PREPOSITIONS SHAPE FRAMES FOR ABSTRACT EVENTS: A QUANTITATIVE STUDY OF SPATIAL-TO-ABSTRACT MAPPING IN ENGLISH SCIENCE WRITING

Christoph Haase

Abstract

Adpositions form closed lexical classes in English which provide cognitive access to representations of complex scenes and events. In academic texts from the domain of natural sciences, these lexical items help build conceptualisations of abstract relationships between phenomena that cannot be perceived directly. As a consequence, any phenomenon that can be hypothesised or measured in the natural sciences can be mediated and transferred into comprehensible processes via linguistic markers by way of "experiential correlation" and subsequent metaphorical extension. This extension is systematically grammaticalised in prepositional phrases.

Key words

corpora, academic English, prepositions, metaphor, cognitive linguistics, conceptualisation, spatial, popular science, image schema, abstractness

1 Introduction

This contribution looks at the distributional properties of the prepositions *at*, *in* and *on* as heads of prepositional phrases that provide a mapping from spatial meanings onto abstract-metaphorical meanings that enable conceptual access to abstract science concepts. We assume that metaphor is an initial expansion of an identity relation between spatial and abstract concepts and that the lexicalization of the metaphor establishes some sort of imagery identity (cf. Lindstromberg 1998).

This quantitative study is based on the Corpus of Specialised and Popular ACademic English (SPACE), which focuses on different sub-disciplines from the natural sciences and contains a wide spectrum of prepositions, as the following example illustrates:

Corpus file: AX0036; Source: arXiv:astro-ph/0402584 v1 25 Feb 2004
AN OBSERVATIONAL SIGNATURE OF EVOLVED OCEANS ON EXTRA-SOLAR TERRESTRIAL PLANETS
For Lyman a, we assume Gaussian line profiles **with** .0 = 8.23 × 109 Hz and .0 = 4.11 × 1010 Hz, corresponding to 1 km s-1 and 5 km s-1, respectively. **With** the column densities given above then t (0) 14 (V = 1 km s-1) and t (0) 0.5 (V = 5 km s-1). As discussed below, such lines are detectable if the velocity shift **of** the star is sufficiently different **from** that **of** the interstellar matter **in** the line-of-sight.

Higher order Lyman lines **in** the circumstellar gas might be detectable even if observations **at** the stellar Lyman a velocity are dominated **by** intervening interstellar absorption. We also note that the absorption lines produced **by** the wind from the planet are likely to be narrower than any structure **in** the Lyman emission lines that are produced **in** the chromospheres **of** main sequence G-type … where … even the thermal width **of** the hydrogen emission is greater than 10 km s-1. We now consider how different levels **of** stellar activity can affect the observability **of** the planet's wind. The mass loss rate **of** hydrogen **from** the planet (.ZH) scales **with** LEUV and thus there is a tendency **for** more active stars to … produce a larger projected column density **of** hydrogen. However, an increase **in** the stellar activity also leads **to** an increase **in** the neutral hydrogen ionization rate (JH) and thus a decrease **in** the size **of** the neutral cloud **around** the planet (R0). … [W]e compute the value **of** the equivalent width **of** an absorption averaged **over** the disk **of** the star when the planet lies **toward** the very center **of** the star. We ignore any limb brightening or darkening **of** the stellar atmosphere. Because there are so many uncertainties, we consider Doppler broadening parameters equal **to** the assumed outflow speeds **of** either 1 km s-1 or 5 km s-1. Considering main sequence stars **of** age, t*, and following Ayres' (1997) study **of** photo-ionization rates and Wood et al.'s (2002a) study **of** stellar winds, we assume that both JH and LEUV scale as t-2…

In the case of *on*, it is obvious on the one hand that the senses of *on* vary between spatial (topological/directional) and metaphorical senses, on the other hand, the example above demonstrates a) that the spatial-core prepositions refer exclusively to abstract entities (*into a lower limit*) and b) that spatial-metaphorical prepositions can refer to concrete entities (*on galaxies*) but not with a spatial interpretation. Instead, the systematic use of metaphor facilitates the building of abstract event frames.

