

Linguistic Issues in Language Technology – LiLT
December 2009

**The Same Semantic Relations Link
Structurally Different Realizations
of Concepts**

Vivi Nastase
Stan Szpakowicz

Submitted, November 2008
Revised, December 2009
Published by CSLI Publications

The Same Semantic Relations Link Structurally Different Realizations of Concepts

VIVI NASTASE, *EML Research gGmbH, Heidelberg, Germany*
(*nastase@eml-research.de*)

STAN SZPAKOWICZ, *University of Ottawa, Ottawa, Canada & Polish Academy of Sciences, Warsaw, Poland*
(*szpak@site.uottawa.ca*)

Abstract

To make sense of an utterance, people identify in its linear linguistic expression the concepts and the connections between them. A concept normally has a lexical realization; connections between concepts often do not, but they are perceived even without the benefit of lexical cues. Making these connections – called *semantic relations* in the field of natural language processing – relies on the form and structure of linguistic expressions, and the concepts these expressions evoke. This implies two levels: the level of the text, the linguistic expression with its form and (grammatical) structure, and the level of the concepts which the speaker wants to convey.

An overview of the literature shows that semantic relations are, for pragmatic reasons, a means to an end – extract information, explain the links between the head of a phrase and its arguments, and so on – and that is why they are analyzed from the perspective of what they link. At the text level, the process of semantic relation analysis is informed by syntactic elements – noun phrases, verbs and their arguments, clauses and so on – thus differentiating semantic relations based on the complexity of the syntactic constructions in

which their arguments appear. At the conceptual level, the same semantic relation is assigned to pairs of concepts, regardless of their surface expression. The process can be said to disregard the implications of having syntactic constructions of various complexity correspond to the concepts linked.

In this article, we propose to put semantic relations first: analyze them, determine what constraints they place on the concepts they connect, and how those concepts can be lexicalized. Lexicalization takes place via expressions of increasing syntactic complexity: phrases, clauses and multi-clause sentences. Next, we show how the linguistic phenomena involved in producing different lexicalizations explain – in a systematic manner – how semantic relations can have instances in syntactic constructions of various complexity. We focus on *binary* semantic relations between concepts/textual elements within sentences.

This kind of analysis leads to a better understanding of the relations themselves and to a systematic account of phenomena related to their occurrence in texts. It reveals some of the assumptions and linguistic gaps people fill when they recognize relations in text. From the computational point of view of text processing, such a solid basis of the analysis of semantic relations adds consistency. Evidence for a particular relation can come from all its instances in a text, regardless of the syntactic form of the concepts it connects. Knowledge of the expected concepts and their syntactic realization may signal the presence of covert or implied information, which we can then work to retrieve. Assigning a semantic relation should be a conscious choice, with the understanding of what implications such a tag has both for the implicitly and explicitly expressed elements of a concept.

1 Introduction

To make sense of a message in a spoken or written utterance, we must understand the concepts mentioned and the connections perceived between them. These connections show how we recognize the contribution of the concepts to the utterance's overall meaning. It is an easier task to identify entities, which are normally lexicalized. Connections between concepts are of various types, and identifying them seldom relies on simple and unambiguous lexical cues, if any. Among the types of connections which should be established between entities for a thorough semantic processing of text, we focus on semantic relations, which describe how entities interact. From a syntactic standpoint, these relations are perceived between the constituents of a sentence – words within a phrase, phrases, clauses. We do not touch upon discourse relations which can span sentence boundaries.

The practical purpose of analyzing semantic relations in text has influenced the view of semantic relations now mainstream in natural language processing: they hold between textual elements of various types. This situates the manifestation of semantic relations in syntax, and may suggest their analysis separately in different syntactic constructions: in noun phrases (Girju et al., 2009), in clauses (Màrquez et al., 2008) or between clauses (Mann and Thompson, 1988, Hovy, 1993).

There is an alternative view: semantic relations should be “above” syntax, and just connect concepts. This is evident in PropBank (Palmer et al., 2005) and NomBank (Meyers et al., 2004), which use similar sets of semantic relations to annotate relations between a verb and its arguments within a clause, and a noun and its modifiers in a noun phrase. From a syntactic point of view, NomBank annotators assume that relations between verbs and their arguments can be (are) the same as relations between a noun and its modifiers. For nominalized verbs, making the case for such a position may not even be necessary (Quirk et al., 1985). The assumption, however, also spills over to situations in which it is necessary to justify that the semantic relation is indeed the same. The relation could be causal (Givon, 1975), temporal (Allen, 1984), spatial (Talmy, 1985) or participatory (Fillmore, 1968). It is assigned to a pair of expressions because their context indicates the type of interaction between the concepts: *because of*, *since*, *due to* may indicate causality, *while*, *since*, *until*, *to* may indicate temporality, and so on.

To explain the purpose of this article, let us consider a few examples.

I will arrive at 11am. ↔ *I will arrive when you arrive.* (1)

I will arrive at a certain time: *at 11am.* or *when you arrive.* A time point can be referred to either explicitly by a temporal entity, or by a punctual occurrence which is used as a temporal reference.

I will meet you in the office. ↔
I will meet you where we met last time. (2)

I will meet you at a certain place: in the office or where we met last time. A location can be referred to explicitly by a location entity, or by an occurrence which unfolds at the spatial coordinates which we want to convey.

Eating sweets before dinner spoils your appetite. ↔
Sweets before dinner spoil your appetite. (3)

One can talk about *sweets* and actually mean *eating sweets*. An occurrence can be referred to directly, or through an entity which, in the given context, suggest the intended occurrence to the hearer.

This article reflects on the implications of taking these two views of a text together – as a linear linguistic expression, and as a manifestation of concepts and their interactions. We analyze the assumptions made when considering that semantic relations link phrases with specific syntactic forms, and the assumptions made about the type of concepts which a semantic relation connects, regardless of the syntactic forms in which they manifest themselves.

We approach semantic relation analysis from the standpoint of natural language processing, and with the practical purpose of a consistent and well justified analytic process. We seek phenomena which explain consistent combinations of constraints which semantic relations impose on the type of concepts¹ they connect, and the syntactic expressions which convey those concepts. We look at the mapping between concepts (occurrences, entities, attributes of occurrences and attributes of entities) and syntactic constructions (clauses, noun phrases, adverbial phrases and adjectival phrases). We pinpoint linguistic phenomena which explain why and how a concept of a certain type, for example an occurrence, can be expressed by a full clause, but also by a noun phrase or even an adjectival phrase.

The article continues in Section 2 with a review of semantic relations and concepts in previous work, and our take on the combination of the two. Section 3 presents phenomena which allow concepts to take various syntactic forms, and the repercussions of those phenomena on “splitting” the information conveyed into overtly and covertly (or implicitly) expressed. The section ends with a brief foray into the way in which concepts and semantic relations manifest themselves in several languages. Section 4 presents general considerations on semantic relations found in the literature; we look at how relations manifest themselves in noun phrases, clauses and multi-clause con-

¹In this work, the term *concept* encompasses occurrences, entities, attributes of occurrences and entities. The term *occurrence* covers all types of events, actions, activities, processes, states and accomplishments – see (Allen, 1984). We consider an occurrence to be not only what is usually expressed by a verb, but the entire situation, including its participants and attributes. What the verb expresses will be the *core* of the occurrence.

structions. Our discussion builds on the lexicalization phenomena (identified in the preceding section) which allow semantic relations to be realized by syntactic constructions of varying complexity. In Section 5 we summarize the work presented, briefly discuss the effect which such a systematic approach to semantic relation analysis can have on text analysis, and draw a few conclusions.

2 Concepts and Relations

Communication via utterances implies giving a linguistic form to a thought or idea. That, in turn, implies (among other things) that concepts are lexicalized in the speaker's mind, and lexicalizations are arranged in a sequence which will allow a listener to retrieve the intended meaning. There may be a broad consensus about what a concept intuitively is, but a precise definition of the term *concept* seems quite elusive. Jackendoff (1989, p. 305) proposes the view that definitions depend on their purpose or context:

Asking a psychologist, a philosopher or a linguist what a concept is is much like asking a physicist what mass is. An answer cannot be given in isolation. Rather, the term plays a certain role in a larger world view which includes the nature of language, of meaning, and of mind. Hence the notion a concept cannot be explicated without at the same time sketching the background against which it is set; and the "correctness" of a particular notion of concept cannot be evaluated without at the same time evaluating the world view in which it plays a role.

To retrieve the concepts expressed is an important element of reaching the meaning of an utterance. It is just as important to understand the connections between them:

The connection is indispensable to the expression of thought. Without the connection we would not be able to express any continuous thought and we could only list a succession of images and ideas isolated from each other, and without any link between them (Tesnière, 1959).

This section continues with an explanation of the definition of *concept* we work with throughout the paper, followed by a brief review of the history of semantic relations in linguistics and natural language processing.

2.1 Concepts

The goal of the work presented here is to find a bridge between concepts and their linguistic expressions. From this perspective, it appears appropriate to adopt a coarse-grained view of concepts, and to distinguish between occurrences (see footnote 1), entities (concrete and abstract objects), and attributes of occurrences and entities. On this view, we can map concepts onto syntactic constructions. We have occurrences, entities and attributes on one side, and constructions on the other: clauses (verb-centred constructions), noun phrases, adverbial phrases and adjectival phrases.

Concept is a loaded term. Margolis and Laurence (2002) comprehensively review theories about concepts, “the most fundamental constructs of the mind”. The articles in the collection show how theories of concepts have appeared and changed, following philosophical, psychological and cognitive explorations into the nature of the mind. Each theory has its strengths and weaknesses, but none of them can by itself answer all the questions about concepts and the way in which people use them. Our coarse-grained view of concepts does not subscribe to, imply or contradict any of these theories. We make no assumptions about the phenomena of concept representation, acquisition, categorization, epistemic justification, analytic entailment and reference determination.

We focus on how concepts can be expressed through linguistic expressions with different syntactic constructions – how, for example, an occurrence is expressed not only by a clause, but also by a noun or an adjectival phrase – and we consider the linguistic phenomena which explain the suitable mappings.

2.2 Semantic Relations

Semantic relations have appeared at four points in the history of linguistics – we will list them shortly – before becoming a mainstream notion in linguistic and in language processing. There has been a shifting perspective. Sometimes the relations themselves were the central concern. Sometimes what mattered more were the entities involved: the focus then was on the *roles* of the entities with respect to a central concept, and the actual relation was implied.² The analysis of semantic relations originated from considerations on grammatical cases, and that is why it is not uncommon to see semantic relations referred to as *cases*; Fillmore’s work (1968) is perhaps the best known example.

1. Relations were first mentioned in the work of Pāṇini, a Sanskrit grammarian, some 2500 years ago³. He identified non-verbal relations, *sesa*, and six verbal relations, *karakas* (including **karta** ‘agent’, **karana** ‘instrument’ and **karma** ‘object’), expressed either by suffixation or compounding, or marked by case endings (Misra, 1966).

2. Semantic relations reappeared only in the 20th century, in the work on syntax of Tesnière (1959). In his view, the verb is the most important element of the sentence, and its arguments are its “dependents”. These arguments can play two roles: **actants** or **circumstances**. Actants and circumstances are tied to syntactic constituents: to subject / direct object / indirect object and to complements, respectively.

