

# Spatial Perspective in Spoken Descriptions of Real World Environments at Different Scales

**Sarah Kriz (kriz@psych.ucsb.edu)**

Department of Psychology, University of California  
Santa Barbara, CA 93106 USA

**Mary Hegarty (hegarty@psych.ucsb.edu)**

Department of Psychology, University of California  
Santa Barbara, CA 93106 USA

## Abstract

This study analyzed the linguistic features found in spoken descriptions of two spatial layouts; participants' homes and the continent of Europe. Specifically, we were interested in whether these environments would be described from different perspectives by the same individuals. We hypothesized that the lexical features and discourse structure of the descriptions would reflect the perspective from which these two environments are most commonly learned—an embedded perspective in the home and an external perspective in the Europe descriptions. The results supported these hypotheses, suggesting that the choice of perspective in linguistic descriptions seems to be affected by the mode of acquisition and environmental features, rather than by a speaker's general preference for one perspective over another.

Keywords: Linguistics, Cognitive Science

When describing how points or landmarks are distributed throughout an environment, a speaker must decide on a perspective or vantage point from which the layout can be described. For example, a description of two adjacent rooms along a hallway can be described from a viewpoint within the environment or from a point external to the environment. A description such as, "You walk out of a doorway into the hallway, and the first door on your right will be the entrance to the next room," places an embodied agent within the environment and describes the layout with verbs of locomotion and terms that indicate landmarks' locations in relation to the agent. On the other hand, a description such as, "There is a hallway with a north-south orientation that has two adjacent rooms on the east wall," describes the same layout, but there is no sense of a person embedded within the environment. Instead, cardinal directions indicate the location of the landmarks, and static verbs mark the existence of the rooms and hallway. In this paper, taking a perspective from within the environment will be called an 'embedded' perspective, while viewing an environment from a point outside of the environment will be called an 'external' perspective.

Linde and Labov (1975) were the first to report these two different styles of describing spatial layouts. They asked participants to describe the apartments in which they resided, and found that two different manners of description

occurred—the "tour" description, in which the speaker seemed to move through the environment, giving a tour of the layout, and the "map" description, in which the speaker described the environment from a stationery point external to the layout. Although both description types were present in their data, embedded (i.e., tour) descriptions accounted for 97% of the data they collected. It seems, then, that there may be some features of an apartment environment that are particularly conducive to an embedded description.

In fact, many empirical studies have shown that aspects of the environment and the manner in which spatial information is acquired can affect the perspective that one uses to encode and describe the layout. For instance, the number of possible paths present in an environment has been found to influence written descriptions of the environment, such that layouts containing a single path are more likely described from an embedded perspective and environments that have multiple paths are usually described using an external perspective (Taylor & Tversky, 1996). Furthermore, Taylor and Tversky (1996) found that embedded descriptions occurred more often when the environment contained landmarks on a single size scale, whereas external descriptions were used to describe layouts that contained landmarks of various sizes. These results suggest that aspects of the environment affect the perspective from which the layout is described.

The manner in which information about an environment is acquired has also been found to have an effect on perspective. Experiment 3 of Taylor and Tversky's (1996) study analyzed the perspective from which students described two areas of the Stanford campus, as well as their home neighborhoods. These three layouts were environments that the students had learned via self-locomotion, and accordingly, the students tended to explain all three environments from an embedded perspective.

The mode of information acquisition has been found to affect the perspective taken not just in spatial descriptions, but spatial memories as well. In a study by Taylor, Naylor, and Chechile (1999), participants were asked to either study a map or navigate through an environment. The participants who studied the map were better able to answer questions tapping external perspective knowledge, whereas participants who directly navigated through the environment

were more accurate in measures testing embedded perspective. Taken together, these results suggest that both features of the environment and the mode of information acquisition have an effect on how perspective is chosen.

Many other findings, however, suggest that the choice of perspective in spatial descriptions may be influenced by factors not related to the environment. For instance, the language of a speaker may affect whether an environment is described from an embedded or external perspective. Emmorey, Tversky, and Taylor (2000) found that American Sign Language signers were more likely to use survey perspective to describe a convention center layout whereas English speakers preferred to use a route perspective to describe the same layout. Second, Taylor and Tversky's (1996, Experiment 2) data suggest that individual preferences may also affect perspective choice. In their experiment, 48 participants were given 4 maps to study and then describe, and of those participants, 27% described all four maps from a single perspective. This finding indicates that some individuals may prefer to describe a variety of spatial layouts from a single perspective.

