Psychonomic Bulletin & Review
2010, 17 (4), 575-580
doi:10.3758/PBR.17.4.575

Function and context affect spatial information packaging at multiple levels

Elena Andonova and Thora Tenbrink
Bremen University, Bremen, Germany

and

Kenny R. Coventry
Northumbria University, Newcastle, England
and Hanse Institute for Advanced Study, Delmenhorst, Germany

In the present study, we examined how context of instruction and information in the visual array to be described affect spatial information packaging across a range of levels of spatial description. Participants described complex scenes containing 3-D dollhouse furniture across two different arrays (functional vs. nonfunctional arrangements of objects) and across instructional contexts (living room context, furniture showroom context, no context). Knowledge about the visual scene and instructional context both had an impact on spatial descriptions, but separately, and at different levels of granularity. The influence of visual context was particularly striking, with marked differences across conditions at multiple levels of information packaging—descriptive trajectories (the order in which objects in the spatial array were described), amount of detail, and explicit mention of atypical object orientation. The importance of visual context as a means of accessing context frames in common ground is discussed.

Experience in the visual world informs our predictions about the identities and typical spatial arrangements of objects that tend to co-occur within particular contexts. For example, a living room usually contains a sofa and a coffee table, and the front of the sofa is normally aligned with the long side of the coffee table. When activated prior to the presentation of information to be processed, such "context frames" (Bar, 2004; alternatively described as schemata, scripts, and frames) have been shown to affect the processing of information in visual recognition, memory, and language understanding. For instance, it is well known that titles presented prior to a text generally facilitate memory for that text (e.g., Dooling & Lachman, 1971). However, less is known about how schemas affect language production choices.

In addition, local information itself can activate background knowledge as processing unfolds in the absence of an initial context frame. For instance, a sofa and coffee table encountered in typical relative orientations may tap into background knowledge, affecting how those objects are conceptualized, irrespective of context. It has been established across a range of domains that visual cues and object information can trigger the activation of higher level reasoning mechanisms and cognitive schemas. For example, in Michotte's (1946/1963) classic studies, animacy and causality were habitually "perceived" in simple patterns of moving abstract shapes. Yet little is known about how context frames activated *in advance of* visual information versus background information activated locally *during* the processing of incoming information affect cognitive processes.

In the present article, we address how initial context presented verbally versus the nature of the visual array to be described affect the communication of spatial information. During communication, speakers are sensitive to the background knowledge they have in common with interlocutors ("common ground"; Clark & Brennan, 1991; Clark & Wilkes-Gibbs, 1986). In a previous study, Ehrich and Koster (1983) found, using purely descriptive analyses, that the production of room descriptions was influenced by the nature of the spatial array being described.

However, in describing a visually presented scene that deviates in some ways from the typical schema settings (e.g., a chair facing away from the table), a speaker may be guided by both visual context and long-term context frames. Speakers commonly rely on general schemas that are part of the culturally acquired knowledge shared with interlocutors, but they cannot rely on shared context frames triggered by visual context when a scene is not visually accessible to their audience. In this case, information-packaging decisions derive from the verbal context, the visual context, or a combination of the two.

The spatial domain is an excellent testing ground for the interaction of visual context and cognitive schemas.

E. Andonova, andonova@uni-bremen.de

Previous research on the comprehension of and memory for texts has shown that spatial information is encoded and retrieved more reliably if it is functional and lies on the "causal path" of a coherent narrative (e.g., Radvansky & Copeland, 2000; Radvansky, Copeland, & Zwaan, 2003; Sundermeier, van den Broek, & Zwaan, 2005). For example, Radvansky et al. established superior memory for functional over nonfunctional spatial relations: Participants remembered the spatial relation of someone standing below an old bridge better than that of someone below a lamppost or of someone in front of an old bridge in the context of a rainy day. In another series of experiments, Sundermeier et al. (2005) showed that, during reading, information concerning spatial relations is reactivated when needed for causal coherence, but less so otherwise. Clearly, functional spatial relations are important because they are involved in more general causal cognitive processes.