3. In the 1960s, Jeffrey Gruber of Noam Chomsky’s group at the MIT introduced six labels to describe the types of entities in a sentence, relative to the main verb (Gruber, 1965): **theme**, **agent**, **goal**, **location**, **accompaniment**

²Further on, when the distinction between *role* and *relation* is important, we write the relation name in SMALLCAPS, and the role in *italics*.

³His dates range from 5th to 7th century BC in various accounts.

and **direction**. Gruber distinguished causative and permissive agents, contingent upon the meaning of the main verb.

4. Starting from a grammatical analysis of cases, as Pāṇini did, Fillmore's analysis (1968) veered off into semantics and led him to the definition of a set of relations between the main verb of the clause and its arguments. He identified a list of "needed" cases: **agentive, instrumental, dative, factitive, locative and objective**⁴. He named the verb-argument relations *cases*. Indeed, they reflect the connection between case markers in inflected Indo-European languages and the arguments of the main verb. In modern English most of these markers have disappeared, but there remain the roles which the arguments play in the semantics of the clause.⁵

Fillmore proposed that cases be regarded as universal relations, whose function is to designate the possible semantic relationships between the main verb of the sentence and its nominal arguments, and that the deep structure advocated by transformational grammarians be replaced with a case structure. The postulated case system is semantic in nature, but the evidence which supports its introduction is surface-syntactic. Cases explain grammatical phenomena – the selection of the subject, the well-formedness of a sentence, and so on.

What a *list* of cases should contain has been a much debated issue. The length varies from several to hundreds. Dowty (1991) brings up a problem of role fragmentation, as illustrated by different refinements in the literature of the agent role: *agent* and *actor* (Jackendoff, 1983); *agent* and *effector* (van Valin, 1990); *volitive, effective, initiative* and *agentive* (Cruse, 1973); fourteen different groups proposed by (Lakoff, 1970). As we already noted, Gruber (1965) identified two possible types of agent: *causative* and *permissive*. *Theme* can also be divided into *incremental, non-incremental* and *holistic* (Dowty, 1991).

The analysis of semantic relations between a verb and its arguments carries over – through the analysis of nominalized verbs – to a noun and its modifiers (Chomsky, 1970), (van Valin and LaPolla, 1997, sec. 7.2). Quirk et al. (1985, sec. 17.51-54) also support the view that the same labels describe the interaction between the head of the phrase analyzed and its arguments, whether it be a noun phrase or a verb phrase.

Schank (1975) in his conceptual dependency representation was the first to introduce relations between acts as a whole, as opposed to relations between an act and its arguments. For example, there is a link between acts representing transfer of money and transfer of possession, implied in verbs such as *buy*. Such links are not established between two specific elements of a sentence, but arise from the representation of words with Schank's conceptual

⁴Fillmore recognized that the list was not comprehensive. Further in the article he mentioned a **benefactive** case, and suggested that others exist.

⁵For the purpose of semantic relation analysis in our work, the distinction between the arguments/modifiers and the adjuncts of a head is secondary: we refer to all of them as arguments.

primitives.

To conclude this overview, let us note that semantic relations help grasp the meaning behind a sentence by explaining the connections between entities and their roles with respect to a central concept. The grammatical notions of *subject*, *object*, *indirect object* were not intended for such semantic interpretations. Subjects, for example, may denote entities actively involved in an event, passive onlookers, or those affected by the event. The traditional notion of case has been the starting point in the development of a set of roles which described the involvement of different entities in the situation captured in a sentence. The idea of semantic relations, initially based on grammatical cases, has appeared in linguistics independently at four different times. From these theories, the idea of semantic relations has expanded first to all arguments of the main verb, not just those for which it subcategorizes, and afterwards to noun phrases, through nominalizations, to clauses and beyond the borders of the sentence to larger units of text.

Present-day research shows the influence of various early theories. PropBank (Palmer et al., 2005) and NomBank (Meyers et al., 2004) use relations in a manner similar to that proposed by Tesnière (1959): there is a set of “core” arguments ArgN (roughly corresponding to Tesnière’s actants) and a set of modifiers (circumstances). Kipper et al. (2008) have frames based on relations in the style of Gruber (1965) and Fillmore (1968).

In the analysis of semantic relations, as in the resources we just mentioned, there is a strong undercurrent of reliance on surface-syntactic forms and of separation based on such forms. Semantic role labelling focuses on relations between the verb and its arguments within a sentence (Márquez et al., 2008); identification of semantic relations between nominals looks at the interaction between two nominals in a sentence (Girju et al., 2009). Information/relation extraction from text and ontology building both search for noun phrases, often simply nouns, and sequences of words which connect them in texts (Hearst, 1992, Yates and Etzioni, 2007, Banko et al., 2007, Van Durme and Schubert, 2008, among many others). Approaching the analysis of semantic relation analysis in this way misses the connection – shared relations, phenomena and data – between the different syntactic constructions. The only correspondence comes from parallels between verbs and their nominalizations.

3 Relations, Concepts, Syntactic Constructions

Semantic relations explain how concepts interact in a given context. They are intuitively recognized, and capture implicit knowledge and experience of the world. For example, a frequently encountered relation CAUSE or CAUSE-EFFECT holds between a cause and its ensuing effect. Experience tells us that something must happen to cause something else to happen. A causal relation, then, will hold between two concepts which are occurrences; in Section 4.1 we will discuss this in more detail and justify our view. Considering all this, it

is easy to explain and understand the causal relation in the following example:

The student was anxious because he was writing an exam. (4)

There are two clear concepts, occurrences overtly expressed as clauses: *he was writing an exam* – Cause, *the student was anxious* – Effect. How, then, does this differ from the following example?

The student was anxious because of the exam. (5)

The same general idea is conveyed and we still perceive a causal relation, but now only the effect is expressed as a clause, whereas the cause manifests itself as a noun which is not even a nominalization. The causal relation requires two occurrences, but the surface form consisting of a clause and a noun phrase seems to refer to one occurrence and one entity. In the context of Example (5), the noun phrase *the exam* conveys more than what it would in isolation. It suggests to the listener some occurrence, of which it is a part – the actual cause of the student’s being anxious. This is the type of analysis on which we will focus in the article: what a semantic relation requires of the concepts it connects, and how various syntactic constructions can convey the required concept.

For this kind of analysis, we adopt a coarse-grained view of concepts, which distinguishes only between occurrences, entities and attributes of occurrences and entities. An occurrence, when fully expressed in language, corresponds to a clause: a syntactic construction centred on the verb (larger than the verb phrase, to include the syntactic subject as well). An entity is usually expressed by a noun phrase. An attribute of an occurrence usually appears as an adverbial phrase, an attribute of an entity – as an adjectival phrase. In the remainder of the section we review and explain examples in which one disregards these “default” assumptions about the type of syntactic constructions through which concepts manifest themselves in language; we identify the linguistic phenomena which explain such situations. Mapping occurrences, entities and their attributes onto expressions with different syntactic constructions explains the ways in which a semantic relation can manifest itself in text.

Meaning has multiple components, some of which are preserved when there is a change in the lexicalization of some of the concepts expressed. For the purpose of this article, we concentrate on the concepts expressed and the connections between them. We let aspect, emphasis, focus and pragmatic implications of different expressions to stay in the background. It is from this restricted point of view that we assert that an idea can have different linguistic expressions. As far as linguistic variation is concerned, we investigate forms which have different syntactic structures, and the phenomena which allow one concept to take such various forms, which in a given context have the

same meaning for the listener⁶.

3.1 Surface Forms

In this section, we catalogue ways in which a concept can be mapped onto different syntactic constructions. We go systematically over all possibilities and illustrate them with examples.

Our purpose is to analyze how occurrences, entities and their attributes can be lexicalized in various syntactic forms. This bridge between the textual and the conceptual level of a text explains how semantic relations can have instances in different syntactic constructions.

Semantic relations have been analyzed with respect to the textual form, based on the kind of syntactic construction to which a relation belongs. Here are the most commonly considered connections:

- between simple clauses in a sentence;
- between a verb and its arguments in a clause;
- between the phrase's head and its modifiers in a phrase.

We drop the distinctions due to the variety of syntactic constructions, and concentrate instead on what the semantic relations link. A semantic relation can connect two occurrences; an occurrence with an entity; an occurrence or an entity with its attribute. For example, the AGENT relation describes the link between an occurrence and an entity which is actively involved in it. The CAUSE-EFFECT relation describes the link between two occurrences, one of which causes the other.

An occurrence can in fact be expressed by a verb, a nominal (gerund, deverbal noun, regular noun) or a deverbal adjective. It is true that different parts of speech stand, by default, for different types of concepts. For example, nouns usually express entities, or rather what we perceive as entities we usually express by nouns. Occurrences are expressed by clauses, whose central element is a verb. There are, however, many situations when these default assumptions do not hold. We analyze and explain those situations.

3.1.1 Syntactic Constructions in Detail

A noun phrase is a structure whose head is a nominal. The relations we assign inside a noun phrase are between the head nominal and its modifiers. The modifiers can be adjectives (*happy* girl, someone *happy*), adverbs (*far-away* place, the people *behind*), participles (*hidden* treasure, *shrinking* glacier), other noun phrases (*examination* board) or even clauses – relative (the man *we saw yesterday*), non-finite (the car *standing outside the station*), appositive (the saying *that absence makes the heart grow fonder*) or premodifying ([he asked] *we don't know how many* people). If the modifier is a prepositional phrase, a semantic relation is assigned between the head noun and the

⁶Nastase (2003) presents an overview of opinions on whether we can express the same thing using different surface forms.

prepositional complement. The preposition is used as a marker (or indicator) for the semantic relation.

A clause consists of a verb and its arguments⁷. A verb's argument may be a noun phrase (*the woman* sang), a prepositional phrase ([the actor] walked *onto the stage*), an adverb ([she] walked *slowly*), an adjective ([the dog] is *brown*) and also a clause ([He] knows *where you hid the presents*).

Suppose that a semantic relation holds between two occurrences, usually expressed by clauses. In order to have the same semantic relation within a clause, one of the occurrences must now appear as a verb's argument – noun phrase, prepositional phrase, adverb or adjective – instead of a full clause. On the surface, the relation which held between two clauses must now hold between a verb and its argument.

3.1.2 Mapping a Concept Onto Different Syntactic Constructions

Showing that a semantic relation can have instances in different types of syntactic constructions is equivalent to showing that a concept can take different surface forms. We consider each mapping between these constructions, and look for phenomena in language which allow a concept to take such distinct syntactic forms.

We consider four types of syntactic constructions: clauses, noun phrases, adjectives and adverbs⁸.

