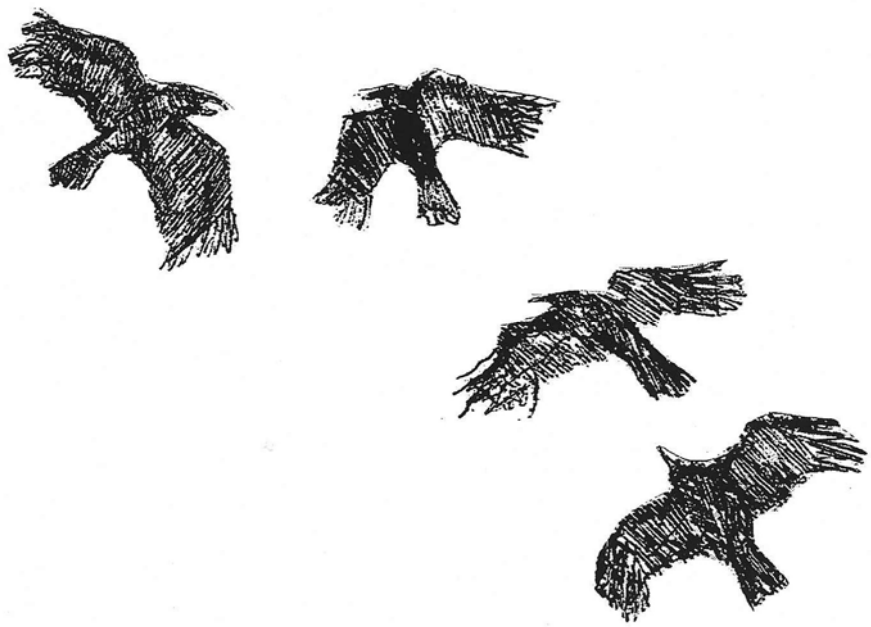
Collocations as a language resource
A functional and cognitive study in English phraseology

Sonja Poulsen



Ph.d. dissertation
Institute of Language and Communication
University of Southern Denmark
2005

Supervisors

Associate professor Fritz Larsen, University of Southern Denmark
Associate professor John M. Dienhart, University of Southern Denmark
Associate professor Alex Klinge, Copenhagen Business School

Front cover

Peter Harder's analogy (1996a: 91) between the functional role that wings play for the survival of birds in biology and the functionality of linguistic constructions inspired me to use birds for my front cover illustration.

My thanks to Kurt Normann for letting me use parts of his etching of flying rooks.

Sonja Poulsen
Institute of Language and Communication
University of Southern Denmark
DK-5230 Odense M

Tel.: +45 6550 2188
E-mail: spo@language.sdu.dk

Contents:

PART THREE

PART FOUR

****

Typographical conventions

Technical terms are marked by single quotes when first introduced and when there is a need to set them apart from the surrounding text: ‘collocation’
Linguistic meanings are also marked by single quotes: *put off* meaning ‘postpone’
Linguistic forms are printed in italics: *record*
Lexical categories are printed in capital letters: BREAK
Citations are marked by double quotes
Titles are printed in italics
Additions to, or omissions from, a citation are marked with square brackets: […]
Key words in the text are printed in bold type
Conceptual metaphors are printed in small capital letters: LIFE IS A JOURNEY

Abbreviations

The British National Corpus	BNC
The Oxford English Dictionary	OED
something (in tables and examples)	sth
somebody (in tables and examples)	sb
occurrence(s) (in tables)	occ.

Abbreviations used only in some sections, are explained where they are used

Tables

Figures

English résumé

This thesis addresses the question of what constitutes a suitable theoretical framework for entrenched collocations, focusing on expressions consisting of a verb and its nominal object. The overall claim is that a functional and cognitive approach is descriptively more adequate and has greater explanatory potential than the traditional approach according to which phraseological expressions are arbitrary, because they are conventional and are moreover characterized as deviating from an assumed norm of full compositionality.

Part 1 includes an outline of the traditional approach to phraseology as well as an overview of definitions of 'collocation' as a technical term. A motivation is given for an alternative functional and cognitive approach that includes a model of categorization based on prototypes rather than criterial features and an account of compositionality according to which the component structures are not building blocks, but dynamic schematic networks which are integrated in a composite structure in a manner that is not predictable from, but motivated by the component structures. Moreover, a functional and cognitive approach allows a diachronic perspective on synchronic evidence, which makes it possible to discuss mechanisms of entrenchment and innovation.

Part 2 discusses the traditional approach to phraseology from a functional and cognitive viewpoint. It starts from the background of practitioners of this approach in foreign language teaching and lexicography and goes on to discuss its theoretical foundations in Firthian linguistics and Russian phraseology. A classical view of categorization as well as structuralist and generative principles are seen as further influences, while a cognitive strand can be detected in recent work. This account is followed by a critical review of the resulting system of categorization. A major problem is the attempt to combine classical categories based on criterial features with a continuum model, which implies gradedness, two notions that are basically incompatible. What defines phraseological combinations in the traditional approach is restricted compositionality, which is defined syntagmatically in terms of the semantic dependence of the verb collocate on the nominal base and paradigmatically in terms of arbitrarily restricted substitutability of

the collocate. Against this it is argued, that the notion of syntagmatic restrictedness as deviation from a norm of full compositionality does not hold up, as lexical items are generally polysemous, and that evidence of restricted substitutability typically reflects factors that are 'extrinsic' rather than 'intrinsic' to a given collocation. In the traditional framework restricted compositionality is furthermore characterized by the absence of features such as predictability, analysability, and literalness, which is also questioned.

Part 3 contains the empirical part of the study, which tests the assumption that a functionally and cognitively based approach will provide a suitable framework for the study of conventional expressions. A discussion of what is functional and what is cognitive about language introduces this part, which goes on to clarify methodological questions relating to the notion of frequency as an input for qualitative analysis and the construal of linguistic data as reflections of cognitive routines and as synchronic evidence of diachronic processes. I chose *break an appointment* for my case study as an example of an entrenched collocation consisting of a 'delexical' verb and a 'deverbal' noun. The data from the British National Corpus include 1000 concordances with *break* (approximately 9.5% of all concordances with *break*) and 908 concordances with *appointment* (all concordances with *appointment*). Only five examples of *break an appointment* were found, all of which are in the data for *appointment*. In the case of *break,* phrasal verbs (here including combinations with adverbs as well as prepositions and adjectives) account for about one third of the concordances.

My research questions address the issues that I see as central aspects of what a theory of collocations should account for: the internal structure of component items and their mode of integration as well as patterns of entrenchment and innovation. My general assumption is that

> … entrenched collocations are speech routines allowing speakers to guide hearers by evoking cognitive routines associated with familiar semantic frames. As linguistic expressions, their function is to further the reproduction, through renewal of connection, of the contexts of situation and the underlying cognitive models, or semantic frames, to which they belong.

While these assumptions reflect a functional and cognitive approach, they are seen as compatible with the usage based Firthian notion of collocation as a level of meaning

which is ‘integral in experience’ both with the level of syntax and with the level of ‘context of situation’, which is abstracted from typical situations of use, and which I compare with the notion of ‘semantic frame’ (Fillmore, e.g. 1982, 1985) in cognitive linguistics.

My research questions were phrased as four claims:

1) The contexts of situation to which entrenched collocations contribute and their underlying frames can be identified by analysing the internal structure of component items and their mode of integration. In this entrenched collocations do not differ from other composite structures.

2) In entrenched collocations consisting of a verb and a nominal object, the noun evokes the dominant frame while the verb profiles a specific aspect of the frame. In collocations that are not entrenched, it is the verb that evokes the dominant frame.

3) Entrenched collocations can be characterized in terms of prototypicality that varies with the schematicity of the verb and its salience in the frame evoked by the noun.

4) The verb in entrenched collocations has a functional, grammaticalized role as support verb.

The first claim relates to what entrenched collocations have in common with other composite structures. It is tested by first analysing the two component items BREAK and APPOINTMENT separately, as complex, polysemous categories. These analyses are preceded by a discussion which (following Taylor, 1992) links the question of how many readings should be posited to meaning at the basic level of categorization and to the number of semantic frames evoked by a word, which is possible in the case of a moderately polysemous category like APPOINTMENT. In the case of BREAK, which is a highly-polysemous category, the number of readings is linked to higher level schemas including frames that are seen as being related in a network model. Each category has been analysed in terms of (1) domains and referential range, (2) image-schemas and event structure, and (3) construction types.

Subsequently, the composite structure *break an appointment* is analysed in terms of the autonomy/dependence alignment of verb and noun (Langacker 1987) with a view to showing that, on the one hand, an entrenched collocation can be analysed like any other composite structure, but that, on the other hand, entrenchment involves a shift in the autonomy/dependence alignment of verb and noun reflecting that it is the noun that evokes the dominant frame, as claimed in my second research question.

The third claim, that entrenched collocations can be characterized in terms of prototypicality, is tested by an analysis which focuses on capturing variability as a dimension of entrenchment. A typology of collocations is posited based on the notions of salience and schematicity. The emphasis is on the continuity between expressions that are entrenched and alternative expressions that are not entrenched, but still perfectly normal, thereby challenging the Saussurean idea that convention in language precludes choice. The range of verbs elaborating a specific substructure of an APPOINTMENT frame are construed as a lexical set with the basic level verb as the prototype and less schematic verbs as alternatives that may or may not be entrenched. This notion of the lexical set differs from Halliday's (1966) notion of a set whose members have 'like privilege of occurrence in collocation'. Rather it is seen as representing a functionally motivated range of variability at the place of the verb and as cutting across the borderline between those combinations that are entrenched and those that are not.

The discussion of the last claim, that the verb in entrenched collocations has a functional, grammaticalized role as support verb, is divided into two parts. In the first part, based on independent evidence from a range of sources, a case is made for a category of support verb constructions whose prototype is a delexical verb with general meaning and a deverbalized noun with a suffix which form a predicative unit. I have argued that entrenched collocations with *break* could be included in such a category, although not as one the most prototypical members. This is supported by the evidence of formal constraints on the use of *break* in the nonsensorimotor domain, which were brought out by the analysis of construction types.

In the second part, a discussion of the grammatical status of *break* and other basic level verbs participating in entrenched collocations concludes that they can be categorized as belonging to a lexico-grammatical prototype category of support verbs. The discussion is based on the view that grammatical structure forms a continuum with the lexicon and that the grammaticalization of a lexical item involves a shift toward the grammatical end of the continuum associated with a more functional role. Following Sweetser (1988), grammaticalization is seen as involving metaphorical extension to an abstract

domain of grammatical meaning. In the case of *break* as a support verb, it is argued that the extension is to the domain of causes and events, where, by analogy with its meaning in the sensorimotor domain of 'artefacts and natural things', *break* expresses object event structure and actionality. Finally, the claim that support verbs are more grammatical than full verbs and that some support verbs are more grammatical than others is discussed in terms of 'level of specificity' and 'overlap between standard and target', two parameters proposed by Langacker (1987). In spite of the fact that break is less schematic, or more specific, than some other basic level verbs, the conclusion is that it can be categorized as a grammaticalized support verb, although it is not the most prototypical example.

Preface and acknowledgements

Like most people, I have learned the conventional collocations of my own language without noticing them much and without any urge to name or categorize them. Later, as a learner of foreign languages, I have struggled to get them right, and, as a teacher of English, I have felt the need to be able to understand and explain them. As a researcher, I have tried to account for them by using the traditional approach to phraseology, which did not, however, prove fully satisfactory. I would therefore like to thank my university for giving me this chance to take a fresh look at collocations from a functional and cognitive perspective. Particularly, I would like to thank my supervisors Fritz Larsen and Alex Klinge for their constructive criticism and sound advice. I am also grateful to John Dienhart, who was my supervisor for a short while, before he died. What he called his 'chicken scratch' taught me a lot of discipline of the detail. It has been a long haul and the interest shown by family, friends and colleagues has been a great encouragement. My very special thanks go to my husband and children: thank you Torben, Maiken and Theis, for listening and for backing me up all the way.

1. Introduction

When puzzled by the nature of a physical object, human beings will turn it over and look at it from a new angle to get an idea of what it is and how it can be used. Also, if we are not quite sure what happened on a special occasion, we typically try to take a different perspective to make better sense of the event. In linguistic research, the same approach works, only at a more abstract level - and it may be harder to reach any definite conclusions. As pointed out by Halliday and Matthiessen, we have to construe the world; "it is not 'given' to us as an established order" (Halliday and Matthiessen 1999: 165). Moreover, certain ways of conceptualizing the nature of a phenomenon – in this case English collocations - tend to get entrenched so that even if it remains puzzling, we may not want to give up a perspective that seems to have yielded useful insights in the past. It is therefore often necessary to reconsider the underlying assumptions of a traditional viewpoint before it is possible to look at a problem from a new perspective. In the case of collocations, the underlying assumptions are mainly about categorization and compositionality, issues which raise basic questions about the nature of language and meaning and over which linguists from different camps are deeply divided (see for example Newmeyer 1998).

Conventional collocations are word combinations which we use in typical situations that we find ourselves in as we go about our lives. Our familiarity with these situations enables us to draw on our experience, so that we know what to expect, what to do and even what to say. I believe that it is generally agreed that, to be accepted as a member of a given society, it is essential to be familiar with the linguistic routines of the language, not only to be able to follow them, but also to deviate from them, if that is what the situation requires. This is why foreign learners of a language work so hard to become acquainted with such expressions, which have been found to be pervasive in language.

