Towards an automatic identification of functional and geometric spatial prepositions

Simon Dobnik (simon.dobnik@gu.se)
Department of Philosophy, Linguistics and Theory of Science
University of Gothenburg, Box 200, 405 30 Göteborg, Sweden

John D. Kelleher (john.d.kelleher@dit.ie)
School of Computing, Dublin Institute of Technology
Kevin Street, Dublin 8, Ireland

Abstract

We examine how corpus-based techniques commonly used to identify collocations and collostructions can be used to distinguish between spatial descriptions that are sensitive to geometric and functional meaning components and whether prepositional functional knowledge can be recovered from instances of preposition use from text corpora.

Keywords: spatial descriptions; geometric; functional; text corpus analysis; entropy; log-likelihood ratio

Introduction

Spatial descriptions with prepositions such as 'The flowers are in the vase' are not only referring to geometric properties of the scene - the arrangement of objects related by the preposition in the Euclidean space - but are also sensitive to some aspects of our world knowledge about these objects, for example the way they interact with each other (Herskovits, 1986). Thus, it is normally not the case that the flowers are entirely contained in the vase but only their stalks are. Furthermore, the contribution of the geometrical and world knowledge component to the meaning of spatial descriptions in not equal with every preposition. For example, (Coventry, Prat-Sala, & Richards, 2001) show in the experiments with human observers of images of a man in the rain holding an umbrella where the umbrella is providing a varying degree of protection from the rain that 'above' is more sensitive to the geometrical component than 'over' and 'over' is more sensitive to the object function component than 'above'. Descriptions of 'the umbrella is over a man' were considered acceptable even in cases where the umbrella was held horizontally but was providing protection from the rain.

Most previous work on generating spatial descriptions grounds the descriptions in contextual models that are either assumed a-priori (Dale & Haddock, 1991) or that are generated from sensor-data (Regier & Carlson, 2001; Dobnik, 2009; J. D. Kelleher & Costello, 2009) using functions that focus solely on the geometric properties of the scene and ignore the functional component of spatial semantics. This is primarily because (i) it is not always straightforward to identify components of functional knowledge a preposition is sensitive to (Garrod, Ferrier, & Campbell, 1999); and (ii) it is challenging to automatically recover and represent such knowledge for every object that we would relate in a description. We would have to develop a complex ontology of object properties and then associate spatial prepositions (either manually or automatically) with rules (see for example (Garrod et al., 1999, p.170)) that pick out certain properties of objects that are relevant in their interaction or for the relation defined by the preposition. In this paper we take a different, less expensive approach. Assuming that a preposition selects for some functional semantic component and that both arguments of this preposition are sensitive to this component we should be able to capture it probabilistically by examining examples of description use in a corpus of text without precisely identify what this component is.

In this paper we describe two corpus studies related to the identification of prepositions that are sensitive to functional knowledge. In the first study we examine whether one can tell whether a particular preposition is more sensitive to functional than geometric knowledge or vice versa and therefore attempt to replicate the results of the aforementioned psychological studies. Having this information during generating descriptions of spatial scenes we can then decide which component - the geometric or the functional one - should be applied for a particular preposition, or ideally in a combined model, given more weight. For grounding spatial descriptions in scene geometry any of the existing models such as (Regier & Carlson, 2001) may be used. However, a model of functional knowledge must be built from corpus. For this we use methods that are common in computational distributional semantics in the analysis of collocations or collostructions (Stefanowitsch & Gries, 2003). A part of the novelty of our approach is that we apply corpus knowledge extraction techniques to generating grounded descriptions of scenes.

Spatial descriptions

As stated before spatial descriptions occur as particular constructions consisting of a spatial preposition, for example 'in', which typically takes two arguments, a *Figure* and *Ground* (Talmy, 1983): 'flowers' and 'vase'. Ground is the object whose location is already known and is therefore integrated in the conversation common ground wheres Figure is the object whose location is under discussion and which is consequently resolved by relating it to the ground. Spatial descriptions may contain simple prepositions such as 'in' or composite prepositions such as 'on the left side of' and alike.