Although the findings on individual differences reported by Taylor and Tversky (1996) suggest the possibility of preference for a certain perspective, their data must be interpreted with caution. The four environments they used in their experiment shared many of the same features, thus, we cannot be sure whether those shared features influenced participants' consistent use of perspective, or if individual preferences accounted for the consistency. Therefore, one of the main objectives of this study was to test whether descriptions of two very different environments varied within participants. If we find that the two descriptions exhibit different perspectives within single participants we may conclude that properties of the environments encourage different perspectives. However, if the participants use the same perspective to describe both environments, this may indicate that individuals have a certain preference for the perspective they take in spatial descriptions, regardless of the features of the environment.

The second objective of this study was to evaluate how speakers linguistically structure their descriptions of real-world environments learned outside of the laboratory from different experiences (maps vs. direct navigation). Because many previous studies on perspective have evaluated descriptions of layouts learned for the purpose of a laboratory study, we were interested in using environments that participants learned through different naturalistic experiences. Although some experiments have been conducted on naturalistic environments, such as an apartment or neighborhood (Linde & Labov, 1975; Taylor & Tversky, Experiment 3, 1996; Ullmer-Ehrich, 1982), they have focused on environments learned from direct navigation. Larger-scale environments learned via experiences outside the laboratory have received little attention.

Our third goal was to assess how perspective may be signaled on a larger discourse level. Previous research

evaluating linguistic features of perspective has focused on a lexical level of analysis (Emmorey, Tversky, & Taylor, 2000; Taylor & Tversky, 1996). However, differences at the discourse level would indicate that the overall organization, not just individual word choice, is influenced by the perspective from which a speaker describes an environment. Therefore, discourse analyses may be able to uncover how perspective choice affects higher-level linguistic planning, such as the organization of clauses.

In this experiment we chose to evaluate two environments that college-educated adults know fairly well—their childhood homes and the continent of Europe. These two environments differ drastically in the scale of space in which they exist, and also in the manner in which spatial information about the layout is acquired. According to Montello's (1993) classes of psychological space, the continent of Europe is considered geographical space, as it is too large to be apprehended through direct locomotion alone. Such spaces are typically apprehended from maps. A house, on the other hand, exists in environmental space and its spatial layout is able to be learned through direct locomotion.

## Method

### Procedure

Seventeen students from the University of California, Santa Barbara were asked to talk about the layouts of two environments—the house in which they were raised and the layout of Europe. Participants were specifically asked by the investigator, "Please describe the layout of (the house you grew up in/Europe) so that I will know where the most important places are." The order of elicitation for the two environments was counterbalanced and participants were given as much time as they needed to describe the environments. Participants were also asked to describe the layout of the UCSB University Center. In this task participants were instructed to describe the layout from either an embedded or external perspective. Due to length limitations, we are not able to include the analysis of this task in this paper. All spatial layout descriptions were videotaped.

After providing the layout descriptions, participants were asked to complete various measures of spatial visualization ability. They were first given the Paper Folding Test (Ekstrom et al., 1976), in order to measure spatial ability. Second, participants completed a questionnaire about their mental imagery preferences and cognitive style. Third, they completed the Santa Barbara Sense-of-Direction Scale (Hegarty et al., 2002), and finally the participants filled out a questionnaire asking about their familiarity with the environments they described.

### Language Coding

Participants' spoken responses were transcribed and the language was coded following Taylor and Tversky's (1996) language coding system. Relational terms, which are used to

locate a landmark in relation to another point, were coded as being either viewer-centered (e.g., *left, right*), environmental (e.g., *north, south*), or landmark-based (i.e., *below* France is Spain). Second, the type of referent used to specify the spatial relation of a newly mentioned landmark was coded as well. For example, a proposition such as, “The kitchen is on your left,” suggests an embodied agent<sup>1</sup> as the referent of the relation. A proposition such as, “And then you have England up in the north,” relates the location of the item to a cardinal direction, and a proposition that names another landmark as the referent (e.g., “On the other side of Belgium you have Switzerland”) was coded as ‘other landmark.’ Finally, the verbs of each proposition were coded according to Taylor and Tversky’s (1996) system. Only the copula (the verb ‘to be’) was considered to be stative, whereas active verbs were any verbs that described motion, including fictive motion (Talmy, 1996).