Here, we address the functionality of multiobject arrays as determined by habitual human use. For example, for a chair and desk to be in a functional arrangement would entail that both the location and mutual orientation of these objects allow for their combined use in a typical human activity. Functional relations between objects—namely, expectations regarding how objects typically interact or how people interact with two objects—have been shown to be an important predictor of lexical choice across a range of spatial relations (for a review, see Coventry & Garrod, 2004), but little is known about the effects of functionality on extended multiobject descriptions.

Our main goal was to ask, for the first time, whether visual context and cognitive schemas exert a combined influence on spatial information packaging in spatial description and, if so, whether their impact can be felt across a range of levels when individuals describe arrays with multiple 3-D objects. Participants saw an arrangement of dollhouse furniture and were asked to provide a description of where objects were in relation to each other. The instructional context told participants that the array represented a secondhand furniture store, told them that it represented a living room, or gave no information about the setting. We expected context frames to affect the global trajectory patterns and the more local description choices for the scene, such as the type and amount of information given. This manipulation was crossed with the nature of the array itself, in which objects were arranged either in a functional way, where objects' mutual positions and orientations were consistent with their typical combined use, or in a nonfunctional way, where objects were positioned in atypical ways with respect to each other. We expected both of these variables to affect spatial description choices at a range of levels, and we wanted to determine the relative influence of instructional context versus visual array context in this process. First, if information consonant with context frames affects spatial description, then main effects of both variables on description choices would be found. Second, the availability of two types of information that can afford access to context frames might necessitate choosing between them. In this case, effects of only one variable might be found, at the expense of the other. Third, both manipulations might conjointly allow access of context frames. In that case, incongruent information across variables might lead to difficulty in accessing the correct frame, resulting in more explicit detail in descriptions to avoid possible inconsistencies in information packaging. It was of great interest to us to learn which of these possibilities would be the case, given that spatial description usually occurs with both verbal and visual context.

METHOD

Participants

Participants were 100 German students (59 female, 41 male, mean age = 21.21 years, age range = 17–57) who received course credit or were paid for their participation.

Design and Stimuli

The design was a 3 (instructional context: living room, furniture store, neutral) × 2 (array type: functional vs. nonfunctional) between-participants design.

Fifteen pieces of 3-D furniture from a dollhouse were placed on a flat wooden dollhouse floor, measuring 71 × 37 cm, and the objects were arranged in a 3 × 5 grid with equal distances between them. The 15 objects included chairs, tables, sofas, shelves, and other similar items, all of which could plausibly be found in a living room and a furniture store. Two arrangements were used: a functional arrangement and a nonfunctional arrangement (Figure 1). Objects remained in the same position in the two arrangements, except for four pairs: a sofa and a cupboard, a table and a chair, a shelf and an armchair, and a serving cart and a chair. These were arranged to form clusters of functionally related objects for the functional array (e.g., a table between two chairs facing toward it) and in a way that minimized typical coplacement patterns for the nonfunctional array.

We also created two linguistic contexts that could be associated with rather different context frames. The arrangement was said to belong to Mr. Meyer's living room ("Herr Meyer's Wohnzimmer") or to Mr. Meyer's secondhand furniture store ("Herr Meyer's Secondhand-Möbelgeschäft"), or no context was given regarding the nature of the location.

Procedure

Testing was conducted in German. Participants were asked to describe the arrangement of the furniture items presented to them so that another person could later arrange the items accurately. There were three versions of the instructions corresponding to the three levels of context manipulation. Descriptions were recorded as audio material and later transcribed. Each participant described one arrangement. The task lasted approximately 10 min.

Response Categories

Participants' descriptions were segmented into utterances (Selting, 2000). We examined the effects of array and context of instruction on the overall volume of information (length of description), on the organization of information found in the descriptive trajectories, and on the type of information (regarding location and orientation), in terms of percentages of use/mention.