In this perspective, it seems puzzling that exactly because of the conventional status of collocations, the traditional approach to phraseology describes them as 'arbitrary' and somehow 'defective'. In this thesis I will look critically at the arguments underlying the traditional approach and try to show that a functional and cognitive approach is descriptively more adequate and has a better explanatory potential.

In my introduction I will briefly outline the theoretical framework of the discipline of phraseology, and introduce the different meanings of 'collocation' as a linguistic term before outlining what I see as problematic aspects of the categorization of collocations used in traditional studies of phraseology. My motivation for choosing a functional and cognitive approach will be given, and I will conclude this part with an overview of the study.

1.1 Phraseology: the traditional approach

One purpose of the present study is to review the notions used for the categorization of collocations in 'phraseology', which according to Howarth (1996: 6) is a term which is becoming generally accepted as the name for the discipline within lexicology that focuses on 'fixed expressions' (Alexander 1978, 1987). I am using 'phraseology' also in a wider sense, as a framework including theoretical as well as practical aspects, as it is used in Cowie (1998b) titled *Phraseology. Theory, analysis, and applications*. The study of fixed expressions reflects a traditional interest by British linguists in the problems that usage patterns pose, especially for foreign language learners, and phraseology can thus be seen to carry on the work of lexicologists and lexicographers like H. E. Palmer and A. S. Hornby in the early part of the 20th century (cf. Cowie 2000: 2). However, according to Howarth (1996: 25), "none of them developed a detailed framework for the analysis of composites".

Phraseology is further indebted to 'Firthian linguistics' in that the notion of 'collocation' as a technical term is attributed to John Rupert Firth. His ideas about the relationship between lexis and grammar, system and instantiation, and systematic

variation, have inspired many linguists, also referred to as 'neo-Firthians'. Among them are John Sinclair, who has done pioneering work in corpus linguistics (Sinclair 1991), and Michael Halliday, a student of Firth's, who has developed his own 'systemic functional grammar' (Halliday 1994). They have made important methodological and theoretical contributions to the study of collocations, and I will draw on their work to the extent that it has contributed to, or may be contrasted with, the traditional approach to phraseology that is the discussed in Part 2 of this study and which is defined by its system of categorization.[1]

Howarth (1996: 30) finds that "Neo-Firthian lexicologists have been primarily interested, not in classifying types of collocation, but in the phenomenon of 'collocation' itself and in what it contributes to linguistic meaning as a whole". Since classification is high on the agenda in phraseological studies, analysts have turned to Russian phraseology, whose central concern is "the establishment of criteria for the precise description of phraseological units and their varying degrees of fixedness" (Howarth 1996: 30). 'Classical' Russian theory has been characterized as "probably the most pervasive influence at work in current phraseological studies" (Cowie 1998b: 2). This approach has been extended and modified without being basically revised, and it constitutes the framework of categorization that this study will evaluate from a functional and cognitive perspective (cf. section 2.1.4 below).

It is interesting to note that a more recent development in Russian phraseology, which focuses on the cultural dimension of collocation and is characterized by Cowie (1998b: 2) as "broadly anthropological", draws on concepts developed in cognitive linguistics over the past two decades (see section 1.4 below for examples of cognitive studies in phraseology). Cognitive notions like 'gradedness' and 'fuzziness' also increasingly appear in studies that are based on the traditional framework of categorization, but so far this has not affected the framework as such.

In Part 2, I will argue that, in addition to the acknowledged influences mentioned above, the system of categorization in phraseology is based on further underlying assumptions

[1] When using the terms 'phraseology, 'phraseological', and 'phraseologists' in this study, I will be referring to this traditional approach (see section 1.3 and Part 2 below).

that can be accounted for in terms of structural and generative influences as well as classical notions of categorization. In the introduction to his anthology, *Phraseology. Theory, analysis, and applications*, Cowie describes phraseology as a discipline that has come of age "[f]ollowing a steady growth of scholarly interest and activity over the last twenty years" (Cowie 1998b: 1). It is characterized as being "by no means purely descriptive" and as having contributed to linguistic theory the notion that "native-like proficiency in a language depends crucially on a stock of prefabricated units – or 'prefabs' – varying in complexity and internal stability" (Cowie 1998b: 1). The past twenty years is also the period in which cognitive linguistics has developed a range of notions for the analysis of word combinations, including a theory of categorization that may provide a better framework for the description and explanation of usage patterns than the one inherited from Russian phraseology. This argument will be developed in sections 1.4 and 1.5 below and constitutes the central claim that I want to test in this study.

1.2 Definitions of 'collocation'

Below are listed five definitions[2] of 'collocation' as a technical term, which all have some form of **co-occurrence** as a central element:

(a) the tendency for lexical items to co-occur in a text, or in a text corpus, whether or not they form a syntactic pattern

(b) the co-occurrence of lexical items in a syntactic pattern, only restricted by general selection restrictions (also referred to as 'free/open collocation', or 'free/open combination')[3]

(c) the tendency for lexical items to co-occur in a syntactic pattern restricted not only by general selection restrictions, but also by usage restrictions on one element (often referred to as 'restricted collocation')

(d) the co-occurrence of lexical items in an unexpected, creative way that conflicts with general selection restrictions and/or usage restrictions (normally referred to as 'creative combination' rather than collocation)[4]

[2] I do not mean to imply that these definitions constitute different senses of *collocation*, rather I construe them as applications of the same reading (cf. section 3.4.1.2 above).

[3] For a discussion of a possible distinction, see section 2.2.3 below.

(e) the tendency for a lexical item to co-occur with a preposition or grammatical structure such as an infinitive or clause (referred to as 'grammatical collocation' in contrast to 'lexical collocation')

Recurrence is equally important in the case of (a) and (c), but not in (b), which is neutral as regards frequency of co-occurrence, nor in (d), which specifically excludes recurrence. The last definition, (e), is special in that it does not refer to co-occurrence between lexical items, but between a lexical item and a grammatical pattern.[5] It will not be discussed in the context of this study. Collocation, as defined under (a), is used in text analysis about the simple tendency for items to co-occur in a text, as a feature of lexical cohesion[6] (Halliday 1994: 333). In sociolinguistic studies, it is used as a key to the cultural connotations of words based on their tendency to co-occur with other words in their context (Stubbs 1996: 172 ff.). Stubbs and others who use text corpora refer to the word whose 'collocability' is being examined as the 'node word' and restrict their search for 'collocates' (other words that it 'collocates with') to a 'span' of four words to the left and right of that word.

Phraseologists, who are mainly interested in collocations as a resource for language production rather than for purposes of text analysis, narrow down the notion of collocation to syntactic combinations, thus combining lexical with syntactic analysis. The semantically autonomous word, whose collocability is being examined, is referred to as the 'base' or 'base word', and its collocates , or 'collocators', are said to constitute its 'collocational range' (cf. section 1.3). The paradigmatic grouping of collocates which are alternatives in a collocation without being fully synonymous, are referred to as a 'set' (Halliday 1966: 153; Cowie 1978: 130). Collocations of the (c)-type, which may reflect lexical as well as grammatical usage patterns, are the ones that form the subject of phraseological studies, as distinct from combinations characterized under (b), " … whose make-up can be explained in terms of general restrictions on co-occurrence, and

[4] J.R. Firth (1968c: 18) in his analysis of "Emily-coloured primulas", a line from a poem by Edith Sitwell, talks about collocation as a "level of meaning which may be personal and idiosyncratic, or normal". (Cf. section 2.1.2).

[5] Cf. footnote 8, this section below.

[6] According to Halliday (1994: 310), lexical cohesion, or continuity, in a text can be established by the repetition of a word, or by using a word that is related to a previous words "either semantically, such that the two are in the broadest sense synonymous, or collocationally, such that the two have a more than ordinary sense to co-occur."

which lie outside the limits of phraseology altogether" (Cowie 1998b: 6). To emphasize that some syntactic transformations are normally possible for (c)-type collocations, e.g. *hard work* and *to work hard* are equally possible the base and collocate may be referred to as 'lexemes'.

In the examples in Table 1 below, the base words are shown in bold type. The structural type that I will focus on is syntagms consisting of verbs and their nominal objects. When I refer specifically to the traditional framework used in phraseology, I will refer to these combinations as 'restricted collocations'. I will use 'collocation' alone to refer to 'conventional', 'institutionalized', or 'entrenched' combinations, which I take to include both the (c)-type and some, but not all, of the combinations under (b), since I am going to suggest (in sections 2.2 and 3.4.4.2 below) that a distinction should be made between 'open collocations', which are institutionalized expressions, and 'free collocations', which are not. I will refer to the creative type (d) as 'creative collocations', since, although they are characterized by deviation from usage patterns, I will argue that they rely on these patterns for their effectiveness (sections 1.3 and 3.4.4).

1. **NOUN** + VERB	*the **water** runs, the **candle** burns, the **gap** widens*
2. VERB + **NOUN**	*make a **deal**, run a **deficit,** drill a **hole***
3. ADJECTIVE + **NOUN**	*a rapid **increase**, a deep **cut***
4. COUNT NOUN + *OF* + UNCOUNTABLE **NOUN**	*a stroke of **luck**, a peal of **thunder***
5. PREMODIFYING NOUN + **NOUN**	*a trade **gap***
6. PREMODIFYING PARTICIPLE + **NOUN**	*a negotiated **settlement,** a fishing **rod***
7. ADVERB + **ADJECTIVE**	*deeply **unhappy***
8. ADVERB + **VERB**	*to **wound** sb deeply, to be deeply **wounded***

Table 1: Syntactic patterns of collocation

Furthermore, I will use the expression 'general selection restrictions' when I refer specifically to the phraseology framework; otherwise I will use the expression 'selection preferences', which is preferred in cognitive linguistics as it suggests a less mechanistic view of language.

In phraseology, usage patterns are typically construed in terms of relative restrictedness (see section 1.2 below), which is seen as both idiosyncratic and arbitrary. One reason probably is that many researchers in phraseology are actively engaged in disciplines like second language pedagogy and lexicography, compiling dictionaries for foreign learners of English (Howarth 1996: 6 f.). These learners are the ones who struggle to learn the idiosyncratic combinations of English, which, most of the time, do not tally with the idiosyncratic expressions of their own language. Native speakers retrieve such combinations easily and do not normally experience any restrictedness, presumably because they have learned them as wholes, in typical situations of usage. Drawing the line between restricted and free collocations has been seen as a challenge for researchers in phraseology. As the number of studies has increased, so has awareness of the problems of drawing this line, and researchers are becoming increasingly interested in the 'fuzzy' area bordering on free collocations, where it becomes harder to argue that the combinations are subject to usage restrictions: "From the late 1980s on, studies of collocations have pushed the boundary that roughly demarcates the 'phraseological' more and more into the zone formerly thought of as free" (Cowie 1998b: 18 f.). In the process, interesting issues of methodology have been raised and the traditional method of categorization has come under strain. In the next section, I will give an outline of the notions reflecting elements of recurrence and usage restrictions in the classification of collocation as defined under (c) above. Recurrence, which involves the idea of frequency, raises important questions about methodology, whereas categorization on the basis of usage restrictions is the main theoretical challenge.

1.3 Methodology and classification in phraseology

Recurrence is an important characteristic. Repeated co-occurrence is one indication that collocations are entrenched in the language and not ad hoc combinations. Typically, the concept of collocation has been narrowed down by adding words like 'habitual', 'regular' or 'significant'. Thus Firth, who introduced collocation as a technical term, gave this frequently cited definition: "The habitual collocations in which words under study appear are quite simply the mere word accompaniment, the other-word material in

which they are most commonly embedded" (Firth 1968a: 180). Jones and Sinclair, (1974: 19) distinguish between collocations in general and 'significant collocations' which are defined as "regular collocation between items". This raises the question of how to establish what are recurrent combinations as well as the more specific issue of the role of qualitative versus quantitative methods of identifying them. In general, what researchers have been doing is to rely on introspection, sometimes combined with, or replaced by, the use of statistical corpus tools. I will return to this discussion about methodology in section 3.2.1, arguing that corpora provide useful evidence, also about relative frequencies, whereas the evaluation of the data requires introspection as well as a suitable theoretical framework, which should be able to account for the categories posited as well as the method of categorization.

It is recognized in phraseology that it is the **conventionality** of collocations that makes them linguistically significant. Howarth (1996: 6) lists the following characteristics:

- They are conventional forms that have a significant role in language production.
- This is partly because they are memorized as ready-made lexical units.
- They are stored together with some indication of their grammatical structure and syntactic and pragmatic function.
- As a result they are recognized as familiar by native-speaker readers/hearers both as regards their form and their associated functions.

