Cultural and Language Influences on the Interpretation of Spatial Prepositions

Mark M. HALL and Christopher B. JONES

Abstract

Culture and language can influence the generation and interpretation of spatial language, which would impact the quality of computational spatial language processing. This paper presents three human-subject experiments aimed at investigating these potential influences on the quantitative interpretations of five spatial prepositions. We show that for the languages (English and German) and cultures investigated (Europe and United States) neither language nor culture have a significant influence.

1 Introduction

Photographic images can almost invariably be regarded as having some geographical context and many images on the web have captions that include some spatial information. That information is usually in the form of spatial natural language that is essentially vague with no explicit geo-referencing. As a consequence, such images cannot be used in a geographic-information-retrieval (GIR) system and are also hard to retrieve using normal key-word based access when the containing place-name is not in the caption. Thus searching for “in London” will not find an image captioned “Near Big Ben”.

Automatically geo-referencing such location descriptions in image captions would enable these images to be indexed spatially and hence included in a GIR system. However such a process requires quantitative models to represent the spatial extent of the spatial prepositions that are frequently employed in image captions. In HALL & JONES 2011 we describe an experiment to acquire that type of quantitative data and a system for automatically geo-referencing location descriptions in image captions.

An issue faced by such a system is that when processing data from the web, the author's first language and culture are in most cases unknown, but might have an impact on the quantitative properties of the cognitive models that they use when deciding upon the appropriate spatial preposition. This paper presents two further experiments aimed at understanding to what degree language and culture influence the quantitative models of vague spatial prepositions.

2 Background

Spatial language as discussed in this paper consists of three basic building blocks. These are *objects* (usually the image's subject), *places* the *object* is in or proximal to, and *spatial* *prepositions* that relate the *objects* to the *places* (LANDAU & JACKENDOFF 1993). This paper will use TALMY's 1983 terminology of *figure* and *ground* for the *object* and *place* linked by the *spatial preposition*. The *figure* is defined as the *object* that is located relative to the *ground object* or *place*. In the phrase “statue near Big Ben”, “statue” is the *figure*, “Big Ben” the *ground* and they are related via the *spatial preposition* “in”.

The difficulty with processing spatial prepositions computationally derives from the dichotomy between the spatial world, which is continuous and quantitative, and the linguistic world which uses a qualitative representation. Because the number of potential spatial configurations is so large and the number of spatial prepositions so small, the spatial prepositions that we use in practice are very flexible in how they can be applied and thus the book can be ``near'' the glass and Reading can also be ``near'' London. The result of this flexibility is that the boundary where something stops being ``near'' becomes vague and heavily dependent on the context involved (see FISHER 2000 or PARSONS 1996). A number of models have been developed to handle this vagueness computationally (EGENHOFER 1991, RANDELL ET AL. 1992, SCHNEIDER 2000, BENNET 2001, LIU ET AL. 2008).

When these models are instantiated with data drawn from people, then the question arises of how strongly language and culture influence spatial cognition. This has been of interest for a long time and evidence is available to support both strong and weak links between the two. KEMMERER & TRANNEL 2000 provide neurological evidence for a weak link, by showing that patients with certain brain lesions can either use spatial language correctly or reason spatially, but not both. On the other hand LEVINSON 2003 describes Aboriginal languages that provide only an absolute reference system (north/east/south/west) and to successfully express spatial information in such a language the brain's spatial cognition system must at all times maintain a complete map of all objects, which is not required in languages that provide other reference systems (KUIPERS 1978). Further evidence for strong links are provided by KLIPPEL & MONTELLO 2007 and MARK ET AL. 2007, while evidence for weaker links comes from LI & GLEITMAN 2002. In the context of general language use STUTTERHEIM 2003 shows that even fluent second-language speakers retain patterns from their first language. This paper will add to this body of knowledge by investigating potential influences of language on quantitative models for spatial prepositions.

Cultural influences on spatial cognition have also been investigated, with FRIEDMAN ET AL. 2005 and XIA & LIU 2007 showing that culture has no significant influence on large-scale distance estimation. Similarly RAGNI ET AL. 2007 investigate topological reasoning and also find no significant differences. None of these have investigated impacts on the use of spatial language and this paper will provide additional knowledge in this area.