Description length. We counted the number of utterances and the number of words in each description. We reasoned that the mere quantity of speech produced would be informative regarding the types of context frames used by participants.

Trajectory type. We examined whether participants adopted a sequential, object-by-object strategy or jumped around the array describing clusters of objects. The order of objects and object clusters (where relevant) in the spatial descriptions were transferred in the form of drawn trajectories onto 2-D pictures of the arrays used (for examples, see Figure 2). *Cluster* was defined as mention of a group of objects (e.g., "There's a table and two chairs").

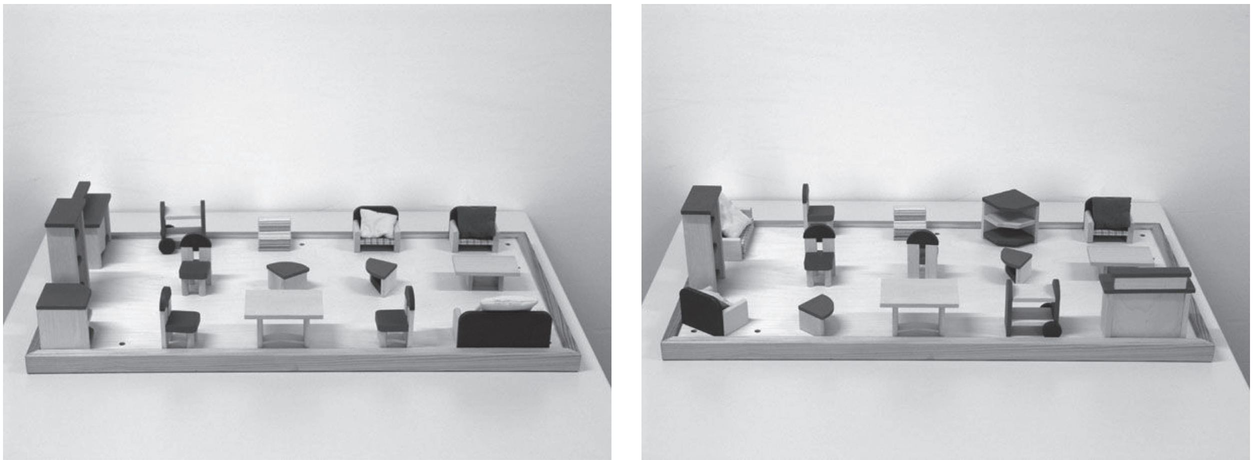

Figure 1. The panel on the left illustrates a functional array. From left to right, the top row includes a cupboard, a serving cart, a fold-out sofa, an armchair with a yellow cushion, and an armchair with a blue cushion. The middle row includes a cupboard with three shelves, a chair, two small corner tables, and a coffee table. The bottom row contains a corner unit, a chair, a dinner table, a chair, and a sofa with a yellow cushion. The panel on the right illustrates a nonfunctional array. From left to right, the top row includes a sofa with a yellow cushion, a chair, a fold-out sofa, a corner unit, and an armchair with a blue cushion. The middle row includes a cupboard with three shelves, two chairs, a small corner table, and a coffee table. The bottom row contains an armchair with a yellow cushion, a small corner table, a dinner table, a serving cart, and a cupboard.

For the drawn trajectories, we counted the number of direct horizontal and direct vertical links between objects, indirect links (including diagonal links and jumps over adjacent nodes on the grid), and clusters of objects. Overall trajectory shapes were also coded as either regular (linear, following the horizontal or vertical lines of the 3×5 grid) or irregular (chaotic, with no discernible order of objects/regions, or cluster based). We hypothesized that the absence of a context frame in which to organize the space and the objects that populate it would trigger more abstract (object- and cluster-independent) linear and regular descriptive trajectories.