In spite of the recognition of these cognitive and functional characteristics of collocations, I will claim that categorization in phraseology is still typically based on criterial features (see 2.1.3.3 below). I will now briefly outline the theoretical notions that are typically used in phraseological studies to identify collocations as being subject to **usage restrictions** and point to the problems involved. Restricted collocations like *foot/pick up/settle the bill* are seen as occupying an intermediate position on a continuum ranging from 'idioms', which are typically defined as noncompositional multi-word units like *fill/fit the bill*, to free combinations like *discuss/tear up the bill*, which are seen as fully compositional. Restricted collocations are defined both syntagmatically, in semantic terms, as " … word-combinations in which one element [...] has a specialized meaning determined by the other element [...]" (Cowie and Howarth 1996: 81), and paradigmatically, in terms of substitutability, or recombinability, (Cowie 1998b: 15):

"Here [in the case of restricted collocations] it is typically not a matter of learning fixed units but of knowing when, and how far, the elements of a collocation are able to recombine with other items".

The assumption underlying the continuum, viz. that full compositionality represents the standard from which restricted collocations and idioms deviate, is not unproblematic. In Part 2 (section 2.2.3), this claim will be contrasted with the stance commonly taken in cognitive linguistics that "… component structures are not the building blocks out of which it [a composite structure] is assembled, but function instead to motivate various aspects of it" (Langacker 1987: 449, 453). According to this view, full compositionality should be regarded as an exception rather than the norm. In line with the idea of full compositionality, studies in phraseology typically see restricted collocations as not only idiosyncratic but also arbitrary. This is in contrast to "perfectly open collocations … formed by general principles of co-occurrence" (Cowie and Howarth 1996: 83 f.). This view will be discussed in Part 2 in the light of the contrasting cognitive view that composite structures (cf. section 1.5 below) are neither rule-bound nor arbitrary. According to Lakoff (1987: 148), the assumption "that the meaning of the whole is a computable function of the meaning of the parts plus the syntactic relationship between the parts" is "simply wrong". However, this assumption seems to underlie the terminology used by Moon (1998: 20 f.). In her typology, the equivalent of restricted collocations is 'defective collocations'. They are a subgroup of 'anomalous collocations' which " ... are problematic in lexicogrammatical terms. They are syntagmatically or paradigmatically aberrant: they cannot therefore be decoded purely compositionally nor encoded freely". Her subclassification is based on "the nature of the anomaly" (see section 2.1.3.2 below). As an alternative, Lakoff offers the concept of 'motivation':[7] "The meaning of the whole is often motivated by the meaning of the parts, but not predictable from them" (Lakoff 1987: 148).

As mentioned above, restricted collocations are mostly construed as consisting of a semantically autonomous element, the 'base' or 'base word', and a semantically

[7] 'Motivated' here means 'neither arbitrary nor rule-bound'; in the literature on phraseology it is regularly used to mean 'analysable', as in Howarth (1996: 24)

dependent element, the 'collocate', sometimes referred to as the 'collocator'.[8] The semantic relationship between the elements is thus construed as hypotactic or hierarchical and in such a way that in a noun + verb, verb + noun, or adjective + noun collocation, the noun will be the autonomous 'base', whereas the other elements are construed as collocates that are dependent on the noun (Hausmann 1985: 119). I will return to this question in terms of what Langacker calls 'autonomy/dependence asymmetry' in sections 2.2.4 and 3.4.4.2, arguing that, although it may be typical for the collocate to be semantically dependent on the base, there may not always be "a significant relation of dependence in one direction or the other" (Langacker 1987: 349). Thus the verb collocate is less obviously dependent on the noun base in *drill a hole* than in *make a difference*, and the noun can also be said to be dependent on the verb to the extent that a salient substructure is elaborated by the verb.

Although, in principle, collocation is understood as being between lexemes and not particular word forms, lexicogrammatical patterns are often found where elements tend to combine in some forms rather than others (*galloping inflation* rather than *inflation galloped* or *inflation gallops*).[9] As a result of corpus studies based on frequency of co-occurrence, Sinclair[10] has found that "many uses of words and phrases show a tendency to co-occur with certain grammatical choices (Sinclair 1991: 112). This is interesting in the context of the discussion of whether lexis (in this case collocational patterns) can be seen as independent from grammar (Gramley and Pätzold 1992: 66 ff.). I will return to this discussion in Part 2 (section 2.2.1) and in Part 3 (section 3.4.4.4), arguing that lexis and grammar should be seen as separate but interdependent levels, which are recruited

[8] Morton Benson, one of the authors of the *BBI Dictionary of English Word Combinations*, has described 'lexical collocations' as usually consisting of two "equal" lexical components – in practice, however, a hierarchical principle is adopted in the BBI, which lists verb or adjective + noun collocations under the noun. In contrast to lexical collocations, 'grammatical collocations' are defined as phrases consisting of a dominant word (noun, adjective, verb) and a preposition or grammatical structure such as an infinitive or clause. (Benson 1985: 61-68).

[9] The Google advanced search engine found 3,200 examples of *galloping inflation*, 52 of *inflation galloped*, and 31 of *inflation gallops*. (http://www.google.dk/advanced_search?hl=da).

[10] Sinclair (1991: 110) has formulated the 'idiom principle', which he contrasts with the 'open choice model', and which involves lexical as well as grammatical choices :"The principle of idiom is that a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analysable into segments.

to express integrated meaning. When it comes to the status of specific items as ‘lexical’ or ‘functional’ (here in the sense ‘grammatical‘ or ‘schematic’), it may not just be a question of ‘either/or’, but rather of ‘more or less’.

As will have appeared from what was said about restrictedness above, it is used to characterize the collocate, not the base word, and it is defined both syntagmatically, in terms of semantic constraints on the meaning of the collocate, and paradigmatically, in terms of lexicogrammatical limitations on the substitutability of the collocate. Categorization is based on distribution in terms of limited and arbitrary variability. Sets of ‘overlapping collocations’, a notion introduced in Cowie (1986: 64), are used as evidence of the arbitrary and restricted nature of collocations. They are typically exemplified by clusters of verbs with figurative senses that are construed as synonymous and which collocate with overlapping sets of nouns, as in Fig. 1. below.

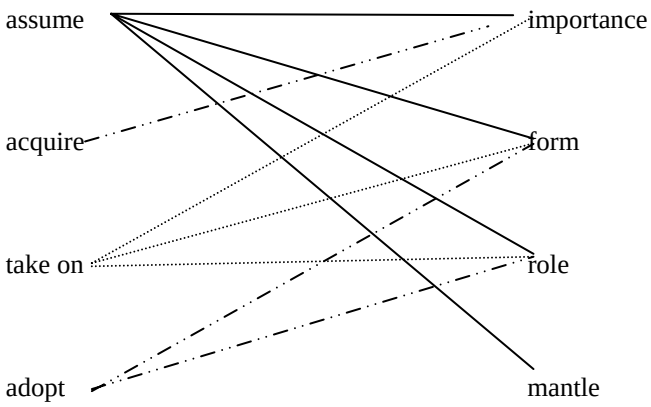


Fig. 1: Overlapping lexical sets (Howarth 1996: 44)

Howarth (1996: 44) argues that “if it can be shown that certain potential collocations in the cluster are arbitrarily blocked, those collocations that are acceptable could be regarded as restricted”. In Part 2 (section 2.2.5), I will pursue the claim that such distributional patterns are not necessarily a good basis for categorization, as they may reflect ‘extrinsic’ factors rather than properties that are ‘intrinsic’ to a word combination (Langacker 1987: 306).

Finally, analysability in terms of **opaqueness** and **transparency** is also used to categorize word combinations; thus Cowie and Howarth (1996: 82) mention opaqueness as an optional criterion in their categorization of restricted collocations, whereas full transparency is seen as a typical feature of free/open combinations/ collocations exemplified by *open the window* and *shut the door*.[11] Full analysability is thus seen as strong, if not conclusive, evidence of full compositionality, against which I am going to argue, in Part 2 (section 2.2.5), that compositionality and analysability should be treated as independent parameters. This is in line with the insight that the notion of analysability "does not refer to the intrinsic complexity of a structure, but rather to a person's awareness of certain aspects of this complexity" (Langacker 1987: 457).

I have outlined what I see as the main elements traditionally used to categorize collocations in studies of phraseology and have pointed to the following claims that I want to discuss:

- Full compositionality is seen as the norm from which restricted collocations deviate, rather than constituting an exception.
- Patterns of collocation are seen as arbitrary rather than motivated.
- Distributional patterns of paradigmatic and syntagmatic restrictedness are used to categorize combinations, whereas these patterns may be merely extrinsic and without relevance for the intrinsic properties of the combinations.
- Analysability, or transparency, is seen as evidence of full compositionality rather than an independent parameter.
- A collocation is seen as a hypotactic structure consisting of a dependent collocate and an autonomous base; a view that does not allow for combinations in which the autonomy/dependence relation is not clearly one-way.

While the existing typology continues to be applied, studies in phraseology will often point to the need for further research. Thus Howarth (1996: 3) refers to "the apparent absence in the literature of theoretical and descriptive studies of this aspect of phraseology" and concludes, "there is no firmly established descriptive framework in

[11] It is interesting that the entry for *door* in *The BBI Dictionary of English Word Combinations* includes both *open* and *shut/close the door* although it is specifically stated in the introduction to the dictionary that "The Combinatory Dictionary does not include free lexical combinations. Free lexical combinations are those in which the two elements do not repeatedly co-occur; the elements are not bound specifically to each other; they occur with other lexical items freely." (Benson et al. 1997: xxx).

which restricted collocations can be placed". Moon (1998: 165) points to typology and categorization as the principal theoretical area for further research and calls for diachronic aspects to be explored. Interestingly, she mentions cognitive notions like 'schemas', 'frames' and 'prototypes' (Moon 1998: 165) as potentially useful concepts in phraseology. Hanks (2000: 3 ff.) specifically encourages lexicographers to utilize the potential of cognitive linguistics as a descriptive framework for phraseology, and in some studies this has already been done (see section 1.3 below). Finally, Cowie (1998b: 8) finds parallels between Mel'čuk's 'lexical functions' and recent work in 'frame semantics'. It should be noted, however, that while the two approaches are not incompatible, they differ in that Mel'čuk's work is in the structural and generative tradition, while Fillmore's analysis of data as evidence of schemas reflecting underlying conceptualizations places his work in a cognitive setting (see sections 1.5 and 3.4.3.4 below).

Studies in phraseology have documented that collocations are an important resource in language production. Thus Cowie and Howarth find in their comparative study of the phraseological competence of native and non-native students "that the percentage of restricted collocations found in an essay, compared with the percentage of combinations of the same syntactic type which were not restricted in any way, is a reliable measure of the written proficiency of its author" (Cowie and Howarth 1996: 84). Also the cognitive and functional status of conventional collocations is acknowledged, but so far this has not influenced the model of categorization used in phraseology.

1.4 Motivation for a functional and cognitive approach

My decision to use a functional and cognitive approach as an alternative framework for the study of collocations is thus partly motivated by theoretical and descriptive work in phraseology, including my own previous study (Poulsen 1991). Furthermore, although restrictedness seems to be an important feature of the behaviour of collocations - especially from the point of view of the foreign learner - it does not in itself offer an explanation of their importance as a language resource, and I will argue that a

satisfactory account should be cognitively and functionally based. Actually, as mentioned above, 'cognitive models' reflecting people's experience of growing up and moving about in the world are beginning to be used in phraseological studies; thus Baranov and Dobrovol'skij (1996) have described the meaning of idioms on the basis of frames and scripts of typical actions, which are examples of cognitive models. Telija, Bragina, Oparina and Sandomirskaya (1994) have looked at cognitive aspects of collocations, and the same authors have used conceptual analysis to investigate the cultural connotations of collocations, as reported by Cowie (1998b: 55 f.). Pauwels (2000) uses a "cognitive linguistic approach to verbal meaning" in his work on the verbs *put, set, lay* and *place*.

Functionalism and cognitive linguistics are typically seen as reactions against structural and generative linguistics, the two dominating theories in 20th century linguistics. Thus Hansen (1998: 159) refers to them collectively as the 'formal paradigm'[12] ('det formalistiske paradigme'), which she characterizes by quoting six so-called 'pre-empirical postulates' listed by Givón (1989: 94 f.). These are postulates which, according to Hansen, both functional and cognitive linguists are expected to disagree with to a greater or lesser extent:

(1) There is an evolutionary discontinuity between pre-human and human communication.

(2) Language is a separate module of the mind/brain, not part of 'general cognition'.

(3) Language structure can be analysed independently of its communicative function.

(4) Some abstract, idealized entity – be it *langue* or *competence* – is the 'object' of linguistic analysis.

[12] Newmeyer (1998: 8) explains 'formal' as follows: "The term is ambiguous between the sense of 'pertaining to (grammatical form)', as opposed to meanings and uses, and the sense of 'formal*ized*', i.e. stated in a mathematically precise vocabulary […] functionalists' criticism of 'formal linguistics' invariably refer to the former."

Lakoff (1987: 486) distinguishes between two types of formalization: he uses the expression 'technical formal systems' to refer to "a special kind of mathematical system of production rules, in which arbitrary symbols are manipulated in an algorithmic fashion without regard to their meaning". According to Lakoff, the claim that grammar is such a system is "the basic metaphor of generative linguistics" . In contrast, formalism in the context of cognitive linguistics may make use of a"practical formal system", which is"a system of principles of some sort expressed precisely, often in a notation which permits one to give precise names to concepts, to state hypotheses in appropriate detail, and to make detailed predictions".

(5) The synchronic facts of language must be studied in strict separation from diachronic facts.

(6) The sign-relation between the linguistic code and its mental designatum is arbitrary, unlike the obvious iconicity seen in pre-human communication.

