

## **UC Merced**

# **Proceedings of the Annual Meeting of the Cognitive Science Society**

### **Title**

A Computational Model of Control Mechanisms in Spatial Term Use

### **Permalink**

<https://escholarship.org/uc/item/330226f3>

### **Journal**

Proceedings of the Annual Meeting of the Cognitive Science Society, 29(29)

### **ISSN**

1069-7977

### **Author**

Schultheis, Holger

### **Publication Date**

2007

Peer reviewed

# A Computational Model of Control Mechanisms in Spatial Term Use

Holger Schultheis (schulth@sfbtr8.uni-bremen.de)

SFB/TR 8 Spatial Cognition, Universität Bremen, Enrique-Schmidt-Str. 5, 28359 Bremen, Germany

## Abstract

The apprehension of spatial terms like “above” or “left” is crucial for communicating about spatial configurations. One important part of apprehension has been shown to be the selection of a *reference frame* (RF). Yet, how this selection is controlled has remained unclear so far. This paper presents a computational analysis and model of the control mechanisms involved in RF selection. Developing the model not only gives a detailed account of the mechanisms involved in RF selection, but also provides new insights regarding the overall sequence of steps involved in spatial term apprehension.

**Keywords:** spatial terms; control; reference frames; computational model; connectionism

## Spatial Term Use

Everyday communication often involves exchanging information about spatial configurations. For instance, to identify the location of some object, say, a fly one might state “The fly is above the table”. Spatial terms like “above” figure prominently in such descriptions and successful communication crucially depends on the communication partners’ abilities to apprehend such spatial terms.

To better understand the processes and representations involved in the apprehension of spatial terms a number of empirical studies (e.g., Burigo & Coventry, 2005; Carlson & Logan, 2001; Regier & Carlson, 2001) have investigated the specifics of human spatial term processing. The empirical results obtained have led to the development of several conceptual (e.g., Logan & Sadler, 1996) and computational (Coventry et al., 2005; Regier & Carlson, 2001) models of spatial term processing. In particular the model of Logan and Sadler (1996) which has been extended and refined in subsequent works (Carlson & Logan, 2001; Carlson-Radvansky & Logan, 1997) has been proven to be a valuable framework for analyzing and accounting for empirical data.

The model assumes that the representations involved in apprehension comprise *reference frames* (RF) and *spatial templates*. RF are assumed to be sets of three orthogonal axes having a distinct origin, orientation, direction, and scale. Spatial templates are thought of as array-like representation structures associated with spatial terms which assign goodness-of-fit values to points in space. Based on these representations the processes engaged in, for example, understanding the utterance “The fly is above the table”, include the following steps: (a) spatially indexing all objects in the scene, (b) identifying the table, (c) imposing multiple RF on the table, (d) aligning the spatial template associated with the term above to the RF, (e) selecting one of the RF (due to steps (d) and (e) the spatial template now assigns certain values to regions of space indicating how well the respective points correspond to the spatial term “above” with respect to the ta-

ble), and (f) identifying the objects which are assigned a high value by the spatial template.

This model constitutes a successful framework for analyzing and accounting for empirical data. However, it is mainly conceptual and, thus, rather coarse-grained regarding the mechanisms underlying human spatial term use. Since the model as a whole is so successful, it seems desirable to unravel the details of mechanisms underlying the different steps. One way to achieve this is to develop computational models from careful analyzes of the available data.

The aim of this contribution is to refine the overall framework by devising a computational model of the mechanisms underlying step (e) of the above sequence, that is, the selection of RF. Like existing empirical research, in developing the model, we will concentrate on the spatial term “above” as used for describing configurations in the plane (see, e.g., Figure 1). Importantly, model development not only refines the framework regarding this particular step but also provides new insights regarding the nature of the overall sequence of steps and specifics of step (c) of the above sequence.

In the following we will first present the computational model / analysis comprising a more detailed description of the RF selection phenomenon, the specifics of the implemented model and its main predictions. Second, the developed model will be evaluated by comparing its behavior with human behavior. Finally, we identify issues for future work and present some speculations as to the applicability / suitability of the devised control mechanisms for other spatial cognition tasks.

