only at function*) achieve the best GOFs for the small functional subset.

The GOF results for the second data set (piggy bank from [15]) and the third data set (petrol can) are depicted in Fig. 10(a) and Fig. 10(b), respectively. The pattern of results are very similar to the results for the first data set: All models are able to closely fit the data, the (f)AVS and *move focus* fit best, and the rest of our proposed alternative extensions result in comparably bad fits, especially the *focus only at function* extension. Note, however, that the fits for the petrol can are generally worse than those achieved on the other two data sets. This may be due to the simplification of the trajectory and the *center-of-mass deviant* functional interaction. Since the data sets in [15] are considerably smaller than the data set from [12], it did not seem reasonable to further reduce the number of data points by employing a functional subset.

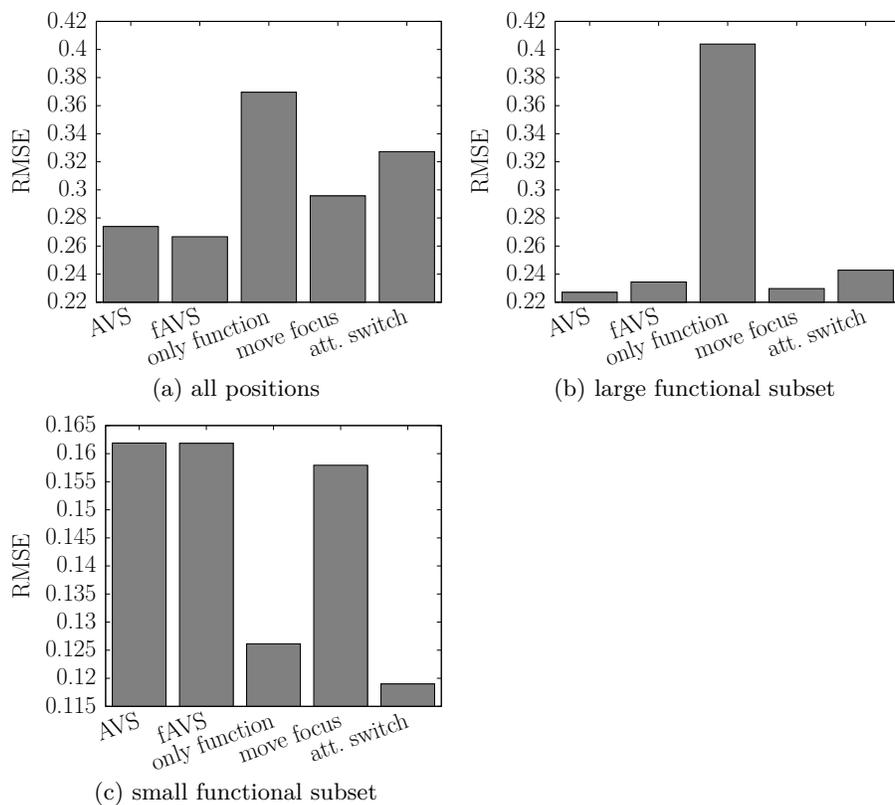


Fig. 9: GOF results for [12]

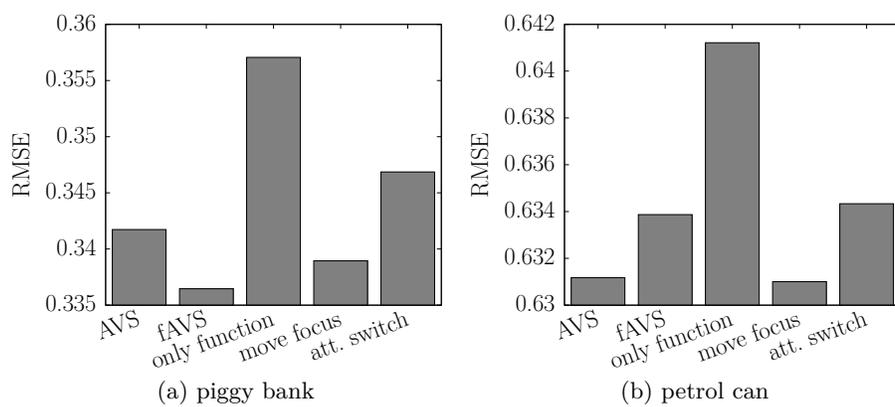


Fig. 10: GOF results for [15]