According to Hansen (1998), functional linguists are especially expected to object to the postulate that language structure can and should be studied independently of the communicative function of language, whereas cognitive linguists would especially object to the claim that language is situated in a special module in the brain, separate from our general cognitive abilities.[13] Although it is convenient to focus on the differences, it is important that most functionalist and cognitive linguists acknowledge the importance of describing the structure of language, although they do not accept claims about autonomous structure.

Functionalism is used as a common denominator for a great diversity of approaches. Van Valin (2001: 329-333), quoting Nichols (1984), distinguishes three broad strands. At one end, 'conservative functionalism' does not challenge generative grammar, but merely adds a functional perspective. At the other end, 'extreme functionalism', exemplified by the work of Hopper (1987), "rejects the validity of any notion of structure other than that of discourse structure and seeks a radical reduction of grammar to discourse" (Van Valin 2001: 331). Between these two general approaches, 'moderate functionalism' sees itself as an alternative to formal theories like generative grammar, but does not try to reduce grammatical structure to discourse structure. In Van Valin's words, moderate functionalists "view grammatical structure as strongly influenced by semantics and pragmatics and undertake to explore the interaction of structure and function in language" (Van Valin 2001: 330).

The functional theory that Van Valin himself subscribes to, viz. Role and Reference Grammar, is given as an example of the moderate brand of functionalism, together with Dik's Functional Grammar (Dik 1989). Halliday's Systemic Functional Grammar (Halliday 1994), which is mentioned as the third of the three "explicitly articulated,

[13] These two postulates reflect different conceptions of 'autonomy': whereas Saussurean linguistics sees language as 'autonomous' in the sense of being a self-contained system, separate from its users, generative linguistics stresses the importance of a human language faculty which is 'autonomous' in the sense of being embedded in the human brain as a separate module.

named theories", is seen as falling between moderate and extreme functionalism, because it "takes a strongly discourse oriented view of language", but "does not deny either the reality of structure in language nor the Saussurean foundations of modern linguistics" (Van Valin 2001: 331-332). This implies that Van Valin considers the sentence to be the normal unit of analysis from which a focus on discourse deviates. As an overall model, Systemic Functional Grammar, with its strong focus on language as a resource rather than just a system of rules, could provide a suitable framework for studying collocations, but collocations are only treated as a feature of text cohesion based on a tendency to co-occur (Halliday 1994: 333); (cf. also definition (a) in section 1.2 above). The functional role of the verb in collocations like *do a dance* (Halliday 1994: 147), and the distinction between expressions like *move the piano* and *play the piano* on the basis of the extent to which the grammatical object can be considered an independent 'participant' (Halliday 1994: 167-168), is compatible with the findings presented in Part 3. Although I will not be able to include it in the present study, I would like to return to this approach later (cf. section 4.2 below).

Although I agree with the view that analysis of natural language cannot be restricted to sentence level, I subscribe to the moderate brand of functionalism, to the extent that I consider paradigmatic and syntagmatic structure as important elements in the description of collocations. However, semantics includes both structural, lexical and contextual elements, and although at any one time collocations are seen as 'entrenched' and relatively stable, allowance has to be made for the diachronic processes of entrenchment and innovation.

Cognitive linguists align themselves with functionalists and in opposition to generative linguistics, e.g. with regard to the distinction between the syntactic, semantic, and pragmatic levels of analysis that cognitive linguists, like functionalists in general, want to break down (Saeed 1997: 299). What sets off cognitive linguistics from functionalism is the emphasis on linguistic knowledge as part of general cognition and on meaning being based on conventionalized conceptual 'structures' such as metaphors, image schemas, and mental spaces (Saeed 1997: 299 f.). A functionalist would emphasize that conceptual structures do not constitute an autonomous system, but interact with social

processes of communication: “Function recruits conceptualizations for a purpose” (Harder: 2001). From a cognitivist point of view, “language is shaped and constrained by the functions it serves and by a great number of related factors: environmental, biological, psychological, developmental, historical, and sociocultural” (Langacker 1999a: 14). Consequently cognitive linguistics embraces a wide range of approaches.

In his cognitive grammar, which is a specific approach within cognitive linguistics, Langacker assumes two basic general cognitive abilities: ‘abstraction’, or ‘schematization’, and ‘categorization’, and claims that “all putative linguistic structures be derivable from primary data via these mechanisms" (Langacker 1999a: 25). The domain of linguistic investigation includes both a ‘descriptivist agenda’ to which generative linguistics tends to give priority, and a ‘functionalist agenda’ (Langacker 1999a: 25).

a. *Descriptivist agenda*

- structure
- sentences
- cognitive representation
- synchronic language structure
- individual languages’ acquired linguistic system
- psychological manifestation
- theory and description

b. *Functionalist agenda*

- function
- discourse
- actual speech behaviour
- language change
- typology and universals
- acquisition process
- neurological basis
- practical application

Langacker argues that both agendas are important and that the dichotomies between structure and function are false from a cognitive point of view. However, what he calls ‘the structural patterns of language’ vary greatly with regard to entrenchment and

stability, and they are dynamic "both in the sense of resting in processing activity (cognitive routines), and also in the sense of being refined, adjusted, and adapted to new circumstances, resulting in complex and ever evolving networks of related patterns" (Langacker 1999a: 21). This does not mean that 'structure' and 'system' are not important: "Moreover, a substantial number of patterns are stable enough, for a long enough span of time, across a large enough population of speakers, that we can take them as constituting a 'linguistic system' susceptible to coherent description" (Langacker 1999a: 21).

An approach that considers both long-term language change, short-term stability and actual flexible speech behaviour, seems suitable for a study of collocations as a language resource. Furthermore, cognitive linguistics tries to provide a natural explanation for language as being based on conceptualization and cognitive routines. Combined with the functional emphasis on the social and communicative role of language, it seems to have greater explanatory potential than a view of conventional expressions as arbitrary combinations, because they fail to combine according to general rules in a fully predictable way.

1.5 What should a theory of collocations be able to account for?

First of all, a model of **categorization** is needed that is suitable for natural language. It seems to me that a basic problem in phraseology has been the conflict between the continuum model, which implies graded category membership, and the attempt to adhere to a traditional, classical model of categorization, in which class membership is based on 'criterial attributes'. To accommodate the great variability of 'fuzzy' natural language data, analysts have had to loosen the criteria for category membership, but that has not solved the problem. Langacker (1987: 16), referring to his own examples, puts it like this: "Unless one alters its basic character, the criterial-attribute model is not equipped to handle such expressions. It can do so only if one loosens the defining criteria, but then there is no non-arbitrary stopping point, and the relaxed criteria hardly serve to distinguish class members from other entities." Applied to collocations that are

neither clearly ‘restricted’, nor clearly ‘free’, a model based on criterial attributes leaves analysts with nothing much but their intuition to fall back on.

An alternative model of categorization is based on the theory of ‘prototypes’ and ‘basic level categories’, developed by Rosch (e.g. 1973, 1975). What distinguishes this model from the classical one is first of all the conception of category membership. Whereas, in the classical model, all members have equal status, prototype theory gives special status to the most representative members on the basis of ‘goodness-of-example’ ratings of their centrality or ‘prototypicality’; a seating instrument with legs is thus judged to be a more prototypical ‘chair’ than one without legs. Moreover, prototype categories cannot be defined in terms of a set of criterial features, and they tend to be blurred at the edges (cf. Geeraerts 1997: 10 ff.). The claim is that prototype effects result from our judgement on the basis of cognitive models, not that they represent inherent category structure (cf. Lakoff 1987: 44 f.). From a functional point of view, the main advantage of prototype categories is that they combine structural stability with flexibility: “new entities and new experiences can be readily associated, perhaps as peripheral members, to a prototype category, without necessarily causing any fundamental restructuring of the category system” (Taylor 1995: 53). Whereas the classical model does not give special status to any particular level of organization, cognitive linguistics furthermore assigns special importance to categorization at the ‘basic level’ in the middle of taxonomic hierarchies. This level has been found to be “cognitively and linguistically more salient than the others” (Taylor 1995: 48).

SUPERORDINATE LEVEL	ANIMAL	FURNITURE
BASIC LEVEL	**DOG**	**CHAIR**
SUBORDINATE LEVEL	RETRIEVER	ROCKER

Table 2: Basic level categories

I will pursue the idea that the entrenchment of collocations can be related to prototypicality at the basic level of categorization, an idea that will be tested in the empirical part of this study. Here I just want to point out that, compared to a classical model, this does not simply imply sorting out language structure according to a different

set of criteria; it means shifting the focus of attention from the categories themselves to the cognitive models underlying categorization.

A theory of collocations should also be able to account for **compositionality**. In Langacker's cognitive grammar, the term 'grammatical construction' "is applied to this entire ensemble: the component structures, their mode of integration, and the resulting composite structure" (Langacker 1987: 277). Component structures are conceptualized as dynamic 'schematic networks' in which the different meanings are nodes linked by 'categorizing relationships' such as 'specialization' or 'extension'. Their mode of integration is described in terms of 'valence relations', involving the combinatorial potential of the component structures. They include correspondences, (often) overlap between shared substructures, and conceptual autonomy and dependence relations. Langacker proposes a synthesis between the prototype model focusing on typical instances of a category, and the schematic network as an abstraction integrating all the members of a category, which in turn elaborate the schema in different ways. Both component structures and composite structures are examples of categories (Langacker 1987: 371). For a theory of collocations, the important thing is that the meaning of the composite structure is not assembled out of component structures in an additive way: "It is more appropriate to say that the component structures *motivate* aspects of the composite structure, and that the degree of motivation is variable (though typically quite substantial)" (Langacker 1987: 292).

Phraseological studies have generally been synchronic, although it is customary to acknowledge a diachronic background for idioms. However, patterns of collocation should also be seen as synchronic evidence of diachronic developments so that mechanisms of **innovation** and **entrenchment** can be accounted for. According to Langacker (1987: 401), when two items are first combined, they are understood in terms of different types of **context**:

1. systemic context: the position of the items in their respective schematic networks
2. syntagmatic context: how the items combine to form a complex expression
3. situational context: the pragmatic circumstances of a usage event

I have argued that judgements about the functionality and acceptability of the new composite structure will be made by hearers/readers in terms of these types of context as well as the Firthian notion of 'context of situation' and the notion of 'co-text' (see sections 2.1.2 and 2.1.5 below). If the new structure is recognized as an extension of the schematic meanings of the component structures and thus motivated by them as well as by the situational context, it will be accepted as an appropriate, or creative innovation. If not, it will be rejected as meaningless (Poulsen 2002: 105). The theory of **'mental spaces'** or **'conceptual integration'** (cf. Fauconnier and Sweetser 1996; Fauconnier 1997; Fauconnier and Turner 1998; Coulson 2000) describes how language users can manipulate reference during discourse by selective projection of input from different input spaces to a blended space, where new meaning may emerge. With its focus on actual usage events, its most obvious use in the context of this study is in the description of creative collocations, as it can be used to show how input from different types of context, including pragmatic context, is recruited and contributes to the meaning of an expression.

With recurrent usage, context-dependency decreases and, through a process of abstraction, the composite structure becomes entrenched or conventionalized. This means that it is no longer understood in terms of the meaning of component structures, although it may still be analysable. Langacker uses the 'scaffolding' metaphor to describe the process: "… component structures are seen as scaffolding erected for the construction of a complex expression; once the complex structure is in place (established as a unit), the scaffolding is no longer essential and is eventually discarded" (Langacker 1987: 461). Although it is not possible to predict exactly which combinations will become entrenched, the theory of 'semantic frames', (Fillmore 1982, 1985; Fillmore and Atkins 1994) can be used to show that entrenchment is both functionally and cognitively based, as already implied above in the section on prototypes and basic level categories. In contrast to mental space theory with its focus on unique usage events, frame semantics focuses on typical situations and actions, which are associated with entrenched cognitive structures combined with linguistic routines. An example is Fillmore's 'commercial event frame', which includes a 'buyer',

a ‘seller’, ‘goods’ and ‘money’ as typical participants, and ‘buying’, ‘selling’ and ‘paying’ as typical actions (Fillmore 1977b: 58 f.). Such frames, according to Lakoff (1987: 68 ff.), are ‘idealized cognitive models’ which we use to “organize our knowledge of the world”. Frames in turn provide structure to mental spaces as an input to a process of innovation, anchoring it in entrenched structure.

