

Words and Patterns: Lexico-Grammatical Patterns and Semantic Relations in Domain-Specific Discourses

Concepción Orna-Montesinos
University of Saragossa
conorna@unizar.es

ABSTRACT

The underlying assumption of this study is the understanding of a specialized term as a summary of disciplinary knowledge, formalized at a textual level in the contextual relations which structure disciplinary lexical knowledge and are therefore essential for the successful interpretation of a text. With that aim this paper carries the analysis of the lexico-grammatical patterns which signal the hyponymy and meronymy relations of the term *building*, a key disciplinary concept in a corpus of construction engineering textbooks, using the WordNet database for reference. The linguistic analysis of the repertoire of lexico-grammatical patterns employed brings to the fore the dual role of hyponymy and meronymy as both semantic and metalinguistic discourse-organizing lexical resources, key in the rhetorical organization of the discourse of this discipline.

1. Introduction

Terminology studies have rightly acknowledged the multifaceted polyhedric nature of terms (Cabr  1998, 2008), which explains why the study of specialized terminology frequently implies an interdisciplinary connection of such fields as linguistics, cognitive science (logic and ontology), information science and computer science. Understanding the meaning of a term such as *building*, the term analyzed in this paper, requires conceiving it as an ontology, as a summary of the network of relations implied by the

term, its concepts, attributes, relations and instances which reflect the expert disciplinary knowledge embedded in it (Montiel-Ponsoda and Aguado de Cea 2010).

Closely related to disciplinary discourse is thus the expert use of domain-specific terminology. Studies of lexis have generally assumed that different discourses structure the world in a different way and that the members of a particular speech community share a common understanding of disciplinary vocabulary (Cabr e 1998; Chung and Nation 2003; Cowie 1988; Coxhead and Nation 2001; Trimble 1985). This common understanding is grounded on the fact that vocabulary choice is intrinsically related to situation and context constraints and thus specific to a particular community, which conditions the semantic and lexical choices available to the members of a particular discourse community to lexicalize their disciplinary world with the purpose of informing about it. It is through the presupposed structuring of disciplinary lexical knowledge that discipline members can interpret the text successfully.

As argued by Montiel-Ponsoda and Aguado de Cea (2008, 2010) disciplinary lexical knowledge is formalized at a textual level not only in its members' choice of lexis but also in their use a repertoire of lexico-grammatical patterns used to express relations. In line with these authors, and drawing on the main postulates of semantics theory (Cruse 1986; Lyons 1977), this paper takes a linguistic perspective to the analysis of the lexico-grammatical patterns signaling the semantic relations of hyponymy and meronymy of the noun *building*.

2. Semantic relations: a multifaceted approach

In recent years the interest drawn by the analysis of natural language has drawn considerable attention to the study of the semantic relations of hyponymy and meronymy. Lexicographers and terminologists as well as Artificial Intelligence and Knowledge Engineering studies have concentrated in the automatic or semi-automatic detection and extraction of lexical, semantic and grammatical information. This interest has resulted in the compilation of lexicons, both of general and domain-dependent character, creating or enhancing dictionaries, thesauri or lexical ontologies, such as WordNet (Fellbaum, 1998) —the reference database for this paper—, therefore becoming a key tool for defining and translating domain-specific terms (Alfonseca and Manandhar 2002; Lindmark, Natt och Dag and Willners 2007). However, different authors (Kozareva, Riloff and Hovy 2008; McNamee, Snow, Schone and Mayfield 2008; Pasca and Harabagiu 2001; Ruiz-Casado, Alfonseca and Castells 2007) have rightly argued about the limitations of existing lexical databases and reports are frequent of notable absences of a number of new relationships of hyponymy and meronymy in databases, for example of named entities or proper nouns, most particularly of specific-domain terms.

As frequently contended, limitations respond to the need for “near human-level language understanding” (Cederberg and Widdows 2003: 114); a view also reinforced by Renouf's (2001) corpus analysis of the lexical signals of word relations, in which she

concludes that “the realizations of the superordinate-hyponym relation which are automatically generated, whilst fascinating from a linguistic-descriptive point of view [...] are largely too unconventional and content-dependent for use in information retrieval, certainly without manual intervention” (Renouf 2001: 38–39). Acknowledging these limitations, this paper proposes a linguistic analysis which attempts to shed light on the use of lexico-grammatical patterns to signal the semantic relations of hyponymy and meronymy in a specialized domain, construction and architecture.

For that purpose, and with a primarily linguistic-oriented approach, this paper draws on the work of the varied disciplines which have addressed the study of semantic relations. Particularly helpful have been the mentioned studies of computational linguistics and ontology engineering, which have created repertoires of lexico-grammatical patterns to identify hyponymic, or type-of, relations (Alfonseca and Manandhar 2002; Bodenreider, Burgun and Rindfleschn 2001; Cederberg and Widdows 2003; Gillam, Tariq and Ahmad 2007; Hearst 1992, 1998; Snow, Jurafsky and Ng 2004) and meronymic, or part-whole, relations (Girju, Badulescu and Moldovan 2003, 2006; van Hage, Kolb and Schreiber 2006). Highly valuable contributions to this paper are also terminology studies, such as Feliú and Cabré’s (2002) extensive catalogue of the prototypical linguistic markers that signal the conceptual relations of similarity (synonymy, similarity, opposition or contrast), inclusion (hyponymy), sequentiality (localization, direction, simultaneity, anteriority, posteriority), causality (cause-effect, process-result), instrument (function), meronymy, and association in specialized texts. Also in this line, Montiel-Ponsoda and Aguado de Cea’s (2008, 2010) linguistic perspective into the development of ontologies has convincingly shown that disciplinary knowledge is formalized in an ontology by creating a repository of the lexico-grammatical patterns used to express semantic relations. These lexico-grammatical patterns are defined by these authors as “linguistic schemas or constructs derived from recurrent expressions in natural language that consist of linguistic and paralinguistic elements that follow a certain syntactic order, and that permit to extract some conclusions about the meaning they express” (2008: 337).

Underlying the mentioned studies are the theoretical postulates of semantic theory (cf. Cruse 1986; Lyons 1977), with which I align myself to view semantic relations as dependant on the network of contextual relations established between words and therefore to argue that the study of vocabulary cannot be separated from the analysis of the context in which words occur. This correlation between lexis and the co-text draws on Meyer’s (2001) concept of “knowledge-rich context”, “a context indicating at least one item of domain knowledge that could be useful for conceptual analysis” (Meyer 2001: 281). A term thus establishes a number of lexical, grammatical or paralinguistic lexico-grammatical patterns, or knowledge patterns as they are also referred to by terminologists. Defining patterns as “linguistic schemas or constructs derived from recurrent expressions in natural language that consist of linguistic and paralinguistic elements that follow a certain syntactic order, and that permit to extract some conclusions about the meaning they express” (Meyer 2001: 337), the analysis of

patterns I attempt to carry in this paper is rooted in Hearst's assumption that "the structure of a language can indicate the meanings of lexical items" (Hearst 1998: 133).