In sum, across all three data sets, the *focus only at function* extension proposed in this article result in relatively bad GOF-values, except for the small functional subset. The fAVS, the AVS, and the two other proposed extensions perform similarly well.

### 3.3 Results: Simple Hold-Out.

As outlined by [18] a good fit is necessary but not sufficient for a “good” model. If the compared models are of different complexity, more complex models may achieve better GOF values solely by virtue of their complexity and not because they provide a better explanation for the observed data [19]. Against this background, it is worthwhile to ask to what extent the obtained GOF results are influenced by model complexity. To investigate the models’ behavior more closely, we employed the *simple hold-out* (SHO) method, which was found to be among the best performing methods that control for model complexity [20].

The key idea of SHO is to use only a part of the data to estimate parameters (or: train the model) and to predict the remaining data with these parameters (or: test the model). This is done several times using different splits of the data and the RMSE of the prediction is saved for each iteration. The median of all obtained prediction RMSEs is used as an evaluation measure: The lower this median prediction error, the better the model is able to account for the data.

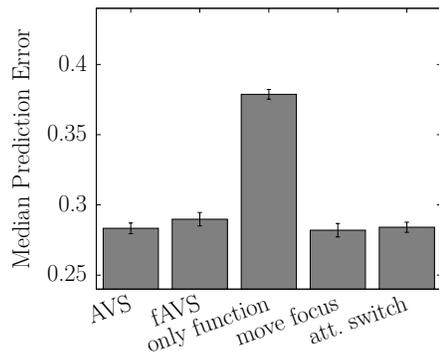
Results for the SHO are shown in Figs. 11 and 12. The shown error bars are the bootstrap standard error estimates computed as stated in [21, p. 47] using 100,000 bootstrap samples.

**Carlson-Radvansky et al.** Simple hold-out results mirror the GOF results, both for all positions (Fig. 11(a)) and for the functional subsets of positions (Figs. 11(b) and 11(c)).

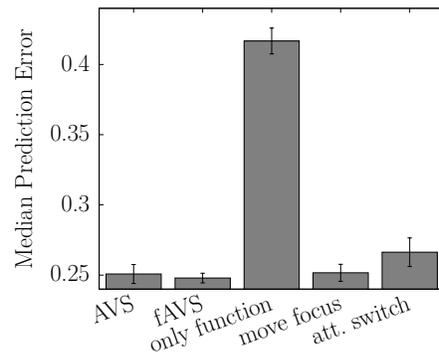
As for the GOF results, the difference between the models’ performance patterns on the small subset versus the other two sets seems noteworthy. On all positions and the large subset all models except the *focus only at function* extension perform similarly while the *focus only at function* extension performs clearly and significantly worse. This pattern is nearly inverted for the small subset: the *focus only at function* extension performs better than the AVS, the fAVS, and the *move focus* extension. The *attentional switch* extension performs best on the small subset.

Accordingly, when the data contains strong functional effects, the *attentional switch* extension provides the best explanation. When the functional effects are embedded in a larger set of ratings, which are not strongly influenced by functional relations, the AVS performs as well as and sometimes even better than all of its functional variants.