To augment Taylor and Tversky’s (1996) coding, we added two additional coding schemes. The first was a variation of Taylor and Tversky’s (1996) verb coding. In our system, four types of verbs were counted: static, including all static verbs such as ‘to have’; dynamic, which was restricted to all verbs that signal movement of an animate being; fictive, including all verbs used to describe fictive motion, as described by Talmy (1996), and see-verbs, which included verbs of perception such as ‘to see.’ The second coding system counted whether each clause specified a landmark’s relation to a physical point, to an embodied agent, or specified no relation. Because counting only specific word types regardless of their function brings some ambiguity into a language analysis, we adopted a coding system that evaluated the use of language on a clausal level. By doing this, we were able to evaluate whether a landmark’s location was specified in terms of the location of an agent or a physical point using the overall meaning of the clause rather than a single lexical item.

At the discourse level the number of choice points and action-existence chains were coded in each description. Choice points, according to Linde and Labov (1975), are junctions in the environment where a person can choose to move in more than one direction. For instance, a “T” intersection allows two choices, a turn to the right or a turn to the left. Choice points in our study were coded as any instance in which a speaker went back to a landmark previously mentioned and presented an alternative route. For example, after describing the upstairs floor of his house, one speaker said, “And then if you instead of going up the stairs, if you make a right, you’ll go into the kitchen.”

Action-existence chains were considered to be a series of two or more clauses in which the first clause(s) specify the action of an agent, and the final clause specifies the existence of a landmark (e.g., “If you turn right, there will be the dining room.”).

---

<sup>1</sup> We use the term ‘embodied agent’ in place of Taylor and Tversky’s (1996) ‘addressee.’

## Results

### Knowledge of the Environments

All 17 participants reported they had seen a map of Europe, and 12 participants had traveled to Europe at least once. On a scale of 1-7, with 7 being very good knowledge, participants’ mean rating of how well they knew European geography was 4.2 (range 2-6).

Three of the 17 participants reported having seen blueprints of their homes. When asked how vivid their memories of their homes were, all participants reported a 6 or 7 rating, indicating their memories were extremely clear. The mean rating was 6.8.

These data suggest that both environments were fairly well-learned, although information about them was acquired via different modalities. Participants’ knowledge of Europe was obtained mostly through map learning. Although a number of participants had traveled to Europe, the scale of the environment does not allow one to acquire full spatial knowledge from direct locomotion. The home environment, on the other hand, was learned via direct locomotion, although three participants had seen blueprints of their homes as well.

### Relational Terms

The number of environmental, viewer-relational, and landmark-based relational terms were counted for each description and then divided by the number of total relational terms found in each description. This yielded a proportion for each relational term. Proportions, rather than raw numbers, were evaluated to control for possible length differences across the two conditions. The proportions calculated for each term were then submitted to a paired samples t-test in order to assess whether the use of each term significantly varied across the two layouts<sup>2</sup>.

As Figure 1 illustrates, viewer-relational terms (*left, right, front, back*) were used significantly more often in the Home descriptions than in the Europe descriptions ( $M=37\%$  v.  $M=1\%$ ;  $t(16)=5.294$ ,  $p<.001$ ), whereas the Europe descriptions contained significantly more environmental terms ( $M=58\%$  v.  $M=1\%$ ;  $t(16)=8.025$ ,  $p<.001$ ). There was no significant difference in the use of landmark-based relational terms across the conditions.

### Verbs

The number of verbs considered stative and active, using Taylor and Tversky’s (1996) coding system, were counted for each description and then divided by the total number of verbs coded. The resulting proportions were not found to be significantly different across the two conditions. The top portion of Table 1 shows the mean percentage of stative and active verbs used in each of the two description types.

---

<sup>2</sup> For ease of discussion the proportions will be reported as percentages.