This choice of syntactic constructions is motivated by what they customarily express. Clauses (centred on the main verb) express occurrences, noun phrases – entities, adjectives – attributes of entities, and adverbs – attributes of occurrences. Clearly, these types of constructions do not cover everything that can appear in a sentence. We do not consider verb phrases. A verb phrase does not correspond to any of the four types of concepts we distinguish; lacking the subject, it cannot represent a complete occurrence. We do not consider prepositional phrases. *Syntactically*, a preposition connects two elements of a sentence (for example, the main verb and a noun phrase). *Conceptually*, a preposition signals a relation between two concepts, one of them represented by the prepositional complement. This view of the role of prepositions is in line with theories of grammar (Quirk et al., 1985, p. 657) and cognitive semantics (Talmy, 1985, vol. 1, p. 349). The preposition is semantically powerful, in that it can force the interpretation of its complement. For example, the clause *you arrive* denotes an occurrence; in the sentence “*We will eat after you arrive.*”, it forces the reader to take into account the temporal dimension

⁷Because we want the structure around the verb to include the subject, we talk of clauses rather than verb phrases.

⁸Although there exist complex adjectival phrases, for the present analysis we will only consider the adjective alone. This is because there are no semantic relations in which an adjective is not connected to a noun or a verb. In the remainder of this article we will refer to this construction as simply *adjective*. Similarly, we will consider the adverb alone. It will appear in relations with nouns and verbs. We will refer to this construction as *adverb*.

of this occurrence, used here to anchor the time of the occurrence *we will eat*. Prepositions can express many types of relations (see (Quirk et al., 1985, ch. 9)). We treat them as connectors, and their influence on the interpretation of their complement is taken into account when mapping a concept onto various syntactic constructions.

We analyze, then, the mapping of a concept onto every possible pair of syntactic constructions we do consider. This will show what surface forms each concept (occurrence, entity, attribute) can take, and what linguistic phenomena explain these form variations. To illustrate the mapping, we must refer to specific parts of the syntactic construction and to elements of the concept. We have adopted a graphical convention which brings out the structural correspondence: the head of a phrase with the concept's core, the arguments⁹ with the concept's attributes.

Table 1 shows all possible pairs of syntactic constructions onto which concepts will be mapped¹⁰. We look at how concepts map onto each element of these pairs and – when such mappings are possible – what phenomena explain them.

TABLE 1 Targets of the mapping of a concept onto a pair of syntactic constructions

clause	noun phrase
clause	adjective
clause	adverb
noun phrase	adjective
noun phrase	adverb
adjective	adverb

One of these mappings is not possible. A concept cannot be mapped onto both a clause structure and an adverb. An adverb serves to clarify some aspect of the occurrence described, but it cannot stand for a complete occurrence. In order for an adverb to conjure up an entire occurrence in the mind of the reader, it would have to be strongly associated with the occurrence lexically or conceptually. In order to have a strong lexical association between an adverb and a verb, some lexical or grammatical phenomena must relate the two words. But no word-formation process accounts for the formation of an adverb directly from a verb phrase (Quirk et al., 1985). Conceptually, adverbs are too general to stand for a particular occurrence.

The sections which follow present the concept mapping onto the five possible pair of constructions from Table 1.

⁹A verb's arguments would be complements, a noun's arguments – modifiers.

¹⁰Though we look at pairs, we do not assume that a concept cannot be mapped onto three types of constructions. Such situations are covered by looking at all possible pairings between the three structures; this is what we do here.

3.1.2.1 Mapping a Concept onto a Clause and onto a Noun Phrase We are interested in expressions of an occurrence, entity or attribute which have different syntactic structure. The meaning which a construction conveys depends on the context, so our analysis is always relative to a larger (relational) context. Figure 1 illustrates the general idea of mapping a concept onto two different syntactic constructions – a clause and a noun phrase.

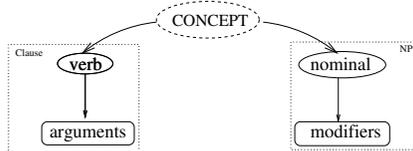


FIGURE 1 Mapping a concept onto a clause and a noun phrase

We consider an occurrence as covering the entire situation it describes, including all participants and attributes. Similarly, we treat an entity as complete with its attributes. From this perspective, concepts are not atomic. In the discussion which follows, we need to look inside the concept and see how its various parts are mapped onto syntactic constructions. We distinguish the *core*, the possible *participants* and *attributes*. The core corresponds to the concept stripped of all information on participants and attributes. The core corresponds to the lexical concept associated with the head of the grammatical phrase.

We have identified four way in which the same concept can be mapped onto a clause and a noun phrase.

1. We can map the core of the concept onto the heads of the two constructions, and its participants and attributes onto the arguments or modifiers respectively (Figure 2).

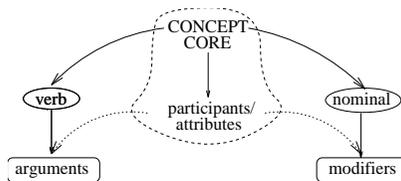


FIGURE 2 Mapping a concept core onto the heads of the constructions

A concept produced as a verb can also be produced as a nominal.

The painter heard the reviewers criticize his work. ↔ (6)
The painter heard the reviewers' criticism of his work.

Figure 3 shows how Example (6) can be illustrated graphically. (Similar instantiations can be drawn for other general graphs in this presentation.)

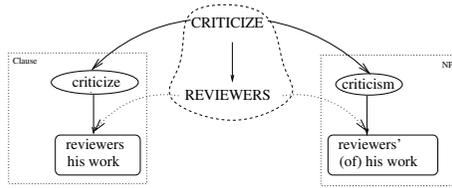


FIGURE 3 Mapping a concept core onto the heads of the constructions: an example

The two linguistic phenomena which account for this possibility are **nominalization** of the verb and **verbalization** of the noun.

2. We can map the concept's core onto the verb and onto at least one¹¹ of the nominal's modifiers, map the related concept's core onto the nominal and onto at least one of the verb's arguments (Figure 4).

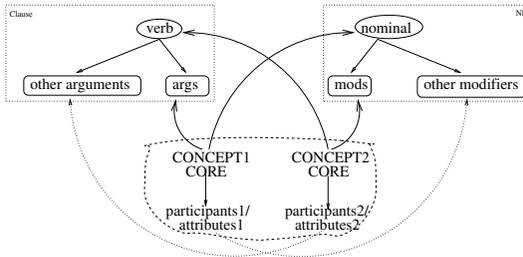


FIGURE 4 Mapping a concept core onto a head and an argument

An event expressed by the verb can be equivalently expressed as a noun's argument, based on the correspondence between a finite verb and a participle:

The ship sank. ↔ *sunk ship* (7)

In Example (7), the relation between the two occurrences of *sinking* corresponds to the relation between a verb and a deverbal adjective.

¹¹While we have not encountered a situation when a concept can be mapped onto a verb and more than one of the noun's modifiers, we do not make the strong claim that the phenomenon is not possible. The same is true for a concept mapped onto a noun and onto one of the verb's arguments.

Generally, we have the mapping of a concept onto a verb and an argument of a noun phrase. The same holds for mapping the same concept onto a noun and a verb's argument.

3. The concept expressed by a combination of the verb and some of its arguments can manifest itself as a noun (Figure 5).

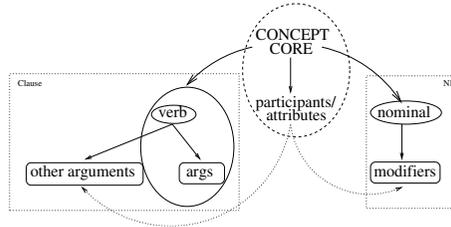


FIGURE 5 Mapping a concept onto a verb (and some of its arguments) and onto a nominal

This situation can be explained by various phenomena of **word formation**. For example:

the one who builds (vb+subj) ↔ *builder* (noun)

the one who was appointed (vb+obj) ↔ *appointee* (noun)

4. The verb has no equivalent expression in the noun phrase (Figure 6). This situation has three explanations.

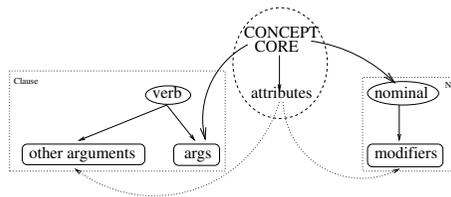


FIGURE 6 Mapping a concept onto verb arguments and onto a nominal and modifiers

Deletion. The verb can be deleted. A combination of its arguments is enough to convey the same entities, related in the same way:

The house was built with bricks. ↔ *brick house* (8)

Metonymy. The whole occurrence expressed by the verb-centred construction can be replaced with one of its arguments:

The student was anxious because he was writing an exam.
The student was anxious because of the exam. (9)

Equivalence. In the given context, a concept expressed by a noun phrase can replace a concept expressed by a clause – without changing the nature of the relation with the context:

We will have tea when you arrive.
We will have tea at 5 o'clock. (10)

Here it is a temporal relation which is maintained despite the change of concept. The concepts *5 o'clock* and *you arrive* are obviously not the same – although they may refer to the same time point – but they both can be used to identify a point in time. The temporal relation captured in Example (10) requires that one concept be an occurrence; the other can be anything, as long as it refers to a specific time point.

When we discussed the mapping of concepts onto various syntactic forms, we said little about the arguments/modifiers in a syntactic construction. There is a difference in the way in which verbs and nouns behave. Verbs have *core* arguments, which must be present for the expression to be grammatical and meaningful, while for all nouns modifiers are optional (Quirk et al., 1985). Therefore, some of the verb's arguments may surface as the noun's modifiers, but there need not be a one-to-one mapping. The arguments give more information, clarify the circumstances and the manner of an occurrence, and so on. We show in the mapping process the arguments and modifiers of a construction which are important for the mapping. Those not shown can be mapped according to the processes described in this section. The arguments in each construction (clause or noun phrase) are themselves clauses or noun, adjectival or adverbial phrases. We analyze all possible mappings of a concept onto pairs of constructions. These mappings also apply to attributes of concepts and participants which surface as the arguments in a verb-centred or nominal-centred construction.

3.1.2.2 Mapping a Concept onto a Clause and onto an Adjective There are two possibilities of mapping a similar concept onto a clause and onto an adjective.

1. We can map a concept onto the verb and onto the adjective (Figure 7).

[The ship] sank. ↔ *sunk [ship]* (11)

2. We can map a concept onto a verb (and some of its arguments) and onto an adjective (Figure 8).

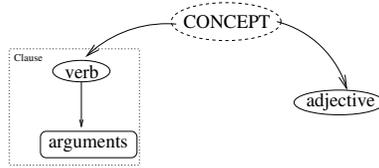


FIGURE 7 Mapping a concept onto a verb and onto an adjective

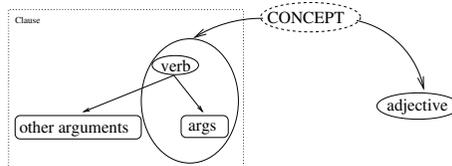


FIGURE 8 Mapping a concept onto a verb and some of its arguments, and onto an adjective

[The liquid] has no colour. ↔ *colourless [liquid]* (12)

[The man] is smoking cigarettes. ↔ *cigarette-smoking [man]* (13)

The phenomena which explain the mapping of a concept onto a clause and onto an adjective have to do with **word formation**, in particular, formation of adjectives from verbal expressions, and formation of verbs from adjectival expressions.