## Computational Analysis and Model

### Reference Frames and the Need for Selection

Based on Levinson (1996), (psycho)linguistic theories often distinguish between three types of RF: absolute, relative, and intrinsic. All of them partition the space into regions which are thought of as above, below, left, right, etc. The essential difference between these three types is the (source of) information that is used to determine where in the perceived space these regions lie. In the case of the absolute RF, environmental information like the experienced gravitational force or the sides of an enclosing room is used, that is, “above” is the opposite direction to the gravitational force. The relative RF, on the contrary, partitions space on the basis of the viewer: The line from head to feet defines the above-below axis. Intrinsic RF are defined only for objects with distinguishable sides, since these define the axis of the RF. A table for example has a top (where things are usually put on) and a bottom (where it usually touches the ground). Similar to the relative RF the above-below axis of the intrinsic RF can then be established as the line from top to bottom.

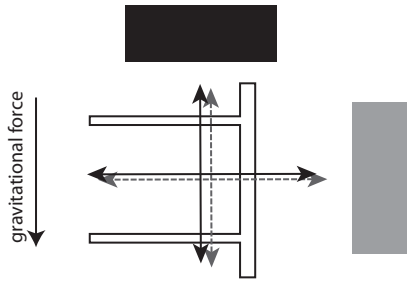


Figure 1: RFs imposed on a table. Absolute / relative RFs (solid black axes) lead to a different partitioning of space than the intrinsic RF (dashed gray axes).

Often these three RF coincide, but in certain situations RFs may partition space differently. Consider, for instance, the scene in Figure 1. Here the object, a table, is not upright, but rotated 90° clockwise. Therefore, the intrinsic RF of the table partitions space differently than the absolute and—assuming an upright viewer—relative RF. “The fly is above the table” could then mean either that the fly is above with respect to the absolute/relative RF (black region in Figure 1) or with respect to the intrinsic RF (gray region in Figure 1). It is such possible disagreement between different types of RF which necessitates selecting one.

To gain a deeper understanding of the mechanisms involved in RF selection we developed a computational model of the selection. The main model components were motivated by three empirical effects reported in the literature: inability of simultaneous activation of several types of RF (Carlson-Radvansky & Irwin, 1994, experiment 3), soft selection (Carlson-Radvansky & Logan, 1997), and negative priming (Carlson-Radvansky & Jiang, 1998). In the following sections we will describe these effects in more detail and present both the model components based on these effects and model predictions arising from the chosen components.

### Simultaneous Activation

In experiment 3 of Carlson-Radvansky and Irwin (1994) participants were informed of the three types of RF and were asked to be prepared to use any type of RF in the experimental task. Interestingly, requiring participants to be equally prepared for any RF did not improve but worsen the participants’ performance. It seems that the participants were not able to prepare for the use of the different types of RF. This inability of preparing for (i.e., activating) several different RF speaks for a direct competition between disagreeing types of RF.

We propose that such direct competition is realized in the form of lateral inhibition. More precisely, our model comprises a number of units which (a) represent the competing entities from which to select, (b) are activated by the information available at the time of spatial term use, (c) feed back positively to themselves, (d) have an activation which is above or equal to zero, and (e) mutually inhibit each other. Assuming  $n$  units, this architecture can formally be described

as follows: The activation function of each unit is the identity for  $x \geq 0$  and zero for  $x < 0$  and the connectivity between the units is given by the weight matrix  $W$  which is defined as

$$W = \begin{pmatrix} 1 & -\epsilon & \dots & -\epsilon \\ -\epsilon & 1 & \dots & -\epsilon \\ \vdots & \vdots & \ddots & \vdots \\ -\epsilon & -\epsilon & \dots & 1 \end{pmatrix} \quad (1)$$

with  $\epsilon$  set appropriately (see below).

Given this activation function and the weight matrix  $W$  consider now a situation in which all of the units are activated to the same degree  $a$ . When representing the activation of the units by a vector  $v$  of length  $n$  this means that  $v_i = a, \forall i: 1 \leq i \leq n$  and, importantly, one iteration  $W * v$  leads to the new activation vector  $\tilde{v}$  with

$$\tilde{v} = \begin{pmatrix} a - \epsilon * (n - 1) * a \\ \vdots \\ a - \epsilon * (n - 1) * a \end{pmatrix}. \quad (2)$$

Thus, (a) all activations decrease and (b) the units’ activations are still pairwise identical. In other words, trying to activate each of the competing units to the same degree will result in none of the units being activated which corresponds to the effect observed by Carlson-Radvansky and Irwin (1994).