3 Experiments

To understand potential language and cultural issues we compare the results of an experiment conducted in the United Kingdom (UK), Austria (AT), and the United States (US). The UK experiment acts as the baseline, with the AT experiment, conducted in German, investigating potential language influences and the US experiment cultural influences. To ensure comparability the experimental setup was the same for all three experiments, with two small exceptions in US experiment.

3.1 Setup

The questionnaire consisted of eight spatial language questions, each accompanied by a black and white map, printed on A4 paper (fig. 1). The pages were held in a ring-binder, so that at each point the question and answer area were displayed on the right hand side, while the map was visible on the left hand side. This guaranteed that the participants only worked on one question at a time and that they were always presented with the correct map for each question. The inclusion of a map was necessitated by the use of very local toponyms, where participant familiarity with the region could not be assumed. While this has an effect on the results (LINDEN & SHEEHY 2004), it was judged that this effect was less problematic than the errors introduced by people guessing where the places were if no map was provided.

Each question consisted of instructions, a primer phrase and a list of answers. The primer phrase was of the structure ``This photo was taken in ___________, which is <spatial preposition> <ground toponym>''. The context of labelling a photo was explicitly mentioned, as the aim is to apply the results to the localisation and generation of image captions and the context influences the interpretation of the spatial relations. Below the primer phrase a list of candidate toponyms for the blank space in the primer phrase was presented. For each toponym an interval rating scale from 1 to 9 was shown, with 1 indicating the toponym did not fit for the primer phrase and 9 a perfect fit. A large number of rating intervals was chosen to elicit detailed results, as a reduction to fewer intervals is always possible during the analysis process (MATELL AND JACOBY 1971). Toponyms and ratings formed a table with toponyms on the vertical axis and the rating scale on the horizontal axis. The toponyms were ordered alphabetically to avoid any ordering bias. The instructions asked the participants to rate how well each of the candidate toponyms would fit into the phrase, if the phrase was used in the spatial configuration shown in the map.

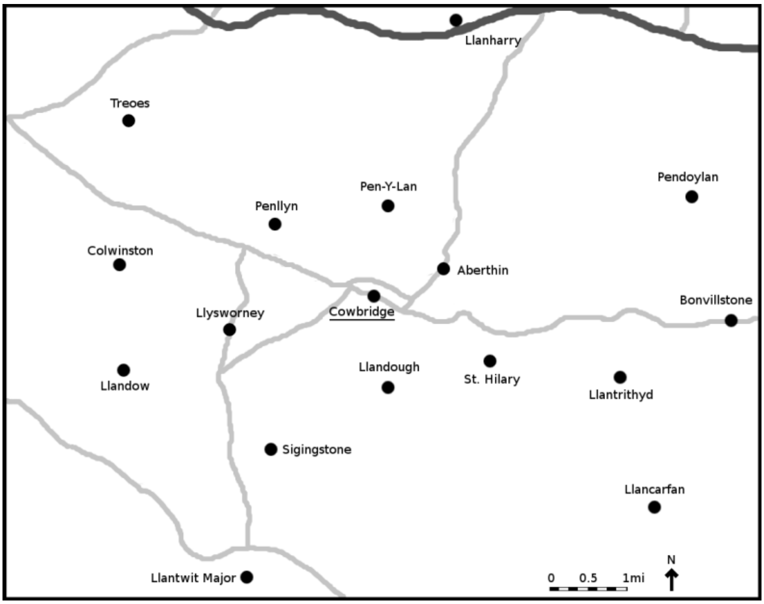

Fig. 1: Map shown to the experiment participants for all five spatial prepositions

On the map all candidate toponyms and the *ground* toponym were displayed as points (fig. 1). The reason for this is that the majority of the available geo-data that will be used to interpret spatial language is point-based, it is thus important that the data acquired on the use of spatial prepositions is also grounded in a point-based representation. Additionally it is known that the relative sizes of the *ground* and *figure* objects influences the interpretation of the spatial preposition, especially of ``near'' (MORROW & CLARK 1988, LUNDBERG & EKMAN 1973). Using a point-based representation normalises the toponyms' sizes, removing that influence from the result. Major roads were also included on the map. At the bottom a scale bar and north arrow were displayed. Participants were explicitly instructed to consult the map when performing the tasks. The maps were printed in grey-scale and while line thickness and strength influence similarity judgements in connected graphs (FABRIKANT ET AL. 2004), the analysis shows no such influence in the results of this experiment.

For the AT experiment the instructions, spatial prepositions and primer phrases were translated into German. No changes were made for the US experiment.