3.1.2.3 Mapping a Concept onto a Noun Phrase and onto an Adjective
Again, there are two possibilities.

1. We can map the same concept onto the head noun and onto the adjective (Figure 9).

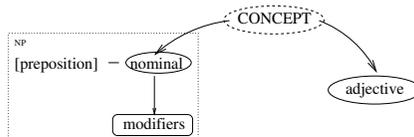


FIGURE 9 Mapping a concept onto a nominal and onto an adjective

The parent's [advice] ... ↔ *parental [advice]* (14)

... [writing] with style ... ↔ stylish [writing] (15)

2. We can map a concept onto the head noun and one or more of its arguments, and onto the adjective (Figure 10). This theoretical possibility seems to have no realization in English.

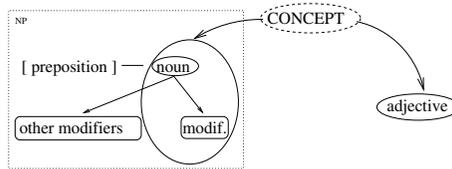


FIGURE 10 Mapping a concept onto a noun and some of its arguments, and onto an adjective

Noun and adjective formation phenomena (adjective nominalization, adjectivalization of the noun, derivation) account for the mapping of a concept onto a noun phrase or an adjective.

3.1.2.4 Mapping a Concept onto a Noun Phrase and onto an Adverb

In a context which expresses direction, we find adverb formation phenomena which allow a nominal expression to surface as an adverb. For example, we can equivalently say *towards home* or *homeward*, *towards north* or *northward*. In all these cases, the noun must be a prepositional complement, and the entire prepositional phrase has an alternative adverbial expression (Figure 11).

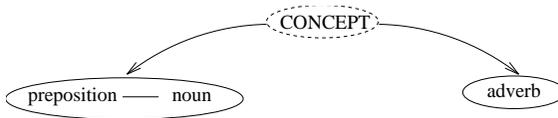


FIGURE 11 Mapping a concept onto a noun with a preposition and onto an adverb

A quantified noun phrase (*every day*) can also be mapped onto an adverb (*daily*). Figure 12 shows this situation.

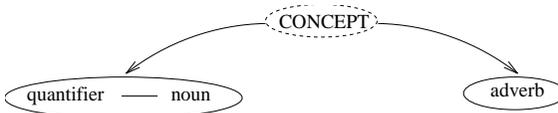


FIGURE 12 Mapping a concept onto a noun-phrase and onto an adverb

3.1.2.5 Mapping a Concept onto an Adjective and onto an Adverb Figure 13 shows the situation when the same concept is mapped onto an adjective and onto an adverb. For example:

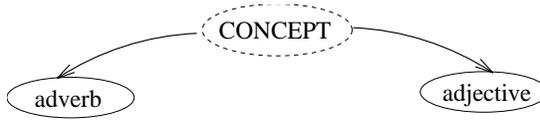


FIGURE 13 Mapping a concept onto an adjective and an adverb

He was driving weekly to the mountains. ↔ *weekly drive* (16)

driving carefully ↔ *a careful drive* (17)

3.2 Surfacing Phenomena

This section presents a systematic review of the linguistic phenomena – mentioned in the preceding discussion – which underlie the syntactically different manifestations of the same concept. Examples illustrate how each of these phenomena explains the mappings discussed in detail in the preceding section.

The analysis of syntactic constructions onto which concepts can be mapped has allowed us to identify phenomena which account for the existence of different surface forms of concepts. In this section we analyze each phenomenon from the standpoint of semantic relations. We show that the existence of these phenomena lends concrete support to the intuition that semantic relations can be the same in different syntactic constructions. While this is not a controversial claim, it is usually made in an *ad hoc* manner, or it is based on the nominalization phenomenon (as in the case of PropBank and NomBank).

3.2.1 Metonymy

Metonymy is a linguistic device whereby a concept is referred to by one of its attributes or by something else (strongly) associated with it. Metonymy allows the substitution of an attribute for the thing, or the other way around. *Synecdoche*¹² better characterizes the phenomena we analyze, but the term is not much used in the literature. In addition, the definition of metonymy has been considered to include synecdoche (Dirven and Verspoor, 1998, Lakoff and Johnson, 1980).

Metonymy is a pervasive phenomenon in language. It shows that understanding is possible in human communication even if the utterance lacks details.

¹²A figure of speech in which a part is used for the whole, the whole for a part, the specific for the general, the general for the specific, or the material for the thing made from it.

Dirven and Verspoor (1998) and Lyons (1995, pp. 266-267) write that the listener's background knowledge makes the interpretation of a text more than the sum of the interpretations of its sentences. Nunberg (1978) and Fauconnier (1985) hold the view that metonymy arises from the associations gathered as a result of cultural and existential factors. Cognitive linguists have adopted this idea and proposed that metonymy is not merely a figure of speech, but a reflection of the conceptual organization of the world (Lakoff, 1987, Johnson, 1987). This idea has been contested: the associationist position supports the view that experiential and cultural factors guide concept formation. Research in linguistics and cognitive development shows that concept formation seems to rely instead on the built-in human capacity for abstraction (Jackendoff and Aaron, 1991). Despite the fact that there has not yet been a consensus on the cognitive foundation of metonymy, psychological research supports the fact that it is a basic human cognitive process.

Briefly, there is more to words in context than their meaning when taken in isolation. Words serve as pointers to a larger body of knowledge, which we access while we interpret a text. From this point of view, words are metonymic: they stand for something larger, more complete – such as a full event or a complex entity. It is a matter of debate what kind of knowledge words “point” to, and how it can be captured. One position is that such knowledge can be contained in dictionaries, thesauri or other lexical resources, as a word's meaning. But this is not always the case. In certain contexts, words refer to something larger (or even different) than (one of) their possible meanings, and may refer to something which is not overtly expressed (Nunberg, 1978, Dirven and Verspoor, 1998).

Insofar as our study of variation in linguistic form and syntactic structure is concerned, metonymy accounts for various expressions of occurrences. The linguistic expression of an occurrence can be realized through one of its participants. Here is an example. We define CAUSE-EFFECT as a relation between two occurrences, the *Cause* and the *Effect*, with the assumption that when the proposition expressing the *Cause* is true, the *Effect* proposition is necessarily true.

The student was anxious because he was writing an exam. (18)

This is an example of the CAUSE-EFFECT relation in a sentence. The *Cause* proposition is *he was writing an exam* and conveys an action. The *Effect* is *the student was anxious*, and conveys a state. Assuming that the whole statement is true, the *Cause* proposition is true, and, as a consequence, the *Effect* proposition becomes true. Syntactically speaking, the *Cause* and *Effect* occurrences are fully expressed through clauses.

The speaker may not know exactly what about the exam makes the student anxious – writing it, answering it orally, thinking about it – or may decide not

to specify this. If so, the speaker might choose the following expression:

The student was anxious because of the exam. (19)

In this situation the relation is still CAUSE-EFFECT. The *Effect* proposition is still expressed by a state occurrence and a corresponding full clause, but the *Cause* is now just a noun phrase – *exam*.

Previously we have postulated that causal relations hold between two occurrences. How does this example fit the requirements of the relation CAUSE-EFFECT? The noun *exam*, in this context, represents not only the concept *exam*, but some action associated with it, possibly writing, that makes the student anxious.

We regard the occurrence as a whole whose parts are its constituents – the main verb, the participants such as *Agent* or *Object*, the qualifiers and so on. In light of this interpretation, metonymy will include the substitution of a part of an occurrence for the whole.

The phenomenon which explains why Example (19) presents a causal relation linking two occurrences is indeed **metonymy**: a part (in this case *exam*, which may play the role of a semantic *Object* in the fully expanded/IMPLIED occurrence) stands for the whole.

We can design experiments to retrieve some of the missing information from corpora, using collocation information. Consider the example:

Coffee after 5 o'clock doesn't let you sleep. (20)

The reader can infer that what is meant is:

Drinking coffee after 5 o'clock doesn't let you sleep. (21)

In certain cases such information can be retrieved, or inferred with good probability, from corpus analysis. We have found collocation information associated with the noun *coffee* extracted from the British National Corpus as part of the WASPS project (Kilgarriff and Tugwell, 2001). Among all the verbs for which *coffee* serves as a syntactic object, the most frequent one is *drink*. There are, however, situations in which corpus analysis does not help, because the event associated with certain word is not a usual event, but something which arises from the context of the sentence. Consider, for example, this sentence and its possible interpretations, adopted from (Pustejovsky, 1995):

Billy began the book. (22)

This sentence by itself could convey the same meaning as the following, more complete sentence:

Billy started to read the book. (23)

If Example (22) follows a text fragment in which it is established that Billy is in fact a goat, then we would rather be inclined to consider Example (23) as equivalent to:

Billy began eating the book. (24)

There may be an alternative explanation of the fact that, in a certain context, a word conjures up the whole occurrence for the reader. This additional information is so closely linked to the meaning of a word that it should be part of the lexicon. Pustejovsky (1995) proposes this approach in his generative lexicon¹³. Information about the meaning of a word in context should be available through a lexicon which comprises explanations about the concept *x* behind each word. The explanations cover four aspects: **constitutive** (what *x* is made of), **formal** (what *x* is), **telic** (what *x*'s function is), **agentive** (how *x* came into being). These four aspects form the *qualia* structure of a lexical item. The *qualia* "constitute the necessary modes of explanation for understanding a word or phrase" (Pustejovsky, 1998), derived from Aristotle's four causes – material, formal, final, efficient (Falcon, 2006). In order to reach the meaning of a word in context, one makes use of the aspects associated with it through the lexicon. For the noun *coffee* in Example (20), the *telic* role would bring the information that *coffee* is for *drinking*.

The generative lexicon approach, however, could not explain the situation presented in Example (22) and Example (24). All information which is incorporated in a lexicon is generic, and it cannot address pragmatic considerations which arise from the word's specific context.

3.2.2 Equivalence

Metaphorical language allows us to conceptualize one thing in terms of another, through analogy and perceived similarities (Lakoff and Johnson, 1980). For example, time and space are conceptualized very similarly: time is viewed as a one-dimensional space. We think about time indicators as points or intervals on this "time line". The activities we perform, the events we participate in, in general all occurrences, have an implicit temporal and spatial dimension. An occurrence takes time to complete, it holds for a certain time interval, or it happens at a certain point in time. Also, an occurrence happens in a region of space, or at a specific location. For this reason, the temporal and spatial dimension of an occurrence may serve as a reference for another occurrence, just as easily as definite time or spatial points or intervals can. The manner in which an occurrence unfolds can also be defined in terms of another occurrence, or by an occurrence attribute (an adverbial expression). From the point of view of concepts and their connections, certain semantic relations – such as temporal and spatial relations – rely on specific dimensions

¹³Talmy (2000) calls it lexicalization.

or attributes of the connected concepts, as opposed to the concepts as such. To clarify, consider Example (10), which we repeat here:

We will have tea when you arrive. ↔
We will have tea at 5 o'clock. (25)

Here, *you arrive* and *5 o'clock* both indicate a specific point in time. The temporal relation captured in these examples (let us call it TIMEAT) imposes the following restrictions on its concepts: one should be an occurrence, the other should indicate a time point when the occurrence takes place (or begins to take place).