To obtain semantic representations of the form *Preposition*(*Figure*,*Ground*) we processed our corpora (see below) with Stanford CoreNLP tools[1] from which we used dependency parses (Marneffe, MacCartney, & Manning, 2006).[2] Unfortunately, the resulting dependency parses are not yet quite what we require nor do prepositions occur in the same dependencies each time. For example, even simple constructions such as (1) 'The flowers in the vase' and (2) 'The flowers are in the vase' result in different dependencies parses:

root(ROOT-0, flowers-2)	root(ROOT-0, are-3)
det(flowers-2, The-1)	det(flowers-2, The-1)
det(vase-5, the-4)	nsubj(are-3, flowers-2)
prep_in(flowers-2, vase-5)	det(vase-6, the-5)
	prep_in (are-3, vase-6)

A series of rules was designed to rewrite these variations into single predicates such as *in(flowers,vase)*. Finally, composite spatial descriptions display a considerable structural variation which may be reflected in their dependency parses, for example 'on the left side of', 'on the left of' and 'on left of'. Such composite descriptions were rewritten as single prepositions 'on_left_side_of' and 'on_left_of' where definite articles were left out. The nominal arguments of spatial prepositions were taken as single word expressions as identified by the dependency parser. Their adjectival (or sometimes adverbial) modifiers and determiners were not taken into account. All words were stemmed and converted to lower case. For each corpus we then obtained frequency counts over the resulting tokens.

Identifying functional and geometric spatial descriptions

Entropy of figure and ground pairs

We hypothesise that prepositions that are sensitive either to geometric or functional knowledge could be distinguished by examining the entropy (Shannon, 1948; Manning & Schütze, 1999, p.61) of their nominal arguments. Entropy is both sensitive to the frequency of tokens and the number of types corresponding to such tokens. It is low if the number of categories is small and their frequencies are high. This is what we expect of prepositions that impose functional constraints on their arguments as these refer to a more restricted set of situations where such functional constraints hold. On the other hand, geometrically sensitive prepositions are not governed by such constraints but mostly by geometrical scene arrangements and the entropy of their figure and ground pair tokens is expected to be higher. Of course, there will also be constraints on the number of objects that can be reasonably related in geometric space or simply the tendencies to prefer certain objects in a particular corpus but we expect that such constraints will be less prominent than the functional constraints.

To calculate entropy we use the standard equation

$$H(X) = -\sum_{x \in X} p(x) log_2 p(x).$$

To make figures comparable the entropy values have been normalised by the maximum attainable entropy $-log_2(n)$ where n is the size of the set (X).

General newspaper text

In our first experiment we selected a collection of periodical texts (52,510,070 sentences) from the British National Corpus (BNC). Overall 406 preposition types were automatically extracted. However, our analysis focused on 6 hand picked prepositions that based on the results in the literature we felt would be an interesting set to examine with respect to the functional versus geometric semantics. The first to prepositions we included in our analysis were the prepositions 'above' and 'over'. The motivation here being that clearly the semantics of these prepositions do overlap, however, at the same time the literature would suggest that 'over' is more sensitive to functional effects then 'above', see for example (Coventry et al., 2001). We also included 'on' and 'in' and the composite prepositions that they prefix, for example 'in the left side of' or 'on the right of'. The literature suggest that 'on' and 'in' have have a functional character in the same way as 'over' (Garrod et al., 1999) while the composite variants are likely to be geometrically defined (J. Kelleher & Ross, 2010). Although the results were promising we identified a number of issues that impacted on the study including: (a) our choice of corpus was problematic because it consisted of newspaper texts and therefore functional prepositions also have many other non-spatial uses, for example temporal uses such as 'in/over three days'; and (b) the number of tokens for some subject-object type were low (a single occurrence) in spite of corpus size and consequently, in these cases, the entropy was dominated simply by frequency. Consequently, we decided to rerun our methodology on more appropriate, spatially focused, datasets, in particular datasets of image descriptions.