3.2 Participants

In the UK and AT experiments participants were recruited from staff and students at Cardiff and Klagenfurt University respectively. In the US participants were selected by an American colleague from colleagues and friends. Across all three experiments the majority of participants were in the 18 – 25 age-group, although both the AT and US experiments had significant groups of older participants. χ^2 tests were used to ensure that the results are statistically significant, but are not reported in detail here (see HALL & JONES 2011 for details). The UK experiment was conducted with 24 participants, the AT experiment with 35, and the US experiment had 49 participants.

3.3 Procedure

Participants were each handed a ring-binder with the questionnaire. They were then verbally instructed to read the instructions in the questionnaire and then, following the instructions, to work carefully through the actual questionnaire. The instructions explained the format of the questionnaire and that the participants were not allowed to consult their answers to previous questions. After completing the core questions, participants were asked for their age, gender and also to rate their familiarity with the regions involved in the questions (also on a scale from 1 to 9). They then returned the questionnaires, were provided with information on the experiment and received payment for their participation.

The experiments were all conducted indoors as RAGHUBIR & KRISHNA 1996 show that in a map based test, whether the test is administered in a lab environment or in the field has no influence on the results.

There was a variation in the US experiment, as participants filled out the questionnaire at a time and location of their choosing, which again might impact the experimental results.

3.4 Results

The basic results for the individual test points on the map are shown in tables 1 and 2, and a detailed analysis of the UK results can be found in HALL & JONES 2011. Here only the core conclusions will be repeated to enable the comparison of the experiments.

Table 1: Median values and inter-quartile ranges for the spatial prepositions "near", "north", and "east". Values are formatted "median / i-q range".

	Near			**North**			**East**		
	UK	**AT**	**US**	**UK**	**AT**	**US**	**UK**	**AT**	**US**
Aberthin	9 / 0.25	9 / 0	9 / 0	6 / 1.5	6 / 1	6 / 2	8 / 1	8 / 1	8 / 2
Bonvillstone	3 / 3.5	2 / 2	1 / 1.25	1 / 1.5	1 / 0	1 / 0	9 / 0	9 / 0	9 / 0
Colwinston	5 / 3	4 / 2.5	2 / 4	1 / 2.5	2 / 2.5	3 / 4	1 / 0	1 / 0	1 / 0
Llancarfan	2 / 2.5	1 / 1	1 / 1	1 / 0	1 / 0	1 / 0	7 / 2	6 / 2	5 / 4
Llandough	8 / 1	8 / 2	8.5 / 1	1 / 0	1 / 0	1 / 0	2 / 3.5	2 / 3.5	2 / 3
Llandow	5.5 / 3	5 / 2	3 / 3	1 / 0.5	1 / 0.5	1 / 0	1 / 0	1 / 0	1 / 0
Llanharry	2 / 2.5	1 / 2	1 / 1	8.5 / 1	9 / 1	8 / 2	2 / 3.5	2 / 3	2 / 4
Llantrithyd	5 / 3	3 / 3	3 / 3	1 / 0	1 / 0	1 / 0	8 / 2	8 / 1	7 / 2
Llantwit Major	2 / 2.5	1 / 1	1 / 1	1 / 0	1 / 0	1 / 0	1 / 0	1 / 0	1 / 0
Llysworney	8 / 2	7 / 1	7 / 2	1 / 1	1 / 0	1 / 0	1 / 0	1 / 0	1 / 0
Pendoylan	3 / 3	2 / 2	2 / 1	5 / 3	5 / 3	4 / 2	7 / 2	8 / 1.5	7 / 2
Penllyn	7.5 / 2	7 / 1	7 / 3	6 / 2	6 / 2.5	6 / 2	1 / 0	1 / 0	1 / 0
Pen-Y-Lan	8 / 2	8 / 1	8 / 1	9 / 0	9 / 0	9 / 0	2 / 3	2 / 4	2 / 4
Sigingstone	5 / 3	6 / 3	5 / 2	1 / 0	1 / 0	1 / 0	1 / 0	1 / 0	1 / 0
St Hilary	8 / 3	7 / 1	7 / 2	1 / 0.5	1 / 0	1 / 0	7 / 1	8 / 2	7 / 2
Treoes	3 / 2	2 / 2.5	1.5 / 2	5 / 2	5 / 2.5	5 / 4	1 / 0	1 / 0	1 / 0

For “near” the primary factor influencing participants' ratings is distance and neither angle nor road connectivity have a statistically significant influence. For the cardinal directions angle is the primary factor and, as with “near”, no other factors (distance, road connectivity) show a statistically significant influence.