When a semantic relation implies such specific constraints on one (or both) of the concepts involved – that it be a point in time, a location in space, the manner in which an event unfolds – it can have instances in different types of syntactic constructions, because the constraints can be satisfied by expressions of different syntactic forms, as in Example (25) above. We refer to this lexicalization phenomenon as **equivalence**. Two different linguistic expressions are equivalent in the sense that they both convey a specific type of concept – time point, location, manner, and so on.

Consider this example:

They practice while others have lunch. (26)

We fix the time when the *practice* occurrence takes place, via another occurrence which unfolds during the same time interval. Or we can choose to name a specific time interval instead:

They practice during lunch hour. (27)

In this case the time interval during which the *practice* occurrence is unfolding is an explicit time interval: the *lunch hour*.

For spatial expressions we encounter the same phenomenon:

The two boats ran into each other where the river
flows into the sea. ↔ (28)
The two boats ran into each other near the delta of the river.

And the same for the manner in which an occurrence unfolds:

He draws as his instructor told him. ↔ (29)
He draws beautifully.

Thompson and Longacre (1985) present the idea that certain dimensions or attributes of concepts can be expressed in various ways. The authors identify clauses which can be replaced by a single word. In particular, adverbs can replace clauses which express time, space or manner. Thompson and

Longacre perform an analysis across several languages (including Isthmus Zapotec from Mexico, Barai from Papua New Guinea, Hausa, Mandarin) to show that this is a pervasive phenomenon in language, not limited to the particulars of English.

3.2.3 Deletion

Sometimes the verb's "job" is only to link two of its arguments. If a link can be inferred without the verb, then the verb can be deleted. Here is an example:

That virus causes the flu. ↔ *flu virus* (CAUSE-EFFECT) (30)

Here, *cause* is a specific predicate which imposes strong constraints on the type of arguments and the link between them. There are, however, very general verbs – such as *be*, *have*, *make* – which do not bring much semantic information to the expression. They can also be deleted without loss of information:

The dog is brown. ↔ *brown dog* (PROPERTY) (31)

The printer is in the lab. ↔ *lab printer* (LOCATION) (32)

There are languages in which the *X is Y* construction reduces to *X Y* (for example, Japanese, Hebrew and Arabic), or is stylistically marked (as in Russian). This supports the fact that a verb such as copular *be* can be deleted without loss of information. *Cause* and *be* are two extreme instances of this phenomenon.

Levi (1978) proposes a list of nine *recoverable deletable predicates*, five of which are verbs – *cause*, *have*, *make*, *use*, *be*¹⁴. When one of these verb predicates is deleted, the expression becomes a noun phrase with similar meaning, for example:

The man has a beard. ↔ *bearded man* (33)

The verb *have* mediates a number of semantic relations, including meronymy/holonymy, possession and property¹⁵. The semantics of the predicate *cause*, on the other hand, is well defined: it can only mediate causal relations.

Research on the analysis of nominal compounds often relies on mining corpora for more explicit paraphrases of the compound through expressions which tend to contain verbs (Turney, 2006, Nakov and Hearst, 2006). This type of analysis shows that verb deletion – which results in concise noun compounds – is not limited to a small set of verbs, but is a rather common phenomenon related to noun phrase formation in English.

The reader may notice that *deletion* is similar to *metonymy* – in both situations only part of an occurrence is lexicalized. The difference comes from what is

¹⁴The other four are prepositions: *in*, *for*, *from*, *about*.

¹⁵"Multi-use" predicates similar to *have* may exist in other languages.

and what is not lexicalized, and how the intended concept is evoked. When there is *deletion*, the entire occurrence is evoked by its lexicalized parts – usually a combination of participants and attributes of the occurrence, but not the core (which corresponds to the main verb). When the phenomenon is *metonymy*, only one part of the occurrence is lexicalized (a participant), and this evokes the intended occurrence through interaction with its context.

3.2.4 Word Formation

When an occurrence surfaces in a syntactic form, it can appear as a clause or just as a word. Metonymy covers one such possibility: when part of an occurrence stands for the whole. The part may be the *Object* or the *Agent*, and it is usually expressed by a noun. Another mechanism to allow for the expression of an occurrence by a word is to let the core of the action come to the surface in some form other than a verb. It may do so as the head noun via nominalization:

The investigator's report was very brief. (34)

or as a modifier via adjectivalization:

The sunk ship was finally discovered. (35)

Word formation processes compress a whole expression into only one word, allowing a concept to take the form of various syntactic structures.

In the process of word formation, a concept expressed using a structure with a head in an open class (noun, verb, adjective or adverb) can also be expressed using a word in another class. The transformations between classes can take the form of prefixation/suffixation, conversion, compounding or other modes. Quirk et al. (1985) present a comprehensive overview of methods of word formation. In the sections which follow we will give examples for each of the open classes and for pronouns (in particular possessive pronouns) which can compress a larger phrase into one word.

3.2.4.1 Nominal Expressions In relations which involve an occurrence and one of its participants, the occurrence is expressed by the main verb. The verb must be there in some form. Nominalization is one of the phenomena which can account for expressing an occurrence through a noun phrase. For example,

The students protested against tuition fee increase. (36)

can be almost equivalently expressed as

student protest against tuition fee increase (37)

The main verb is expressed using a nominal form, and the whole clause is changed into a noun phrase.

Consider Example (19), repeated here:

The student was anxious because of the exam. (38)

This parallels the noun phrase *exam anxiety*: the *Effect* proposition is a nominal expressing a state-type occurrence. The nouns *exam* and *anxiety* are connected by causality, but the information about who is in a state of anxiety has been omitted.

The genitive case can be included in this discussion of nominal expressions. Nouns inflected for the genitive case can be paraphrases of, or be paraphrased as, clauses which display various types of semantic relations (Quirk et al., 1985), including:

POSSESSION

Mrs. Johnson's passport ↔ *Mrs. Johnson has a passport.* (39)

AGENT

the boy's application ↔ *The boy applied for ...* (40)

OBJECT

the boy's release ↔ *... released the boy.* (41)

MEASURE

ten days' absence ↔ *The absence lasted ten days/ [She] was absent for ten days.* (42)

The similarity between a clause and its nominalized version is largely recognized in linguistics (Quirk et al., 1985) and cognitive linguistics (van Valin and LaPolla, 1997).

In verb nominalization, the transformation can affect the arguments of the verb as well, as in:

The parents refused ... ↔ *parental refusal* (43)

In this situation the *Agent* of the occurrence has been adjectivalized, while the head verb has been nominalized.

Adjectivalization of a noun can imply a deleted verb:

The nation has a debt. ↔ *national debt* (44)

In this case, *national* implies/incorporates the predicate *has*.

The verb can also be mapped onto a noun's modifier in the form of a deverbal adjective:

the treasure vanished ↔ *vanished treasure* (45)

There are, then, two different phenomena of interest: the adjectivalization of the noun, and the adjectivalization of the verb.

The participial form of a verb can be used as an adjective. There are adjectives which are not derived from verbs, although they have the same suffix as participles (for example *talented*, *diseased*, *unexpected*). They are paraphrases of larger verb phrases, usually with a deletable predicate *have* or *be* as head: ... *has talents*, ... *has a disease*, ... *was not expected*.

According to Quirk et al. (1985), participial adjectives with the suffix *-ed* which have a corresponding transitive verb usually have a passive meaning. So, the attributive form corresponds to a passive paraphrase:

lost property ↔ *property which has been lost* (46)

In an analysis of non-predicating adjectives, Levi (1978) claims that denominal adjectives inherit the thematic roles of the original noun. Raskin and Nirenburg (1995) also share this view, which they support with the following examples (p. 13):

agentive *presidential refusal, editorial comment*

objective *constitutional amendment, oceanic study*

locative *marginal note, marine life*

dative [genitive]/possessive *feminine intuition, occupational hazard*

instrumental *manual labour, solar generator*

This property is interesting, because it shows that semantic relations are preserved during nominalizations of the main verb and adjectivalization of its arguments.

3.2.4.2 Verb Expressions Verbs can be formed from nouns and adjectives. Denominal verbs have an implicit relation. One of the entities in this relation is the noun at the origin of the denominal verb, and the other is an unspecified action. Let us consider an example:

hammer – implicit relation INSTRUMENT

We hammered the nail into the board. ↔

We hit the nail repeatedly with the hammer to drive the nail into the board. (47)

It is implied that a *hammer* was used, and is actually the *Instrument* in some unspecified action. The link between the noun and the denominal verb can be retrieved, as we will see in the following discussion.

According to Kelly (1998), there are two types of denominal verbs.

Rule-derived: The sense of the verb can be inferred from the semantic category of the original noun. The relation between the verb and the noun is captured in word definitions. Here is an example:

We hammered the two boards together. (48)

to hammer

LDOCE (Procter, 1978): (sense 1) to hit something with a hammer in order to force it into a particular position or shape.

Collins (Hanks, 1986): (sense 12) to strike or beat (a nail, wood, etc.) with or as if with a hammer.

WordNet 3.0: (sense 1) beat with or as if with a hammer.

Idiosyncratically derived: The sense depends on idiosyncratic aspects of the original noun. Example:

He pigged out at the buffet. (49)

We understand that *to pig out* means *to eat in a certain manner, like pigs do*. The relation between the verb and the noun is not explicit in the verb definition:

LDOCE (sense 1) to eat a lot of food.

Collins (sense 1) to gorge oneself.

WordNet 1.6 (sense 2) eat greedily. (WordNet 3.0 has a better gloss: “overeat or eat immodestly; make a pig of oneself;”.)

For the purpose of detecting semantic relations, denominal verbs of the second type are a more concise expression of a longer paraphrase. The paraphrase is replaced by a word because of some analogy: a characteristic associated with an entity expressed by a noun matches the ideas conveyed by the paraphrase. Without the appropriate knowledge resource, such denominal verbs are hard to analyze.

The rule-derived denominal verbs are, however, transparently connected to the original nouns. This is interesting from the perspective of semantic relations/role analysis. It means that the verb expresses an occurrence where the entity denoted by the noun plays a specific role.

3.2.4.3 Adverbial Expressions Adverbs can be formed from adjectives by suffixation, mainly by the suffix *-ly*:

slow ↔ *slowly*

Adverbs expressing direction can be formed by adding the suffix *-ward* to the noun which shows the direction:

towards home ↔ *homeward*

towards north ↔ *northward*

Temporal expressions can also be concisely expressed by an adverb:

every week ↔ *weekly*

every day ↔ *daily*

3.2.4.4 Pronouns While pronouns mainly stand for larger structures (anaphorically), they can also allow a mapping. Specifically, a possessive pronoun can replace a structure centred on a verb such as *have* or *belong to*. Those verbs express such varied relations as, for example, POSSESSION or AUTHORSHIP:

She has a car. ↔ *her car* (50)

The car belongs to her. ↔ *her car* (51)

The relation between *car* and a person (*she/her*) is POSSESSION in these situations.

3.3 Across Languages

In this section, the phenomena discussed in the preceding section are revisited in the context of three European languages typologically different from English.

To strengthen the argument that certain linguistic phenomena account for how semantic relations can surface in different syntactic form, we look at a few other languages. Do the phenomena described in Section 3.2 account for semantically equivalent forms in languages other than English? We chose three European languages – Russian, Romanian and Hungarian – two of them from the Indo-European family. For the *equivalence* phenomenon, we cite studies in more exotic languages.