In addition to accounting for the preparedness effect, the proposed structure implies three properties of RF selection. One property is related to RF representation the other two are related to the model’s dynamics.

**Reference Frame Representation** The above presented model structure implies that—contrary to common assumption (e.g., Burigo & Coventry, 2005; Carlson-Radvansky & Logan, 1997)—it is not the different types of RF which are activated and selected. If units would represent different types of RF, this would mean that whenever there is evidence for one of the RF types (i.e., an object has intrinsic sides, gravity is present, etc.) the corresponding unit would be activated. In particular, even if all RF were aligned they would activate different units which would lead to competition.

Since competition between aligned RF seems highly implausible, we propose that it is not RFs, but their parametrization which is activated, competing, and selected. Thus, in our model the units competing with each other represent parameter values. This approach allows basing the observed competition on more plausible grounds; only if the different types of RF have disagreeing parameter values, competition will arise.

Furthermore, assuming a competition of parameter values instead of RFs yields a cognitively more efficient account of the processes involved in spatial term apprehension. If only the parameters values compete there is no need to activate multiple RF, impose multiple RF on a scene, activate multiple spatial templates, and align all the latter with the former; only the activation and selection of parameter values is necessary. After a set of parameters has been selected only one RF and

one spatial template needs to be imposed on the scene. Thus, a cognitive system realizing competition via parameter values utilizes fewer complex representations (just one RF and one template) and less processing steps (just imposing and align one RF and one template) than a system realizing competition via RF. In presuming that the human cognitive system has evolved to employ its restricted resources as efficiently as possible, a cognitive model implementing parameter competition seems to be cognitively more plausible.

Besides being more plausible this account of the selection process implied by the proposed structure is supported by recent findings of Carlson and van Deman (in press). Consequently, in our model the competing units represent certain parameter values. More precisely, since for “above” only orientation and direction are important, the model contains four competing units representing the orientations and directions as given by the vectors  $(0, 1)^t, (1, 0)^t, (0, -1)^t, (-1, 0)^t$  in a standard Cartesian coordinate system.

**Model Dynamics** The implications regarding the model’s dynamics concern how initial unit activation is related to model behavior/output.

First, the unit receiving the highest initial activation will win the competition. Due to the nature of the lateral inhibition (see Equation 1) the amount of activation of a unit after one iteration of the model is directly proportional to its activation before that iteration. As a result, the unit with the highest activation will remain the unit with the highest activation until the end of the competition and, thus, win the competition.

Second, the number of iterations necessary to determine the winner depends on the differences between the initial activations of the units: The smaller the initial differences the more iterations are necessary for the competition. If environmental (absolute RF), object (intrinsic RF), and person (relative RF) information all indicate the same direction, only one of the units will be activated and, thus, competition will need zero iterations. If, on the other hand, different sources of information indicate different directions, different units will be activated leading to competition. The precise number of iterations for this competition will depend on the relative magnitude of the different units’ initial activations.

Assuming that the number of iterations in the model is proportional to the time it takes participants to select one particular direction, the model makes the following predictions: (i) spatial term use is fastest if all information sources indicate the same direction; (ii) the weaker the initial evidence for the finally winning direction compared to the evidence for competing directions the longer the reaction times. Regarding the latter, for instance, reaction times should be longer for selecting the absolute direction when it is not the same as the intrinsic direction than for selecting the absolute direction when it is the same as the intrinsic direction.

### Soft Selection

One important aspect of controlling parameter selection is when to terminate the selection process, that is, when to end

the competition. In the extreme case one could argue (e.g., Carlson, 1999) that competition is not finished before there is unequivocal evidence for only one of the competing parameter settings. Alternatively, competition could stop when the indication for one set of parameters is sufficiently high compared to the other possible parameter settings without requiring that there is strict evidence for only one set of the competing parameters.

Empirical evidence regarding these two possibilities is ambiguous. Some experiments speak for a strict selection whereas others suggest a non-strict selection. In particular, the experiments by Carlson-Radvansky and Logan (1997) show that there seem to be strong interindividual differences regarding how strict parameter selection is. Taken together these results indicate that (a) non-strict selection happens to occur and (b) the strictness of the selection varies across situations and individuals.