The results of the three experiments have been compared on a point-by-point basis using a χ^2 test. Those points where the answer distributions are statistically significantly different between the experiments are listed in table 3.

Table 2: Median values and inter-quartile ranges for the spatial prepositions "south" and "west". Values are formatted "median / i-q range".

	South			West		
	UK	**AT**	**US**	**UK**	**AT**	**US**
Aberthin	1 / 0	1 / 0	1 / 0	1 / 0	1 / 0	1 / 0
Bonvillstone	1 / 2	1 / 2	2 / 2.5	1 / 0	1 / 0	1 / 0
Colwinston	1 / 0	1 / 0	1 / 0	9 / 1	9 / 1	9 / 1
Llancarfan	5 / 6	5 / 3	5 / 3	1 / 0	1 / 0	1 / 0
Llandough	9 / 0	9 / 0	9 / 0	1 / 2	1 / 1.5	1 / 0
Llandow	5 / 4	5 / 3	4.5 / 3	8 / 1	8 / 1.5	8 / 2
Llanharry	1 / 0	1 / 0	1 / 0	1 / 0	1 / 0	1 / 0
Llantrithyd	5 / 2.25	3.5 / 3	3 / 4	1 / 0	1 / 0	1 / 0
Llantwit Major	8 / 1	8 / 2	8 / 2	3 / 3.5	4 / 4	3 / 5
Llysworney	3 / 3.25	2 / 3	3 / 3	9 / 0	9 / 0.5	9 / 0
Pendoylan	1 / 0	1 / 0	1 / 0	1 / 0	1 / 0	1 / 0
Penllyn	1 / 0	1 / 0	1 / 0	7 / 2	7 / 2	7 / 2
Pen-Y-Lan	1 / 0	1 / 0	1 / 0	1 / 1	1 / 2	1 / 0
Sigingstone	8 / 1	8 / 1	8 / 1	5.5 / 3	5 / 3	5 / 5
St Hilary	7 / 1.5	6 / 2	6 / 3	1 / 0	1 / 0	1 / 0
Treoes	1 / 0	1 / 0	1 / 0	7 / 2	7 / 1	6 / 2

To ensure that the results are not based solely on geometric reasoning on the map, they have been compared to a number of models. For “near” three models were tested:

- *banded*: A broad-boundary model splitting the space evenly into concentric discs “near” (predicted value - 9), “somehow near” (predicted value - 5), and “not-near” (predicted value - 1) spaces;
- *road-distance*: Distance from the *ground* place via the local road network. Values are scaled linearly between 9 for the closest point and 1 for the most distant point;

- *distance*: Linear distance from the ground toponym. Values are scaled linearly between 9 for the closest point and 1 for the most distant point.

For the cardinal directions only “north” and “south” were tested with two models

- *banded*: A broad-boundary model consisting of the centre-cone +/- 45° from the primary direction (predicted value - 9), the area between +/- 45° and +/- 90° from the primary direction (predicted value – 5), and the area outside of +/- 90° from the primary direction (predicted value – 1);
- *angle*: Angles between 0° and +/-90° are given an angle linearly scaled from 9 at 0° to 1 at +/-90°. Angles above +/-90° are given a predicted value of 1.

Table 3: Measurement points where a χ^2 test shows a statistically significant difference between the experiments.

	UK - AT	**UK - US**	**AT - US**
Near	Llancarfan ($p < 0.05$) Llantrythid ($p < 0.05$)	Bonvillstone ($p < 0.01$) Colwinston ($p < 0.01$) Llancarfan ($p < 0.01$) Llandow ($p < 0.05$) Llantrythid ($p < 0.01$) Treoes ($p < 0.01$)	Treoes ($p < 0.05$)
North	-	Llandow ($p < 0.05$) Llysworney ($p < 0.05$)	Llandow ($p < 0.05$)
East	-	-	Aberthin ($p < 0.05$)
South	-	-	-
West	-	-	Pen-Y-Lan ($p < 0.05$)

The values predicted by the models for each point are compared to the experimental results using the median-absolute-deviation (MAD) between the model prediction and the experimental results. A model with a *good fit* is defined as a model where all MADs ≤ 2 and the sum of the MADs for all points ≤ 12. With sixteen prediction points per model, a *good fit* model would have at most six points (approximately one-third) with a MAD of 2 (the maximum differences allowed for a *good fit* model) or at most 12 points (approximately three-quarters) with a MAD of 1 (indicating only minimal difference).