Among the phenomena we analyzed, metonymy, equivalence and deletion seem to be rather language-independent, and more conceptual in nature. If that is the case, we should discover their existence by way of examples similar to examples in English which gave clues about these phenomena in the first place.

Word-formation processes are more language-dependent, but they appear in all languages. We expect a varying extent to which words from different parts of speech are related to one another through word-formation phenomena.

For each of the phenomena identified, we present examples in English (EN), Russian (RU), Romanian (RO) and Hungarian (HU). We chose the examples so that the expressions are equivalent across the languages.

Metonymy

EN

Sweets before dinner spoil your appetite. ↔ (52)

Eating sweets before eating dinner spoil your appetite. (53)

HU

Az édesség vacsora előtt elrontja az étvágyat.
the sweet dinner before spoil the appetite

Az édesség fogyasztása vacsora előtt elrontja az étvágyat.
the sweet consumption dinner before spoil the appetite

RO

Dulciuri înainte de cină strică apetitul.
sweets before of dinner spoil appetite_{def.art.}¹⁶

Dacă manânci dulciuri înainte de a cina îți strici apetitul.
if eat_{2sg} sweets before of to dine to-you spoil_{2sg} appetite_{def.art.}

RU

Сладости перед обедом портят аппетит.
sweets before dinner spoil appetite

Если вы едите сладости перед обедом, вы портите свой аппетит.
if you eat sweets before dinner, you spoil of-self_{2sg}¹⁷ appetite

The slight variations in these expressions are due to the fact that literal translations of Example (53) did not sound natural in Russian, Hungarian and Romanian. The phenomenon of metonymy, however, is clearly captured.

Equivalence**EN**

They practice while students have lunch. ↔ (54)

They practice during lunch hour. (55)

HU

Gyakorolnak, miközben a tanulók ebédelnek.
practice_{3pl} sometime-in-the-interval the students lunch_{verb}

gyakorolnak ebédszünetben
practice_{3pl} in-the-lunch-hour

¹⁶def.art. = definite article

¹⁷2sg = second person singular

RO

Ei exersează în timp ce studenții iau prânzul.
 they practice_{3pl} in time of students_{def.art.} take lunch_{def.art.}

Ei exersează în timpul orei de prânz.
 they practice in time_{def.art.} hour_{GEN} of lunch

RU

Они репетируют когда студенты едят обед.
 they practice when students eat lunch

Они репетируют во время обеденного перерыва
 they practice in time lunch_{adj.GEN} break_{GEN}¹⁸

Thompson and Longacre (1985) present examples in Hausa and Mandarin, among others. The languages they chose allow for temporal, locative and manner clauses to be replaced by single words, while the semantic relations remain the same. We reproduce a few examples in Isthmus Zapotec, an Otomanguan language of Mexico (the italics emphasize the clauses and words which express temporal, spatial and manner relations; the emphasis is the authors')¹⁹:

Time

- a. Kundubi bi yánaji
 is blowing wind today
 'It's windy today'
- b. *Ora geeda-be zune ni.*
 when (POT)come-he (FUT)do I it
 'When he comes I'll do it'

Locative

- a. Nabeza Juan rarí
 dwells John here
 'John lives here'
- b. *Ra zeeda-be-ke nuu ti dani*
 where is coming-he-that is a hill
 'Where he was coming along, there was a hill'

¹⁸GEN = genitive case

¹⁹POT = POTENTIAL; FUT = FUTURE; COMPL = COMPLEMENT

Manner

- a. *Nageenda* *biluže-be*
 quickly finished-he
 ‘He finished quickly’
- b. *Gu’nu* *sika ma* *guti-lu*
 (POT)do you like already (COMPL)die-you
 ‘Act as if you’re dead’

Deletion

English allows for easy formation of base noun phrases²⁰ through the deletion of a predicate. Deletion is possible in other languages as well, but it does not produce as concise an expression as it does in English.

EN

the virus which causes the flu ↔ *flu virus* (56)

HU

a *vírus, ami az influenzát okozza* ↔ *influenza vírus*
 the virus which the flu causes ↔ flu virus

RO

virusul care cauzeaza gripa ↔ *virusul gripei*
virus_{def.art.} which causes flu_{def.art.} ↔ *virus_{def.art.} flu_{GEN}*

RU

вирус вызвавший грипп ↔ *вирус гриппа*
 virus having-caused flu ↔ virus flu_{GEN}

Word Formation

Other languages form more easily nouns to name entities according to their properties, which English does not allow:

EN

the hanged man (57)

RO

omul spânzurat ↔ *spânzuratul*
man_{def.art.} hanged ↔ *hanged_{def.art.masc.sg}*

²⁰noun phrases composed only of a noun and one (simple, not phrasal) modifier

RU

повешенный человек ↔ *повешенный*
 hanged human ↔ hanged_{masc.sg}²¹

Whereas in English we may say *the hanged*, this will denote a class of people as opposed to a specific individual. Hungarian behaves as English in this case, because we cannot refer to an entity via one of its attributes. A more comprehensive study is required to decide whether we simply could not find an example, or whether none exists.

The examples illustrate nominalization in which a noun-adjective compound is nominalized, but the core of the nominalization is the adjective rather than the head noun.

Here is an instance of verb nominalization, across the languages we look at:

EN

The students protested against tuition fee increase. ↔ (58)

student protest against tuition fee increase (59)

HU

A tanulók tüntettek a tandíj emelése ellen.
 the students protested the tuition-fee increase against

tanuló tüntetés a tandíj emelése ellen
 student protest the tuition-fee increase against

RO

Studentii au protestat împotriva creșterii taxelor de școlarizare.
 students_{def.art.} have protested against increasing fees_{GEN} of tuition

protestul studenților împotriva creșterii taxelor de școlarizare
 protest students_{GEN} against increasing fees_{GEN} of tuition

RU

Студенты протестовали против повышения оплаты за обучение.
 students protested against increase_{GEN} pay_{GEN} for tuition

студенческий протест против повышения оплаты за обучение
 student_{adj.} protest against increase_{GEN} pay_{GEN} for tuition

²¹masc.sg. = masculine singular

Nominalization is not always as successful in compressing expressions in languages other than English:

EN

The student was anxious because of the exam. ↔ *exam anxiety* (60)

HU

A tanuló izgult a vizsga miatt. ↔ *vizsgaláz*
the student was-anxious the exam because-of ↔ exam-fever

RO

Studentul era agitat din cauza examenului. ↔
student_{def.art.} was anxious of cause exam_{GEN}

agitație din cauza examenului
anxiety of cause exam_{GEN}

RU

Студент волновался перед экзаменом ↔
student was-anxious before exam_{def.art.}

пред-экзаменационная лихорадка
pre-exam_{adj.} fever

The alternative expressions are all noun phrases, but in Romanian there is no base NP like in English. The predicate indicating causality could not be deleted from the phrase in Romanian as easily as it was deleted from the English version. This observation suggests an interesting experiment, should multilingual corpora be available. In a parallel multilingual corpus, we could find parallel expressions of the same concepts in different languages. Some languages need to keep some indicators of the relation between the concepts presented in the utterance, like *cauza* (RO). The parallel expressions richer in such information could then be used to assign semantic relations to expressions in other languages.

4 A Unified View of Semantic Relations: A Case Study

This section shows how the “machinery” developed in Section 3 works. The effect of a systematic analysis of semantic relations is a list which does not distinguish between the three types of syntactic constructions – phrases, clauses and multi-clause sentences.

A concept can manifest itself in different syntactic forms. Since relations describe interactions between pairs of concepts, there should be the same links between the concepts' different surface expressions. In analyzing semantic relations, we start from the relation's *name*, chosen presumably to suggest certain phenomenon, and try to understand its meaning. This leads to uncovering what a relation implies in terms of describing the interaction/link between two concepts.

Consider, for example, the temporal relation TIME THROUGH. It describes the relative position on the time axis of a time interval and the duration of an occurrence. The occurrence unfolds all through the specified time interval, but the time interval can be expressed in many ways. It can be explicit (*five minutes*), or implicit, defined by another occurrence which satisfies the requirements of the relation – here, an occurrence bound in time (*the soprano sang an aria*).

Based on a review of semantic relations used or mentioned in related research, in this section we talk about six classes of relations: **causality**, **temporality**, **spatiality**, **conjunctive**, **participant** and **quality**. We review how they were previously regarded in the literature, and how can they be considered from the perspective presented in detail in this paper.

To help the reader follow the syntactic phenomena responsible for the expression of similar concepts in different forms, we keep the examples presented for each relation as consistent as possible. That is to say, we aim to have examples which display a concept in all types of syntactic constructions. This restriction is relaxed when it produces examples with questionable wording, or when we want to make a point about certain type of surface expression.

4.1 Causal relations

Semantic relations in this class describe the causal interaction between two occurrences, one of which influences in some way (causes, opposes, enables) the other.

The idea that causality is a relation between two events was adopted from research in philosophy (Davidson, 1967) and psychology (Schank, 1973, Miller and Johnson-Laird, 1976). In linguistics there are three views of this issue. Causality is regarded as a relation between an agent and an event (Jackendoff, 1990, Rappaport and Levin, 1988), as a relation between two events (Dowty, 1979, Givon, 1975), or as a relation between two entities, an *Agent* and a *Patient* (Lakoff and Johnson, 1980). The first two views arise from the analysis of verb meanings. Lakoff and Johnson's view arises from an analysis of causation as a basic, but decomposable, concept of human activity. Causality is decomposed in terms of separate actions which the *Agent* undertakes in order to achieve a change of state in the *Patient*.

We also view causality relations as holding between two occurrences. Nonetheless, relations in a clause are sometimes better explained by considering cer-

tain constituents of the clauses other than the head verb. For some relations, the subject of P_c is not always the *Agent* of the occurrence which causes P_e , but it can also be the *Object*.

Relations grouped under **causality** describe how two occurrences influence each other. Both the *Cause* and *Effect* propositions in a causal relation are occurrences, but not always fully specified. An occurrence can take a clausal, noun-phrase or adjectival form by means of metonymy, deletion or nominalizations.

4.2 Temporal relations

This class contains relations which describe the position of an occurrence on the time axis relative to a time denoted by a (temporal) expression. A point or interval of time can be conveyed through an explicit time expression or by an occurrence which can express a point in time (*the clock struck midnight*) or a time interval (*he filled the bottle with tap water*) introduced by a marker which has a temporal reading in that particular context (such as *when* or *since*).

In the literature on temporal relations, the focus is on the representation of the flow of events in order to facilitate certain tasks rather than on the analysis of linguistic phenomena for the discovery of temporal relations among concepts. Allen (1984) introduced a theory of temporal relations and representation of events, to address issues concerning different types of occurrences²². The temporal relations most frequently found in the literature revolve around the analysis of events and their relative position on the time axis, using the temporal aspects which the main verb of the event provides. The focus is on positioning occurrences on the time axis relative to time indicators in the sentence; we do not analyze the verb for particular temporal attributes.