Entrenchment of lexical categories may be part of a process of lexicalization, as in the case of compounds like *blackbird* or idioms like *fit/fill the bill*, or it may involve the development of lexical categories into more functional categories in a process of grammaticalization. An example often mentioned in the literature is the development of English *go* from a lexical verb of motion to a free grammatical morpheme indicating future tense (cf. section 3.4.4.4 below). Grammaticalization involves semantic change over a long period of time and has been described as including “metaphorical extension and semantic and morphological reduction leading to greater generality of meaning and wider use and accompanied by infusion of meaning from the context” (Bybee et al. 1994: 6). Drawing on studies of grammaticalization, a theory of collocations should provide a framework for a discussion of whether collocates have in some cases developed a role that is (more) functional. An assumption to this effect seems to be implied by the way in which restricted verb + nominal object collocations are sometimes characterized semantically on the basis of the verb, as belonging to one of three types (Aisenstadt 1979: 71):

(1) The verb is ‘grammaticalized’ , or ‘delexicalized’, as in *make progress; take a turn.*

(2) The verb is used in a figurative sense, as in *launch a campaign; lift a restriction.*

(3) The verb is used in a specialized or ‘technical’ sense, as in *compile a dictionary; drill a hole.*

Verbs in the first group are also sometimes referred to as ‘semi-auxiliaries’, ‘auxiliaries’, or ‘light’ verbs (Mel'cuk, 1998: 30, 36). Baron and Herslund (1998), in their article on ‘verbo-nominal predicates’, which they also refer to as ‘support verb constructions’, discuss the relationship between verb and object in transitive constructions of the type *make an improvement*, consisting of a delexical verb and a so-called ‘object effectum’, a ‘deverbal’ noun denoting the result of an action. In Part 3 (section 3.4.4.4), I will discuss the theoretical aspects of grammaticalization in relation

to collocations, as part of the evaluation of the empirical evidence, and make tentative suggestions as to how verb collocates might be characterized in functional terms.

1.6 Overview

Having introduced my general research aims, I will conclude Part 1 with an overview of the thesis. Part 2 will outline the theoretical and practical background of phraseology while Part 3 includes the empirical part of my study, in which I will test the assumption that a functionally and cognitively based approach will provide a suitable framework for the study of collocations. Theoretical and methodological aspects will be discussed, before I give a detailed description of my research aims and motivate the design of my case study of the composite structure *break an appointment*. After presenting my findings and summarizing them, the suitability of the methods will be evaluated. In Part 4, I will present my overall conclusions as regards the descriptive adequacy and explanatory potential of a functional and cognitive approach as compared to the traditional approach used in the framework of phraseology, and areas for possible future research will be suggested.

2. The foundations of a traditional approach to phraseology

In this part of my study (section 2.1), I will begin by tracing the theoretical influences on phraseology, both those that are routinely acknowledged by phraseologists and underlying assumptions that are not specifically acknowledged and may even be disclaimed. The British lexicography tradition, which I relate to the tradition of teaching English as a foreign language, as well as Firthian linguistics and Russian phraseology are in the former group, while structural and generative notions and the principle of classical categorization are in the latter. Finally, I will trace a strand of cognitive linguistics in phraseology, arguing that it is still on the sideline, waiting to conquer the central ground.

In section 2.2, I will discuss the notions used to categorize collocations in phraseology, developing the claims put forward in section 1.3 above. I will question the descriptive and explanatory adequacy of a framework that is based on a criterial-attribute model of categorization and which characterizes entrenched collocations as arbitrary deviations from a standard of full compositionality. I will pursue the argument that distributional patterns of paradigmatic and syntagmatic restrictedness may be extrinsic and therefore not suitable as a basis for categorization, which should take into account the intrinsic properties of component structures. Furthermore, I will argue that the analysability of composite structures in terms of transparency versus opacity, or literal versus figurative meaning, should be seen as an independent parameter rather than evidence of full compositionality. Finally, it should not be assumed that collocations, by definition, consist of a dependent collocate and an autonomous base; rather allowance should be made for collocations in which the autonomy/dependence relation is not clearly one-way.

Cowie's 1998 anthology of phraseology, referred to in section 1.1. above, as well as his history of English dictionaries for foreign learners (Cowie 1999) have been important sources, as have major recent studies in phraseology (Howarth 1996; Cowie and Howarth 1996; Moon 1998). Not only do these researchers provide documentation for the pervasiveness and complexity of the phenomenon of collocation, they also include in-depth discussions of the problems encountered in categorization and of the theoretical influences inspiring their work.

2.1 Theoretical influences on phraseology

As indicated, the influences of British lexicography, Firthian linguistics and Russian phraseology are specifically acknowledged and discussed in contemporary writing on phraseology, and, as mentioned in section 1.1 above, it is the Russian tradition with its systematic approach to categorization that sets the agenda of theoretical discussion (cf. section 2.1.4 below). Solutions to problems thrown up by the increasing volume of data that recent corpus studies have made available are sought by extending and modifying this framework rather than by discussing alternative ones. To understand this approach, it is necessary to appreciate that, for those who originally developed the concept of collocation, the main challenge was to teach the English language to foreign learners.

2.1.1 A practical concern: teaching English as a foreign language

The Empire did more than supply Britain with useful commodities like cotton, mutton and tea and provide an outlet for her industrial products; by creating a worldwide demand for the English language, it set the stage for the development of the concept of collocation. Adventurous linguists got the opportunity to travel, paying their way by teaching their mother tongue to eager foreign learners. Thus, in 1922, Harold Edward Palmer was invited to Japan as a linguistic adviser to the Japanese Ministry of

Education to examine and report on the teaching of English in secondary schools.[14] A year after his arrival there, in 1923, he was appointed director of a newly established Institute for Research in English Teaching (IRET), whose aims were to encourage "'reform' methods of English teaching, research and experiment in linguistics, and the training of teachers" (Cowie 1999: 5). Whereas the institute did not quite achieve its ambitions for Japan, its significance as a centre for research into the problems of teaching English, according to Cowie, was worldwide. It was during his period at IRET, in 1927, that Palmer started the compilation of a list of collocations, a project in which he was later joined by another Englishman, Albert Sydney Hornby, who came to teach English literature, in which he had a degree from University College, London, but found language teaching to be a more pressing need (Cowie 1999: 8).

Although it had not yet won recognition as a technical term, 'collocation' had been used in the context of linguistics before; thus *the Oxford English Dictionary* (OED) has citations going back to the 18th century:

- **1750** HARRIS *Hermes* II. iv. Wks. (1841) 197 The accusative. In modern languages ... being subsequent to its verb, in the collocation of the words.
- **1751** JOHNSON *Rambler* No. 88, 5 The difference of harmony arising ... from the collocation of vowels and consonants.
- **1873** EARLE *Philol. Eng. Tongue* (ed. 2) §630 All languages use greater freedom of collocation in poetry than in prose.

Awareness of the fact that usage patterns caused problems for learners had been reflected, in different ways, in English dictionaries, especially bilingual ones, dating back to the 16th century (Moon 2000).[15] However, according to Cowie (1999: 52), Palmer and Hornby's collocation project was the first large-scale analysis of phraseology with the needs of the foreign learner in mind. The findings were published in 1933 as the *Second Interim Report on English Collocations* (Palmer 1933a). The

[14] The information about Harold E. Palmer and Albert S. Hornby in this section is based on their personal histories, which are briefly chronicled in the introduction to Cowie 1999.

[15] In his article on the challenges that collocations present for lexicographers, Hausmann (1985) reviews specialized French dictionaries of collocations, the oldest one being *Les epithetes de M. de La Porte, Parisien* from 1571. Specialized dictionaries of English collocations did not appear till the 1980s (Cowie 1999: 77 ff.).

report treats collocations as specific instances of syntactic patterns (definition (b) in section 1.2 above); the question of whether they should be analysed at a separate, lexical level had not yet come up (cf. section 2.1.2 below).

While acknowledging the pioneering work of Palmer and Hornby, phraseologists tend to comment critically on their use of collocation as an inclusive term covering expressions that phraseologists today would typically classify as idioms. As an example of an idiom that he finds is wrongly categorized as a collocation in the *Second Interim Report*, Cowie (1999: 56) gives *to hold one's tongue*, which "has developed a sense ('be silent') which is no longer straightforwardly relatable to the literal holding of one's tongue (and is to that extent an 'idiom')."[16] According to Cowie (1999: 56), Palmer objected to the term 'idiom', which he felt was too loosely defined; instead he chose to use 'collocation', which was not yet established in linguistics, and which was to be defined as a theoretical term by John Rupert Firth 20 years later (see section 2.1.2 below).

This discussion about the borderline between idioms and collocations is a recurrent theme in the literature on phraseology, but it tends to get less attention now that the problem of distinguishing between restricted and free collocations has turned out to be more difficult to deal with. In section 2.2 below, I will return to the discussion of categorization. At this point, let me suggest that a lasting influence of Palmer and Hornby's work has been to fix the perspective of studies in phraseology to that of the foreign learner and the teacher of English as a foreign language. It may be a consequence of the commitment to this point of view that theoretical issues still tend to be addressed with an eye to the application of findings in foreign language teaching and in dictionaries for foreign learners. This may involve a bias in favour of an approach that sees collocations as an inherently arbitrary phenomenon that is best addressed by strict categorization. I will pursue this claim in section 2.1.3.3 below; for now I will let Palmer have the last word on the *Second Interim Report*, as reported by Cowie (1999: 52 f.): "It will tend to confirm his [the language teacher's] impression that it is not so much the words of English nor the grammar of English that makes English difficult, but that that vague and undefined obstacle to progress in the learning of English consists for

[16] The relevance of the literal/figurative distinction is discussed in section 2.2.5.

the most part in the existence of so many odd comings-together-of words" (Palmer 1933a, 1933b).

Palmer and Hornby were presumably the first ones to use the term 'collocation' in the context of phraseology. They used it to denote syntactic word combinations reflecting usage patterns, but covering a narrower range than the established term 'idiom', which they also held to include proverbs, sayings, and figurative expressions (Cowie 1999: 54). However, as mentioned in section 1.1 above, phraseologists will say that they have taken the theoretical notion of collocation not from Palmer and Hornby but from Firth. In the next section, I will discuss Firth's concept of collocation, arguing that, whereas he did not provide the categories that phraseologists felt were needed, his theoretical notions are in many respects compatible with a functional and cognitive framework.

2.1.2 Firthian linguistics

Trying to explain why Britain did not have a chair in linguistics before 1944, the American linguist J. C. Catford (1969: 248) noted that "The British are reputed to be more concerned with practical things – with applications – than with theories". It is true that, because of their commitment to the point of view of the foreign learner, people like Palmer and Hornby took a practical approach to language, and they used the term collocation mainly as a label for a phenomenon that had to be taught, and for which no established term existed. Firth, on the other hand, introduced collocation as a technical term to denote one of the levels of meaning posed in his context and usage based theory.

Firth argued that although a language event was basically to be considered as "a whole" which was "integral in experience", for the purposes of linguistic description it should be construed as being dispersed to four interdependent and 'mutually congruent' **levels of meaning**, using techniques suitable for each (Firth 1968a: 176).

(1) context of situation

(2) collocation

(3) syntax including colligation (relationships between categories rather than words as such)

(4) phonology and phonetics

According to Firth, to focus attention on one level of analysis at a time would help the linguist deal with the complexity and heterogeneity of language, especially if, at the same time, the analysis was limited to the language belonging to a specific area of experience, a so-called 'restricted language' (Firth 1968b: 98). Firth explained that a statement of meaning could not be achieved "by one analysis at one level, in one fell swoop" (Palmer 1968: 5). In order to emphasize that he saw the relationship between the levels as one of interdependence, Firth is reported to have used the metaphor of a lift moving freely between levels "without giving priority to any one and without proceeding in any one direction" (Palmer 1968: 5). I will discuss the three first levels, leaving out the level of phonology and phonetics, which is not directly relevant to collocations:[17]

Firth made a clear distinction between the analytical levels of **collocation** and **syntax**: whereas the collocations of a word amounted to "statements of the habitual or customary places of that word in collocational order but not in any other contextual order and emphatically not in any grammatical order", meaning at the grammatical level was stated "in terms of word and sentence classes or of similar categories and of the interrelation of those categories in colligations" (Firth 1968a: 180 f.). This distinction has been used as the basis for an approach that treats lexis as an independent level and studies collocations without considering the syntactic combinations that words enter into (cf. definition (a) in section 1.2 above). Paradoxically, as mentioned in section 1.3, an important finding of that approach has been that grammatical and lexical choices are in fact closely related. Thus Stubbs (1996: 37) refers to the phenomenon of 'co-selection' of lexis and grammar, using Sinclair's example of the noun *lap*, which is

[17] Prosody can of course add meaning to a collocation, for example by a 'marked' stress pattern: *You have got to* '*take an* '*interest*, where stressing the verb draws attention to the analysability of the collocation. Body language could be added as a further level of meaning, which was not considered by Firth, however.

more likely to occur in a prepositional phrase in adjunct position than to occur as the subject or object of a clause (Sinclair 1992: 14). Perhaps as a reaction to the view of lexical items as 'slot fillers' in syntactic structures, some linguists, turning the tables, will claim that "each word has its own grammar" (for example Stubbs 1996). In contrast to those who give priority to either syntax or lexis, Firth finds meaning at all levels and does not give priority to any one of them. By comparison, in cognitive linguistics the levels of analysis are not clearly separated, and both the lexicon and the grammar are studied as "inventories of meaningful units" (Geeraerts 1997: 9).