Descriptions of images

The IAPR TC-12 Benchmark corpus (Grubinger, Clough, Müller, & Deselaers, 2006)[3] contains 20,000 images and their corresponding linguistic descriptions. We only use the text part of this corpus, the descriptions from complete English annotations entered under the <description> xml tag. This contains a description of the scene on the photograph, for example in 25251.eng we find 'four adults are sitting and playing *with* small children *on* a carpet *on* a red floor; four children are sitting *at* a table *in* the foreground; a dark blue wall and a wooden shelf *with* many toys *on* it *in* the background;'. As demonstrated by this example, sentences (separated by semi-colons for convenient parsing) contain a high count of spatial prepositions and therefore this corpus is particularly suitable for our task.

We also use the linguistic descriptions from the 8K ImageFlickr dataset (Rashtchian, Young, Hodosh, & Hockenmaier, 2010)[4] which contains 8108 images. Each image is accom-

[1] http://nlp.stanford.edu/software/corenlp.shtml

[2] http://nlp.stanford.edu/software/lex-parser.shtml We use collapsed dependencies with processed coordinating conjunctions.

[3] http://imageclef.org/photodata

[4] http://nlp.cs.illinois.edu/HockenmaierGroup/8k-pictures.html

panied by simple sentences describing entities and actions in images. Each sentence is written by a different annotator. The authors of the corpus assessed the annotations and ensured their quality. For example ‘The three children are playing on the rails’, ‘Three children climb *on* a livestock fence’ and ‘Three children stand and climb *on* a fence.’, etc. All sentences thus describe the same scene in a single sentence but from a slightly different describer’s viewpoint.

From a manual analysis of examples from both corpora we get an impression that IAPR TC-12 Benchmark corpus gives us more examples of prepositional use that we are looking for than the 8K ImageFlickr dataset since because in the IAPR TC-12 Benchmark corpus the annotators seem more likely to describe relations between objects on the picture (we hypothesis that this is because they can use many sentences) whereas in 8K ImageFlickr corpus, being restricted to a single sentence, they describe events that are taking place in the picture which our dependency rewrite rules cannot capture completely. For the analyses described below we merged both corpora and treated them as a single corpus to obtain preposition, figure and ground counts. Note that only words that have been extracted successfully are used in the calculations of the statistical measures and hence the success of extraction only has an effect on the number of examples extracted and does not bias the measures calculated.

Entropy results

We processed the combined IAPR TC-12 Benchmark and 8K ImageFlickr corpus using the same process as we described earlier: Stanford CoreNLP analysis with stopwords excluded and all items stemmed and written to lowercase; followed by the application of our hand-crafted rewrite rule to standardise the analysis of prepositions into single predicates and to merge composite spatial descriptions. This process extracted extracted 98 preposition types, including both simple and composite prepositions, with a total of 96,512 token occurrences. The list was manually checked and types that were not prepositions but were identified as such by the automatic processing were excluded. Similarly we also excluded all prepositions that could not have a spatial usage, for example ’with’. This resulted in a list of 77 preposition types (54,428 token occurrences). The entropy of preposition figure-ground arguments is biased by low token counts of the latter. For example, if a preposition occurs only with one figure-ground argument type which also has a token count 1 (at_gray_bottom_of, in_different_front_of, on_front_of) , the entropy is 0. If a preposition occurs, say with two or more figure-ground argument types where each type has a single token count (to_left_of (2), alongside (4) and close_to (5)), the entropy is 1.[5] Clearly, such cases are not very informative of a preposition’s figure-ground selection properties and all prepositions which occur with less than 10 figure-ground tokens were removed from the dataset. There were 34 such cases or 108 token occurrences. Note that here we consider and count both figure and ground as token pair and as a result the counts are lower than if figure and ground were counted independently. This is because we assume that a functional spatial relation is defined between both a figure and a ground. Finally, we end up with 43 preposition types (54,320 token occurrences) for which we calculate normalised entropies of their figure ground pairs as shown in Table 1.