Evaluating all models using this metric shows that none of them provide a good fit for the observed results. In the case of “near” all models violate the maximum total MAD value constraint, while all cardinal direction models violate the maximum individual MAD constrain. To enable some conclusions to be drawn from the models, we define the *best fit* model as the model with the lowest total MAD for that spatial preposition (tab. 4).

Table 4: *Best fit* models for the spatial prepositions "near", "north", and "south".

	UK	**AT**	**US**
Near	road-distance	distance	banded
North	banded	banded	Banded
South	angle	angle	angle

4 Discussion

Across all three experiments the large inter-quartile (i-q) ranges for most points that are located between the areas where the spatial preposition definitely holds and where it definitely does not hold, strengthen the conclusion drawn from the UK experiment that the quantitative interpretations are very personal. This is enforced by the model testing, which shows that due to the large amount of variation in the answers no model fits the data well.

The second major conclusion is that language has no significant influence on the quantitative interpretations of the cardinal directions. For "near" there is a small difference in that there are two points that are statistically different and the *best fit* models differ between the UK and AT experiments. However, as the two points are in a part of the map with few roads indicated and the *best fit* model is *distance* instead of *road-distance*, we conclude that the differences between the two data-sets are most likely due to the lack of local knowledge on the part of the AT participants and not a language influence.

The cultural influence shows a similar pattern. No influence is found for any of the cardinal directions, but for "near" there is significant difference between the UK and US results for six of the 16 points. While this could indicate that there are cultural influences on "near", it is more likely that as with the AT results, the lack of local knowledge is the primary factor causing the differences. This is strongly supported by the fact that the differences are not consistent. For example Bonvillstone is significantly different, but Pendoylan is not, although they are both a similar distance from the centre-point and in the UK results their median values are the same and i-q ranges differ only minimally. Similarly Llandow is significantly different, but Sigingstone is not. Thus either the cultural influence is very nuanced, which given the participants reported lack of local knowledge is very unlikely, or the difference is primarily caused by the lack of local knowledge.

The results indicate that the lack of local knowledge leads to a simpler interpretation of the spatial preposition, as indicated by the fact that the *best fit* models are simpler for the AT and US results (*distance* and *banded* respectively). Also for both the AT and US results the i-q ranges are lower for half the measurement points, implying a greater amount of agreement between participants, which is most likely caused by the participants falling back to the simpler models that do not take local knowledge into account.

One caveat is that it is also possible that the UK – US differences are solely due to the different participant selection process for the US experiment, however due to the differences not being that large we believe this to be unlikely.

Interestingly the lack of local knowledge has no significant influence on the cardinal directions. This is slightly surprising as the original analysis showed that amongst the UK participants higher local knowledge led to slightly different interpretations. The relatively small number of UK participants did not make it possible to determine whether these differences were significant, but combining the results with the two other experiments clearly shows that the changes caused by local knowledge are minimal and lie within the normal variation between participants irrespective of their local knowledge.

5 Conclusion

This paper described three experiments conducted to investigate potential influences of language and culture on quantitative interpretations of five vague spatial prepositions ("near", "north", "east", "south", and "west"). The experiments were conducted in the United Kingdom as a baseline, in Austria, translated into German, to investigate language influences, and in the United States to investigate cultural influences.

The results show that neither language nor cultural factors have a statistically significant influence on "near" or the cardinal directions. However, we find that local knowledge influences the results for "near", but not for the cardinal directions. The lack of local knowledge causes the participants to tend towards using simpler interpretations for "near", although there is still a lot of variation in how they apply these simpler interpretations that cannot be explained by simple geometric models.

For automated geo-referencing the results are positive, as they indicate that at least for the investigated cultural sphere the geo-referencing algorithms do not have to take into account language or cultural influences, although further studies are necessary to validate that this is true for a larger set of languages and cultures. The strong influence of local knowledge is however something that needs to be incorporated and further study is necessary to test how such information should be integrated into the geo-referencing process. Also other potential factors such as task context and technological socialisation should be investigated.

References

BENNET, B. (2001), Application of Supervaluation Semantics to Vaguely Defined Spatial Concepts Spatial Information Theory. Foundations of Geographic Information Science : International Conference, COSIT 2001, 108-123.