We hypothesize that quantitative arguments for mapping can be obtained from texts in which the more direct and literal mappings occur in the "softer" sciences like biology as opposed to "hard" sciences like physics (cf. Schmied 2002). An investigation of prepositional phrases and an assignment of roles to the parts of the construction (commonly: landmark, localizer and modalizer) render a complex picture of prepositional usage. The landmarks (usually NPs) serve as locative or abstract complements with direct or peripheral import for the construal of the frames (cf. Lopez Rúa 2005).

Thus, a quantitative profile of spatial-to-abstract mappings can generate comparable text-type specific signatures. We can also show that a continuum of abstractness exists on which concrete spatial and purely abstract notions form the poles and hypothetical-concrete and imperceptible-concrete notions occupy middle ground.

2 Spatial prepositions: Conceptualization issues

2.1 The Cognitive-Linguistic Stance

The spatial experience overrides all other experiences and the domain of space is an empirically rich background against which other experiences, especially those related to abstractness, are mapped. Cognitive linguistics tries to answer the question, "whether space is conceived as an abstract entity described by geometry or topology or as an indispensable host for our experience of the world" (Vandeloise 2006: 139). According to conventional description schemes within the cognitive linguistic paradigm, any embodied experience of speakers gives rise to conceptual structures which are – as long as spatial experience is concerned – anchored in trajectory (TR)-landmark (LM) configurations (cf. Tyler & Evans 2003: 23). When these configurations are abstract and therefore not accessible to direct experience, the abstractness is conceptualized by extending the embodiment, or: "the foundation of conceptualizing capacity is the image-schema in which a spatial structure is mapped into a conceptual structure" (Mandler 1992: 591). This means that spatial prepositions form semantic polysemy networks which act like a mental coordinate system (Rice 1993). A direct mapping can be observed in the experiential contiguity between a vertical scale and a decrease in material (i.e. water dripping from a tube; thus emptying its container from the bottom) which is applied to abstract negative experience, e.g. "My luck went <u>down</u> the drain" or to abstract positive experience by e.g. the amelioration of a state of affairs, even in fully idiomatized forms: "Let's spruce <u>up</u> these buildings". In this way, we can find systematic homologue lexicalization patterns between direct and abstract experience, commonly captured in the metaphorical use. The prepositional system, that according to Langacker (2008) has lexical redundancy, shows therefore, quite the opposite: it is a versatile system in which the lexical items have a spectrum of meanings. The following table summarizes distinct meanings of *over* in a flat ontology adapted from the LDOCE:

Spatial			**Temporal**		**Metaphorical**	
SL	locative		**TS**	temporal	**MQ**	quantity
SC	cover	2dim (surface)	**TD**	during	**MC**	causal
SD	directional	1dim (line)			**MS**	control
			<u>**Lexical**</u>			
SB	bilateral				**ME**	evaluation

			X	lexical		
SS	down				**MM**	medium
SR	range				**MD**	directional
			L		**MR**	range

Table 1: Extended ontology for *over*

Over has therefore three main domains of meaning, spatial, temporal and metaphorical which themselves have a more refined 1-tier subsystem of senses. In a polysemy graph, these senses can be represented as follows:

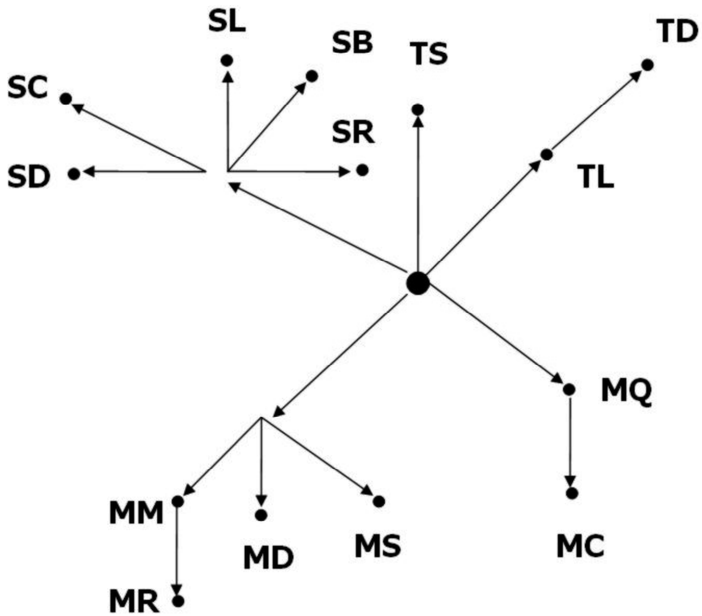

Figure 1: Polysemous graph of a flat ontology for *over*, cf. also Tyler and Evans (2003: 80), ontological categories adapted from LDOCE

For lexical semantics, this poses a challenging question as the polysemy of spatial prepositional phrases makes it difficult to find recurrent mapping patterns of iconicity in abstractness because for every situation, abstractness is differently conceptualized (cf. also van der Gucht, Willems & De Cuypere 2007). This problem is also evidenced by the following table: it represents the distribution of the senses in a corpus of academic texts from the physical sciences. It further serves as a pilot study for the full-scale investigation of all senses in all corpora.

PHYSICS	Spatial												Temporal				Metaphorical												
	SL		SC		SD		SB		SS		SR		TS		TD		MQ		MC		MS		ME		MM		MD		MR
0001AX																													7
0003AX02																			1										
0004AX01											4	fjlm			6	cdegkn													4
0004AX02															1	a													2
0004AX03															1														
0004AX04																													5
0005AX															1														
0006AX01											1	a									1	b							8
0006AX02																	1	b											1
0007NS													1																
0008AX																													1
0013AX															3	cde	1	a											1
0013NS															3														
0015AX															3	bcd													1
0016AX																							2						
0017AX															2	bd													5
0017NS															1														
0018AX															1	b	1	a											
0019AX															1														
0020AX																					2	ad							2
0021AX																													1
0021NS																													1
0022AX											2	ad							1	b									1
0023AX															2	ab													
0023NS					3	abc													1	d									
0025AX																													9
0026AX	1	a															1	c			1	b							
0027AX															2														
0027NS																	1												
0028AX															1	a													3
0030AX															1	b					1	a							1
0031NS															2														
0032AX															1														
0032NS					1																								
0033AX											2																		
0035AX											7																		
0036AX																	1	a											1
0036NS															1														
0037AX															3	ace													2
0038AX											1	c			3	abd													
0039AX																	1												
0040AX																													2
0041AX																													2

Table 2: Distribution of senses for *over* in the physical sciences (SPACE corpus)

Without detailing the pilot results, the usage patterns in the table show remarkable clustering for the temporal field (mid section) and metaphorical range (MR, rightmost column). The study (described below) was designed to determine the metaphorical use of spatial prepositions in order to establish marked profiles in different academic domains.

3 Methodology and data

3.1 Data material

All prepositional occurrences and their collocations were retrieved from the Corpus of Specialised and Popular ACademic English (SPACE), compiled from 2007 at Chemnitz University of Technology and available for registered users in the 2009 version online. The corpus has a binary structure in which part 1 is comprised of academic texts, part 2 of parallel popular texts. This means that a text in part 1 corresponds to a particular text in part 2. Articles in popular science publications such as *Scientific American* or *New Scientist* are commonly based on current publications in academic journals or preprint servers. These research results are routinely made accessible to general-interest readers. As the original article is usually referenced in the popular article, we started compiling both original and popularized versions of the articles. Some of the data evaluation is presented in Haase and Schmied (2010 forthc.).