With his study of the word *get*, Firth demonstrated that he did not mean for analysis at the collocational level to stand alone; rather it should "be completed by a statement of the interrelations of the syntactical categories within the collocation" (Palmer 1968: 23). In his article, "Lexis as a linguistic level", Halliday (1966: 159) also emphasized that analysis at the collocational and syntactic levels should be combined so that it would be possible to compare descriptions based on the 'lexical' and the 'lexicogrammatical' method. Halliday acknowledged that levels were separated only for the purpose of analysis and that a well-formed statement had to "observe lexical usage restrictions as well as grammatical rules of syntax; i.e. it must be '**lexicogrammatical**'" (Halliday 1966: 161). Mitchell, also writing in the Firthian tradition, agreed with Halliday that distribution should be accounted for in lexical as well as grammatical terms (Mitchell 1971: 50 f.). He described collocations as "particular members of generalized classes of associations that we have labelled 'colligations'", but did not foresee the phenomenon of 'co-selection' later noted by Sinclair. Rather he emphasized that collocation should be seen as not of 'words', but of 'roots' or 'lexemes' with a certain 'scatter' of forms, e.g.: *he works hard, a hard worker, hard-working,* and *hard work.*

Firth's concept of '**context of situation**' reflects his 'monistic' view:[18] "As we know so little about mind and as our study is essentially social, I shall cease to respect the duality of mind and body, thought and word, and be satisfied with the whole man, thinking and

[18] 'Monism' is contrasted with the 'mind-body dualism' associated with the thinking of the 17th-century French philosopher and mathematician René Descartes. In his theory, to which cognitivists strongly object, mind, and body refer to different kinds of entities that interact causally: the mind causes the body to do things, and the body causes the mind to feel things.

acting as a whole, in association with his fellows" (Firth 1968a: 170). This implied that the language should not be explained by referring to "inner mental happenings" that could not be observed anyway, but by observing language in use. Unlike Malinowski (1935), the anthropologist from whom he took the idea, Firth construed 'context of situation' as an abstraction, "a schematic construct for application especially to typical 'repetitive' events in the social process", but, at the same time, it also ensured the continuity with actual language events, providing "an insurance that a text is attested as common usage in which the occasional, individual and idiosyncratic features are not in the focus of attention" (Firth 1968a: 176). This element of continuity between language use and language system is not found in the Saussurean version of structural linguistics (cf. section 2.1.3.1 below).

Firth's ideas about habits in language conflicted with the notion of competence later developed by generative linguists, according to which creativity was based on the operation of rules of syntax (see section 2.1.3.2 below). By acknowledging usage constraints on language, Firth seemed to be adopting a regressive positivist and behaviouristic position. Thus Langendoen (1968: 3) dismissed Firth's ideas as "of no interest at all for the study of meaning", mainly because "Firth's view is based on the opinion that language is not 'creative' and that a person is totally constrained essentially to say what he does by the given social institution". From a functional and cognitive point of view, however, there is no conflict between creative innovation and the existence of entrenched language routines abstracted from concrete situations of use; rather the former is construed as building on the latter (cf. section 1.5 above).

Actually, Firth's theory of context of situation is functionally motivated, and it seems to have a great deal in common with Fillmore's theory of frame semantics, as Firth's notion, too, associates typical situations with linguistic routines (Firth 1957: 28):

> Conversation is much more of a roughly prescribed ritual than most people think. Once someone speaks to you, you are in a relatively determined context, and you are not free just to say what you please. We are born individuals. But to satisfy our needs we have to become social persons, and every social person is a bundle of roles or personae; so that the situational and linguistic categories would not be unmanageable. Many new categories would arise from a systematic observation of the facts.

This shows that for Firth, as for functional and cognitive linguists today, language categories reflect cognitive processes of abstraction from actual usage events: “There are no brute facts. [...] An isolate is always an abstraction from the language complex, which is itself abstracted from the mush of general goings-on” (Firth 1968a: 199). It should be noted that Firth’s abstract notion of context of situation differs from Langacker’s notion of situational context (see section 1.5 above), which refers to the pragmatic circumstances in a particular usage event - ‘the mush of general goings-on’ - and which constitutes an additional level of meaning not included in Firth’s model. Today, this level has its own models of description and analysis, for example in the theory of mental spaces and conceptual integration mentioned in section 1.5 above.

Firth described **collocations** as words in their ‘familiar’, ‘habitual’, or ‘usual’ company: “In this connection, I would like to put forward the concept of collocation […]. This is the study of key-words, pivotal words, leading words, by presenting them in the company they usually keep - that is to say, an element of their meaning is indicated when their habitual word accompaniments are shown” (Firth 1968b: 106, 1968a: 179 f.). The fact that Firth talks about ‘an element of their meaning’ shows that he sees words as polysemous, and not as ‘building blocks’, which is also apparent from his statement that “Words must not be treated as if they had isolate meaning and occurred and could be used in free distribution” (Firth 1968c: 18). Firth stated that collocation was not to be interpreted *as* context, but, on the other hand it seems that collocations are conceived of as being *in* a specific context: “the *placing* of a *text* [that is actual words] as a constituent in a context of situation contributes to the statement of meaning since situations are set up to recognize use” (Firth 1968c: 18) and, quoting Wittgenstein (1953) “the meaning of words lies in their *use*” (Firth 1968a: 179). It appears that even if meaning by collocation adds meaning to a context of situation, the habitual co-occurrence of words in that context is what created that meaning in the first place, and it is only activated in that context: “You shall know a word by the company it keeps!” (Firth 1968a: 179). I will pursue the argument, in section 3.4.4 below, that whether a collocation is ‘familiar’, ‘usual’ or ‘habitual’ cannot be ascertained quantitatively, by measuring statistical frequency of co-occurrence in a corpus, but should be assessed qualitatively, in relation to the specific contexts of situation. This amounts to an

extension of Halliday's notion of lexicogrammatical 'well-formedness' to include the contextual level.

Firth's ideas of polysemous word structure and language as use are compatible with the cognitive notion of compositionality according to which component structures are conceived of as dynamic schematic networks of meaning (see section 1.5). The following remark seems to support this line of argument: "The structures attributed to 'texts' […] are schematic. Only within such limited systems can commutation provide the basis of a functional or meaning value, and substitution not amounting to commutation, the absence of such value" (Firth 1968c: 18). The following examples from Firth's study of *get* (ibid. 20) can be used to illustrate (1) 'commutation',[19] which changes the meaning of *get*, and (2) 'substitution not amounting to commutation', which does not:

(1) *he's got the blues* ⇒ *he's got the sack*

(2) *he's got the blues* ⇒ *he's got the measles*

It is important to note Firth's point that substitutions in collocation were not to be regarded as synonyms (Firth 1968c: 23) and that he did not discuss substitution in terms of the usage restrictions that are a central concern in phraseology (cf. sections 1.3 and 2.2).

The distinction between **syntagmatic structure** and **paradigmatic system** was stressed by Firth as being "the first principle of analysis" (Firth 1968a: 200): "Structure consists of elements in interior syntagmatic relation and these elements have their places in an order of mutual expectancy. […] Systems of commutable terms or units are set up to state the paradigmatic values of the elements". Halliday (1966: 152) argued, in support of Firth's distinction between a lexical and a grammatical level of analysis, that, on both the syntagmatic and the paradigmatic axes, lexical patterns differed from the grammatical patterns of 'structure' and 'system' and should therefore be referred to by different terms:

[19] The Danish linguist Louis Hjelmslev (1899-1965) is credited for the notion of 'commutation' and the method of using a 'commutation test' to separate the senses in a paradigm (Hansen 1998: 38).

> First, in place of the highly abstract relation of structure […], lexis seems to require the recognition merely of linear co-occurrence together with some measure of significant proximity, either a scale or at least a cut-off point. It is this syntagmatic relationship which is referred to as 'collocation'.

Studying collocations independently of syntax poses the problem of how to define 'significant proximity'. In corpus linguistics it is now defined rather narrowly and, it would seem, somewhat arbitrarily as "a maximum of four words intervening" (Sinclair 1991: 170), whereas Firth himself indicated that the analysis should include complete sentences and, in the case of conversation, might even be extended to include the utterances of preceding and following speakers (Firth 1968b: 106). Mitchell (1971: 53) also emphasized that collocation was not mere juxtaposition and could cut across sentence boundaries, thus providing evidence of the "essentially on-going nature of language": *He didn't want the job. I don't think he even applied* [my emphasis]. Assuming that habits with words are likely to be formed in habitual situations, I suggest, as a principled way of determining cut-off points, that the amount of actual text included for the purposes of analysis at the collocational level should be enough to identify the abstract context of situation involved.

On the paradigmatic axis, as pointed out by Halliday (1966: 152), lexical and grammatical patterns also differ:

> Similarly in place of the 'system' which, with its known and stated set of terms in choice relation, lends itself to a deterministic model, lexis requires the open-ended 'set', assignment to which is best regarded as probabilistic. [...] Collocational and lexical set are mutually defining as are structure and system: the set is the grouping of members with like privilege of occurrence in collocation.

Halliday uses the following examples with *strong* and *powerful* to illustrate how lexical sets vary with the collocation in question:

a strong argument	*a *strong car*	*strong tea*
a powerful argument	*a powerful car*	**powerful tea*

As Halliday put it "the paradigmatic relation of *strong* to *powerful* is not a constant but depends on the syntagmatic relation into which each enters" (Halliday 1966: 150). Halliday thus defines a 'lexical set' as follows:

(1) it is 'open-ended'
(2) assignment to the set is best regarded as 'probabilistic'
(3) its members have 'like privilege of occurrence in collocation'

The first feature, open-endedness, needs to be modified to allow for the third feature according to which the paradigmatic set is defined by the syntagmatic relationship of collocation. As Halliday's examples of collocations with *car* and *tea* show, it cannot be assumed that sets can be freely extended to include synonyms. On the other hand, the set including *strong* and *powerful*, which is defined by its collocation with *argument*, can be extended to include adjectives like *convincing*, *compelling* and *potent.*[20] The second feature, which refers to 'probabilism' only indicates that membership of a set cannot be predicted with any certainty, and does not point to any principle that accounts for the limits to substitutability. I would therefore characterize this feature as the placeholder for a theory that offers an explanation for the nature of the set. As regards the third feature, 'like privilege of use', it is not to be expected that the set will consist of synonyms. If collocations are 'habits with words' reflecting the context of situation in which these habits were formed, I would claim that the set is best construed as a prototype category or schematic network that mirrors habitual variations in use (cf. the empirical studies in Part 3 (sections 3.4.2 and 3.4.3).

In line with my previous arguments, I suggest that the level of context of situation should be recognized as having a special place in the theory, for the following two reasons: Firstly, it differs from the levels of syntax and collocation as it cannot be analysed in terms of syntagms and paradigms in quite the same way; rather it consists of specific processes and role configurations determined by abstraction from actual events. Secondly, it provides the functional framework within which the other levels are integrated. I find it plausible that, in an actual language event, conceptualization should start from the abstract context of situation and recruit the associated lexicogrammatical patterns from there. Fig. 2 below illustrates the integration of levels in a series of language events: the occasions on which categories "find application in renewal of connection with the sources of the abstractions" (Firth 1968a: 200).

[20] The examples are from Collins's Corpus Concordance Sampler accessed at (http://www.collins.co.uk/Corpus/CorpusSearch.aspx

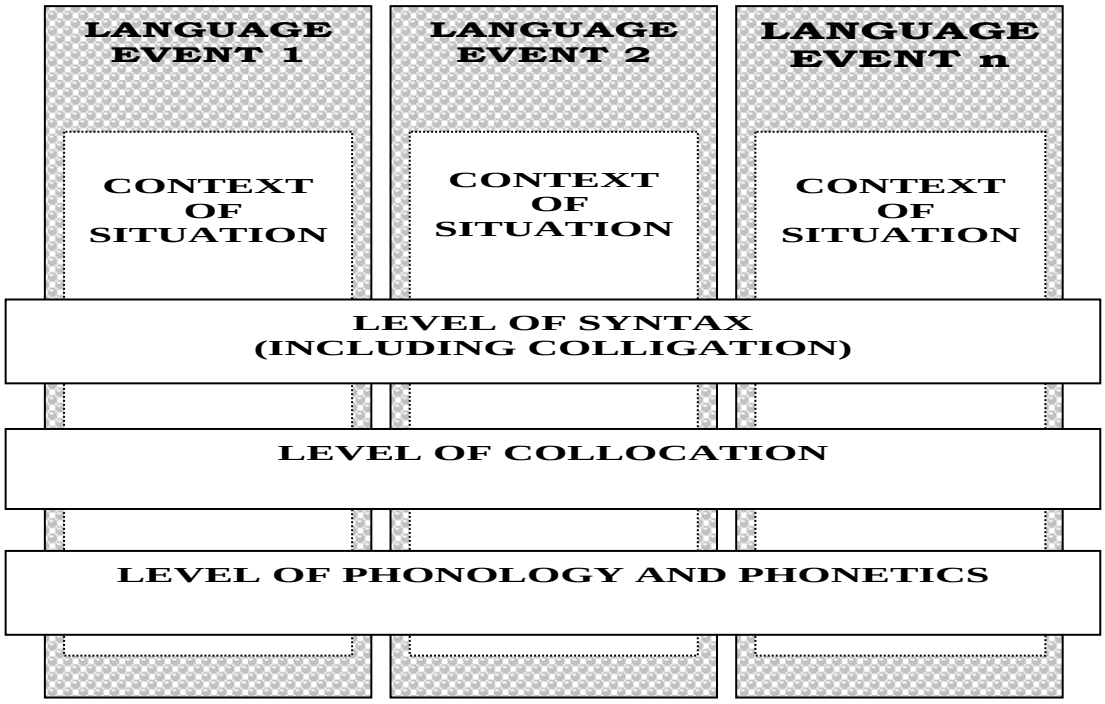


Fig. 2: Integrated levels of meaning (based on Firth 1968a).