EGENHOFER, M. (1991), Reasoning about Binary Topological Relations. Second Symposium on Large Spatial Databases, Springer-Verlag, 525, 143-160.

FABRIKANT, S., MONTELLO, D., RUOCCO, M. & MIDDLETON, R. (2004), The Distance-Similarity Metaphor in Network-Display. Spatializations Cartography and Geographic Information Science, 31, 237-252.

FISHER, P. (2000), Sorites paradox and vague geographies. Fuzzy Sets and Systems, 113, 7-18.

FRIEDMAN, A., KERKMAN, D., BROWN, N., STEA, D. & CAPPELLO, H. (2002), Cross-cultural similarities and differences in North Americans' geographic location judgments. Bulletin & Review, 9, 1054-1060.

HALL, M.M., JONES, C.B. (2011), Interpreting Spatial Language in Image Captions. Cognitive Processing 12 (1), 67-94.

KEMMERER, D. & TRANEL, D. (2000), A Double Dissociation between Linguistic and Perceptual Representations of Spatial Relationships. Cognitive Neuropsychology, 17, 393-414.

KLIPPEL, A. & MONTELLO, D. (2007), Linguistic and Nonlinguistic Turn Direction Concepts. Spatial Information Theory, 8th International Conference, COSIT 2007, 354-372.

KUIPERS, B. (1978), Modeling Spatial Knowledge. Cognitive Science: A Multidisciplinary Journal, 2, 129-153.

LANDAU, B. & JACKENDOFF, R. (1993), ``What'' and ``where'' in spatial language and spatial cognition. Behavioral and Brain Sciences, 16, 217-238.

LEVINSON, S. (2002), Space in language and cognition: Explorations in cognitive diversity. Cambridge: CUP.

LI, P. & GLEITMAN, L. (2002), Turning the tables: language and spatial reasoning. Cognition, 83, 265-294.

LINDEN, M. & SHEEHY, N. (2004), Comparison of a Verbal Questionnaire and Map in Eliciting Environmental Perceptions. Environment and Behavior, 36, 32-40.

LIU, Y.; GOODCHILD, M.; GUO, Q.; TIAN, Y. & WU, L. (2008), Towards a General Field model and its order in GIS. International Journal of Geographical Information Science, 22, 623-643

LUNDBERG, U. & EKMAN, G. (1973) Subjective Geographic Distance: A Multidimensional Comparison. Psychometrika, 38, 113-122.

MARK, D., TURK, A. & STEA, D. (2007), Progress on Yindjibarndi Ethnophysiography. Spatial Information Theory, 8th International Conference, COSIT 2007, 1-19.

MATELL, M. & JACOBY, J. (1971), Is There an Optimal Number of Alternatives for Likert Scale Items? Study 1: Reliability and Validity. Educational and Psychological Measurement, 31, 657-674.

MORROW, D. & CLARK, H. (1988), Interpreting words in spatial descriptions. Language and Cognitive Processes, 3, 275-291.

PARSONS, S. (1996), Current Approaches to Handling Imperfect Information in Data and Knowledge Bases. Knowledge and Data Engineering, 3, 353-372.

RAGHUBIR, P. & KRISHNA, A. (1996), As the Crow Flies: Bias in Consumers' Map-Based Distance Judgments. Journal of Consumer Research, 23, 26-39.

RAGNI, M.; TSEDEN, B. & KNAUFF, M. (2007), Cross-Cultural Similarities in Topological Reasoning. Spatial Information Theory, 8th International Conference, COSIT 2007, 32-46.

RANDELL, D., CUI, Z. & COHN, A. (1992), A Spatial Logic Based on Regions and Connection. KR'92 Principles of Knowledge Representation and Reasoning, 165-176.

SCHNEIDER, M. (2000), Finite Resolution Crisp and Fuzzy Spatial Objects. International Symposium on Spatial Data Handling, 3-17.

STUTTERHEIM, C. (2003), Linguistic structure and information organisation: The case of very advanced learners. EUROSLA Yearbook 3, 183-206.

TALMY, L. (1983), How Language Structures Space Spatial Orientation, New York: Plenum, 225-282.

XIAO, D. & LIU, Y. (2007), Study of Cultural Impacts on Location Judgments in Eastern China. Spatial Information Theory, 8th International Conference, COSIT 2007, 20-31.