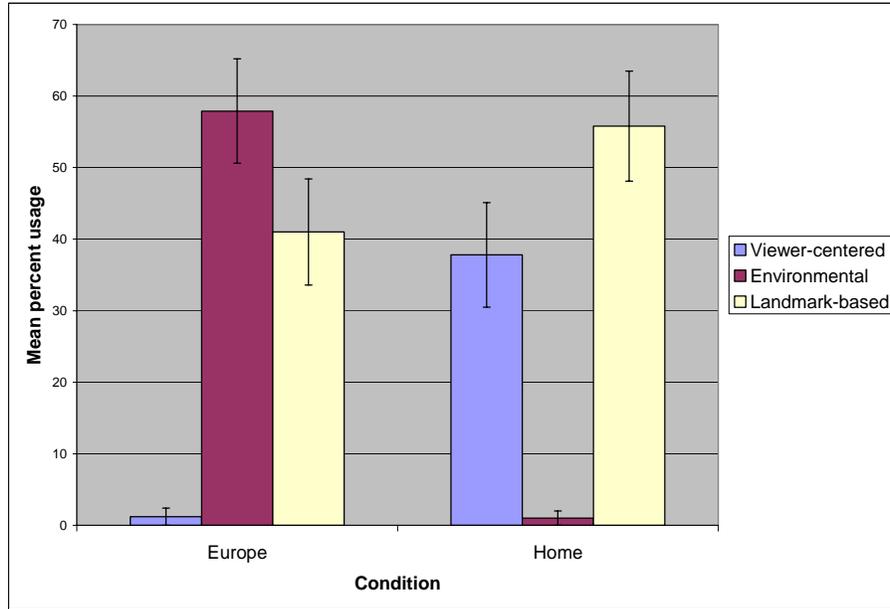


Figure 1: Mean percent usage of relational terms.

However, when active verbs were divided into dynamic motion of an agent and fictive motion, as they were in our coding scheme, a significant difference across the two conditions was found. As the bottom half of Table 1 shows, verbs that expressed dynamic motion of an agent were used more often in the Home descriptions ( $M=30\%$ ) than the Europe descriptions ( $M=12\%$ ;  $t(16)=2.573$ ,  $p=.020$ ). However, fictive motion was more prevalent in the Europe descriptions ( $M=17\%$ ), and used fairly infrequently in the Home descriptions ( $M=6\%$ ;  $t(16)=2.246$ ,  $p=.039$ ).

Table 1: Distribution of Verb Types.

	Europe	Home
Taylor & Tversky's Scheme		
Stative	62%	60%
Active	38%	40%
Kriz & Hegarty's Scheme		
Static	67%	63%
Dynamic*	12%	30%
Fictive*	17%	6%
See	4%	1%

\*significant at  $p < .05$

## Referents

The first mention of a landmark within each description was coded for the manner in which it was located within the environment, either by reference to a cardinal direction, an embodied agent, or in relation to another landmark. Again, each referent type was counted and divided by the total number of referent occurrences in the discourse, and these proportions were used to test for significant differences across the two conditions. T-tests revealed that the use of

reference by cardinal direction was significantly different across the two conditions. As Table 2 shows, specifying a landmark's relation to a cardinal direction occurred more often in the Europe descriptions ( $M=30\%$ ) than in the Home descriptions ( $M=1\%$ ;  $t(16)=4.000$ ,  $p=.001$ ). Conversely, using an embodied agent to specify the location of a landmark was more frequent in Home descriptions ( $M=38\%$ ) than in the Europe descriptions ( $M=4\%$ ;  $t(16)=4.788$ ,  $p<.001$ ). Relating the location of one landmark to that of another landmark within the environment was the most frequent method of specifying landmark locations, but the relative frequency of using another landmark as a referent did not differ between the two conditions.

Table 2: Distribution of Referents Used to Specify Location.

	Europe	Home
Taylor & Tversky's Scheme		
Other landmark	74%	61%
Embodied Agent *	4%	38%
Cardinal Direction *	22%	1%
Kriz & Hegarty's Scheme		
Physical Point *	53%	39%
Embodied Agent *	3%	20%

\*significant at  $p < .05$

We created a second method of coding the locations of landmarks specified within participants' layout descriptions. In this system, we coded the relationship specified in each clause as either relating a landmark's location to a physical point in the environment or to the position of an embodied agent. All clauses that did not specify a relation were not included in the analysis. The results closely patterned with

the differences described above. Fifty-three percent of the clauses in the Europe descriptions related the location of a landmark to a physical point within the environment, whereas only 39% of the clauses in the Home descriptions were physical point relations ( $t(16)=2.267, p=.038$ ). However, clauses relating a landmark's location to the position of an embodied agent were infrequent in the Europe descriptions ( $M=3\%$ ), but significantly more frequent in the Home descriptions ( $M=20\%$ ;  $t(16)=4.448, p<.001$ ). These differences are illustrated in Table 2.