Within the framework of systemic functional linguistics, Halliday (1985) and Martin (1992) apply a discourse semantics perspective to the analysis of the text, very helpful for the understanding of the rhetorical role of semantic relations I carry in this paper. Agreeing with these authors I conceive the text as a semantic unit, as a network of relationships rather than as a group of sentences without connection. Cohesion can thus be understood as a semantic concept. By developing a network of cohesive ties in which an element is dependent on the reference to another to be interpreted and effectively decoded, the text acquires its texture. Cohesive chains of reference, substitution, ellipsis, conjunction or lexical cohesion between elements make the text a complete meaningful semantic unit. At the theoretical basis of the study of contextual relations proposed in this work is Halliday and Hasan's (1976) analysis of the role lexical cohesion in text and their view of the use of lexical items (general nouns, synonyms, near synonyms, hyponyms and meronyms) in providing the text with lexical cohesion. This network-relation view of discourse is also shared by Hoey (1991), who sees text as an "interrelated package of information" (p. 48) in which lexical and non-lexical cohesive links are used to connect and organize the informational flow of the text and to make it meaningful. The semantic relations of synonymy, antonymy, hyponymy and, to a lesser extent, meronymy (Carter 1987; Carter and McCarthy 1988; Martin 1992; McCarthy 1991; Salkie 1995), the semantic prosody of lexical items and the associations (Hoey 2005; Louw 1993; Sinclair 1996) and expectations built in their meaning and the collocational and phraseological patterns of certain lexical items (Stubbs 2001) also contribute to make the text a cohesive unit.

Drawing on the multifaceted picture of this myriad of theoretical perspectives my goal in this paper is to analyze the semantic relations of hyponymy and meronymy established by the noun *building* in the discourse of architecture and construction engineering textbooks and their formalization in a number of lexico-grammatical patterns which signal the relations. This analysis will attempt to confirm the insights and expectations of the various authors mentioned with respect to the use of semantic relations and of their lexico-grammatical patterns in text. The formal analysis of these patterns will be the starting point to discover how the relations of hyponymy and meronymy contribute to the rhetorical structure of specialized discourse.

3. The Construction Textbooks Corpus: A Corpus-Based Analysis

This paper presents the results of a corpus-based study which explored 176 samples (ranging between 665 and 31,096 words, and with approximately one million total words) of some of the best-known specialized textbooks recently published in the fields of construction engineering and architecture. The *Construction Textbooks Corpus* (CTC henceforth) is thus a domain-specific corpus which covers the various knowledge areas of the discipline, including professional topics such as materials, building services,

theory of architecture, design and planning, history of architecture, sustainability, urban design, civil engineering, conservation and restoration, management or structural engineering, among others.

Reflecting the disciplinary complexity of buildings for the construction and architecture profession, the frequency analysis of the corpus also confirmed the linguistic predominance of the noun *building* (the concordancing analysis of the corpus, carried with *WordSmith Tools 4.0* (Scott 1999), yielded 4,750 occurrences). The building is thus a key product and process of this profession: these professionals make buildings and they certainly write about buildings. This prevalence, both textual and disciplinary, requires, in my view, to approach the semantic complexity of the term *building*, in the understanding that the concept integrates the knowledge about the concepts, attributes, relations and instances of buildings. With this aim, I widened the focus to include the hyponyms (the types of buildings) and meronyms (the parts of the building) of the noun *building*, thus assuming that semantic relations form a network of relations in which meaning is acquired, which in the case of a specialized domain, like the domain of construction and architecture that this paper deals with, plays the further role of unveiling the specific disciplinary semantic connotations acquired by the textual and discursal relations established.

Frequency findings were then contrasted with the WordNet lexical database (Fellbaum 1998) in search of the hyponyms and meronyms of *building* employed in the corpus. As the following chart shows, 132 of the 569 hyponyms and 167 of the 456 meronyms compiled in WordNet appear in the CTC. Hyponyms are used 1,781 times and meronyms are used 3,364 times.

	CTC			WordNet
	#	Frequency	%	#
Hyponyms	132	1,781	23.20%	569
Meronyms	167	3,364	36.62%	456

Table 1. WordNet and CTC hyponyms and meronyms.

Hyponyms and meronyms were then manually searched for the contextual relations established in the corpus with the aim of analyzing the recurrent lexico-grammatical patterns which signal the existence of a hyponymic or meronymic relation, eventually attempting to establish a correlation between lexis and the co-text in which those hyponyms and meronyms appear. It must be pointed out that the criterion for the study of patterns was to limit the analysis to the sentence and paragraph boundaries. Although hyponymy and meronymy have been found to be frequent cohesion devices used consistently throughout the text, only those hypernym-hyponym holonym-meronym chains in the context of one paragraph were considered for analysis.

4. Lexico-grammatical patterns of hyponymy and meronymy in the CTC

The CTC findings are consistent with previous studies of hyponymy and meronymy relations, particularly those of Snow et al. (2004), who claim that patterns do not usually occur within the sentence boundary but rather throughout the paragraph. The mixed use of hypernyms/hyponyms, general nouns, holonyms/meronyms, and even false hyponyms and meronyms, particularly in enumerations, stresses the contradiction between disciplinary reasoning of expert thinking and the linguistic reasoning of the database design.

4.1. Hyponymy patterns in the CTC

The 132 hyponyms found in the CTC were manually searched for their relation with the reference hypernym *building* as well as for the relation with any of its possible hypernyms listed in the hierarchical tree of the noun *building* drawn in the database, from *entity*, the most general hypernym, to *abbey*, the lowest level of hyponymy. In 124 cases the hyponymy relation is signaled by the following lexico-grammatical patterns:

Pattern		Frequency		%
such (as)	building(s) + such as + [hyponym, hyponym, ... and hyponym]	12	33	26.6
	building(s) + such as + [hyponym, hyponym, ... or hyponym]	10		
	building(s) + such as + hyponym	5		
	such buildings as + [hyponym, hyponym, ... and hyponym]	3		
	building(s) + such as + [hyponym, hyponym, ... etc.]	2		
	hyponym + such buildings	1		
other	[hyponym, hyponym, ...] and other buildings	11	19	15.3
	[hyponym, hyponym, ...] or other building	4		
	other + hypernym + ... hyponym	2		
	[hyponym, hyponym, ...] and others	1		
	buildings + one + hyponym + and + the other	1		
example	building(s) + example (of this) + be + hyponym	4	19	15.3
	example of + building + be + hyponym	2		
	hyponym + be + example (of) + building	2		
	examples of + building + (be)	3		
	building + for example + [hyponym, hyponym ... (and/or) hyponym]	2		
	building + hyponym + for example	3		
	hyponym + for example + building	1		
	for example + [hyponym, hyponym ... and hyponym]	2		
apposition	building, hyponym	6	11	8.9
	building (hyponym)	2		
	building – [hyponym, hyponym (and) hyponym]	3		