According to experimental studies (Garrod et al., 1999; Coventry et al., 2001) prepositions ‘in’, ‘on’ and ‘over’ have been identified as being influenced by a functional component and we hypothesised that this should reflect in the lower entropy of their figure and ground arguments. Examining Table 1 the prediction seems to be confirmed as these prepositions are ranked as 3, 10 and 6 respectively, thus in the top quarter of the preposition list. Also, (Coventry et al., 2001) contrast the functional ‘over’ (rank 6) with the geometric ‘above’ (rank 23) which are ranked considerably differently. Unfortunately, we lack experimental results to confirm that other prepositions that occur with these prepositions are influenced by a functional component. What we can generalise from Table 1 and the preceding observation is that not all spatial descriptions are equally selective of their figure and ground arguments, those with lower entropy are more selective then those with higher entropy. Such selectiveness may be linked to the requirement of a functional semantic component between the figure and ground arguments. We can therefore not make a clear-cut distinction between functional and geometric descriptions but we can make conclusions about their tendencies. In terms of natural language generation the figures in Table 1 could be used in determining the weight the geometric and functional model should be given while generating a description with a particular preposition.

A well known difficulty of a shallow-corpus based approach to natural language semantics is that it is not straightforward to distinguish between the effects of other relations existent in the text that may be at play simultaneously. For example, the above selectional tendencies of prepositions could be an artefact of a bias in the corpus; for example, the corpus may be biased towards a certain kind of description or because the annotators prefer to use one preposition over another. Although both factors were reasonably minimised in our case by using specialist corpora we did notice that ’on_left_side_of’ which comes on the top with rank 1 is possibly because of an annotator bias towards this description: the same description is repeated in several consecutive sentences. On the other hand, such bias may be annotators functional knowledge about the preposition usage and hence the annotator bias and functional knowledge may not be straightforwardly distinguishable.

We hypothesised that only simple prepositions are functional but their composite variants are geometric. We assumed this is because their components such as ‘side’, ‘front’

[5] Many of these were composite spatial relations that were created by us by merging dependency relations. No doubt tweaking the rules further would provide more optimal compositions with better token counts.

Rank	Preposition	FG-T	Tks	Norm FG ent
1	on_left_side_of	5	31	0.35448
2	underneath	31	74	0.65535
3	**in**	7584	34846	0.6714
4	onto	49	86	0.79109
5	down	83	142	0.81099
6	**over**	440	736	0.83106
7	at	1393	2726	0.83148
8	on_top_of	61	87	0.83409
9	against	50	68	0.85171
10	**on**	4897	10085	0.852
11	on_side_of	46	63	0.87644
12	into	78	110	0.88426
13	around	156	220	0.89393
14	at_bottom_of	13	16	0.89445
15	on_back_of	9	11	0.89489
16	through	179	245	0.89738
17	in_front_of	1278	1938	0.90998
18	after	28	35	0.91088
19	behind	318	437	0.91235
20	at_top_of	46	59	0.91715
21	before	25	30	0.91848
22	under	167	220	0.92096
23	above	145	190	0.9228
24	to	114	138	0.94187
25	between	129	153	0.9508
26	below	13	14	0.96248
27	towards	66	74	0.96354
28	near	608	692	0.97035
29	by	87	95	0.97195
30	across	76	81	0.97906
31	along	119	127	0.98112
32	next_to	168	178	0.98384
33	outside_of	29	30	0.98641
34	outside	70	73	0.98672
35	inside	58	60	0.98871
36	beneath	10	10	1
37	out_of	11	11	1
38	amidst	13	13	1
39	past	13	13	1
40	out	15	15	1
41	among	21	21	1
42	toward	33	33	1
43	beside	34	34	1

Table 1: Prepositions ranked by the normalised entropy (Norm FG ent) of their figure-ground (FG) arguments. FG-T shows the number of the figure-ground types that a preposition occurs with, Tks is the number of figure-ground tokens which is also equal to the preposition tokens. The total number of tokens is 54,320.