Types of spatial information. Utterances were classified into three types. One class described the location of an object without reference to its orientation, including sequential information (e.g., "There is a chair, then a table," etc.) and projective term usage that described location when a direction was specified (e.g., in front of, behind, left of, etc.). Another class of utterances described the orientation of an object with or without information on its location (e.g., "The chair is pointing to the left," or "The table in front of the chair pointing to the left."). We examined the relative presence of orientation information in descriptions as a percentage of all spatial reference utterances. We expected more orientation references in descriptions of the nonfunctional array, in which some objects were oriented in an atypical, schema-incongruent way. Finally, there were utterances that did not refer to location or orientation information (e.g., "There is a yellow table," "The chair has a backrest," etc.).

All transcriptions and trajectory transformations were undertaken by coders who were blind to the purpose of the study. There was 96.46% intercoder agreement on location utterances and 92.04% intercoder agreement on orientation utterances, based on matching coding choices for 12% of the data.

RESULTS

The data were analyzed using a 2 (array type: functional, nonfunctional) $\times$ 3 (instructional context: furniture store, living room, neutral) between-participants ANOVA, unless otherwise stated ($\alpha = .05$). Follow-up analyses were performed with LSD tests across the response categories outlined above.[1]

Description Length

For the number of words overall, no main effect of context was found [$F(2,94) = 1.256$, $p = .290$], but the effect of array type was significant [$F(1,94) = 4.86$, $MS_e = .061$, $p = .030$, $\eta_p^2 = .049$]. Overall, the mean number of words produced for the nonfunctional condition was 324 per description; for the functional condition, it was 254 words. The context $\times$ array interaction was not significant.

For the number of utterances overall, there were reliable main effects of context [$F(2,94) = 3.18$, $MS_e = .034$, $p = .046$, $\eta_p^2 = .063$] and array type [$F(1,94) = 5.85$, $MS_e = .034$, $p = .018$, $\eta_p^2 = .059$], but the interaction was not reliable. Overall, the mean number of utterances produced in the nonfunctional condition was 23.06 (SD = 14.25), as compared with a mean of 18.11 (SD = 7.83) in the functional condition. The mean numbers of utterances for the living room, furniture store, and neutral context conditions were 20.06, 17.69, and 25.31, respectively. The furniture store context produced reliably fewer utterances than did the neutral context ($p = .012$).

Trajectory Type

The mean numbers of types of links and clusters by array type are displayed in Table 1, and two examples of trajectories are shown in Figure 2. Analyses of the trajectories of scene description examined the ratio of the number of direct links (vertical and horizontal), of indirect links (diagonals and jumps), and of object clusters, to the number of utterances for each description. There was a main effect of array type on the number of clusters per utterance [$F(1,94) = 41.948$, $MS_e = .258$, $p < .001$, $\eta_p^2 = .309$]. Clusters of furniture items were mentioned considerably more frequently in descriptions of the functional arrangement (16.19% of all utterances) than in descriptions of the nonfunctional arrangement