Accordingly, the model presented here has been equipped with a non-strict selection mechanism the strictness of which is assumed to be one of the main parameters of the model. More precisely, the selection mechanism is realized by a single unit, the *gating unit*, which receives input from all four competing units and feeds forward to four output units. This gating unit is activated only if  $\text{activation}_i / (\sum_{j \neq i} \text{activation}_j) > t$  holds for one of the competing units  $i$ , where  $t$  is the selection threshold of the model. Once this criterion is reached the activation of the competing units is fed to the output units unchanged. Activation of the output units indicates that a parameter setting has been selected. The selected parameter setting is computed as the weighted sum of the vectors represented by the four competing units, where the activation of the output units are taken as the weights for the corresponding vectors. If, for example, the threshold  $t$  were set to 5 and after competition the output units would be activated with the values 5, 0.5, 0, 0 the resulting parameter setting would be given as  $5 * (0, 1)^t + 0.5 * (1, 0)^t + 0 * (0, -1)^t + 0 * (-1, 0)^t = (0.5, 5)^t$ .

An interesting prediction arising from this selection mechanism is the covariation of reaction time and the extent to which the selected parameter setting is a combination of the competing parameter values. If the gating criterion is set to a high value this will—all else being equal—result in more iterations, that is, in longer reaction times. At the same time a high criterion value will enforce large differences in the activations of the output units and, thus, enforce a parameter setting which is influenced strongly by only one of the directions. Accordingly, in a production task such as the one employed in experiment 3 of Carlson-Radvansky and Irwin (1993) reaction times should be comparatively fast in cases where participants put the to be located object into a position indicating a combination of intrinsic and absolute directions.

### Negative Priming

The last effect to be considered stems from the study by Carlson-Radvansky and Jiang (1998). In their experiments they used a negative priming paradigm to further elucidate

the control dynamics involved in RF selection. The major result of these experiments was that some aspect of the control mechanisms underlying parameter selection hampers activating and selecting parameter settings which previously have not been selected.

This implies that there needs to be some form of memory accounting for the—at least recent—history of the control system. We propose that this memory is realized by four extra units. Each of these units is a *shunting model* as developed by Grossberg (1982). Basically, a shunting unit is a time delayed store of the difference of the excitatory and inhibitory inputs it is receiving. In our model each shunting unit receives excitatory input from one of the competing units and inhibitory input from all other competing units and feeds back only to the competing unit from which it receives excitatory input. Due to this setup the net competition signal (i.e., the difference between excitatory self-facilitation and inhibition by the rivaling units) will be fed back to each competing unit indirectly via the corresponding shunting unit and in particular, the shunting unit will accumulate the net competition signal. After competition ends the competition signal stored in the shunting units, will decay to zero over time, but will partly still be available in subsequent situations. Since the net competition signal of losing units will be negative subsequent activation and selection of these units will be more effortful resulting in the negative priming effect observed by Carlson-Radvansky and Jiang (1998).

As the other two model design decisions—and partly in combination with them—implementing negative priming by shunting models leads to specific predictions. First, competing parameter values imply that negative priming will be observed when the not selected orientation and direction in the first situation is the same as in the second situation. In particular, negative priming should also occur if the negatively primed direction was activated by object information in the first situation but is activated by person information in the second situation. RF competition, on the contrary, would predict no priming effect in such a setting. Second, the net competition signal of the winning unit will be positive. As a result, subsequently activating and selecting the same direction should be easier than activating and selecting this direction for the first time. Put differently, positive priming of the previously selected direction should also be observable in human behavior.

The overall architecture resulting from the above considerations is depicted in Figure 2.