and 'bottom' appear to introduce a geometric frame.[6] However, this is not confirmed by Table 1 where the composite relations tend to rank equally high as their simple variants. Furthermore, it even appears that composite spatial relations cluster with their simple variants. For example, 'on_left_side_of' (1), 'onto' (4), 'on_top_of' (8), 'on_side_of' (11) and 'on_back_of' (15) are close to 'on' (10). Slightly less well, 'into' (12) and 'in_front_of' (17) cluster with 'in' (3) and 'at_bottom_of' (14), 'at_top_of' (20) cluster with 'at' (7) in the same half of the table. On the other hand, in the bottom half of the table we get 'outside_of' (33), 'outside' (34), 'out of' (37) and 'out' (40); 'next_to' (32), 'by' (29) and 'near' (28); 'towards' (27), 'across' (30), 'along' (31) and 'to' (24). 'amidst' (38) is close to 'among' (41). Antonyms also appear to have similar ranks. For example, 'under' (22), 'above' (23) and 'below' (26); 'over' (6) pairs with 'underneath' (2) at the top of the list but not 'beneath' (36); 'outside_of' (33), 'outside' (34) and 'inside' (35); 'after' (18) and 'before' (21). Overall, it appears that the entropy of a preposition's arguments does reflect semantic similarity between prepositions.

To further test the relationship between composite and non-composite prepositions we separated them into two lists preserving their ranking order and calculated Pearson's correlation coefficient on their normalised entropies of figure-ground tokens. There are only 9 composite spatial descriptions ('on_left_side_of', 'on_top_of', 'on_side_of', 'at_bottom_of', 'on_back_of', 'in_front_of', 'at_top_of', 'outside_of', and 'out_of') and so only the first 9 corresponding simple prepositions were considered in the calculation ('underneath', 'in', 'onto', 'down', 'over', 'at', 'against', 'on', and 'into'). There is high correlation between the lists 0.80476, $p = 0.00587$ confirming that the correlation is statistically significant. Given that the lists were built by preserving the order of items from the original ranking list shown in Table 1 this shows that composite relations are equally distributed along the entropy spectrum with non-composite ones; and since we only consider the items up to 'into' from the original list, the composite preposition correlate highly with the simple prepositions from the first half of the list. Thus, they show a functional character.

Although in the aggregate there is a high correlation between composite and non-composite spatial prepositions, if we focus our analysis on a particular non-composite preposition and its variants we see that this correlation breaks down. Correlating the entropies of the list ('on_left_side_of', 'on_top_of', 'on_side_of', 'at_bottom_of', 'on_back_of', 'in_front_of', 'at_top_of', 'outside_of', and 'out_of') with the list ('on', 'on', 'on', 'at', 'on', 'in', 'at', 'outside', 'out') gives us a low Pearson's correlation coefficient 0.18470 which is statistically non-significant ($p = 0.63096$). The lack of correlation between particular composite and non-composite prepositions points to the fact that the

[6]Following (Garrod et al., 1999) we can see that functional knowledge for 'in' and 'on' may introduce geometric constraints, in this case location control.

composite prepositions do have a distinctive semantics from their non-composite counterparts (despite the fact that our aggregate analysis of the correlation between non-composite and composite suggests that non-composite prepositions do have a functional aspect to their semantics).

Generating descriptions with functional spatial prepositions

We argue that the entropy analysis method described previously allows us to identify the degree to which a preposition is sensitive to functional knowledge. However, we also have to build a model of such knowledge. We assume that functional knowledge is encoded implicitly in the choice of the figure-ground arguments that a preposition takes. Therefore, to build a (shallow) model of prepositional functional knowledge one can evaluate the strength of association between the figure and ground pair and the preposition.

Log likelihood ratio

Log likelihood ratio (Dunning, 1993; Manning & Schütze, 1999, p.172) is a well known measure used in the analysis of collocations and collostructions. It is a log ratio of the likelihood of Hypothesis 1 (H_1) over Hypothesis 2 (H_2), where Hypothesis 1 states that words in a particular bigram $w_1 w_2$ are independent and Hypothesis 2 states that they are dependent:

$$log\lambda = log\frac{L(H_1)}{L(H2)}$$

Log likelihood ratio therefore tells us how many times more likely H_2 is compared to H_1. Another useful property of the log likelihood ratio is that $-2log\lambda$ approaches asymptotically the χ^2 distribution which means that its values can be tested for statistical significance using the χ^2 distribution.