As a closed lexical class, prepositions show a very diversified frequency spectrum from extremely frequent to extremely rare. In this study, only the most frequent 33 prepositions were considered, of which the ranks 1-10 are displayed below. The frequency spread for the full sample was from *of* (20,884 occurrences, 42 per 1,000 words) to *beneath* (8 occurrences, 0.016 per 1,000 words). The figures for all prepositions for all domains are given below.

Prep	Specialised Physics	ratio	Specialised BioScience	ratio	Popular Physics	ratio	Popular BioSc.	ratio	total	ratio
	001AX-046AX	per 1,000	047PN-107PN	per 1,000	001NS-046NS	per 1,000	047NS-107NS	per 1,000		per 1,000
of	8,258	51.02	10,199	38.18	1,330	32.68	1,097	35.97	20,884	41.75
in	3,563	22.01	5,557	20.80	577	14.18	619	20.30	10,316	20.63
for	2,145	13.25	2,195	8.22	208	5.11	202	6.62	4,750	9.50
with	1,513	9.35	2,056	7.70	144	3.54	173	5.67	3,886	7.77
by	1,489	9.20	1,845	6.91	162	3.98	144	4.72	3,640	7.28
at	1,118	6.91	1,221	4.57	204	5.01	164	5.38	2,707	5.41
on	1,039	6.42	1,126	4.22	159	3.91	158	5.18	2,482	4.96
from	914	5.65	1,607	6.02	163	4.01	185	6.07	2,869	5.74
as	870	5.37	870	3.26	181	4.45	154	5.05	2,075	4.15
between	272	1.68	463	1.73	37	0.91	27	0.89	799	1.60

Table 3: Distribution of the ten most frequent prepositions in the SPACE corpus

This considerable lexical spread needs delimitation by focusing on a small field of prepositions that can be tested for their membership in one or several lexical classes.

3.2 Hypotheses and quantitative arguments

Deviating for the canonical cognitive linguistic view expressed in Section 2, we feel the necessity to introduce what can be called a caveat of metaphorical use:

The metaphorical use in academic texts very frequently refers to representations of abstractness, as captured in data readouts, scales, diagrams, graphs, tables etc. These representations can be perceived directly, not metaphorically, and are therefore called in our terminology, second-order metaphors.

This is illustrated by a category mapping for the example AX0036 given in the introduction.

locative-spatial, 2nd order metaphor		abstract-spatial, 1st order metaphor	
item	**type**	**item**	**type**
for	abstract direction	*from*	locative direction
with	abstract inclusion – inside	*in*	locative inclusion
of	abstract inclusion – outside	*around*	locative motion
over	abstract motion	*toward*	locative direction
by	abstract causal	*by*	abstract correlating

Table 4: Metaphor types for prepositions in AX0036

However, for these items, not only configuration of the spatial prepositions is important, also the construal of the nominal items in these configurations. The following possible combinations for the three most basic spatial prepositions can be determined:

(1) P_{abstr} + $N_{concrete}$ → *in a way*
(2) $P_{concrete}$ + $N_{concrete}$ → *at Planck scale*
(3) P_{abstr} + N_{abstr} → *on noncommutative geometry*

Based on this terminology, metaphor typology and observation, we can thus formulate several hypotheses that can be verified or falsified using the SPACE corpus:

1) The representational level of discourse makes academic texts comprehensible to experts as well as to laypersons
2) Two types of metaphor need to be isolated:
   first-order metaphor: involves a mapping from the perceptible to the abstract
   second-order metaphor: involves a mapping from the abstract-perceptible to the abstract-imperceptible
3) Academic discourse predominately bases on second-order metaphor
4) Non-abstract usage dominates popular science discourse
5) Abstract usage dominates specialized science discourse
6) Domain-specific signatures characterize different branches of hard and soft sciences