### Overall Model

Assume that some input arrives at the control model (indicated by the sketched input connections in Figure 2) and initially activates the competing units (labeled “competition” in Figure 2). If the input arises from a situation with disagreeing parameter values, the gating unit (labeled “gating” in Figure 2) will not be activated and, thus, competition takes place. In each iteration of the competition the competing units feed their activation both back to themselves and to

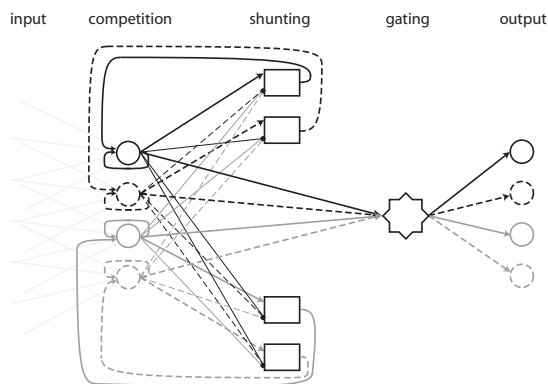


Figure 2: The overall model. See text for details.

all shunting units (labeled “shunting” in Figure 2) and the shunting units feed the resulting net competition  $net^c$  back to the corresponding competing units. Thus, after one iteration the new activation  $act_i(k+1)$  of a competing unit  $i$  is given as  $act_i(k+1) = act_i(k) + net_i^c$ . Since  $net_i^c$  approximately amounts to  $act_i(k) - \epsilon * \sum_{j \neq i} act_j(k)$ ,  $act_i(k+1)$  can be written as  $2 * act_i(k) - \epsilon * \sum_{j \neq i} act_j(k)$ . Accordingly,  $\epsilon \in [1/3, 2/3]$  must hold, because if  $\epsilon$  were below  $1/3$ , convergence of the competition would not be guaranteed, and if  $\epsilon$  were above  $2/3$ , the process might converge with selecting none of the competing parameter values. The model presented here uses  $\epsilon = 0.4$ . By such choosing  $\epsilon$  it is guaranteed that after a finite number of iterations the gating unit will feed the current activation of the competing units to the output units (labeled “output” in Figure 2).

### Evaluation

As detailed above the developed model not only accounts for a number of empirical effects observed in a range of studies, but also generates several new predictions. Unfortunately, there are no experiments or published data which allow testing the model in its entirety. In particular, the predictions arising from the gating and shunting mechanisms cannot be tested with currently available data. However, pertinent data is available regarding the predictions stemming from the lateral inhibition structure. Of these available data we chose to model the experiment 2 of Carlson-Radvansky and Irwin (1994), since this experiment is one of the few studies available where—in some conditions—all three types of information sources are in disagreement.

In their experiment Carlson-Radvansky and Irwin (1994) had their participants work on a sentence-picture verification task; after reading a description indicating that a fly is above an object a picture was shown and the participants had to indicate as quickly as possible whether the description was correct with respect to the picture. This verification task was repeated under several conditions which were essentially generated by three manipulations. The first manipulation concerned the orientation of the object. The object could be either upright (*canonical*) or rotated by  $90^\circ$  clockwise (*non-*

*canonical*). The second manipulation concerned the orientation of the participants. Some participants conducted the experiment upright (*no tilt* group), some reclining with their head to the right (aligned with *top* group), and some reclining with their head to the left (aligned with *bottom* group). The last manipulation concerned the location of the fly which could be either above the object with respect to one type of RF or not above with respect to all types of RF.

For succinctly expressing the condition and the corresponding subject response we will use the following general notation format: XX-XXX<sub>yes, no</sub>, where the 'X' are variables to be replaced by A (absolute), R (relative), I (intrinsic), C (canonical), and NC (non-canonical). The first two capital letters indicate whether the object was rotated or not. The second three letters signify both which RF are disagreeing and where the located object is. The subscript shows whether the participants responded with "yes" or "no". For example, C-ARI<sub>yes</sub> signifies that the object was not rotated, all three RF were agreeing, the fly was above according to these RF, and the participants responded "yes". Two special cases are C<sub>no</sub> and NC<sub>no</sub>. The former means that the fly was not above with respect to all types of RF and subjects responded "no". The latter means that the fly was located below the object according to the intrinsic RF and the participants responded "no".

**Model Application** Since the determination of the input activations is currently not part of the model, the input in the different conditions had to be approximated. Assuming that the proportions with which the participants used the different RF are directly related to the activation which is input to the competing units, we employed the proportions reported by Carlson-Radvansky and Irwin (1994) to derive the input activations. Based on the acquired proportions we set the amount of activation stemming from absolute, relative, and intrinsic sources to be 7.2, 0.7, 2.1, respectively. Thus, for instance, in the condition NC-AR<sub>yes</sub>, the initial activation of the competing units was set to 7.9, 2.1, 0, 0.