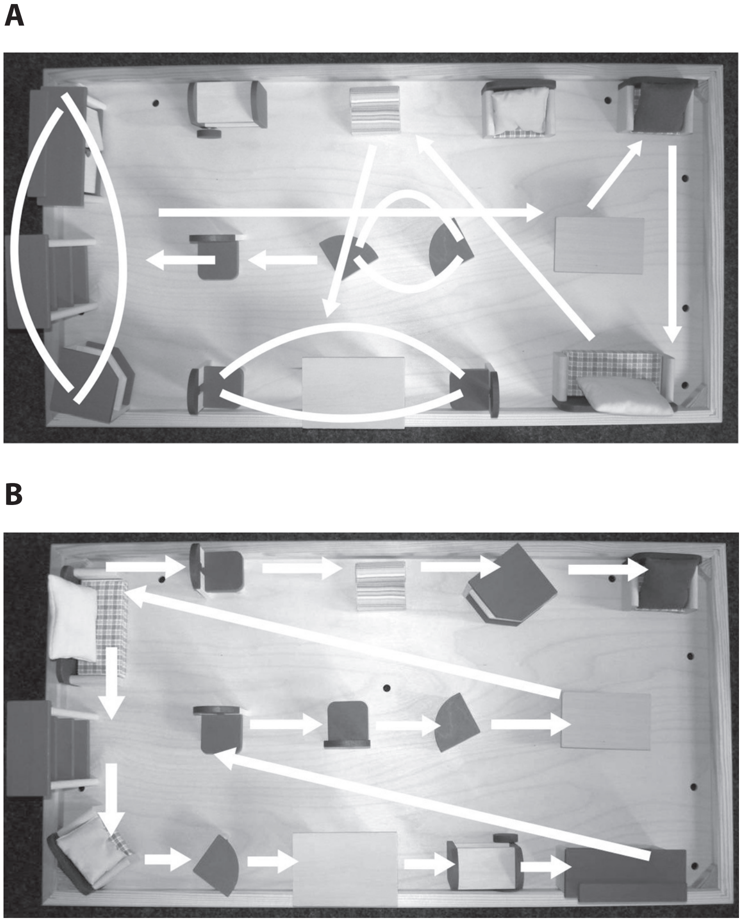

Figure 2. (A) The functional array and an example of a cluster-based, irregularly shaped trajectory. (B) The nonfunctional array and an example of a linear, regularly shaped trajectory.

(3.82% of all utterances). No other main effects or interactions were reliable.

Table 2 shows the frequency of regular versus irregular trajectories by condition. Regular trajectories described the array mostly as either a zigzag pattern or a figure six pattern. A 2 × 2 (array × trajectory shape) chi-square analysis showed a significant effect of array type on the frequency of trajectory choices [$\chi^2(1) = 4.76$, $p = .029$]. Participants' descriptive trajectories were more likely to have a regular shape in the nonfunctional array condition (62.26%) than in the functional arrangements (40.43%). A 3 × 2 (context × trajectory shape) chi-square analysis revealed no significant effect of context type on frequency of trajectory choices [$\chi^2(1) = 0.0142$, $p = .99$].

Spatial Information Type

We first examined the mean percentages of utterances containing spatial versus other information about objects. Nonspatial (*other*) utterances typically described spatially irrelevant properties of objects, such as color, size, parts, or partitioning, and general organization of the visual array (e.g., "There are long and short walls," etc.). Spatial information utterances either did or did not contain information about object orientation. Table 3 shows the mean percentages of utterances describing object location but not orientation, utterances describing object orientation, and other (nonspatial) utterances. An analysis of the percentages of utterances referring to object orientation produced a significant main effect of spatial array [$F(1,94) = 5.969$, $MS_e = .449$, $p = .016$, $\eta_p^2 = .060$]. A total of 30.96% ($SD = 23.57$, range = 0–80) of utterances contained information on orientation in nonfunctional array descriptions, as compared with 17.97% ($SD = 18.41$, range = 0–77.78) in the functional array.

Table 1
Mean Number of Types of Links and Clusters by Array Type

Links/ Clusters	Functional Array			Nonfunctional Array		
	M	*SD*	Range	*M*	*SD*	Range
Horizontal	6.4	2.8	2–12	7.5	3.8	0–13
Vertical	3.3	1.8	0–8	3.8	2.7	0–9
Indirect	2.3	1.5	0–7	2.3	1.9	0–9
Clusters	2.3	1.7	0–7	0.7	1.3	0–5

Note—*Horizontal*, direct horizontal links; *vertical*, direct vertical links; *indirect*, diagonal links or jumps across the grid; *clusters*, objects described in a cluster.

Table 2
Number of Regularly Versus Irregularly Shaped Trajectories by Context and Array Type

Context	Functional Array		Nonfunctional Array	
	Regular	Irregular	Regular	Irregular
Living room	8	9	10	8
Furniture store	6	11	13	6
Neutral	5	8	10	6

An analysis of the percentages of *other* utterances containing neither location nor orientation information produced a significant main effect of context [$F(2,94) = 4.374$, $MS_e = .325$, $p = .015$, $\eta_p^2 = .085$]. A total of 14.96% of all utterances ($SD = 14.18$, range = 0–50) made no reference to the location or orientation of objects in the neutral condition, as compared with 7.45% ($SD = 9.10$, range = 0–31) in the furniture store context condition and 8.27% ($SD = 9.53$, range = 0–32) in the living room context condition. The neutral condition differed significantly from the furniture store ($p = .007$) and living room ($p = .025$) context conditions.