include	building + include + [hyponym, hyponym ... and hyponym]	5	9	7.3
	buildings + including + [hyponym, hyponym ... and hyponym]	4		
type	building types + [hyponym, hyponym, ... hyponym]	3	6	4.8
	[hyponym, hyponym, ...] + building type(s)	1		
	types of + buildings	2		
superlative	hyponym + superlative + building	3	6	4.8
	superlative + building + hyponym	3		
be a	hyponym + be + building	3	3	2.4
whether ... or	building + whether + hyponym + or + hyponym	2	3	2.4
	hypernym + whether + [hyponym, hyponym, ... or hyponym]	1		
especially	building + especially + [hyponym, hyponym, ... and hyponym]	2	2	2.4
like	buildings + like + [hyponym, hyponym, ... and hyponym]	2	2	1.6
compare	compare + building + with + [hyponym, hyponym, ... and hyponym]	1	2	1.6
	building + compared to + hyponym	1		
among	buildings + among them + [hyponym, hyponym, ... and hyponym]	1	1	0.8
as	hypernym + as + hyponym	1	1	0.8
e.g.	building + e.g. + hyponym	1	1	0.8
for instance	building + [hyponym, hyponym, ... and hyponym] + for instance	1	1	0.8
i.e.	building + i.e. + [hyponym, hyponym, ... and hyponym]	1	1	0.8
mostly	building + mostly + hyponym	1	1	0.8
or the many	[hyponym, hyponym, ...] + or the many + building	1	1	0.8
sort of	hyponym + sort of + building	1	1	0.8
the like	hypernym + [hyponym, hyponym, ...] + and the like	1	1	0.8
				124

Table 2. Hyponymy patterns of *building* in the CTC.

The corpus analysis validates the prevalence of Hearst's patterns (1992) (*such as*, *and/or other*, *including*, *especially*) and of three of the four patterns added by Snow et al. (2004) (*like*, *is a*, apposition) — the fourth pattern detected by these authors (NP_Y called NP_X) does not, however, apply to the noun *building* in the corpus. As also noted in the aforementioned studies, the most frequent lexico-grammatical pattern is *such as* (33 occurrences, 26.6%) (e.g. *large **building** facilities such as **shopping centers, schools, hospital complex***),¹ followed by *other* (19 occurrences, 15.3%) (e.g. ***churches and other places of worship***), appositions (11 occurrences, 8.9%) (e.g. *the new Royal **Observatory**, the **building** that was to house the meetings of the Academy of Sciences*), *including/include* (9 occurrences, 7.3%) (e.g. *several important community **buildings**, including **schools, pubs and clubs***), *is a* (3 occurrences, 2.4%) (e.g. *a **temple** is considered a measured sacred space and is an **edifice**, perceived as the residing place of ...*) and *like* (2 occurrences, 1.6%) (e.g. *other **building** typologies like **offices and shopping centres***).

However, although these patterns account for 64.5% of all cases in which a pattern signals the presence of a hyponymic relation, the CTC analysis has also helped to detect a number of new patterns used. Particularly productive among these new lexico-grammatical patterns found in the CTC is the use of *(for) example* (with 19 occurrences) to signal the hyponymic relation (e.g. *examples of institutional construction are medical clinics and hospitals, schools and universities**, *recreational centres and athletic stadiums**, *governmental buildings and houses of worship and other religious buildings*). Also used in the corpus are other new lexical markers such as prepositions (*like*), adverbs (e.g. *i.e.*) and prepositional phrases (*for instance*), all of them semantically related to exemplification. Other new patterns include the noun *type* (6 occurrences, 4.8%; with one instance of the more colloquial pattern *sort of*) (e.g. *the technology and structure of various types of buildings (skyscrapers, wide-spanning structures, etc.) developed in various ways*), the use of the superlative (e.g. *St. Peter's Cathedral in Rome, the most important building of the period*) or the use the subordinator *whether ... or ...* (e.g. *A roof, for example, does the same job on every building. Whether it covers a garden shed or a steel mill**).

4.2. Meronymy patterns in the CTC

Table 3 below lists the different patterns used in the CTC to signal the relation between the 287 meronyms and its holonym, *building*.

The meronymy relation is not frequently explicitly stated in the corpus with the use of patterns such as *part (of)* (which could be expected as paradigmatic of the relation), although seven instances are found (e.g. *all the main load-bearing parts of a building*). Other noun patterns found in the corpus are *component*, *element*, *feature* and a metaphorical use of *slice*. As previous studies have also shown (Girju et al., 2003, 2006; van Hage et al., 2006), the patterns 'Noun Phrase part + Prepositional Phrase whole' (e.g. *the walls of a building; each room in a house*) and 'Noun Phrase whole + Prepositional Phrase part' (e.g. *the house with masonry walls and timber floor and roof structures*) are the most frequent lexico-grammatical patterns of the meronymy relation, (143 occurrences, 49.8%). These prepositional phrases can be re-phrased as full relative clauses with the verbs *have*, to stress the meaning that the whole has, contains or is made up of certain parts (e.g. *modern homes* with kitchens, bathrooms and central heating* → *homes* which have kitchens, bathrooms and central heating*). They are also frequently used in the corpus as circumstance adverbials which refer to the position or location of the part in the whole (e.g. *a courtyard at the rear of the building*).

Another phrase level pattern, the Noun + Noun (N+N) pattern (59 occurrences, 20.6%) with two variants: the Npart + Nwhole (e.g. *courtyard houses*) or the Nwhole + Npart (e.g. *building extension*), is the second most frequent meronymic pattern, although it must be pointed out that 26 of those occurrences correspond to the sequence *story/storey + building(s) / house / block* / house*, etc.. N+N patterns are used in the corpus to refer to the composition of the building (e.g. *a three-bedroom house* → a house (whole) which has three bedrooms (parts)); to the location of the parts in the

building (**basement room** → *a room* (whole) *which is located in the basement* (whole)); or to express the partitive relation, i.e. to identify the parts of the building (e.g. **theatre steps** → *the steps* (part) *are part of the theatre* (whole)). The *s*-genitive pattern, a semantically parallel structure, is statistically less important for signaling meronymic relations (5 occurrences) (e.g. *the **building's** most dramatic internal feature; the **house's** sloped roof*).

Pattern		Frequency	%
NP part + PP whole NP whole + PP part	(building +) of (43) / in (33) / with (19) / within (5) / on (4) / above (3) / at (3) / inside (3) / around (2) / from (2) / to (2) / for (1) / next to (1) / up to (1) / PP + PP (12) / other prepositional phrases (9) (+ building)	143	49.8
N part + N whole	part + building	33	20.6
N whole + N part	building + part	14	
	part + part	12	
NP part + verb + NP whole NP whole + verb + NP part	have (14) / be (7) / contain (4) / arrange (2) / consist (2) / find (2) / set (2) / surround (2) / there + (be) (2) / compartment (1) / comprise (1) / distinguish (1) / exist (1) / extend (1) / finish (1) / form (1) / incorporate (1) / install (1) / join (1) / lead (1) / locate (1) / offer (1) / regard (1) / rest (1) / run (1) / serve (1) / shape (1) / situate (1) / stand (1)	57	19.9
NOUN + whole whole + NOUN	part (7) / component (6) / element (6) / feature (3) / slice (1)	23	8.0
whole + 's + part	building + 's + part	5	1.7
			287

Table 3. Meronymy patterns of *building* in the CTC.