Given the thus determined input values the only free parameter of the model, that is, the gating criterion was estimated from the empirical data and found to be 27.6. Using this criterion the model was run for each condition investigated in experiment 2 of Carlson-Radvansky and Irwin (1994). The results of the simulation are shown in Table 1. For each condition (column 1) the table shows the reaction times as observed in the experiment (column 2), the model iterations (column 4), and the reaction times predicted by the model based on a linear regression of the human data on the model iterations (column 3).

Several things seem noteworthy with respect to the results. First of all, as exhibited by the correlation of  $r = 0.71$  the model generally accounts well for the empirical data—especially if considering that only one parameter was fit for the 20 data points. Furthermore, the effects predicted by the lateral inhibition structure are present in the data. For example, the condition where all information sources agree on the direction (C-ARI<sub>yes</sub>) has the smallest reaction time.

Table 1: Empirical and model results with respect to experiment 2 of Carlson-Radvansky and Irwin (1994).

Condition	Study Results	Model Results	
	ms	ms	iterations
<i>no tilt group</i>			
C-ARI <sub>yes</sub>	642	760	0
C <sub>no</sub>	781	760	0
NC-AR <sub>yes</sub>	815	886	60
NC <sub>no</sub>	834	932	82
NC-I <sub>no</sub>	893	886	60
NC-I <sub>yes</sub>	1019	1108	166
<i>aligned with top group</i>			
C-AI <sub>yes</sub>	704	804	21
C <sub>no</sub>	839	846	41
C-R <sub>no</sub>	846	804	21
NC <sub>no</sub>	854	969	100
NC-A <sub>yes</sub>	857	943	87
NC-RI <sub>no</sub>	919	943	87
NC-RI <sub>yes</sub>	1035	1041	134
<i>aligned with bottom group</i>			
C-AI <sub>yes</sub>	725	804	21
C-R <sub>no</sub>	898	804	21
NC-A <sub>yes</sub>	917	894	64
C <sub>no</sub>	992	825	31
NC-R <sub>no</sub>	1040	928	80
NC-I <sub>no</sub>	1094	913	73
NC-I <sub>yes</sub>	1209	1063	144

In addition, comparing the reaction times of C-ARI<sub>yes</sub>, NC-AR<sub>yes</sub>, and NC-A<sub>yes</sub> shows that—as predicted by the model—reaction times increase the more competing sources of information rival the selection of the winning direction.

There are also some effects in the data posing problems to the model as it currently is. For example, no-responses seem to take longer than yes-responses regardless of the selection situation (compare, e.g., C-ARI<sub>yes</sub> with C<sub>no</sub>). Yet, no-responses have commonly and across various tasks been found to take longer than yes-responses in psychological studies. This suggests an effect which is not specific to reference frame selection and, thus, by definition not in the scope of the presented model. A second main effect the model has difficulties with is the fact that reaction times across all conditions are higher in the bottom group than in the top group than in the no tilt group. It is not clear what the reason for this increase of reaction times across groups is. It could be, however, that reclining increased the perceived difficulty of the experimental task. If this was the case, and a higher perceived difficulty led to a more conservative selection criterion (i.e., a higher gating threshold), the model would predict exactly such a reaction time difference between the groups. Nevertheless, we refrained from modeling the data including this assumption, since this seemed too ad-hoc and, moreover, would have resulted in three free parameters.

Summing up, although the model realizes a single step of the processes hypothesized to be involved in spatial term use, it captures the general trend in pertinent experimental data quite well. In addition, human reaction time differences between separate conditions in the modeled experiment were as predicted by the model. Finally, due to its design the model can also account for additional empirical effects not investigated in the modeled empirical data such as negative priming (Carlson-Radvansky & Jiang, 1998) and the fact that spatial templates can be aligned to a direction which is intermediate to the directions competing with each other (Carlson-Radvansky & Logan, 1997).