DISCUSSION

Using natural descriptions for spatial scenes containing multiple 3-D objects has provided us with an insight for the first time into the importance of functional object relations and descriptive context—two means of potentially activating context frames—for spatial description across a range of levels of analysis. This range includes the level of explicitness and detail included (the volume of information that is made explicitly available), the specification of object location and orientation (variation in the kinds of information that are explicated), and the overall descriptive strategy (the sequential organization of this information).

At the coarsest grain level—that of the amount of information given in a description (number of words and number of utterances)—there were main effects of both array type and context of instruction, but these variables did not interact. More fine-grained levels of analyses elucidate why this is the case.

With respect to the strategies employed by participants to describe their way around the array (i.e., their descriptive trajectories), there was a main effect of array alone in the predicted direction. When the relative positions of objects suggested a functionally familiar context frame, descriptions reflected that organization (i.e., arrays were not typically described linearly from top left to bottom right). The results strongly suggest that participants first scanned the whole array, trying to map any relative object positions onto background knowledge if possible, thereby providing hooks to background knowledge also for future interlocutors, who had to arrange the arrays from the descriptions. If a context frame cannot be found as a structuring principle for object information, speakers opt for a regular and transparent spatial sequencing in object reference to aid future hearers. This finding is consistent with the findings of Ehrich and Koster (1983).

Turning to the finest grain level of analyses—the content of the utterances—we again found effects of array and instructional context working independently of each other. Instructional context, but not array, influenced the degree to which local nonspatial information was included in the descriptions. In the neutral condition, where a clear contextual schema was not evoked in the instructions, more such utterances were produced than in the schema-supported instructional contexts. Spatial array, on the other hand, affected the extent to which participants referred to the orientation of objects in the arrangements.

For the functional array, where the relative orientations of objects can be inferred from a background context frame, there is less need to spell out how individual objects are oriented. In contrast, for the nonfunctional array, the absence of plausible relations between objects necessitates specific orientation information. This is related to the need to refer to atypical visual and spatial features more than to typical default characteristics that are part of the schema (Lockridge & Brennan, 2002) and to the tendency for speakers to make information explicit when not guided by the assumption that it is part of a shared cognitive schema or the common conversational ground (Clark & Brennan, 1991).

The overall pattern of results is revealing regarding the role of context frames in spatial information packaging across a number of levels, an issue that the present study addressed for the first time. The weaker effects of instructional context and the lack of interaction between array and context indicate the dominance of the visual spatial array in activating frames. Participants were guided by their assumptions of shared context frames (as types/prototypes) when describing the specific array within a specific context instruction (as a token/instantiation in an episodic scene).

Table 3
Mean Percentages of All Utterances That Carry Location Information Only, Orientation Information, and Other Nonspatial Information, by Context of Instruction (Living Room, Furniture Store, Neutral) and Array Type (Functional, Nonfunctional)

Information Carried	Living Room		Furniture Store		Neutral	
	Functional	Nonfunctional	Functional	Nonfunctional	Functional	Nonfunctional
Location only	74.01	56.64	79.60	64.35	66.70	67.20
Orientation	16.46	36.27	14.48	26.83	15.57	20.09
Other	9.53	7.09	5.92	8.82	17.73	12.71

Note—*Location only*, percentages of utterances describing objects' location but not orientation; *orientation*, percentages of utterances describing objects' orientation; *other*, percentages of utterances describing neither the location nor the orientation of objects.

However, although the *instruction* itself was general enough to activate a general *type* (e.g., a living room), this was nevertheless relatively underspecified (e.g., whose living room, where, how large, how tidy, etc.). The *array* itself provided more specific information (closer to a *token*) regarding the specific objects and their relative positions.