Prepositional collostructions

In natural language generation we typically (but not always) start with figure and ground objects which we want to relate with a preposition. Translating this to the bi-gram scenario of the log likelihood ratio this means that a particular figure-ground pair is taken as $w_1 = fg$ and we want to evaluate dependence of the preposition $w_2 = prep$ on it. To calculate the log likelihood ratio for a particular figure-ground pair and a preposition we need to find the occurrence count of that fg pair $C(fg)$, the occurrence count of that preposition $C(prep)$ and the occurrence count where both of them occur together $C(fg, prep)$. Overall, we extracted 32,462 figure-ground types which occur with various prepositions 96,749 times in total. 17,575 figure-ground types occur only once, 26,140 less than 5 times.

Table 2 gives some examples of the calculated log likelihood ratios. For example, it shows that the most likely preposition to relate the figure 'boy' with the ground 'shirt' is 'in', another slightly less likely possibility is 'with'. The third column of the table from the right gives the p value with which H_1 may be rejected and the last columns tells us how many times a given bigram (*fg, prep*) is more likely under H_2 than H_1. For example, it is 5,18 $\times 10^{52}$ more likely that 'boy*shirt' are used with 'in' rather than any random word.

In the second example we examine what figures and prepositions the ground 'umbrella' is most strongly associated with (for the sake of brevity we exclude some examples). The bottom half of Table 2 shows that 'umbrella' is most strongly associated with the ground 'people' and the preposition 'with'. 'child*umbrella' is more strongly associated with 'under' than 'with' although it occurs with 'under' only once and two times with 'with'. This is because 'over' occurs less frequently overall. However, this is not the case for 'woman*umbrella' which is more strongly associated with 'with' than 'under' but the difference is very small.

When generating spatial descriptions a system needs to decide the order it which it will check the appropriateness of different spatial relations. This checking can be computationally expensive, particular in a situated system where visually processing routines may need to be invoked. We argue that the log likelihood ratio between the a target figure-ground pair and the set of candidate prepositions can provide a useful guide in ordering the way the system evaluates the prepositions for inclusion in the output. The functional relations between objects and prepositions captured by this corpus study gives an important insight how humans view situations. However, basing generation purely on functional knowledge is not enough for successful descriptions of visual scenes. The functional component may interact with a geometric component: for example, when do we choose 'over' and when 'under'.[7] It appears that geometric and functional knowledge must be somehow integrated.

Conclusion and further work

In this paper we discuss two studies of identifying functional relations that exist between spatial prepositions and their argument. In the first study we show that prepositions with stronger functional component have more restricted argument selection properties than prepositions more strongly influenced by a geometric component. However, there seems to be no clear division between the two kinds of prepositions but different degrees of tendency. In the second study, we attempted to build a model of functional knowledge between prepositions and their arguments using techniques from collocation/collostruction analysis and discuss how such knowledge could be used in natural language generation.

The current work could be extended in many different ways. It appears that most composite descriptions containing 'left' and 'right' did not make it to Table 1. A plausible reason for this is that our dependency extraction rules could not capture all the occurrences. No doubt, further tweaking of these rules and getting to know the corpora better would give us more examples. Another important direction is to integrate

[7] Interestingly, the corpus does not contain any occurrences of 'over' with 'umbrella'.

$-2log\lambda$	$C(fg)$	$C(prep)$	$C(fg, prep)$	fg	$prep$	p	$p < 0.05$	H_2 vs. H_1
242.76	282	34,887	229	boy*shirt	in	9.86×10^{-55}	yes	$5{,}18 \times 10^{52}$
26.98	282	30,717	51	boy*shirt	with	2.06×10^{-7}	yes	721,655.5
2.85	282	34	1	boy*shirt	without	0.091461204	no	4.155
1.51	282	74	1	boy*shirt	underneath	0.219177447	no	2.127
16.06	7	30,717	7	people*umbrella	with	6.13×10^{-5}	yes	3076.878
12.16	1	222	1	boy*umbrella	under	0.000488543	yes	436.788
9.39	2	222	1	table*umbrella	under	0.002180673	yes	109.447
8.35	3	222	1	child*umbrella	under	0.003859083	yes	65.006
6.88	3	30,717	3	sculpture*umbrella	with	0.008696962	yes	31.25
6.83	6	30,717	5	woman*umbrella	with	0.008960309	yes	30.428
6.78	6	222	1	woman*umbrella	under	0.009244184	yes	29.592
4.59	2	30,717	2	girl*umbrella	with	0.032172078	yes	9.921
2.29	1	30,717	1	man*umbrella	with	0.129822168	no	3.15
1.53	3	30,717	2	child*umbrella	with	0.215491662	no	2.153