This model does not imply, however, that language events, which correspond to Langacker's situational context (cf. section 1.5 above, and this section below), are constrained to the activation of entrenched patterns; on the contrary, creative strategies are perfectly possible and normal at all levels. Actually, one of the few examples that Firth analyses is not a 'habitual', 'usual' or 'familiar' collocation, but a creative one from a poem by Edith Sitwell [my emphasis]:

For spring is here, the auriculas
And the Emily-coloured primulas
Bob in their pinafores on the grass

To help define the meaning, or 'value'[21] of *Emily*, Firth set up a number of paradigms of habitual collocations including *-coloured*, and *Emily* was related to the one in which

[21] Palmer (1968: 7) comments: "Firth took from de Saussure (while rejecting much of de Saussure's theory) the notion of value (valeur). There was commutation of the terms in each system and the values of the terms were derivable from the system itself". Firth disagreed with the structuralist view of language as a single system, especially, according to Palmer, he objected to the addition by Antoine Meillet, student of Ferdinand de Saussure, that it was a system "où tout se tient", i.e. a self-contained system where everything depends on everything else.

coloured was preceded by nominals like *rose, coffee, chocolate,* and *plum*. This method, which combined analysis on the collocational and the syntactic level, was extended on the level of collocation to include the wider context: "part of the meaning of *Emily-coloured primulas* is collocation with *Bob in their pinafores on the grass*. This level I have termed meaning by collocation, which may be personal and idiosyncratic, or normal" (Firth 1968c: 15 ff.). Meaning by collocation, according to Firth, as reported by Howarth (1996: 27), was first and foremost "an abstraction at the syntagmatic level" (Firth 1957: 194) and he was not concerned with the identification of subtypes of collocation based on restricted substitutability. His examples cover a wide spectrum, and some of them are unlikely to be classified as collocations by phraseologists today: *silly ass* (figurative idiom); *Emily-coloured* (nonce formation), *British Army* and *English literature* (free collocations); (Firth 1968a: 179, 1968b: 106, 1968c: 15).

In my view, what Firth's concept of meaning by collocation amounts to is a theory of compositionality based on the polysemous nature of words and the enormous meaning potential of word combinations, manageable only because words are used in habitual company in familiar contexts of situation. Firth's multi-level 'polysystemic' framework seems compatible with the nonreductive approach found in cognitive linguistics, especially the notions of semantic frames and of dynamic schematic networks, and I suggest that, between them, these notions should be able to provide a functionally and cognitively based framework for the explanation of the phenomenon of collocation.

Whereas Firth was more interested in the meaning potential of collocations than in restrictions on substitutability, Halliday (1966) discussed the idiosyncratic nature of the set, and Mitchell (1971: 54) commented on the encoding problems that it posed: "We are probably all aware of the operation of […] collocational constraints as we search for the 'right' choice among, say, *achieve, accomplish, effect, execute, implement, realize,* etc. to associate with *plan* or *project* or *proposal* or *ambition* or *object* or *objective,* and a certain inescapable 'prescriptivism' informing language choices is perhaps worthy of note in passing". Phraseologists, on the other hand, see restrictions on substitutability as a central problem, and they find that the theoretical framework developed by Firth and

his followers does not lend itself easily to the strict categorization of collocations that they believe is required (cf. section 1.1 above).

2.1.3 Underlying assumptions

In this section I will claim that, in addition to the theoretical sources discussed so far, structuralist and generative tenets have influenced the system of categorization used in phraseology. This means that there is an inbuilt tension between a polysystemic, usage-based approach, represented by Firthianism, and monosystemic approaches represented by the more strictly synchronic Saussurean version of structural linguistics and arguably also by generative linguistics. Paradoxically, since the subject of study is language use, the usage-based approach seems to have lost out. Moreover, while the functional and cognitive significance of collocations is acknowledged, this is mainly done to legitimize the subject as a worthwhile area of study from an applied viewpoint, and the explanatory potential of a functionally and cognitively based approach has not been explored.

2.1.3.1 Structuralist dichotomies

In the structuralist framework, language is conceived of as a single, self-contained system, where the 'values' of linguistic entities are defined by their relations to other entities within the system. Ferdinand de Saussure used the metaphor of a game of chess, where rules define the way in which the pieces can be moved in relation to each other. The book *Cours de linguistique générale*[22] "is credited with turning the tide of linguistic thought from the diachronic […] orientation which had dominated nineteenth-century linguistics to interest in the synchronic […] study of language" (Campbell 2001: 95 f.). Today nearly all approaches to linguistics can be said to be structuralist in some sense, as Campbell points out, but at the same time functional and cognitive approaches reflect

[22] *Cours de linguistique générale was* was based on students' notes from Saussure's course in general linguistics at the University of Geneva and published in 1916, three years after his death.

a renewed interest in the continuity of language (see section 1.4), whose synchronic structure is seen as resulting from ongoing diachronic processes of usage.

Whereas Firth also took a structuralist approach to the analysis of language and believed that language should be studied from a synchronic point of view, he saw no discontinuity between language use and language system. His approach was polysystemic and usage based in the sense that the abstract contexts of situation were construed as being related "in renewal of connection", i.e. by instances of use, to "an observable and justifiable grouped set of events in the run of experience" (Firth 1968a: 175). Firth thus added an important element of continuity to structural linguistics, but this element is not present in phraseology, which seems to have embraced a more purely synchronic, Saussurean type of structuralism instead.

Categorization in phraseological studies relies on features like 'arbitrary restrictedness', 'analysability' or 'motivation', and 'freedom' of collocation (cf. section 1.3 above), which appear to be based on the Saussurean notion of the arbitrary sign and the dichotomy between the socially shared language system, 'langue', and 'parole', the actual utterances of individuals, or 'speaking'.[23] The linguistic sign, according to Saussure (1966: 5), unites a concept, the 'signified', and a sound-image, the 'signifier'. The relationship between them is assumed to be 'arbitrary' in that it is 'unmotivated', meaning that there is no natural connection between the concept and the sound-image, for example between "the idea of 'sister'" and "the succession of sounds *s-ö-r* which serves as its signifier in French" (Saussure 1966: 67). What arbitrariness is not meant to indicate is that individuals have a choice; once established, the signifier is "fixed, not free, with respect to the linguistic community that uses it" (Saussure 1966: 71).

> We say to language: "Choose!" but we add: "It must be this sign and no other." No individual, even if he willed it, could modify in any way at all the choice that has been made; and what is more, the community itself cannot control so much as a single word; it is bound to existing language.

This principle, which Saussure himself referred to as 'the stacked deck', shows that his characterization of the approach as 'static linguistics' (Saussure 1966: 101 f.) is to be

[23] 'Speaking' is used as the equivalent of 'parole' in the English translation of *Cours de linguistique générale* (Saussure 1966).

taken quite literally (Saussure 1966: 71). Because of his strictly synchronic point of view, Saussure equates motivation with analysability; a view also found in phraseology, e.g. in Howarth's study (1996: 18), where he talks about "the 'motivation' or analysability of idioms". Such a view of motivation blocks a functional explanation, since convention is recognized only as an arbitrary restriction on present freedom in language: "At every moment solidarity with the past checks freedom of choice. […] Because the sign is arbitrary, it follows no law other than that of tradition, and because it is based on tradition, it is arbitrary" (Saussure 1966: 74). The same view predominates in phraseology today, thus Howarth (1996: 20) specifically relates arbitrariness to conventionality.

Saussure (1966: 131 f.) did allow for variations in the degree of arbitrariness, however; thus composite expressions, exemplified by *dix-neuf*, could be relatively motivated by the systematic relations that each item separately entered into: " ... dix-neuf suggests its own terms and other terms associated with it (e.g. dix 'ten', neuf 'nine', vingt-neuf 'twenty-nine', dix-huit 'eighteen', soixante-dix 'seventy', etc.)". The degree of arbitrariness, then, depended on "the ease of syntagmatic analysis and the obviousness of the meaning of the subunits present" (Saussure 1966: 132). Saussure further emphasized that the value of the whole term was never equal to the sum of the value of the parts, implying that (as in a functional and cognitive framework) the component items are not regarded as building blocks, and their meaning is therefore not expected to be predictable from that of the component items, but neither would it be entirely independent of their meanings, or 'values' in other contexts.

Saussure posed absolute arbitrariness as the main principle, because although one could imagine how at some point in the past a signifier was first combined with a signified, such an act was never recorded, and the question was not worth asking anyway since "the only real object of linguistics is the normal, regular life of an existing idiom" (Saussure 1966: 71 f). The question is: if the original choice had not been well motivated, would the sign still be used? I find it safer to assume that the main principle is relative arbitrariness, of which evidence abounds in the products of extension, composition and blending causing one signifier to be associated with ever growing

networks of signifieds. Polysemy thus provides synchronic evidence of diachronic processes, showing that although these processes are not predictable, they are not arbitrary either, but 'motivated' (in the sense in which the word is used in a functional and cognitive approach). As far as collocations are concerned, it can be concluded that they do not have to be construed as absolutely arbitrary even in a Saussurean framework, and my claim is that construing them as relatively arbitrary, or motivated, has greater explanatory potential.

Interesting in the context of this study is the place of collocations in the langue/parole dichotomy. Since the object of study is langue, (the language system), which can, however, only be accessed via parole (speaking), the theory needs to address the problem of separating the two. In his discussion of the status of syntagmatic relations, Saussure (1966: 124) granted that there was "no clear-cut boundary between the language fact, which is a sign of collective usage, and the fact that belongs to speaking and depends on individual freedom". The sentence, he argued, belonged to speaking, not to language, because it was "characterized by freedom of combinations"; conventional expressions, on the other hand, belonged to language, because usage prohibited change. In addition to conventional expressions, which were often characterized by irregularities of syntax or semantics, language also included expressions that were constructed regularly if they exemplified a pattern. The principle used here to distinguish between langue and parole is also found in phraseology to distinguish between restricted collocations and free collocations: syntagms that can be constructed freely without any usage restrictions do not belong to langue, and they should not count as restricted collocations. However, the problem is where to draw the line.

2.1.3.2 Generative principles

While Saussure had given usage a central place as part of the language system (langue), and Firth acknowledged usage both at the level of collocation and at the level of context of situation, generative linguistics shifted the focus completely from what was seen as

external behaviour to knowledge of language, lodged in the mind of the individual native speaker as an innate capacity and, at the same time, reflecting universal principles applying to all languages. The 'Chomskyan revolution' (Campbell 2001: 100) gained momentum in the late 1950s after Noam Chomsky had published the first version of his theory, *Syntactic structures* (1957), challenging the behaviourist approach then dominating American structural linguistics. This approach was associated with linguists in the Bloomfieldian tradition, who held that in order to secure scientific objectivity it was best to start from the assumption that children's minds were blank slates, and that language learning depended on external stimuli. The linguistic responses to such stimuli then constituted the behaviour to be studied by the linguist, who could draw parallels to stimulus/response experiments conducted on rats and pigeons, as described in B. F. Skinner's book *Verbal behaviour* (Skinner 1957). Chomsky wrote a lengthy review containing damaging criticism of Skinner, till then a highly respected experimental psychologist and leader of the behaviourist movement (Gardner 1995: 1).

A linguistic school that started as a reaction against an approach relying exclusively on usage and which is mainly interested in universal principles of syntax is not the most obvious framework for the study of collocations. In fact, as already mentioned (section 2.1.2), Firth's notion of meaning by collocation was deemed to be of no interest at all by Langendoen, a student of Noam Chomsky's. According to Stubbs (1996: 43), " ... a contrast is inevitable with the Chomskyan position, where the emphasis is on creativity, but where routine can be conceived of only negatively, with behaviourist connotations, as mere habit formation". What Stubbs refers to as "one of Firth's major insights", namely "the need to achieve a balance, in theories of human behaviour, between freedom and constraint, variation and routine, individual and social" (Stubbs 1996: 43) was not appreciated by generative linguists. Nevertheless, I will argue that the central notion used to categorize collocations in phraseological studies, viz. restrictedness, is influenced by the generative idea of creative linguistic competence. This idea might not in itself be incongruent with the notion of meaning by collocation if it had not been for the fact that it was coupled with a formal language of description, which assumes classical categories based on discrete features (see section 2.1.3.3 below).

The distinction in generative linguistics between competence and performance is sometimes presented as being a parallel of the Saussurean langue/parole dichotomy, as when Newmeyer (1998: 289) refers to the former as "more current parlance" for the latter. However, a comparison between the two sets of terms reveals differences that are significant in the context of a study of collocations. In both cases, the idea is to make a distinction between the language system and actual acts of speaking, but whereas Saussure focused on the system of language as a socially shared phenomenon that did not allow any choice for the individual (cf. section 2.1.3.1), the focus in generative linguistics is on the creative competence of the individual native speaker to generate an infinite number of sentences based on a finite number of principles. Moreover, while for Saussure freedom of combination was associated with parole, for generative linguists the principles behind this freedom are at the very heart of creative linguistic competence. Since structure dependency is seen as the central property of human language, the central area of study is syntax, which takes the form of universal principles with 'parameter settings' for each language. In addition, native speaker competence includes knowledge of the features of lexical items of a given language, and projection rules[24] matching lexical items with the relevant slots in the syntactical structure, so that only grammatical sentences are generated.