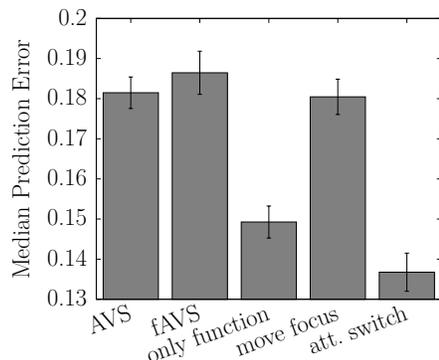
**Hörberg.** The pattern of SHO results is less clear for the data from [15]. Although the relative performance of the models for the piggy bank data (Fig. 12(a))



(a) all positions



(b) large functional subset



(c) small functional subset

Fig. 11: SHO results for [12]

is quite similar to the one observed with GOF values, the large standard errors render the obtained SHO differences inconclusive.

The SHO results for the petrol can data are shown in Fig. 12(b). As for the GOF results, all models have comparatively great difficulties in capturing the petrol can ratings. In contrast to the GOF results, however, the SHO results suggest that there are no substantial differences in how well the models can account for the petrol can data.

The reason for this might be the simplification of the trajectory to a simple point in all tested models. This provides further evidence that the shape and the functional parts of the trajectory are more important than currently assumed in the AVS. An important step towards improving the AVS should, therefore, consist of a better implementation of how shape and functional parts of the trajectory are taken into account (see also discussion in [15]).

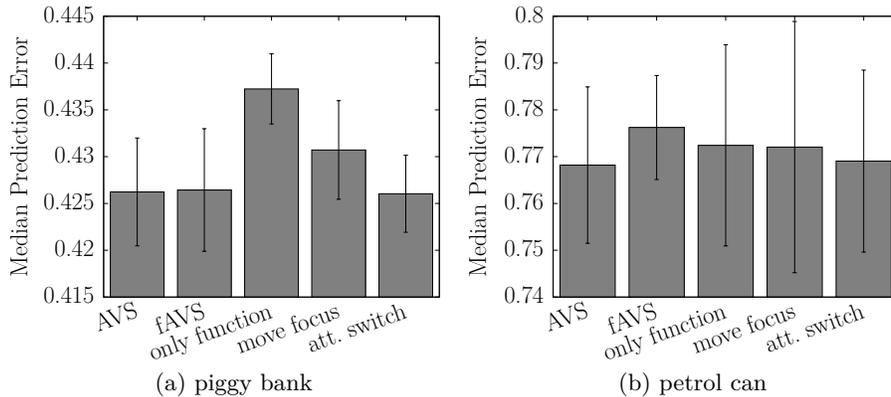


Fig. 12: SHO results for [15]

## 4 Conclusion

The *attentional switch* extension appears to constitute a more plausible functional extension of the AVS than any of the three other extensions. Accordingly, our simulations lend more support to the assumption of a uni-focal instead of a multi-focal distribution of visual spatial attention during spatial term use involving functionally related objects.

At the same time, the advantage of the *attentional switch* extension is restricted to a data set that contains strong functional effects. For all other data sets, the AVS accounted for the modeled human ratings as well as all functional extensions, although computation of ratings in the AVS ignores any functional relationship between landmark and trajectory. This calls into question whether any of the functional extensions appropriately captures the impact of functional

relations on spatial language use. All three considered data sets contained statistically reliable effects of functional relations on acceptability ratings. Even though some of these effects may be considered small, an appropriate functional extension of the AVS would be expected to – at least slightly – better account for the data containing these effects than the AVS. This pattern of results suggests that either the AVS extensions did not capture the functional effects properly or that the extensions captured them at the cost of rating accuracy for other trajectory positions.

Against this background, our future work intends to further investigate possible functional extensions of the AVS in a number of ways. First, we will more closely examine the nature of the differences in the ratings generated by AVS and its extensions. Second, we will consider data related to the use of spatial terms that are assumed to be more strongly influenced by functional relations (e.g., *over* and *över*). Third, we plan to explore possibilities to appropriately deal with spatially extended trajectory positions such that differences found between *center-of-mass aligned* and *center-of-mass deviant* relations can be captured. Ultimately, this work is aimed at devising a refined version of the (f)AVS.