Verbs, a sentence level meronymic pattern, are less frequently used than other patterns (57 occurrences, 19.9%). *Have* is the most frequent verb (14 instances) to signal the part or parts of the whole; it is also the most basic and explicit expression of a meronymic relation (e.g. *downtown **hotels** have as many **tennis courts***, **pools*** and **saunas** as **resorts***). Also frequent is the verb *be*, the part-whole relation being then more implicit (e.g. *the **Pentagon building** suffered only partial collapse — this was due to its being an in situ reinforced concrete **frame***). The array of other verbs used to convey meronymy falls into three categories: those which express a partitive relation, such as *contain*, *consist*, *find*, *there (be)*, *comprise*, *distinguish*, *incorporate*, *regard* and *exist*; a further group of verbs which refer to the way the *building* is designed, such as *compartment*, *finish*, *form*, *install*, *join*, and *shape*; and a final group which includes the

verbs which express the position and location of the parts in the *building*, such as *set* and *surround*.

5. A pattern-based analysis: from lexis-level to text-level findings

The analysis of the corpus has brought to the fore the mixed used of a new typology of markers which has shown to draw a multi-dimensional connection of perspectives involving boundaries, patterns, variants and combinations of markers. One very significant of these perspectives is the mixed used of patterns, with combinations of two or more patterns appearing together in the same paragraph thus reinforcing the signaling of the relation. Examples of mixed types include *such as* and *and other* together in “*other building concepts such as one-stop primary care centers, walk-in centers, polyclinics, superclinics, health kiosks, drop-in facilities, local health care resource centers, and other community based health facilities*”; *type, other* and *such as*, which are combined in “*many other commercial building types such as shops*, restaurants or pubs*”; and *type* and *for example* mix in “*components specifically designed for and applicable to a particular building type, for example schools or housing**. These patterns, typical markers of the hyponymic relation, are also found as markers of the relation of co-meronymy (e.g. *other building components, such as beams*, floors*, roofs, and other walls*). Meronymy patterns also appear together in some sequences: a prepositional pattern together with a verbal pattern in “*a wall of a residential building having one or two storeys*” and a verb pattern, a prepositional pattern and a genitive pattern are found in “*house forms in such climates are either long and thin or have a courtyard, or light well, in the centre of the house, to maximize the building’s wall area*. It is also interesting to note that the hyponymic relation between *house* and *building* is also used in combination with the meronymic relation of the house to its parts (*courtyard* and *wall*) in this last example.

It is not only lexico-grammatical typified patterns that are used to signal semantic relations. General nouns (*example, type, sort*), verbs (*include, compare*) or adverbs (*mostly, especially*) are interlinked with these studied patterns as signaling devices to tacitly imply the existence of other members of the same category (*other, or the many, among*), as, for example, the superlative or the verb *compare* semantically mark the implicit existence of other members to which a particular building is compared. The use of a hyponymy pair can be expected as a cohesion device which avoids repetition in *St. Peter’s Cathedral in Rome, the most important building of the period*, in which the hypernym *building* avoids the excessive repetition of *cathedral*.

The appearance of the hyponym can be signaled by the use of the hypernym *building* altered by a syntagmatic modifier. In *industrial buildings (...) such as cast nodes of the Renault Centre* the use of the adjective *industrial* accompanying *building* (a hyponym of the more general noun *building*) establishes yet a further taxonomical level of hyponymy: **building** → **industrial building** → **centre**. Other syntagmatic modifiers include nouns (e.g. *several important community buildings, including*

schools, pubs and clubs), prepositional phrases (e.g. *a **building** of architectural or historic importance (...)* *such as Georgian terrace in Charlotte Square*) or clauses (e.g. ***buildings** that require a great deal of repetitive activities, such as high rise offices buildings, apartments and hotels*). Numerals are also used, sometimes combined with semi-determiners, quantifiers or demonstrative pronouns as cataphoric referents, whose meaning typically anticipates the enumeration that follows:

the Jingu National Stadium, first built in 1958, was extended for the occasion (Figure 1.8) but, as in Rome, two smaller fully-enclosed halls caught international attention. These were Kenzo Tange's Swimming Arena and Sports Arena seating 4000 and 15000 spectators respectively

Lexical markers, such as *including* or *following*, are used to guide the reader in the text to the enumeration that can be expected to follow (e.g. *several important community buildings, including schools, pubs and clubs*). As the meaning of a large number of the patterns found in the corpus (*such as, among, include, and the like, consist, comprise*) clearly anticipates, the concepts of hyponymy and meronymy cannot be separated from the use of enumerations of co-hyponyms (e.g. *of major civic buildings — city halls, opera houses, museums**) and co-meronyms (*building elements — stairs, roofs, ceiling finishes, etc.*). The use of lexico-grammatical patterns and the contextual analysis involved yielded a first level of lexico-semantic findings, i.e. the appearance of a number of new hypernyms and hyponyms. The analysis of these patterns helped to detect a number of hyponyms and meronyms of *building* which were not listed by WordNet. For example, in the pattern “*the number of different buildings used in a day, a week and a year including houses, schools, shops*, factories*, libraries, sports centers, cinemas, concert halls and churches*” ‘*buildings ... including ...*’ clearly signals *shops* and *factories* as types of buildings, and thus as hyponyms of *building*. The case of *shop*, and of its synonyms *store* and *department store*, and *factory* are examples of hyponyms not listed in WordNet. Other notable absences are *university, station, museum, art gallery, granary, warehouse, steel mill, prison* and *auditorium*. Also absent in the CTC are synonyms of *house* such as *dwelling, home, housing, accommodation*.

The CTC revealed examples of named entities not included by WordNet such as the names of famous cathedrals other than Chartres (e.g. *the success of St. Denis, other cathedrals were soon begun*). Findings thus corroborate claims about the incomplete and insufficient coverage of named entities, or proper nouns, in WordNet and other lexical ontologies (McNamee et al., 2008; Pasca & Harabagiu, 2001), which can be justified because of the dynamic and evolving nature of this category. It can thus be argued that recognition of named entities will depend on the user's knowledge of the subject. Although knowledge of Spanish and German would be required to understand *Torre* as a translation of *skyscraper* and *Hofburg* as an equivalent of *Imperial Palace*, the use of hyponymy patterns facilitates assimilation of similar structures and therefore understanding in “*reconstruction, repair or renovation of buildings which may be of supreme historical importance, such as Windsor Castle (...), the Wiener Hofburg* (...)* *or Torre Windsor* (Madrid)*”. Domain knowledge needs to be employed to identify the

Petronas Towers in “constructing the tallest building in the world, as Malaysia has done with the **Petronas Towers*** (Fig. 1.4) and China is now doing with the **World Financial Centre** in Shanghai”.