Miller and Johnson-Laird (1976) observe that words in all major classes can be used to express temporality (verbs: *end, precede*, nouns: *day, month, tomorrow*, adjectives: *former, present, successive*, adverbs: *eventually, often, soon*, prepositions: *at, during*, conjunctions: *as soon as, before, until*). This is in addition to tense, which is explicitly represented in the form of the verb. Some temporal markers may have different functions, and introduce different types of temporal expressions, as shown in the following sentences:

It happened before noon. ↔
It happened before he left. (61)

In the first sentence, the preposition *before* introduces a prepositional phrase which has an adverbial function. In the second sentence, the conjunction *before* connects two clauses. The temporal relation (which we call PRECEDENCE) is the same.

²²Those are actions which involve non-activity, actions which cannot be defined by decomposition into sub-actions, and actions which occur simultaneously and may interact.

All **Temporal** relations share the following property: one of the entities they connect is an occurrence, the other serves as a reference on the time axis. The reference may be direct, through an explicit time interval, or indirect, through another occurrence. It may be also the case, for the TIMEAT relation, that both entities linked are time references. *Equivalence* and *word formation* are the phenomena which allow for a temporal concept to surface in different syntactic forms.

4.3 Spatial relations

Spatial relations describe the position of an occurrence or an entity with respect to a point, area or volume in the three-dimensional space. The location in space can be expressed by an explicit spatial location (*near the river*) or by an occurrence introduced by a spatial preposition (*where the two rivers meet*). We extend the notion of locality to encompass both the location of an occurrence and that of an entity. The expression of a **Spatial** relation inside a noun phrase whose head is a regular noun (does not have a verbal reading) can be regarded as an expression from which a verb, possibly the existential *be*, was omitted.

In their analysis of spatial relations, Miller and Johnson-Laird (1976) make this point: if an occurrence has a spatial location, so do its participants. In the sentence:

The plumber worked in the kitchen. (62)

the adverbial phrase *in the kitchen* shows where the *work* occurrence takes place. In this example, the *Agent* of the *work* must be located at the same place. One can say, then, that locating the occurrence is equivalent to locating the *worker*, who in this case is the *Agent*. So, the expressions “*The plumber worked in the kitchen*”, “*The plumber was in the kitchen*”, and “*the plumber in the kitchen*” are all instances of the same spatial relation.

In analyzing spatial relations, Talmy (1985) concentrates on the entities involved in a relation. Both closed-class elements (prepositions) and open-class elements (verbs) may indicate the relative position of the entities involved. Deciding on a relation between objects imposes a point of reference. This will cause one or more of the entities in the scene to become *Grounds* (referential objects), while one will become the *Figure* (the object which will be located with respect to the others). Talmy opposes Fillmore’s proposal for **Spatial** relations – Locative, Source, Path, Goal – because they concentrate on direction rather than on the entities involved. In his analysis, Talmy identifies a *small set of primitive station/motion formulas – ones which seem to underlie all more complex characterization of stasis and movement in language*. Talmy refers only to location of entities. In our analyses, we have encountered situations in which either Talmy’s or Fillmore’s view can be more appropriate. When locating an occurrence (with respect to another occurrence or to an en-

tity), Fillmore's relations apply, whereas Talmy's relations fit best when one entity is located with respect to another.

Spatiality relations name the link between an occurrence or an entity and a region in space. Occurrences can be expressed by clauses or nominal phrases, following *nominalization* derivations.

4.4 Conjunctive relations

The relations in this class describe the conjunction or disjunction of occurrences, entities or attributes. The concepts which interact in such a manner must be of the same type. The concepts can both be

- occurrences:

She rides her bike every day and some days she goes swimming.

- entities:

Books and chocolates are her favourite things.

- attributes of occurrences:

He writes fast and stylish.

- attributes of entities:

a red and juicy apple

That is why these relations can have instances only in multi-clause sentences and noun phrases. These two relations cannot have instances in a clause, because they are assigned to paratactic constructions (coordinates). Relations in a clause – originally called cases – describe the connections between a superordinate (the verb) and a subordinate (the argument). Cases are always assigned to hypotactic constructions (subordinates).

It can be argued that CONJUNCTION and DISJUNCTION in a noun phrase have a deeper meaning than what they can have in a clause: the entities in a noun phrase are in a CONJUNCTION or DISJUNCTION relation because of something they share, whereas in a clause this is not necessarily true. One can construct instances of these two relations in a multi-clause sentence such that the occurrences therein are unconnected, though this is not normally so in real texts. Occurrences are connected because the speaker intended to make them so, according to some reason which might or might not be evident. They are sometimes connected because they appear to express causality, temporal ordering, but if this connection is not clear, a more generic CONJUNCTION/DISJUNCTION relation can be assigned. If the occurrences are linked for a more pragmatic reason (for example, the speaker was reminded of both occurrences at the same time), the assignment of a CONJUNCTION relation is not a weaker assignment but the only one possible. In a noun phrase, the connection between concepts is more obvious.

We can consider that **Conjunctive** relations link two occurrences. A conjunctive relation can be found in a noun phrase, as in this example:

I like apples and pears.

(63)

This coordinate noun phrase has a distributive reading: since the *like* occurrence applies to both *apples and pears*, it applies to each conjunct separately. This results in a conjunction of two *like* occurrences, one whose *Object* is *apples* and one whose *Object* is *pears*. Generally, a coordinate noun phrase in a subject or object position has either a collective or a distributive reading, depending on the semantics of the phrase and the main verb (Hobbs, 2000).

In the case of **Conjunctive** relations, the occurrences linked can be expressed by full clauses, or just by noun phrases. The phenomena which account for the different syntactic forms of concepts connected by CONJUNCTION or DISJUNCTION are *metonymy* and *word formation*.

4.5 Participant relations

Participant relations describe the interaction between an occurrence and each of its participants. The occurrence can be expressed by a verb, deverbal noun or deverbal adjective. A participant can be an entity or an occurrence.

Participant is one of the most common classes of semantic relations discussed in the literature. It encompasses the relations spun out of grammatical cases such as AGENT, OBJECT, INSTRUMENT.

The number and type of semantic relations in this class vary according to the goal pursued. We encounter the relations in this class only in clauses and noun phrases, because by definition they hold between the occurrence and one of its participants.

In a different view of semantic relations, Barker and Dowty (1993) actually search for thematic roles between a noun and its modifier which cannot appear between a verb and one of its arguments. Their general view of the thematic roles is different, and they group all verb-argument roles into two cluster-concepts – *Proto-Agent* and *Proto-Patient*. In order to find roles between a noun and its modifier which are exclusive to noun phrases, they look at *ultra-nominals* – *nouns which are least plausibly derived from verbs, and those denoting an object*. The nominal thematic roles found are grouped into two other cluster-concepts: *Proto-Part* and *Proto-Whole*. Each is characterized by certain properties, and the position of each of the arguments in the noun phrase is determined by the *nominal argument selection principle*.

The first scholar whom we know to have analyzed relations of a semantic nature, Pāṇini, has identified (besides the *karakas* – the verbal relations) non-verbal relations which he grouped under the name *sesa*. Bhartṛhari who studied Pāṇini's work in the 5th century AD further analyzed this group of relations. He affirmed that although these relations do not represent *karaka* relations, they may involve or be preceded by one of the *karakas*. For example, the possessive case is supposed to have been preceded by some action: the possessive expression “king's man” implies that there was the king's action

which has led to the establishment of a master-servant relation. In expressions like “branch of the tree” and “father’s son”, the relations like part-whole and procreator-offspring are “the results of previous actions not mentioned in the sentences, actions in which these objects were accessories. That previous status lingers somewhat in the present status and that is why the present status is looked upon as a kind of *karaka*, though its relation with the action expressed in the sentence is rather remote” (from (Iyer, 1969) cf. (Manjali, 1997)).

Participant relations do not appear between clauses, because they are relations which involve participants in occurrences. There are, however, instances of the participant expressed by an occurrence, such as the *Object* of predicates like *know*, *say*:

I know what you did last summer. (64)

While the participant of an occurrence can also be identified through an occurrence, it will be a nominal clause that serves to further specify a (generic) concept:

The man who was painting the walls got sick. (65)

One of the arguments of a **Participant** relation must be an occurrence. It must surface somehow or be implied in the syntactic expression of an occurrence-participant pair.

Among the **Participant** relations from the literature, PART and WHOLE occur only in a noun phrase, and they link two entities, one of which represents the whole and the other the part.

The phenomena which account for the different surface realizations of an occurrence are *word formation (nominalizations, adjectivalizations)*, *deletion*, and *metonymy*. The participant may also surface in different syntactic structures. *Adjectivalization* is a phenomenon which allows for a concept to surface as a modifier in a noun phrase.

4.6 Quality relations

The **Quality** class groups relations which describe different attributes associated with concepts: Manner, Material, Measure, Order, Content, Container. MANNER, for example, captures the relation between an occurrence and its attribute which describes how the occurrence unfolds:

She writes beautifully. ↔ *beautiful writing* (66)

The grouping of relations into the **Quality** class has not been introduced in other research – in contrast with the more coherent and commonly agreed-upon classes Causal, Temporal, Spatial, Participant and Conjunctive. Whereas there is no consensus on a list of semantic relations (for any type of syntactic

constructions), researchers tend to agree on the first five classes of relations. There are a number of relations of the sixth class in the literature. Quirk et al. (1985, p. 557) give examples of MANNER relations introduced by various adjuncts:

They began arguing loudly. (67)

You should write as I tell you to. (68)

Thompson and Longacre (1985) also show various expressions of the MANNER relation through various forms in different languages (see Section 3.2.2). PROPERTY captures the relation between an entity and its properties. We use properties to identify specific objects and to separate them from similar objects, based on some distinguishing characteristics: *blue book*, *hard project*, and so on. It is one of the basic relations associated with the verb *be*, which can be deleted to obtain a more concise phrase (Levi, 1978).

TYPE is the name we gave the hyponym-hypernym relation (IS-A) – *maple leaf*. This is probably the relation most employed in research in natural language processing and artificial intelligence: it is the backbone of any ontology. It has been mined successfully (Hearst, 1992), and it underlies the organization of WordNet (Miller et al., 1995).

MATERIAL relates an entity with the “stuff” this entity is made of. Vanderwende (1994), for example, includes a relation called *Made of what* in her list, which would correspond to our MATERIAL relation. MATERIAL is also included by Leonard (1984), and so is EQUATIVE.

Quality relations hold between an occurrence and one of its arguments. The argument shows aspects of the occurrence – manner, type, measure, etc. The occurrence must surface in the syntactic expression. The phenomena identified in the verbalization of concepts connected by this type of relations are *metonymy* and *word formation* processes.

5 Conclusions

Semantic relations are essential in establishing the meaning of a text. They show how individual pieces are connected, to express the idea behind larger or smaller fragments of texts. While they are not physically present in texts, the relations are perceived through our effort of making sense of what has been said.

The motivation for this work has been the state of semantic relation analysis in the natural language processing community. For various practical reasons, relations are seen as connecting forms with specific syntactic structures – nominals and their arguments, verbs and their arguments – or concepts (expressed by nominals) in a sentence. The view explored in this article assumes that relations mean something in themselves, so they place certain constraints on the concepts they link. As long as these constraints are satisfied, the con-

cepts can take linguistic expressions pertaining to various types of syntactic constructions. We have explored linguistic phenomena which allow for such varied realisations to occur, and their effect on finding instances of semantic relations in different types of syntactic constructions. Such an analysis allows us to detect situations where information is implicit or covert. This information can then be specifically mined, for example in aid of knowledge acquisition. This study provides a solid basis for the semantic analysis of text, which can help improve our computational approaches to the task. It brings text analysis closer to reaching the essence of what the speaker wanted to convey.

