

Generating and Interpreting Referring Expressions with Vague Spatial Language

Adam Richard-Bollans¹[0000-0003-1980-0107]

University of Leeds, UK
mm15alrb@leeds.ac.uk

Keywords: Lexical Semantics · Spatial Prepositions · Situated Dialogue.

1 Introduction

Spatial prepositions are ubiquitous in situated dialogue and can be used to denote a vast array of configurations which greatly diverge from any typical meaning. There is much discussion on the semantics of spatial prepositions regarding how their meanings are shaped and understood. Despite the great interest in the topic, there is not a substantial body of data that includes the myriad of potentially salient features that are discussed. Moreover, there is much work relating to situated dialogue and the challenges of referring expression generation and comprehension, however most of this avoids expressions that are either *relational* or *vague*.

Our work is aimed initially at acquiring data on spatial preposition usage that allows for rich feature extraction and therefore a more detailed analysis of the semantics and pragmatics of spatial language usage. In order to do this we are setting up an annotation environment containing multiple tasks. Our current proposal is based on an earlier preliminary study, which we discuss in [10].

2 Related Work

Initial attempts to understand and model spatial language naturally focused heavily on geometry. However, as has been recognized in the past couple of decades, spatial constraints are not enough to fully characterize spatial prepositions [1, 2, 5, 6]. The use of prepositions is determined by geometric, functional *and* conventional considerations, as evidenced in [1, 3, 5].

Following [9], we therefore believe that a full semantic account of spatial prepositions, particularly for *on* and *in*, ought to include distinct functional & conventional components in order to (1) be closely aligned with human usage and understanding and (2) aid automated interpretation and generation of spatial expressions.

In order to address this we first require data that relates spatial prepositions and the various features that influence their use. There exist some wide-ranging annotated datasets, see [4, 12], however the annotations are not restricted to

spatial language and many of the datasets are image-based so extracting meaningful features is extremely difficult. Many experimental studies have been conducted over the past couple of decades into particular aspects of spatial prepositions; however these are either image-based such as [7], or in real environments such as [11] where meaningful feature extraction is difficult; or in very constrained/simple environments such as [8].

Overall, we find have found that there is a lack of detailed geometric, functional and contextual data which hinders the capacity to properly investigate the semantic complexity of spatial prepositions and provide pragmatic analysis on how they are used to achieve communicative success.

3 Approach

The overall aim of our research is to create a system which is able to generate and interpret natural and effective spatial expressions in complex environments. We are concerned primarily with the English language but hope that some of our methodology and analysis will be transferable to other languages.

We recently conducted a preliminary study with a basic annotation framework¹, which we discuss in [10]. This study comprised virtual environments in which participants are given a first person view of a scene which they can navigate using a mouse and keyboard. In these environments we set up two distinct tasks — a Selection Task and a Description Task. The Selection Task was designed to efficiently collect large amounts of data regarding the semantics, with minimal pragmatic considerations. Participants were given a preposition on screen and asked to select all figure-ground pairs in a given scene which fit the preposition. The Description Task aimed to provide more focused data to aid pragmatic analysis and test models of figure selection. In this task objects were highlighted and participants were able to type in a spatial description of the object. We asked participants to give descriptions of the object locations using a definite description, in the format *figure + preposition + ground* e.g. ‘the guitar by the bookshelf’.

Given the results of this initial study we are developing and expanding our annotation framework to collect more meaningful data. We hope to improve the possibility of semantic analysis by focusing the provided annotations and asking participants to categorise and compare specific configurations rather than freely select configurations. This will provide data for the refinement and testing of a semantic model. The comparisons in particular will give us information regarding the *typicality* of spatial prepositions, which will be a central notion in our dialogue system.

The semantic model will be incorporated into a prototype dialogue system which will be used to generate and interpret written spatial expressions. Our framework will also include a game environment where users will have to communicate with our dialogue system in order to complete a simple task, e.g. collecting

¹ <https://github.com/alrichardbollans/spatial-preposition-annotation-tool-blender>

objects in a scene, thus providing an environment for testing the communicative success of the dialogue model.

4 Further Work

Building on our preliminary study, we are improving on the methodology to overcome some of the issues that arose. We hope the new set up, which is currently online², will allow a more meaningful and robust analysis. On top of this we will build a prototype implementation of a dialogue system into the data collection, both as a means to collect data on and inform pragmatic strategies and also to allow a ‘gamification’ of the data collection process. Moreover, we hope that our data collection platform can be used more widely by other researchers in the field.

References

1. Coventry, K.R., Carmichael, R., Garrod, S.C.: Spatial prepositions, object-specific function, and task requirements. *Journal of Semantics* **11**(4), 289–309 (1994)
2. Coventry, K.R., Garrod, S.C.: Saying, seeing and acting: The psychological semantics of spatial prepositions. Psychology Press (2004)
3. Feist, M.I., Gentner, D.: Factors Involved in the Use of In and On. *Proc Annual Meeting of the Cognitive Science Society* p. 7 (2003)
4. Ferraro, F., Mostafazadeh, N., Ting-Hao, Huang, Vanderwende, L., Devlin, J., Galley, M., Mitchell, M.: A Survey of Current Datasets for Vision and Language Research. arXiv:1506.06833 [cs] (2015)
5. Garrod, S., Ferrier, G., Campbell, S.: In and on: investigating the functional geometry of spatial prepositions. *Cognition* **72**(2), 167–189 (Sep 1999)
6. Herskovits, A.: *Language and spatial cognition*. Cambridge University Press (1987)
7. Kelleher, J., Sloan, C., Mac Namee, B.: An investigation into the semantics of English topological prepositions. *Cognitive Processing* **10**(S2), 233–236 (Sep 2009)
8. Platonov, G., Schubert, L.: Computational Models for Spatial Prepositions. In: *Proc 1st International Workshop on Spatial Language Understanding*. pp. 21–30 (2018)
9. Richard-Bollans, A.: Towards a Cognitive Model of the Semantics of Spatial Prepositions. In: *ESSLLI Student Session Proceedings*. Springer (2018)
10. Richard-Bollans, A., Gómez Álvarez, L., Bennett, B., Cohn, A.G.: Investigating the Dimensions of Spatial Language (Under Review) (2019)
11. Tenbrink, T., Andonova, E., Schole, G., Coventry, K.R.: Communicative Success in Spatial Dialogue: The Impact of Functional Features and Dialogue Strategies. *Language and Speech* **60**(2), 318–329 (2017)
12. Wu, Q., Teney, D., Wang, P., Shen, C., Dick, A., Hengel, A.v.d.: Visual Question Answering: A Survey of Methods and Datasets. arXiv:1607.05910 [cs] (2016)

² See http://adamrichard-bollans.co.uk/spatial_language_project.html