Missing from the meronymy database are also meronyms. Some of these new meronyms are not listed as meronyms of *building*, but they do appear as meronyms of some hypernyms of *building* like *structure* or *construction*. These include structural elements, such as *partition*, *façade*, *foundation* or *beam* (e.g. *the interior of a building usually is compartmented into spaces or rooms by horizontal dividers (floor-ceiling or roof-ceiling systems) and vertical dividers (e.g. interior walls and partitions*)*; *this can be beneficial for the privacy of ground floor dwellings* but can also create bland facades* if raised too high*). Also missing from the meronymy database are some parts of the building which might have a parallel role to that of rooms: *corridor*, *porch* *landing* or *balcony* (e.g. *the corridor* enabled private activities to evolve and the house took on the form of an internal street, with rooms arranged in an orderly form along either side; the building was oblong and consisted of three rooms: the porch*, the main room, and the Holy of Holies in which the Ark rested*). A particular absence from the database is the case of *floor*. Although *floor* (meaning ‘a structure consisting of a room or set of rooms at a single position along a vertical scale’) is listed as a meronym of *building*, its sense “the inside lower horizontal surface, as of a room, hallway, tent, or other structure”, as in “*other building components, such as beams*, floors*, roofs, and other walls*”, does not appear.

However, the lexico-grammatical patterns described also revealed instances of ‘false hyponyms’. In other words, authors mix other nouns with a number of hypernyms, hyponyms and meronyms of *building* in the enumeration. For example, a dubious case of hyponymy is the consideration of *health kiosks* as a building as would be suggested by the pattern “*other building concepts such as one-stop primary care centres, walk-in centres, polyclinics, superclinics, health kiosks, drop-in facilities, local health care resource centres, and other community based health facilities*”. A *churchyard* is not a building although its status is made equivalent to that of *church* in “*all churches, churchyards and other ecclesiastical buildings*”. A further issue is the mixed use of hyponyms and meronyms in enumerations, such as the meronyms *flat* and *apartment* in “*buildings for human habitation, including single-family dwellings*, condominiums, multifamily townhouses, flats and apartments and high-rise apartment buildings*”. As evidenced from the corpus analysis, this pattern of ‘false hyponymy’ is more common when the hypernym is not *building*, but another noun, whose wider meaning might refer to both buildings and other constructions.

Although *building* is the most common hypernym in the corpus, there are instances of other hypernyms of *building* used, such as *structure*, *construction* or *complex*. Other nouns, not considered by WordNet as hypernyms of *building*, are also used in the CTC, the reason being that these new hypernyms include profession-specific nouns such as *project*, *program*, *brief*, *design* or *development*. Implicit in their meaning these hypernyms contain a reference to the building: *project* means ‘project for a building’, *design* implies ‘design of a building’, a *development* is a ‘group of houses or buildings’,

etc. Therefore, *project*, *program*, *brief*, *design* and *development* were considered as hypernyms of *building* in the sense that they would refer not only to a building but to other types of constructions, such as bridges or dams.

Nouns referring to the social function of buildings replace *building*. *Facility*, ‘a building or place that provides a particular service or is used for a particular industry’ or *venue*, ‘the scene of any event or action (especially the place of a meeting)’ are used as co-hyponyms of *building*. While these hypernyms add a feature of specific use to the meaning of *building*, general nouns such as *place* or *space* (Halliday & Hassan, 1976) add a very general, even vague, referential meaning only interpretable in terms of the context or by the use of modifiers, as in “**places of assembly, such as theaters, concert halls, auditoriums, and stadiums**”; “**places where people go when away from their homes such as: (...) churches and other places of worship**”.

The relation between *building* and its hyponyms is not always clearly stated and the hyponyms do not refer directly to the hypernym but to other nouns. The pattern *such as* does not refer directly to *buildings* but to *use of buildings* in “*adaptive use of buildings, such as utilizing a mediaeval convent in Venice to house a school and laboratory for stone conservation, or turning an eighteenth century barn into a domestic dwelling, is often the only way that historic and aesthetic values can be saved economically and historic buildings brought up to contemporary standards*”.

Developing from this semantic analysis, a further level of analysis emerges: the use of semantic relations in rhetorically organizing the discourse of construction engineering textbooks. Hyponymy and meronymy pairs are used in some of the key rhetorical techniques of scientific and technological discourse (Trimble, 1985): description, definition, classification and exemplification. Very frequently the part-whole relation of meronyms and *building* is used for general descriptive purposes since a full description of any object or artifact, or building in the case of the CTC, requires a description of its parts (e.g. *these monumental buildings had strong horizontal layering, mansard roofs and classical elements*). A common rhetorical type of description specifies the location and position of parts in wholes, i.e., of parts in buildings (e.g. *toilets in shops; rooms within a building*). The prevalence of prepositional phrases, frequently in homosemantic series — prepositional phrases of the same category (Biber et al. (1999: 814) — signaling different position or location relations between parts and wholes (e.g. the *front room of a two-room painting studio* on the second floor of an old brick building*) and of location verbs (e.g. *a building that extends above the ground level*) as lexico-grammatical patterns signaling meronymy further corroborates that tendency.

The frequently argued (c.f. Hoey, 1983; Trimble, 1985) close relation of semantic relations, particularly hyponymy and exemplification is ratified by the high frequency of the patterns *such as / for example / i.e / e.g.* Conversely, this also stresses the importance of this rhetorical function in specialized discourses, frequently chosen to organize information in patterns moving from general to specific information, or as Hoey (1983) has rightly shown, in the Preview-Detail and Generalization-Example patterns, which writers employ to provide their text with a logical sequence of ideas and

eventually to clarify the presentation of information, therefore facilitating effective readers' comprehension of the texts: A general abstract statement about a *building* (or another hypernym), perhaps too difficult to understand, is then exemplified with a hyponym or an enumeration of hyponyms:

GENERALIZATION [Architects have used the sunlight effect in **buildings** to create a specific atmosphere,] **EXAMPLE 1** [as for example the shafts of light entering the south side of our great *cathedrals*;] **EXAMPLE 2** [and on a much smaller scale the use in *houses* of daylight and sunlight entry from above to provide necessary functional light to interior areas, where otherwise little natural light would be available.]

or meronyms: the idea expressed by the holonym is exemplified with the specific mention of the parts:

GENERALIZATION [Affinities, on the other hand, indicate activities that share something besides circulatory convenience, and thus may tend toward one another in a **building** for reasons of performance or constructability.] **EXAMPLE 1** [Here, a good residential example is *kitchens* and *bathrooms*.]

But exemplification does not simply mean providing a random list of building types or parts. It also implies the classification of buildings according to their type, as in example 1, or according to their parts, as in example 2:

The occupancy group to which a **building** official assigns a **building** depends on the use to which the **building** is put. Typical classifications include one- and two-story *dwellings**; *apartment buildings*, *hotels*, *dormitories*; industrial **buildings** with noncombustible, combustible, or hazardous contents; *schools*; *hospitals* and nursing *homes*; and **places** of assembly, such as *theaters*, concert *halls*, *auditoriums*, and *stadiums*.