Not all these research questions will be addressed by the data discussed here. A survey of first-order metaphor and second-order metaphor for the preposition *over* offers the following distribution:

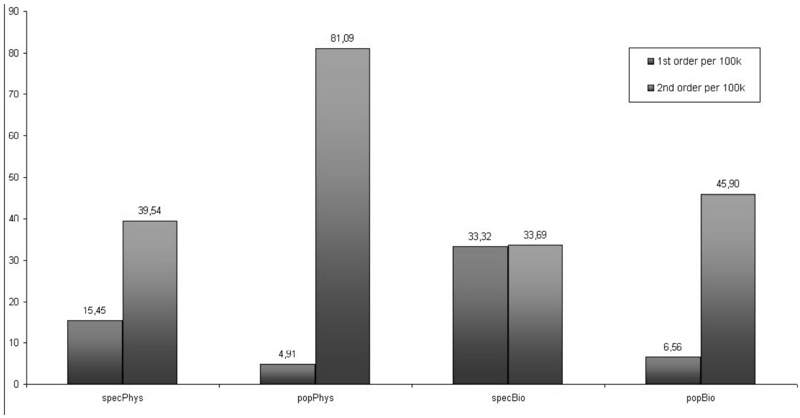

Figure 2: Cumulated data for first-order and second-order metaphor for *over*

The simple frequency counts show relatively low figures for first-order metaphor (left-hand side bars) which is extremely striking for the popular sciences (popPhys and popBio). Only in the domain of the specialized biosciences, there is no significant difference for the two types of metaphor. The picture gets more interesting when we move away from *over* and turn to *at, in* and *on*. First, we establish baseline figures in which we can determine the overall distribution of *at, in* and *on*:

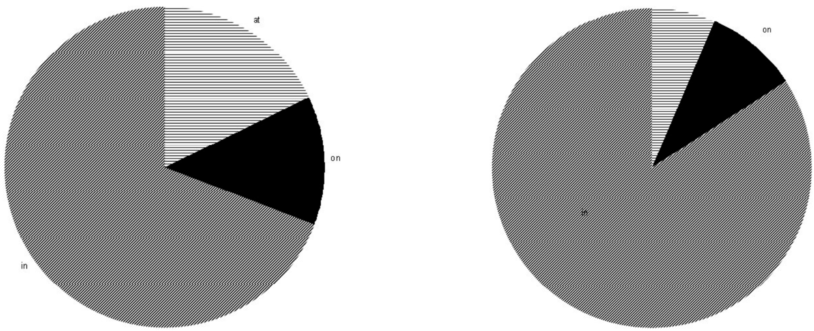

Figure 3: Overall distribution of *at* (bright grey area), *in* (dark grey), *on* (black) – biosciences (left pie chart) and physics (right)

These frequencies establish that *in* is the most frequent preposition for both domains. It further proves the point that the metaphorical system that is built up by *in* is especially sophisticated, starting from Lakoffian CONTAINER-metaphors to more refined applications.

The metaphorical productivity in the usage patterns of *in* is evidenced by the following profiles of categorical membership across a range of texts from the SPACE corpus. It is shown here for the range of texts that constitute the domain of genetics within the biosciences:

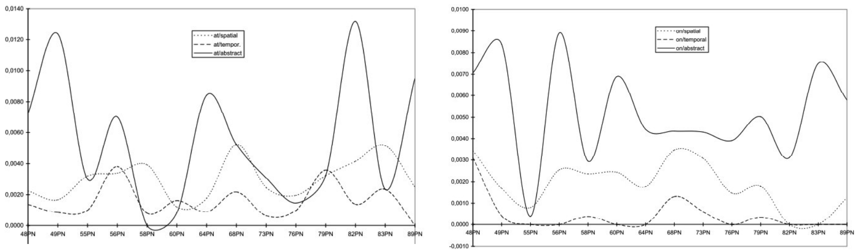

Figure 4: Profile graphs for *at* and *on* in specialized genetics (normalized)