## Discourse

As Table 3 illustrates, both of the discourse features were used significantly more often in the home descriptions than in the Europe descriptions. On average, speakers structured their home descriptions with more choice points than their Europe descriptions ( $M=.75$  v.  $M=0, t(15)=2.666, p=.018$ ). Their home descriptions also contained significantly more action-existence chains than the Europe descriptions ( $M=3.06$  v.  $M=.69, t(15)=4.173, p=.001$ ).

Table 3: Mean Number of Discourse Features by Description Type.

	Europe	Home
Choice points*	0	.75
Action-Existence Chains*	.69	3.06

\*significant at  $p<.05$

## Individual Differences

Neither the Paper Folding Test scores (Ekstrom et al., 1976) nor Sense of Direction measure (Hegarty et al., 2002), correlated with the frequency of any linguistic features measured in this study. Furthermore, participants' self-reported knowledge of European geography did not correlate with the use of any linguistic features in their Europe descriptions. Because the all participants reported having extremely vivid memories of the homes they described, correlations with linguistic output in the Home descriptions were not performed.

## Discussion

This study was designed to examine how people describe two environments that they have learned via different experiences outside of the laboratory. We were interested in determining whether these environments would be described using linguistic features that reflect the way in which the spatial information was acquired. While there have been previous studies evaluating the descriptions of people describing the homes in which they reside (Linde & Labov, 1975; Ullmer-Ehrich, 1982), very few studies have analyzed the descriptions of naturalistically-acquired geographic-scale spaces, such as the layout of countries on a continent.

One objective of this study was to evaluate whether speakers generally prefer to structure linguistic spatial descriptions from only one perspective. Many previous

studies have suggested that the type of environmental features and the manner of information acquisition may affect whether a speaker formulates an embedded or external description about a layout. However, because these experiments used a between-subjects design, they could not rule out the possibility that certain speakers may prefer a certain perspective. Thus, our study tested how the same speakers talked about these two different environments.

The results clearly indicate that speakers change their descriptions in order to convey an embedded perspective in the Home descriptions and an external perspective in the Europe descriptions. Many linguistic features were used significantly more often in the Home descriptions—dynamic verbs, viewer-centered relational terms, clauses relating a landmark to the location of an embodied agent, as well as larger discourse features such as choice points and action-existence chains. On the contrary, fictive motion, environmental terms, and clauses relating a landmark to a physical point in the environment were significantly more frequent in the Europe descriptions.

While the trend of linguistically differentiating between layouts was found to be present in the data, this does not suggest that all speakers gave external descriptions of Europe and embedded descriptions of their homes. In fact, two of the 17 participants of the participants structured their Europe descriptions from an embedded perspective. For both of these participants, dynamic verbs occurred frequently in their descriptions, and one of the participants used an embodied agent to specify location in 36% of the clauses in his Europe description. This number is extremely high, compared to the mean of 3% shown in Table 2. These data suggest that while most participants show a clear distinction between the two layouts, there are participants who chose an embedded perspective for both of their descriptions. In sum, we feel that while the general trend shows a preference towards differentiating between these two layouts, the participants who do not follow the trend should not be dismissed. Unfortunately, none of the data we collected can explain why these two participants chose an embedded perspective for both of their layout descriptions. Further studies researching this point more carefully may be able to assess whether a small number of participants rely on describing layouts from only one perspective.

A second objective of this study was to evaluate how speakers structure their descriptions of environments learned outside of the laboratory, in particular whether the different manners in which the information was acquired affected perspective choice. The results of the study suggest that participants clearly distinguished between the two environments in the linguistic devices they employed, and their descriptions mirrored the way in which they acquired information about the space, either by locomoting within the environment or by reading a map.

It is essential to note that the way in which spatial information was acquired cannot be the only factor that determines which perspective is used in spatial descriptions. Both Emmorey, et al. (2000) and Taylor and Tversky (1996) found that participants used route perspective in some environmental descriptions although they learned about these environments from maps. Therefore, we propose that

it is not the acquisition process, per se, that affects the linguistic structure of the description, but rather the speaker's assumption about how to best package information for an addressee. One way in which a speaker may assess an addressee's need for the information is by evaluating how the information is acquired in a non-linguistic learning situation.