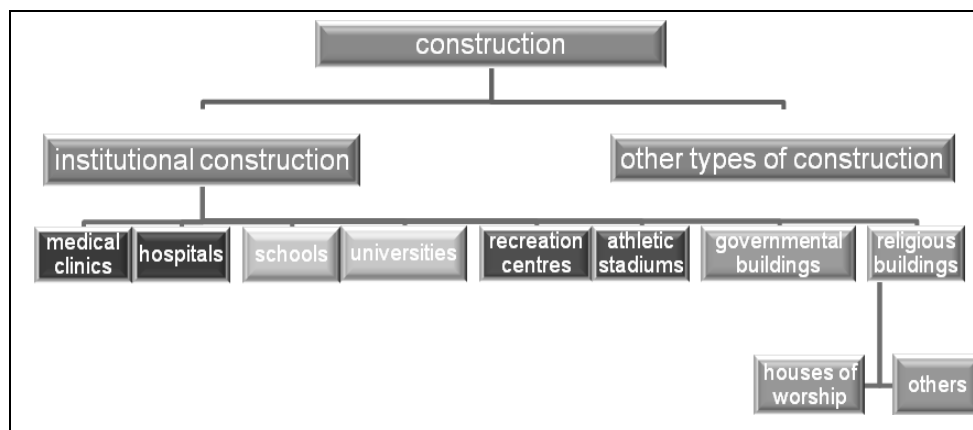
Price is no indicator — expensive **hotels** may have small *rooms*, while budget **hotels** have larger, better appointed *rooms*.

The semantic relations of hyponymy and meronymy thus prove to play a key role in creating taxonomies of buildings. The following example

Examples of institutional **construction** are [medical *clinics* and *hospitals*], [*schools* and *universities*], [recreational *centres* and athletic *stadiums*], *governmental buildings* and [*houses of worship* and other religious **buildings**].

illustrates how co-hyponyms are structured into different sub-sets according to their meaning similarity. Because these buildings are not at the same level, co-hyponyms are necessary to label the right level in the hierarchy; the use of *and* makes it clear the unequal relation. The hypernym *institutional construction* is a tacit class, which implicitly suggests that there are other types of construction. At the same time it implies the existence of a further level of hypernymy, *construction*. *Houses of worship and other religious buildings* uses a signaling marker which Darian (1997: 826) calls the

“etcetera factor”, the words, overtly stated or implied, that “indicate there are other unstated members of a class”. The taxonomical relation can be illustrated in the following diagram:



Creating taxonomies responds to the aim of ordering and explaining, defining and classifying the world is a common purpose of the techno-scientific community, which as has been pointed out in the literature (Martin 1993; Wignell, Martin and Eggins 1993) pervades their discourse and requires the creation of the specific vocabulary to designate and label it; in other words, the scientific process is realized through the linguistic process involved. The construction of hierarchies with categories and subcategories of types and parts, and the use of suggested tacit classes, requires command of the necessary abstract knowledge about their characteristics; knowledge only shared by expert members of the discipline. Definition and classification allow architects to produce taxonomies which, although originally verbal, are transformed into taxonomies of building types and parts, architectural styles, functions or users. If we assume that specialized texts respond to the need to transmit expertise, particularly explicit in examples like the one before, for a construction and architecture expert, it is particularly relevant to clarify the unequal relation between the buildings, the totally different implications in terms of design, functionality or physical features. Although to the naked eye they might all be categorized generically with the hypernym ‘construction’, not all of them are the same for these professionals, who have the disciplinary knowledge to be able to group them in adequate subcategories.

Although this paper has focused on the use of lexico-grammatical patterns to signal the appearance of semantic relations, it must be pointed out that corpus findings show that hyponymy and meronymy relations, though marked in a number of cases, as seen above, most frequently appear in free text, thus relying on the semantic relation between hyponymic and meronymic pairs to provide lexical cohesion:

First, the skills of medical doctors and engineers are especially needed in the immediate post-crisis environments to treat the wounded and traumatized, refit *hospitals* and health *centres*, repair roofs and walls to make **buildings** habitable, and ensure vital services of water purification and wastewater treatment.

Since the co-hyponyms *hospital* and *health centres* inherit the meaning of the hypernym *buildings*, their use means a repetition of meaning, rather than a repetition of the same word. Similarly, the mention of the meronyms *roofs* and *walls* implicitly carries a mention to the building of which they are part. Beyond the purely lexical role of identifying the particular type or part of a building, the semantic relations of hyponymy and meronymy, whether signaled by any type of lexico-grammatical pattern, undoubtedly acquire an organizing role in the discourse of construction engineering because hyponym/hypernym pairs, and, to a lesser extent, meronym/holonym ones are a frequent resource to make the text a cohesive whole. The taxonomical relation established between the pairs provides semantic and lexical cohesion to the text while avoiding repetition (Halliday 1985; Halliday and Hasan 1976; Hoey 1983, 1991; Martin 1992; Salkie 1995). Cohesion relies on the semantic power of hyponymic and meronymic pairs to connect sentences and also to interpret the relation between them. It also displays expert knowledge of the subject, which is addressed to the uninitiated reader, who needs clarification and guidance, on the one hand, but it is also addressed to inter and intra-specialists (Cloître and Shinn 1985), to whom command of the content and rhetorical practices of the discipline is addressed.

6. Some concluding remarks

Drawing on the approach to semantic relations of such fields as semantics, ontology engineering or terminology, and focusing on the analysis of the noun *building* in a corpus of construction engineering and architecture textbooks, this paper has taken a linguistic perspective which has sought to provide a view on the use of lexico-grammatical patterns to signal the appearance of hyponymy and meronymy relations in the discourse of this disciplinary community. The lexico-grammatical patterns markers of the hyponymic and meronymic relations of *building* have proved to play a dual role in the corpus. The patterns have a semantic function, that of signaling hyponyms as types of building and meronyms as parts of the building. At the same time they fulfill a metalinguistic, discourse-organizing function, equivalent, as also claimed by Renouf (2001), to the lexical signaling of logical relationships at a clause level, marking some of the most important rhetorical functions of the techno-scientific discourse (Trimble, 1985), such as exemplification of an abstract, general or complex concept, definition and classification.