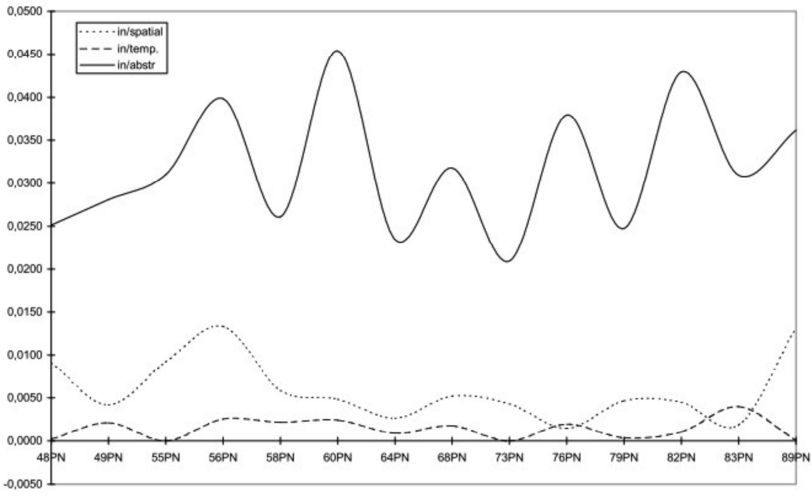

Figure 5: Profile graphs for *in* in specialized genetics (normalized)

In all graphs displayed above, the continuous line represents the metaphorical impact of the preposition. The vertical elevation stands for the relative frequency in relation to the specific text (horizontal scale). The metaphorical application of the respective preposition comes out highest for all three prepositions but is most significant for *in*. As expected, the dotted profile line (spatial application), however, is not significantly different to the dashed line (temporal application).

In the following analysis of variance, we looked at the difference between the domains in order to eliminate the unwanted bias that e.g. genetics does not vary from the other groups. As is evidenced by a *F*=7.23 for the biosciences, this is clearly not the case.

p<0.005	**bioSc**	**phys**	**astrophys**	**part.phys**	**quantphys**	**biochem**	**genetics**	**microbio**
Popular	30,499	40,694						
Specialised			17,407	44,125	59,076	33,700	112,176	28,297
Df1	46	61	6	5	7	6	14	6
Df2	6	6	6	6	6	6	6	6
F	7.23	8.00	11.56	9.9	10.05	7.89	6.4	8.12

Table 5: Data integrity in a mixed-design ANOVA for science domains

In a further refocusing of the approach, we look at the data for the prepositions *at* and *in* which is due to the overall lower counts of *on*. It also allows for a closer look at the dependence of abstractness on the different science domains. An intuitive rule here would be that more abstract fields of science are reflected in a higher use of abstract prepositions. In this respect, the physical sciences employ modes of higher abstraction up to the point where phenomena are posited that have only theoretical ramifications in the real world, i.e. all phenomena at the quantum level of consideration. For that end, we have compared the abstract vs. the spatial and temporal use of prepositions in three domains of physics, astrophysics (17,000 words), particle physics (44,000 words) and quantum physics (59,000 words; all figures rounded). The results of the partial study follow below:

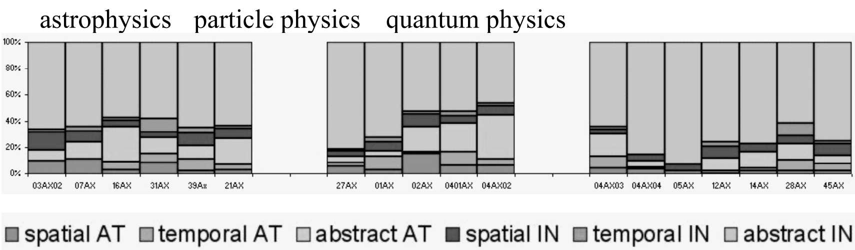

Figure 6: Distribution of *at* and *in* usage in physics sub-domains

The bar charts display the distribution of different usage patterns of prepositions for *at* (bottom of the chart) and *in* (top of the chart) with the brightest areas representing the metaphorical use, the darkest representing spatial use. The abstractness of quantum physics in contrast to e.g. astrophysics can be considered especially significant.