In using naturalistic environments that were learned outside of the laboratory, we encountered some confounds that may make our results harder to interpret. Because the scale of space and mode of learning were confounded in this study, we cannot conclude that mode of acquisition alone determined the type of spatial perspective used. Features of the environment may promote a certain acquisition method, thus the two are not inseparable in studies conducted on learning "in the wild." In laboratory studies, the two factors can be untangled, and future laboratory studies may be able to provide more conclusive findings regarding the effect of mode of acquisition and environmental features on perspective choice.

Another shortcoming of this experiment is the difference in participants' familiarity with the two environments. These speakers, although very familiar with Europe, had much more exposure to their childhood homes. Thus, the differences found between the two types of descriptions may have been due to speakers' familiarity with the environment. However, the fact that participants' reported knowledge of European geography did not correlate with any of the linguistic features found in their descriptions suggests that familiarity of the environment did not influence how participants structured their descriptions. If the use of embedded perspective is simply a function of high familiarity, we would expect speakers with better knowledge of European geography very well to structure their descriptions in a more embedded manner. That was not the case.

Our final objective was to analyze how embedded and external descriptions are structured on a larger discourse level. Previous studies of perspective in spatial descriptions have only evaluated the use of language on a lexical level. By extending our analysis to discourse, we provide even stronger evidence that these two types of perspective are differentiated in linguistic descriptions. Moreover, the discourse organization of the two description types suggests that speakers structure their descriptions in a manner that mirrors the information acquisition process. Action-existence chains structure language so that information is presented as it was perceived, in a series of agent movements followed by noting the existence of a landmark. It is not surprising that these were common in the Home descriptions, but few were found in the Europe descriptions.

In comparison with previous studies, our data corroborate most of the linguistic findings. However, there was one difference between our data and descriptions from previous studies that should be brought to light. Landmark relational terms and the use of other landmarks as referents occurred frequently across both environment descriptions in our data. Taylor and Tversky (1996) reported that landmark relational terms and other landmark referents were used very rarely in the written descriptions they collected.

It is possible that the well-known nature of the environments caused speakers to use more landmark relational terms. Every culture has shared knowledge about what landmarks are expected to be included in a home or a continent. The reliance on relating one landmark to another may be due to the perceived ease in which the addressee is able to understand what types of landmarks are likely contained within the environment. The difference in the use of landmark relations between our study and previous studies may be due to the environments that were evaluated. However, the majority of our findings support previous work in this area, and the discourse analyses bring a novel contribution to the field of perspective in spatial descriptions.

### Acknowledgments

This research was funded by a University of California, Santa Barbara, Humanities and Social Sciences Research Grant awarded to Sarah Kriz.

### References

- Ekstrom, R., French, J., Harman, H., & Derman, D. (1976). *Kit of Factor-Referenced Cognitive Tests*. Princeton, NJ: Educational Testing Service.
- Emmorey, K., Tversky, B., & Taylor, H. (2000). Using space to describe space: Perspective in speech, sign, and gesture. *Spatial Cognition and Computation*, 2, 157-180.
- Hegarty, M., Richardson, A., Montello, D., Lovelace, K., & Subbiah, I. (2002). Development of a self-report measure of environmental spatial ability. *Intelligence*, 30, 425-447.
- Linde, C. & Labov, W. (1975). Spatial structures as a site for the study of language and thought. *Language*, 51, 924-939.
- Montello, D. (1993). Scale and multiple psychologies of space. In A. Frank & I. Campari (Eds.), *Spatial Information Theory: A theoretical basis for GIS*. Berlin: Springer-Verlag Lecture Notes in Computer Science 716.
- Talmy, L. (1996). Fictive motion in language and "caption." In P. Bloom, M. Peterson, L. Nadel, & M. Garrett (Eds.), *Space and language*. Cambridge, MA: MIT Press.
- Taylor, H., Naylor, S., & Chechile, N. (1999). Goal-specific influences on the representation of spatial perspective. *Memory & Cognition*, 27(2), 309-319.
- Taylor, H. & Tversky, B. (1996). Perspective in spatial descriptions. *Journal of Memory and Language*, 35, 371-391.
- Ullmer-Ehrich, V. (1982). The structure of living space descriptions. In R. J. Jarvella & W. Klein (Eds.), *Speech, place, and action* (pp. 219-249). New York: Wiley.