As regards the use of lexico-grammatical patterns, the contextual analysis of the network of relationships established between *building* and its hyponyms and meronyms has shown the recurrence of a number of lexico-grammatical patterns, in which the semantic reference component is frequently signaled at a textual level. Findings have corroborated the validity of previous studies of hyponymy and meronymy patterns (Girju et al. 2003, 2006; van Hage et al. 2006; Hearst 1992; 1998; Snow et al., 2004) and extended the repertoire with further, though more lexico-grammatically complex,

patterns. However, the difficulty of applicability of the new patterns for the automatic or semi-automatic extraction of the semantic relations of hyponymy and meronymy cannot be ignored. Although some of the patterns are quite straightforward, many of the new patterns added by this study show a higher level of lexico-grammatical complexity, which might lead to erroneous or ambiguous interpretations, particularly if the criteria of frequency of occurrence, accuracy, simplicity, popularity and productivity, frequently signaled as the criteria for the acceptance and reliability of a pattern (Hearst 1992; Kozareva et al. 2008) are to be fulfilled. These lexico-grammatical patterns should then be manually searched and filtered to avoid what has been referred to as “noise” (“contexts that the system ‘thinks’ illustrate the specified conceptual relation, but which really do not” Meyer 2001: 289), and linguistically-informed restrictions should be applied to their extraction if meaningful results are to be achieved. As this paper has attempted to show, a linguistic intervention would be even more necessary in the case of specialized domains in which disciplinary complexity adds to the lexico-grammatical complexity of discourse.

Findings have also highlighted a multidimensional blend of types of markers involving combination of patterns, applied not only to the term *building* itself but also to hypernyms of it, thus crossing boundaries and very frequently showing the ambiguity of both semantic relations and markers. This complexity of relations and pattern search can hardly be denied; their textual realization is probably equally unpredictable since, as corroborated by Halskov and Barrière (2008: 24) “they are part of natural rather than controlled or artificial language. There is virtually no limit to the creativity with which human beings express themselves, even when conveying specialized knowledge to each other”. To this respect two issues are worth mentioning: first the varied character of form and content of free text and, second, the fact that few pattern-based semantic relations of *building* appeared within the sentence, but rather throughout the whole text, as they are frequently used as cohesion devices (cf. Cruse 1986; Halliday and Hasan 1976; Lyons 1977). The semantic relations that the noun *building* establishes with its hyponyms and meronyms, though sometimes marked by lexico-grammatical patterns which signal the relation, most frequently appear in free text, thus relying on the semantic relation between hyponymic or meronymic pairs to provide lexical cohesion. The mixed use in enumerations of hypernyms, hyponyms, general nouns, meronyms, and even false hyponyms and meronyms of *building*, stresses the contradiction between disciplinary reasoning of expert thinking about buildings and the linguistic reasoning of the database design.

The analysis has thus helped to draw some formal observations about the occurrence of lexical markers and the anaphoric devices employed in the patterns but also some semantic observations, about the limited coverage of the specialized discourse of construction engineering in the database. The absence of some notable hyponyms and meronyms of *building* has confirmed claims about the general character of the WordNet database (Bodenreider et al. 2001; Cederberg and Widdows 2003; Kozareva et al. 2008; McNamee et al. 2008; Pasca and Harabagiu 2001; Ruiz-Casado et al. 2007), understandably lacking the specificity required to cover domain-specific terminology.

The CTC, the corpus employed for the work presented, is certainly a specialized corpus of the domain of engineering construction and architecture, therefore, a compendium of its members' knowledge, not always shared by members of other discourse communities, including those linguists involved in the development of large lexical databases such as WordNet. The particular nature of a specialized domain like that of the construction and architecture profession and their dynamic and ever changing nature would explain the difficulty of full coverage in the database. At the same time, recognition of the lowest levels of specificity, the single particular building, requires deep knowledge of the discipline, only available to insiders, and thus not compiled by WordNet. With this paper I have attempted to contribute a lexical insight into the specialized domain of construction engineering and architecture which should help to enlarge the lexicon of databases or dictionaries by adding certain proper nouns, new words, or new technical and sub-technical senses to general words.

In sum, the use of hyponyms and meronyms suggests that specialized writing requires specialized lexis to designate buildings. The semantic relations of hyponymy and meronymy, with their taxonomical organization of buildings and their parts offer the specificity demanded by the specialized discourse of the community of construction engineers. The CTC analysis has shown that the ontological knowledge about a building is constructed with the knowledge about different types of building and the parts that form them. This knowledge has proved to be specific to the domain of the construction engineering profession, for which only certain buildings and certain parts are relevant and others do not, hence reflecting at the textual level the disciplinary knowledge about buildings. As corroborated by the present study, linguistic research of specialized domains can fruitfully contribute to enlarge the lexicon of databases or dictionaries by widening the range of proper nouns and named entities, and by adding new specialized lexis or new technical and sub-technical senses to general terms.

Note

1. building and hypernyms of building

hyponyms / meronyms

* hyponyms / hypernyms not listed in WordNet

meronyms not listed in WordNet

lexico-grammatical patterns

References

- Alfonseca, Enrique and Sureh Manandhar (2002): "Improving an Ontology Refinement Method with Hyponymy Patterns". *Proceedings of the 3rd International Conference on Language Resources and Evaluation*. Las Palmas, Spain, 235–239.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan (1999): *Longman Grammar of Spoken and Written English*. Harlow: Longman.
- Bodenreider, Olivier, Anita Burgun and Thomas Rindfleschn (2001): "Lexically-suggested

- hyponymic relations among medical terms and their representation in the UMLS". *Proceedings of the Terminology and Artificial Intelligence Conference*. Nancy, France, 11–21.
- Cabré, M. Teresa (1998): *Terminology: Theory, Methods and Applications*. Amsterdam/Philadelphia: John Benjamins.
- (2008): "El principio de poliedricidad: la articulación de lo discursivo, lo cognitivo y lo lingüístico en Terminología (I)". *Ibérica* 16: 9–36.
- Carter, Ronald (1987): *Vocabulary: Applied linguistic perspectives*. London: Allen & Unwin.
- Carter, Ronald and Michael McCarthy (eds.) (1988): *Vocabulary and language teaching*. London: Longman.
- Cederberg, Scott and Dominic Widdows (2003): "Using LSA and Noun Coordination Information to Improve the Precision and Recall of Automatic Hyponymy Extraction". *Proceedings of the 7th Conference on Natural Language Learning*. Edmonton, Canada, 111–118.
- Chung, Teresa and Paul Nation (2003): "Technical Vocabulary in Specialised Texts". *Reading in a Foreign Language* 15, 2: 103–116.
- Cloître, Michel and Terry Shinn (1985): "Expository Practice: Social, Cognitive and Epistemological Linkages". In T. Shinn and R. Whitley, eds., *Expository science: Forms and functions of popularization*. Dordrecht: Reidel Publishing Company, 31–60.
- Cowie, Anthony (1988): "Stable and Creative Aspects of Vocabulary Use". In R. Carter, and M. McCarthy, eds., *Vocabulary and Language Teaching*. London: Longman, 126–139.
- Coxhead, Averyl and Paul Nation (2001): "The Specialised Vocabulary of English for Academic Purposes". In J. Flowerdew, ed., *Research Perspectives on English for Academic Purposes*. Cambridge: Cambridge University Press, 252–267.
- Cruse, Allan (1986): *Lexical Semantics*. Cambridge: Cambridge University Press.
- Darian, Steven (1997): "The Language of Classifying in Introductory Science Texts". *Journal of Pragmatics* 27, 6: 815–839.
- Feliú, Judith and M. Teresa Cabré (2002): "Conceptual Relations in Specialized Texts: New Typology and an Extraction System Proposal". *Proceedings of the 6th International Conference on Terminology and Knowledge Engineering*. Nancy, France, 45–49.
- Fellbaum, Christiane (ed.) (1998): *WordNet. An Electronic Lexical Database*. Cambridge, MA: The MIT Press.
- Gillam, Lee, Mariam Tariq and Khurshid Ahmad (2007): "Terminology and the Construction of Ontology". In F. Ibekwe-SanJuan, A. Condamines and M. T. Cabré, eds., *Application-driven Terminology Engineering*. Amsterdam/Philadelphia: John Benjamins, 49–74.
- Girju, Roxana, Adriana Badulescu and Dan Moldovan (2003): "Learning Semantic Constraints for the Automatic Discovery of Part-Whole Relations". *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology*. Vol. 1. (Edmonton, Canada). Morristown, NJ: Association for Computational Linguistics, 1–8.
- (2006): "Automatic Discovery of Part-Whole Relations". *Computational Linguistics* 32, 1: 83–135.
- van Hage, Willem, Hap Kolb and Guus Schreiber (2006): "A method for Learning Part-Whole Relations". *Proceedings of the Semantic Web Conference*. Athens, GA, 723–735.
- Halliday, Michael (1985): *An Introduction to Functional Grammar*. London: Edward Arnold.
- Halliday, Michael and Ruqaiya Hasan (1976): *Cohesion in English*. London: Longman.
- Halskov, Jakob and Caroline Barrière (2008): "Web-Based Extraction of Semantic Relation