### Conclusion

In this contribution we presented a computational model of the control mechanisms underlying RF selection in spatial term use. The model was validated by showing its ability to reasonably fit pertinent experimental data. Consequently, the model constitutes a refinement of the framework initially proposed by (Logan & Sadler, 1996). In particular, the model does not just add a new isolated part to this framework, but connects rather naturally to existing models (e.g., Regier & Carlson, 2001) developed in this framework. The model by Regier and Carlson (2001) does explain how humans arrive at a judgment of the appropriateness of the location of an object with respect to a spatial term. Put differently, this model elucidates from which mechanisms the observed spatial templates do arise. As one crucial prerequisite to produce the ratings, the model by Regier and Carlson (2001) needs a reference direction. Such a reference direction is exactly what the model presented in this contribution has as its result. Combining this observation with the above considerations of RF representation yields a new sequence of steps for the overall framework: steps (c) – (e) (see above) should be replaced by (c) selecting parameter settings, and (d) computing the spatial template as demanded by the current task based on the model by Regier and Carlson (2001). As an additional contribution, our model generates a number of new predictions which can be easily tested empirically.

Although the model was developed concentrating on the term “above”, it can be assumed to be valid beyond this domain. Previous research indicates considerable similarities between the apprehension of “above” and other terms (cf. Carlson, 1999; Logan & Sadler, 1996). Thus, there is good reason to believe that the proposed control mechanisms are suitable for all spatial terms. Moreover, the proposed control model might also be accurate for other spatial cognition tasks where the use of RF has been assumed (see Schultheis, in review; Allen, 1999). Yet, to which extent the control mechanisms proposed in this contribution transfer to other spatial cognition task is an issue for future work.

### Acknowledgments

This paper presents work done in the project R1-[ImageSpace] (SFB/TR 8 Spatial Cognition). Funding by the Ger-

man Research Foundation (DFG) is gratefully acknowledged. We also thank the reviewers for their helpful comments.

### References

- Allen, G. L. (1999). Children’s control of reference systems in spatial tasks: Foundations of spatial cognitive skill? *Spatial Cognition and Computation*, 1(4), 413 - 429.
- Burigo, M., & Coventry, K. (2005). Reference frame conflict in assigning direction to space. In C. Freksa, M. Knauff, B. Krieg-Brückner, B. Nebel, & T. Barkowsky (Eds.), *Spatial cognition IV: Reasoning, Action, Interaction*. Berlin: Springer.
- Carlson, L. A. (1999). Selecting a reference frame. *Spatial Cognition and Computation*, 1(4), 365 - 379.
- Carlson, L. A., & van Deman, R. (in press). Inhibition within a reference frame during the interpretation of spatial language. *Cognition*.
- Carlson, L. A., & Logan, G. D. (2001). Using spatial terms to select an object. *Memory & Cognition*, 29(6), 883 - 892.
- Carlson-Radvansky, L. A., & Irwin, D. E. (1993). Frames of reference in vision and language: Where is above? *Cognition*, 46, 223-244.
- Carlson-Radvansky, L. A., & Irwin, D. E. (1994). Reference frame activation during spatial term assignment. *Journal of Memory and Language*, 33, 646 - 671.
- Carlson-Radvansky, L. A., & Jiang, Y. (1998). Inhibition accompanies reference-frame selection. *Psychological Science*, 9(5), 386 - 391.
- Carlson-Radvansky, L. A., & Logan, G. D. (1997). The influence of reference frame selection on spatial template construction. *Journal of Memory and Language*, 37, 411 - 437.
- Coventry, K. R., Cangelosi, A., Rajapakse, R., Bacon, A., Newstead, S., Joyce, D., et al. (2005). Spatial prepositions and vague quantifiers: Implementing the functional geometric framework. In C. Freksa, M. Knauff, B. Krieg-Brückner, B. Nebel, & T. Barkowsky (Eds.), *Spatial cognition IV: Reasoning, Action, Interaction*. Berlin: Springer.
- Grossberg, S. (1982). *Studies of mind and brain*. Boston: Reidel Publishing.
- Levinson, S. (1996). Frames of reference and molyneux’s question: cross-linguistic evidence. In P. Bloom, M. Peterson, M. Garrett, & L. Nadel (Eds.), *Language and space*. MA: M.I.T Press.
- Logan, G. D., & Sadler, D. D. (1996). A computational analysis of the apprehension of spatial relations. In P. Bloom, M. Peterson, M. Garrett, & L. Nadel (Eds.), *Language and space*. MA: M.I.T Press.
- 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.
- Schultheis, H. (in review). A control perspective on imaginal perspective taking. In *Proceedings of the 8th International Conference on Cognitive Modeling*.