As attention is directed around the array, the information in focus is the specific part of the spatial array being attended to. In that sense, the visual array also dominates because this information sits within an attentional window as spatial description unfolds. Visual cues and associated contexts appear to have the upper hand because they provide more information at a more specific level than general context frames do.

The finding that the visual array had a greater impact on speakers' choices than the context of instruction did suggests that information that is attended to and that changes as attention unfolds may generally dominate the access to background knowledge. In dialogue, however, verbal context may be revisited more frequently and thus may facilitate and reinforce context frames more reliably. Furthermore, to what extent the patterns found in monologic descriptions would also hold in a dialogic interaction remains to be studied in the future—for instance, whether the diverging strategies we established in the analyses of descriptive trajectories (regular vs. irregular) would persist or whether and how they may be modified across the span of a dialogic interaction.

The main conclusion of this research concerns the relative and combined contributions of long-standing knowledge in terms of cognitive schemas and visual context in spatial information packaging. Our findings reveal no evidence of close integration of the two constraints. It appears that these two information sources may be in competition, with visual context and contextual frames affecting different levels of information packaging in a relatively independent way. Finally, this study shows, for the first time, how context frames provide sets of expectations that can guide not only perception and memory but also language production of extended descriptions designed for a generic audience.

AUTHOR NOTE

This research was partially supported by a fellowship from the Hanse Institute for Advanced Studies awarded to K.R.C., by funding from the German Research Foundation for the SFB/TR8 Transregional Spatial Cognition Research Center (Project I5-[DiaSpace]) to E.A. and T.T., and by funding from the Volkswagen Foundation for the Tandem Project on Wayfinding Strategies in Behavior and Language granted to T.T. We thank our student assistants for their help, reviewers of an earlier version of this study with a smaller data set (Cognitive Science Society Conference Proceedings, 2008), and two other reviewers for helpful comments. The order of authorship of the present article is arbitrary; the three authors contributed equally to the work reported. Correspondence concerning this article should be addressed to E. Andonova, SFB-TR 8 Transregional Spatial Cognition Research Center, Bremen University, Bremen, Germany (e-mail: andonova@uni-bremen.de).

REFERENCES

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*, **5**, 617-629.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127-149). Washington, DC: American Psychological Association.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, **22**, 1-39.

Coventry, K. R., & Garrod, S. C. (2004). *Saying, seeing, and acting: The psychological semantics of spatial prepositions*. New York: Psychology Press.

Dooling, D. J., & Lachman, R. (1971). Effects of comprehension on retention of prose. *Journal of Experimental Psychology*, **88**, 216-222.

Ehrich, V., & Koster, C. (1983). Discourse organization and sentence form: The structure of room descriptions in Dutch. *Discourse Processes*, **6**, 169-195.

Lockridge, C. B., & Brennan, S. E. (2002). Addressees' needs influence speakers' early syntactic choices. *Psychonomic Bulletin & Review*, **9**, 550-557.

Michotte, A. (1963). *The perception of causality* (T. R. Miles & E. Miles, Trans.). New York: Basic Books. (Original work published 1946)

Radvansky, G. A., & Copeland, D. E. (2000). Functionality and spatial relations in memory and language. *Memory & Cognition*, **28**, 987-992.

Radvansky, G. A., Copeland, D. E., & Zwaan, R. A. (2003). Aging and functional spatial relations in comprehension and memory. *Psychology & Aging*, **18**, 161-165.

Selting, M. (2000). The construction of units in conversational talk. *Language in Society*, **29**, 477-517.

Sundermeier, B. A., van den Broek, P., & Zwaan, R. A. (2005). Causal coherence and the availability of locations and objects during narrative comprehension. *Memory & Cognition*, **33**, 462-470.

NOTE

1. Dependent measures were logarithmically transformed by a 10-base logarithm with one added, $\log_{10}(x+1)$, to avoid the undefined $\log_{10}(0)$.

(Manuscript received September 21, 2009;
revision accepted for publication January 22, 2010.)