- Instances for Terminology Work". *Terminology* 14, 1: 20–44.
- Hearst, Marti A. (1992): "Automatic Acquisition of Hyponyms from Large Text Corpora". *Proceedings of the Fourteenth International Conference on Computational Linguistics*. Nantes, France, 539–545.
- (1998): "Automated Discovery of WordNet Relations". In C. Fellbaum ed. *WordNet. An Electronic Lexical Database*. Cambridge, MA: The MIT Press, 131–151.
- Hoey, Michael (1983): *On the Surface of Discourse*. London: George Allen & Unwin.
- (1991): *Patterns of Lexis in Text*. Oxford: Oxford University Press.
- (2005): *Lexical Priming: A New Theory of Words and Language*. London: Routledge.
- Kozareva, Zornitsa, Ellen Riloff and Eduard Hovy (2008): "Semantic Class Learning from the Web with Hyponym Pattern Linkage Graphs". *Proceedings of Association for Computational Linguistics: Human Language Technologies Conference*. Columbus, OH, 1048–1056.
- Lindmark, Kerstin, Johan Natt och Dag and Caroline Willners (2007): "Lexical Semantics for Software Requirements Engineering – A Corpus-Based Approach". In Roberta Facchinetti, ed., *Corpus Linguistics 25 Years On*. Amsterdam: Rodopi, 365–385.
- Louw, Bill (1993): "Irony in the Text or Insincerity in the Writer? The Diagnostic Potential of Semantic Prosodies". In M. Baker, G. Francis and E. Tognini-Bonelli, eds., *Text and Technology: In Honour of John Sinclair*. Amsterdam/Philadelphia: John Benjamins, 157–176.
- Lyons, John (1977): *Semantics*. Cambridge: Cambridge University Press.
- Martin, James (1992): *English Text: System and Structure*. Philadelphia: John Benjamins.
- (1993): "Literacy in Science: Learning to Handle Text as Technology". In M. A. K. Halliday and J. R. Martin, eds., *Writing Science: Literacy and Discursive Power*. Pittsburgh: University of Pittsburgh Press, 166–202.
- McCarthy, Michael (1991): *Discourse Analysis for Language Learners*. Cambridge: Cambridge University Press.
- McNamee, Paul, Rion Snow, Patrick Schone and James Mayfield (2008): "Learning Named Entity Hyponyms for Question Answering". *Proceedings of the Third International Joint Conference on Natural Language Processing*. Hyderabad, India.
- Meyer, Ingrid (2001): "Extracting Knowledge-Rich Contexts for Terminography. A Conceptual and Methodological Framework". In D. Borigault, C. Jacquemin and M. C. L'Homme, eds., *Recent Advances in Computational Terminology*. Amsterdam: John Benjamins, 279–302.
- Montiel-Ponsoda, Elena and Guadalupe Aguado de Cea (2008): "Using Natural Language Patterns for the Development of Ontologies". In P. Sánchez-Hernández, P. Pérez-Paredes, P. Aguado-Jiménez and R. Criado-Sánchez, eds., *Researching and teaching specialized languages: new contexts, new challenges. Proceedings of the AELFE 2008 Conference*. Editum, Universidad de Murcia, 332–345.
- (2010): "Using Natural Language Patterns for the Development of Ontologies". In V. Bhatia, P. Sánchez-Hernández and P. Pérez-Paredes, eds., *Researching Specialized Languages*. Amsterdam: John Benjamins, 211–230.
- Pasca, Marius and Sanda Harabagiu (2001): "The Informative Role of WordNet in Open-Domain Question Answering". *Proceedings of the North American Chapter of the Association for Computational Linguistics Workshop on WordNet and other Lexical Resources*. Pittsburgh, PA.
- Renouf, Antoinette (2001): "Lexical Signals of Word Relations". In M. Scott and G. Thompson, eds., *Patterns of Text: In Honour of Michael Hoey*. Amsterdam: John

- Benjamins, 35–54.
- Ruiz-Casado, Maria, Enrique Alfonseca and Pablo Castells (2007): “Automatising the Learning of Lexical patterns: An Application to the Enrichment of WordNet by Extracting Semantic Relationships from Wikipedia”. *Data & Knowledge Engineering* 61, 3: 484–499.
- Salkie, Raphael (1995): *Text and Discourse Analysis*. London: Routledge.
- Scott, Mike (1999): *Wordsmith Tools Version 3*. Oxford: Oxford University Press.
- Sinclair, John (1996): “The Empty Lexicon”. *International Journal of Corpus Linguistics* 1, 1: 99–119.
- Snow, Rion, Daniel Jurafsky and Andrew Ng (2004): “Learning Syntactic Patterns for Automatic Hypernym Discovery”. In L. Saul, Y. Weiss and L. Bottou, eds., *Advances in Neural Information Processing Systems, 17. Proceedings of the 2004 Conference*. Cambridge, MA: The MIT Press, 1297–1304.
- Stubbs, Michael (2001): *Words and Phrases: Corpus Studies in Lexical Semantics*. Oxford: Blackwell.
- Trimble, Louis (1985): *English for Science and Technology. A Discourse Approach*. Cambridge: Cambridge University Press.
- Wignell, Peter, James Martin and Suzanne Eggins (1993): “The Discourse of Geography: Ordering and Explaining the Experiential World”. In M. A. K. Halliday and J. R. Martin, eds., *Writing Science: Literacy and Discursive Power*. Pittsburgh, PA: University of Pittsburgh Press, 136–165.