Table 2: Log likelihood ratios for 'boy*shirt' and 'umbrella' as ground

the model of functional knowledge in a natural language generation system, for example for describing images that are a part of the corpora from which the sentences were extracted and asking human evaluators to judge the appropriateness of such descriptions. It would be interesting to examine how accurately a system can generate descriptions by using purely functional knowledge without resorting to help of geometric models. Finally, we are interested how to integrate the model of functional knowledge with a geometric model for spatial descriptions as outlined in (Dobnik, Cooper, & Larsson, 2013).

References

Coventry, K. R., Prat-Sala, M., & Richards, L. (2001). The interplay between geometry and function in the apprehension of Over, Under, Above and Below. *Journal of Memory and Language*, *44*(3), 376–398.

Dale, R., & Haddock, N. (1991). Generating referring expressions involving relations. In *Proceedings of the EACL '91* (pp. 161–166). Stroudsburg, PA, USA: ACL.

Dobnik, S. (2009). *Teaching mobile robots to use spatial words*. Unpublished doctoral dissertation, University of Oxford, United Kingdom.

Dobnik, S., Cooper, R., & Larsson, S. (2013). Modelling language, action, and perception in Type Theory with Records. In D. Duchier & Y. Parmentier (Eds.), *Constraint solving and language processing (CSLP 2012).* Berlin, Heidelberg: Springer. (forthcoming)

Dunning, T. (1993). Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*, *19*(1), 61-74.

Garrod, S., Ferrier, G., & Campbell, S. (1999). In and on: investigating the functional geometry of spatial prepositions. *Cognition*, *72*(2), 167–189.

Grubinger, M., Clough, P. D., Müller, H., & Deselaers, T. (2006). The IAPR benchmark: A new evaluation resource for visual information systems. In *Proceedings of OntoImage 2006 at LREC 2006.* Genoa, Italy: ELRA.

Herskovits, A. (1986). *Language and spatial cognition: an interdisciplinary study of the prepositions in English*. Cambridge: Cambridge University Press.

Kelleher, J., & Ross, R. (2010). Topology in composite spatial terms. In D. N. Rapp (Ed.), *Spatial Cognition 2010: Poster sessions.* Mt. Hood/Portland, Oregon, USA: Cognitive Science Society.

Kelleher, J. D., & Costello, F. J. (2009). Applying computational models of spatial prepositions to visually situated dialog. *Computational Linguistics*, *35*(2), 271–306.

Manning, C. D., & Schütze, H. (1999). *Foundations of statistical natural language processing*. Cambridge, Mass.: MIT Press.

Marneffe, M.-C. D., MacCartney, B., & Manning, C. D. (2006). Generating typed dependency parses from phrase structure parses. In *Proceedings of LREC 2006* (pp. 449–454). Genoa, Italy: ELRA.

Rashtchian, C., Young, P., Hodosh, M., & Hockenmaier, J. (2010). Collecting image annotations using Amazon's Mechanical Turk. In *Proceedings of the NAACL HLT 2010 Workshop on creating speech and language data with Amazon's Mechanical Turk.* Los Angeles, CA: NAACL.

Regier, T., & Carlson, L. A. (2001). Grounding spatial language in perception: an empirical and computational investigation. *Journal of Experimental Psychology: General*, *130*(2), 273–298.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*(3), 379–423.

Stefanowitsch, A., & Gries, S. T. (2003). Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics*, *8*(2), 209–243.

Talmy, L. (1983). How language structures space. In H. L. Pick Jr. & L. P. Acredolo (Eds.), *Spatial orientation: theory, research, and application* (pp. 225–282). New York: Plenum Press.