6 Acknowledgements

Many thanks to Professor Andrew Brook for his detailed and hugely helpful comments on a previous version of the paper. We also thank Marina Sokolova, Istvan Hernádvölgyi, Ella Huszti and Zsuzsanna Racz for helping with the examples in Russian and Hungarian. Thanks to the reviewers for insightful comments from varied points of view.

References

- Allen, James F. 1984. Towards a general theory of action and time. *Artificial Intelligence* 23(2):123–154.
- Banko, Michele, Michael J. Cafarella, Stephen Soderland, Matt Broadhead, and Oren Etzioni. 2007. Open information extraction from the Web. In *Proceedings of the 20th International Joint Conference on Artificial Intelligence*, Hyderabad, India, 6–12 January 2007, pages 2670–2676.
- Barker, Chris and David Dowty. 1993. Non-verbal thematic proto-roles. In A. Schafer, ed., *NELS 23*, vol. 1, pages 49–62. Amherst: GSLA.
- Chomsky, Noam. 1970. Remarks on nominalizations. In R. Jacobs and P. Rosenbaum, eds., *Readings in English Transformational Grammar*, pages 184–221. Waltham, MA, USA: Ginn and Co.
- Cruse, Alan D. 1973. Some thoughts on agentivity. *Journal of Linguistics* 9:1–204.
- Davidson, Donald. 1967. Causal relations. In *Essays on Actions and Events*, pages 149–162. Oxford: Clarendon Press. second edition.
- Dirven, Rene and Marjolijn Verspoor. 1998. *Cognitive Exploration of Language and Linguistics*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Dowty, David. 1979. *Word Meaning and Montague Grammar*. Dordrecht: Reidel.
- Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67:547–619.
- Falcon, Andrea. 2006. Aristotle on causality. In *Stanford Encyclopedia of Philosophy*. <http://plato.stanford.edu/entries/aristotle-causality/>.
- Fauconnier, Gilles. 1985. *Mental Spaces: Aspects of Meaning Construction in Language*. Cambridge, MA: MIT Press.
- Fillmore, Charles. 1968. The case for case. In E. Bach and R. T. Harms, eds., *Universals in Linguistic Theory*, pages 1–88. Holt, Rinehart and Winston.

- Girju, Roxana, Preslav Nakov, Vivi Nastase, Stan Szpakowicz, Peter Turney, and Deniz Yuret. 2009. Classification of semantic relations between nominals. *Language Resources and Evaluation* 43(2):105–121.
- Givon, Talmy. 1975. Cause and control: On the semantics of interpersonal manipulation. *Syntax and Semantics* 4:59–89.
- Gruber, Jeffrey. 1965. *Studies in Lexical Relations*. Ph.D. thesis, MIT, Cambridge, MA. Reprinted in Jeffrey Gruber. 1976. *Lexical Structures in Syntax and Semantics*. Part I. North-Holland Publishing Company, Amsterdam.
- Hanks, Patrick, ed. 1986. *Collins Dictionary of the English Language*. London and Glasgow: Collins.
- Hearst, Marti A. 1992. Automatic acquisition of hyponyms from large text corpora. In *Proceedings of the 15th International Conference on Computational Linguistics*, Nantes, France, 23-28 August 1992, pages 539–545.
- Hobbs, Jerry. 2000. *Syntax and Metonymy*. <http://www.isi.edu/~hobbs/metsyn/metsyn.html>.
- Hovy, Eduard. 1993. Automated discourse generation using discourse structure relations. *Artificial Intelligence: Special Issue on Natural Language Processing* 63:341–385.
- Iyer, K. A. Subramania. 1969. *Bhartrhari. A Study of Vakyapadiya in the Light of Ancient Commentaries*. Poona: Deccan College Postgraduate Research Institute.
- Jackendoff, Ray. 1983. *Semantics and Cognition*. Cambridge, MA: The MIT Press.
- Jackendoff, Ray. 1989. What is a concept, that a person may grasp it? *Language* 4(1-2):68–102.
- Jackendoff, Ray. 1990. *Semantic Structures*. Cambridge, MA: MIT Press.
- Jackendoff, Ray and David Aaron. 1991. Review of Lakoff and Turner (1989). *Language* 67(2):320–338.
- Johnson, Mark. 1987. *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. Chicago, IL: University of Chicago Press.
- Kelly, Michael H. 1998. Rule and idiosyncratically derived denominal verbs: Effects on language production and comprehension. *Memory and Cognition* 26:369–381.
- Kilgarriff, Adam and David Tugwell. 2001. WORD SKETCH: Extraction and display of significant collocations for lexicography. In *Workshop on Collocation: Computational Extraction, Analysis and Exploitation, 39th ACL & 10th EACL*, pages 32–38. Toulouse, France.
- Kipper, Karin, Anna Korhonen, Neville Ryant, and Martha Palmer. 2008. A large-scale classification of english verbs. *Language Resources and Evaluation Journal* 42(1):21–40.
- Lakoff, George. 1970. *Irregularity in Syntax*. New York: Holt, Rinehart and Winston.
- Lakoff, George. 1987. *Women, Fire and Dangerous Things*. Chicago: Chicago University Press.
- Lakoff, George and Mark Johnson. 1980. *Metaphors we live by*. Chicago: University of Chicago Press.
- Lakoff, George and Mark Turner. 1989. *More than Cool Reason: A Field Guide to Poetic Metaphor*. Chicago, IL: University of Chicago Press.

- Leonard, Rosemary. 1984. *The Interpretation of English Noun Sequences*. Amsterdam: North Holland.
- Levi, Judith N. 1978. *The Syntax and Semantics of Complex Nominals*. New York: Academic Press.
- Lyons, John. 1995. *Linguistic Semantics: An Introduction*. Cambridge University Press.
- Manjali, Franson D. 1997. Dynamic models in semiotics/semantics. Cyber Semiotic Institute. <http://www.chass.utoronto.ca/epc/srb/cyber/manout.html>.
- Mann, William C. and Sandra A. Thompson. 1988. Rhetorical structure theory: Toward a functional theory of text organization. *Text* 8(3):243–281.
- Margolis, Eric and Stephen Laurence, eds. 2002. *Concepts: core readings*. Cambridge, Massachusetts, USA: The MIT Press.
- Marquez, Lluís, Xavier Carreras, Kenneth C. Litkowski, and Suzanne Stevenson. 2008. Semantic role labeling: An introduction to the special issue. *Computational Linguistics* 34(2):145–159.
- Meyers, A., R. Reeves, C. Macleod, R. Szekely, V. Zielinska, B. Young, and R. Grishman. 2004. The nombank project: An interim report. In A. Meyers, ed., *HLT-NAACL 2004 Workshop: Frontiers in Corpus Annotation*, pages 24–31. Boston, Massachusetts, USA: Association for Computational Linguistics.
- Miller, George A., Christiane Fellbaum, Derek Gross, Katherine Miller, Richard Beckwith, and Randee Teng. 1995. *Five Papers on WordNet*. Princeton.
- Miller, George A. and Philip N. Johnson-Laird. 1976. *Language and Perception*. Cambridge, Massachusetts: The Belknap Press of Harvard University Press.
- Misra, Vidya Niwas. 1966. *The Descriptive Technique of Panini*. The Hague: Mouton.
- Nakov, Presley and Marti Hearst. 2006. Using verbs to characterize noun-noun relations. In *Proceedings of the 12th International Conference on Artificial Intelligence: Methodology, Systems, Applications (AIMSA 06)*, pages 233–244. Varna, Bulgaria.
- Nastase, Vivi. 2003. *Semantic Relations Across Syntactic Levels*. Ph.D. thesis, School of Information Technology and Engineering, University of Ottawa.
- Nunberg, Geoffrey. 1978. *The Pragmatics of Reference*. Bloomington, Indiana: Indiana University Linguistics Club.
- Palmer, Martha, Paul Kingsbury, and Daniel Gildea. 2005. The proposition bank: An annotated corpus of semantic roles. *Computational Linguistics* 31(1):71–106.
- Procter, P., ed. 1978. *Longman Dictionary of Contemporary English*. Essex, UK: Longman Group Ltd.
- Pustejovsky, James. 1995. *The Generative Lexicon*. Cambridge, MA: MIT Press.
- Pustejovsky, James. 1998. The semantics of lexical underspecification. *Folia Linguistica*.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London and New York: Longman.
- Rappaport, Malka and Beth Levin. 1988. What to do with theta roles. In W. Wilkins, ed., *Syntax and Semantics (21): Thematic Relations*, pages 7–36. New York: Academic Press.

- Raskin, Victor and Sergei Nirenburg. 1995. Lexical semantics of adjectives: A microtheory of adjectival meaning. Memoranda in Computer and Cognitive Science M CCS-95-288, Computing Research Lab, New Mexico State University.
- Schank, Roger C. 1973. Identification of conceptualizations underlying natural language. In R. Schank and K. Colby, eds., *Computer Models of Thought and Language*. San Francisco, CA: W.H. Freeman.
- Schank, Roger C. 1975. *Conceptual Information*. Amsterdam: North-Holland Publishing Company.
- Talmy, Leonard. 1985. How language structures space. In L. A. H. L. Pick Jr., ed., *Spatial Orientation: Theory, Research and Application*. London and New York: Plenum Press.
- Talmy, Leonard. 2000. Lexicalization patterns. In *Toward a Cognitive Semantics*. Cambridge, MA, USA: The MIT Press.
- Tesnière, Lucien. 1959. *Éléments de syntaxe structurale*. Paris: C. Klincksieck.
- Thompson, Sandra A. and Robert E. Longacre. 1985. Adverbial clauses. In T. Shopen, ed., *Language Typology and Syntactic Description (vol.2): Complex Constructions*, chap. 4, pages 171–234. Cambridge: Cambridge University Press.
- Turney, Peter. 2006. Expressing implicit semantic relations without supervision. In *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics (Coling/ACL-06)*, pages 313–320. Sydney, Australia.
- Van Durme, Benjamin and Lenhart Schubert. 2008. Open knowledge extraction through compositional language processing. In *Proceedings of the Conference on Semantics in Text Processing (STEP) 2008*, pages 239–254.
- van Valin, Robert and Randy LaPolla. 1997. *Syntax: Structure, Meaning and Function*. Cambridge Textbooks in Linguistics. Cambridge University Press.
- van Valin, Robert D. 1990. Semantic parameters of split transitivity. *Language* 66:221–260.
- Vanderwende, Lucy. 1994. Algorithm for the automatic interpretation of noun sequences. In *Proceedings of the 15th International Conference in Computational Linguistics (COLING 94)*, pages 782–788. Kyoto, Japan.
- Yates, Alexander and Oren Etzioni. 2007. Unsupervised resolution of objects and relations on the web. In *Proceedings of Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics*, Rochester, N.Y., 22–27 April 2007, pages 121–130.