4 Conclusion

The results of the study confirm the overall assumption that academic texts, independently of discipline and readership, employ metaphorical means to map abstract phenomena into domains directly accessible to human perception. Furthermore, we argued for the necessity of a terminological split between metaphors of two different kinds: first-order metaphor (or: the canonical type from the cognitive-linguistic paradigm) and second-order metaphor. The latter refers not to abstract phenomena per se but to the representations of these phenomena in items of data and observation. These items are directly accessible to human perception. The figures show for a selected sample of features (prepositions, narrowed down to spatial *at* and *in*) that it is especially this type of metaphor that dominates the argumentation structure in the sciences. Clearly, a direct dependence of the metaphorical usage patterns on the degree of abstractness of the scientific discipline can be determined. More abstract sciences rank higher in their metaphorical import than less abstract disciplines.

References

Capelle, B. (2004) 'The particularity of particles, or why they are not just 'intransitive prepositions.'' *Belgian Journal of Linguistics 18*, 29-57.

Evans, V. and Tyler, A. (2004) 'Rethinking English 'prepositions of movement': The case of *to* and *through*.' *Belgian Journal of Linguistics 18*, 247-270.

Goldberg, A. (1995) *Constructions. A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Haase, C. and Schmied, J. (2010 forthc.) 'Conceptualising spatial relationships in academic discourse: A corpus-cognitive account of locative-spatial and abstract-spatial prepositions.' In: Roszkowski, S. (ed.) *Proceedings of the Conference on Practical Applications of Language and Computers, PALC 2009*. Frankfurt, New York: Lang.

Langacker, R.W. (2008) *Cognitive Grammar*. Oxford: Oxford University Press.

Lindstromberg, S. (1998) *English Prepositions Explained*. Amsterdam: John Benjamins.

Lopez Rúa, P. (2003) *Birds, Colours and Prepositions. The Theory of Categorization and its Applications in Linguistics*. München: LINCOM.

Mandler, J.M. (1992) 'How to build a baby II: Conceptual Primitives.' *Psychological Review 99*, 587-604.

Rice, S. (1993) 'Far afield in lexical fields: The English prepositions.' *ESCOL 92*, 106-117.

Schmied, J. (2002) 'Prototypes, transfer and idiomaticity: An empirical analysis of local prepositions in English and German.' In: Rabade, L.I. and Suarez, S.M.D. (eds) *Studies in Contrastive Linguistics.* Universidade de Santiago de Compostela. 947-959.

Tyler, A. and Evans, V. (2003) *The Semantics of English Prepositions: Spatial Scenes, Embodied Meaning and Cognition.* Cambridge: Cambridge University Press.

Thomas, E. (2004) 'On 'syntactic' versus 'semantic' telicity: Evidence from *in* and *on*.' *Belgian Journal of Linguistics 18*, 145-166.

Trim, R. (2005) *Metaphor Networks. The Comparative Evolution of Figurative Language*. New York: Palgrave.

Vandeloise, C. (2006) 'Are there spatial prepositions?' In: Hickmann, M. and Robert, S. (eds) *Space in Languages: Linguistic Systems and Cognitive Categories.* Amsterdam, Philadelphia: John Benjamins. 139-154.

Van der Gucht, F., Willems, K. and De Cuypere, L. (2006) 'The iconicity of embodied meaning. Polysemy of spatial prepositions in the cognitive framework.' *Language Sciences 29*, 733-754.