## Acknowledgement

This paper presents work done in the project R1-[ImageSpace] of the Transregional Collaborative Research Center SFB/TR 8 Spatial Cognition. Funding by the German Research Foundation (DFG) is gratefully acknowledged. We also thank the reviewers for their insightful and constructive comments.

## References

1. Posner, M.I.: Orienting of attention. *Quarterly Journal of Experimental Psychology* **32** (1980) 3–25
2. Eriksen, C.W., Yeh, Y.Y.: Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception and Performance* **11** (1985) 583–597
3. McCormick, P.A., Klein, R., Johnston, S.: Splitting versus sharing focal attention: Comment on Castiello and Umiltà (1992). *Journal of Experimental Psychology: Human Perception and Performance* **24** (1998) 350–357
4. Yap, J.Y., Lim, S.W.H.: Media multitasking predicts unitary versus splitting visual focal attention. *Journal of Cognitive Psychology* **25** (2013) 889–902
5. Eimer, M., Grubert, A.: Spatial attention can be allocated rapidly and in parallel to new visual objects. *Current Biology* **24** (2014) 193–198
6. Jans, B., Peters, J.C., De Weerd, P.: Visual Spatial Attention to Multiple Locations at Once: The Jury Is Still Out. *Psychological review* **117**(2) (April 2010) 637–84
7. Carlson, L.A., Logan, G.D.: Attention and spatial language. In Itti, L., Rees, G., Tsotsos, J., eds.: *Neurobiology of attention*. Elsevier, San Diego, CA (2005) 330–336
8. Carlson, L.A., Logan, G.D.: Using spatial terms to select an object. *Memory & Cognition* **29** (2001) 883–892
9. Franconeri, S.L., Scimeca, J.M., Roth, J.C., Helseth, S.A., Kahn, L.E.: Flexible visual processing of spatial relationships. *Cognition* **122** (2012) 210–227

10. Regier, T., Carlson, L.A.: Grounding Spatial Language in Perception: An Empirical and Computational Investigation. *Journal of Experimental Psychology: General* **130**(2) (2001) 273–298
11. Coventry, K.R., Garrod, S.C.: *Saying, seeing, and acting: The psychological semantics of spatial prepositions*. Psychology Press, New York (2004)
12. Carlson-Radvansky, L.A., Covey, E.S., Lattanzi, K.M.: "What" Effects on "Where": Functional Influences on Spatial Relations. *Psychological Science* **10**(6) (November 1999) 516–521
13. Carlson, L.A., Regier, T., Lopez, W., Corrigan, B.: Attention Unites Form and Function in Spatial Language. *Spatial Cognition & Computation* **6**(4) (December 2006) 295–308
14. Logan, G.D., Sadler, D.D.: A computational analysis of the apprehension of spatial relations. In Bloom, P., Peterson, M., Garrett, M., Nadel, L., eds.: *Language and Space*. M.I.T Press, MA (1996) 493–529
15. Hörberg, T.: Influences of Form and Function on the Acceptability of Projective Prepositions in Swedish. *Spatial Cognition & Computation* **8**(3) (September 2008) 193–218
16. Hörberg, T.: *Influences of Form and Function on Spatial Relations: Establishing functional and geometric influences on projective prepositions in Swedish*. Magister thesis, Stockholm University (2006)
17. Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E.: Equation of state calculations by fast computing machines. *The journal of chemical physics* **21**(6) (1953) 1087–1092
18. Roberts, S., Pashler, H.: How persuasive is a good fit? A comment on theory testing. *Psychological review* **107**(2) (2000) 358–367
19. Pitt, M.A., Myung, I.J.: When a good fit can be bad. *Trends in cognitive sciences* **6**(10) (2002) 421–425
20. Schultheis, H., Singhaniya, A., Chaplot, D.S.: *Comparing Model Comparison Methods*. In: *Proceedings of the 35th Annual Conference of the Cognitive Science Society*. (2013)
21. Efron, B., Tibshirani, R.J.: *An Introduction to the Bootstrap*. Chapman & Hall, New York (1993)