Generative linguistics also includes Chomsky's notion of 'pragmatic competence' which "places language in the institutional settings of its use, relating intentions and purposes to the linguistic means at hand" (Chomsky 1980, in Cook & Newson 1996: 23). The discussion of the different meanings of terms like 'formal', and 'structuralist' by Newmeyer (1998: 7 f.) shows acute awareness of the way in which language use influences meaning, but generative linguistics does not consider such knowledge to be part of linguistic competence as such, communication being only one of the uses to which language is put (Cook 1988: 13 f.).

[24] The 'minimalist program, which is the most recent version of Chomsky's generative theory, discards the earlier distinction between deep and surface structure as well as notions like the projection principle, which had been accounted for in terms of these levels. The current thesis is that "the faculty of language engages other systems of the mind/brain at two 'interface levels'"; a language representation thus has a 'phonetic representation' legible to the sensorimotor system as well as a 'semantic representation' legible to the conceptual system (Chomsky 2000: 9 f.)

So where does this leave collocations? As conventional expressions, they belong to langue in the Saussurean system, as described in section 2.1.3.1. Likewise, in a generative framework, they are part of a native speaker's competence, but they do not fit smoothly into a formal system which presumes a clear-cut distinction between rules of syntax and a lexicon consisting of items that can be described in terms of general, discrete features (for example '± noun', '±animate', '± human'). The challenge is similar to that presented by idioms, which Howarth (1996: 17) describes as follows: "On the one hand, idioms are semantically irregular, not analysable by normal lexicogrammatical rules and therefore belong in the lexicon. On the other hand, they are syntactic units, conforming to a variety of regular syntactic combinatorial rules, in which case, by placing them in the lexicon, we introduce a great deal of redundant information into the grammar".

In a generative framework, then, collocations, like idioms, are construed as deviating from a standard which assumes full compositionality and free substitutability once the general rules of syntax have been complied with and the general selection restrictions of lexical items have been met. Referring to Weinreich (1969), Moon (1998: 14 f.) sums up this view of idioms and collocations as follows:

> ... studies of FEIs [fixed expressions and idioms] from the perspective of transformational or transformational-generative grammar (TG) begin with syntax. The syntactic or grammatical aberrance or anomalousness of strings leads to their classification as noncompositional units. FEIs are regarded as exceptions to syntactic rules, or as unique realizations of rules. Because they are non-productive, they cannot be generated freely, and productivity is part and parcel of TG-models [my emphasis].

A number of generative linguists quoted by Moon talk about "exceptions that prove the rule" (Katz 1973: 359) and "'idiomatic material' outside the rules of the language system" (Harris 1991: 43). Chafe (1968: 111), according to Howarth (1996: 18), also treated idioms as 'anomalous', but at the same time emphasized that because of their frequency they could not be "relegated to the status of curios". This dilemma follows from the priority traditionally given to syntax within the generative framework and from a system of categorization that is based on discrete features. It is an approach that pushes idioms and collocations to the very periphery of the language system, and I would therefore claim that it is not suitable for categorizing them. I will return to this

line of argument (introduced in section 1.3 above) in the more detailed discussion of the notions used to categorize collocations in section 2.2.

2.1.3.3 Classical categories

Whereas the notion of the arbitrariness of conventional expressions goes back to Saussure, I have argued above that phraseology has adopted from generative linguistics the view of collocations as deviations from a standard of full compositionality. This implies what Lakoff (1987: 148) has referred to as "a rule-based system generating composite structures from component structures in a predictable way". Predictability requires the operation of general principles of syntax and categorization of lexical items based on selection restrictions that can be stated in terms of discrete features or criterial attributes. In other words, what is required is the classical model of categorization. This model is 'classical' in the sense that it goes back to Greek antiquity and Aristotle (1933), who distinguished between the 'essence' of a thing, that which determines what a thing is, and its 'accidents' or incidental properties. For instance, the 'essence' of man was said to be represented by the two features 'two-footed' and 'animal', whereas other properties, such as the colour of a person's skin, might be true of individuals but irrelevant in determining whether an entity was a man. The model can also be said to be classical in the sense that it has dominated especially structuralist and generative linguistics for much of the 20th century (Taylor 1995: 22).

Classical categories have clear boundaries, all members have equal status, and they are defined in terms of necessary and sufficient features, which are binary in the sense that an entity either has or does not have a certain feature. Phonologists have developed the Aristotelian notion of features, assuming that they are 'primitives' which cannot be further subdivided, and that they are universal. Furthermore, features are conceived of as 'abstract': they represent human speech capability rather than physical acts of speech, and they are conceived of as being 'innate' rather than learnt by imitation (Taylor 1995: 22 f.). The advantages of such rigidly defined categories seem obvious also for semantic description, and especially from the point of view of generative linguistics. As Taylor

puts it: “The ‘empirical’ justification of a feature approach to semantic categories appeals essentially to the same kinds of argument that were used in phonology, namely, the fact that features enable the linguist to make economical and insightful statements about the structure of language” (Taylor 1995: 30). As an example, features can be used to state the ‘selection restrictions’ of lexical items in terms of natural classes, such as ‘human’ and ‘animate’, which are assumed to be universal semantic primitives (Taylor 1995: 33). The problem is, however, that it is not possible to reduce all possible meanings to a finite set of primitives. A distinction between ‘markers’ (features expressing general semantic properties) and ‘distinguishers’ representing idiosyncratic elements of meaning, has been suggested in generative linguistics (Katz and Postal (1964: 14, in Taylor 1995: 33), but it does not make up for the inability of this framework to capture the principle of dynamic meaning creation to meet any imaginable language requirement.

An attempt to subject collocations to classical categorization coupled with a generative, rule-based notion of creativity is doubly problematic. In a generative framework collocations, as conventional expressions, can only be categorized in terms of their deficiencies. This is most obviously reflected by category labels such as ‘anomalous collocation’ and ‘defective collocation’ used by Moon (see section 1.3 above), but the more widely used term ‘restricted collocation’ reflects the same view. Moreover, since classical category membership is an all-or-nothing affair, a given combination should possess all the ‘necessary and sufficient anomalies’, not just some of them and not just to a certain degree. Phraseologists have found, however, that it is not easy to reduce anomaly to a manageable number of features or attributes, much less to classify combinations in terms of them, and to accommodate the fluidity of natural language, they have adopted from Russian phraseology the notion of the ‘continuum’, which with its implication of gradedness is not compatible with a model based on discrete features.

2.1.4 Russian phraseology

In the cold-war period of isolation of East from West following World War 2 in Europe, Russian scholars carried out extensive work on phraseological categories, which became accessible in the West only with considerable delay through the work of Eastern European scholars like Klappenbach (1968) and Weinreich (1969). The appeal of the work of Russian phraseologists like Vinogradov (1947) and Amosova (1963) was that it offered a comprehensive framework covering the whole spectrum of phraseology. The purpose of this section is to consider the descriptive and explanatory potential of the Russian model, which proposes to integrate a systematic description of phraseological categories into a continuum of word combinations. This is relevant as a background for the discussion in section 2.2 below of the problems encountered in categorizing collocations today by analysts who are building on this framework.

Author	**General category**	**Opaque, invariable unit**	**Partially motivated unit**	**Phraseologically bound unit**
Vinogradov (1947)	phraseological unit	phraseological fusion	phraseological unity	phraseological combination
Amosova (1963)	phraseological unit	idiom	idiom (not differentiated)	phraseme, or phraseoloid
Cowie (1981)	composite	pure idiom	figurative idiom	restricted collocation
Howarth (1996)	composite unit	pure idiom	figurative idiom	restricted collocation

Table 3: Subcategories within the phraseological spectrum (based on Cowie 1998b: 7)

The approach includes the analysis of both sentence-length units like sayings and pragmatic formulae, and of less than sentence-length units; only the latter group will be considered here, the emphasis being on verb + nominal object collocations. I will use Cowie and Howarth's equivalents for the Russian terms, drawing on the table provided by Cowie (1998b: 7), which furthermore includes the terms used by Gläser (1988) and Mel'čuk (1988) that have been excluded here. It should be noted, however, that Amosova did not consider phraseoloids to be part of the phraseological spectrum, which she defined more narrowly than Vinogradov and phraseologists today. In the Russian

model, restricted collocations form the middle part of a **continuum**, or a 'scale of idiomaticity', ranging from "the most freely co-occurring lexical items and transparent combinations to […] the most cast-iron and opaque idiomatic expressions" (Howarth 1996: 32). For the first time, collocations are placed in a comprehensive framework that sets them apart both from idioms, whose meaning is not derivable from the meaning of the parts, and free combinations, whose meaning is seen as fully compositional and which are consequently excluded from the phraseological spectrum altogether.

CATEGORY	COMPOSITIONAL	ANALYSABLE	PHRASEOLOGICAL COMBINATION
free combination	yes	yes	no
restricted collocation	partly	yes	yes
figurative idiom	no	yes	yes
pure idiom	no	no	yes

Table 4: Russian categories

Within the spectrum, three subcategories are distinguished: (1) restricted collocations which are only partly compositional, one element having a meaning that is restricted by the other, (2) figurative idioms which are units of meaning and therefore not compositional, but which have a literal equivalent and are therefore analysable, and (3) pure idioms which are neither compositional nor analysable as they do not have any known non-figurative equivalent. Referring to a system of categories as a 'continuum' implies that the categories have graded membership and fuzzy boundaries, and that there is a unifying principle holding them together. So what is the principle unifying phraseological expressions, what are the criteria used to identify the categories along the continuum, and in what sense is category membership graded?

From what has already been said, it appears that the main unifying principle is deviation, to a greater or lesser extent, from a standard of full compositionality, the deviation being related to decreasing analysability and increasing fixedness in terms of restricted substitutability. The main distinction is between restricted combinations that are 'phraseological' and free combinations that are not. The continuum represents a

strictly synchronic view; thus Howarth (1996: 23) refers to the term ‘phraseological’ as usefully indicating “a property that expressions have to varying degrees” and suggests that the continuum model may be taken to represent the “degrees of stability with which expressions are stored in the mental lexicon”, which gives the model “great descriptive value and perhaps psychological validity”.

While I would agree that the continuum usefully illustrates that the mode of integration of composite structures shows considerable variation in terms of analysability (see also section 2.1.5, about the ‘decomposition’ hypothesis), I find that its psychological validity is limited as long as expressions are placed on the continuum, not on the basis of functional and cognitive criteria, but on the basis of distributional features and deviation from an assumed norm of full compositionality. Furthermore, the attempt to categorize composite structures in terms of specific features implies classical categories, which, as already indicated, are not compatible with the idea of a continuum. On the other hand, if such categories were replaced by prototype categories, the continuum model would gain in psychological validity and, at the same time, many of the problems of categorization might be resolved.

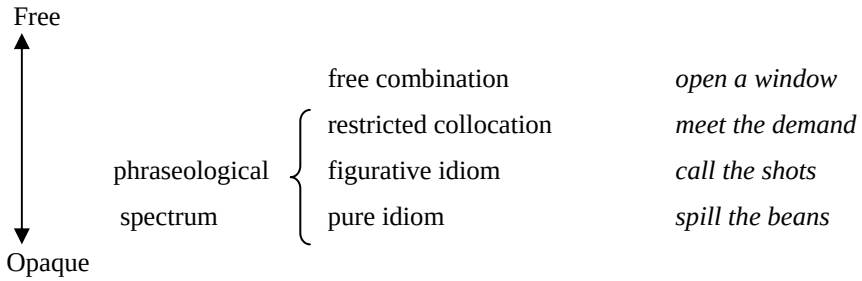


Fig. 3: A scale of idiomaticity (based on Cowie 1999: 71)

In Fig. 3, I have used examples given by Cowie to illustrate “the shading-off from free to opaque” of verb + noun combinations placed along the continuum, which he refers to as a ‘scale of idiomaticity’. I will return to these examples in the discussion of the fuzzy boundaries and internal gradedness of categories in section 2.2.5 below, arguing that they illustrate quite well that word combinations cannot be strictly classified in terms of criterial features. The basic problem of distinguishing between what is phraseological

and what is not has been inherited from the Russian phraseologists: whereas Vinogradov's equivalent of restricted collocations, 'phraseological combinations', would allow the sense of, for instance, the verb in a verb + nominal object combination to be 'phraseologically bound' by a single word or a limited set of words, Amosova's 'phraseme' definition required a single determining word, such as *teeth* in *grind one's teeth* (Cowie 1998b: 5 f.). Cowie's example *meet the demand*, in which *meet* is determined not only by *demand*, but also by nouns like *need, requirements, and request*, shows how phraseologists today are taking their lead from Vinogradov rather than Amosova. The formal requirements are relaxed even further in response to the needs of foreign learners, who turn out to have greater difficulties in coping with collocations that allow quite a lot of variation than with the most restricted ones. According to Howarth (1996: 30),"The contentious issue - given that we are on the borders of phraseology - is the degree to which a collocation can be varied and still remain 'restricted'".

A diachronic interpretation of the continuum, related to analysability, is found in Cowie et al. (1983: xii f.): "Historically, pure idioms form the end-point of a process by which word-combinations first establish themselves through constant re-use, then undergo figurative extension and finally petrify or congeal". Also Howarth (1996: 24) discerns "a diachronic scale from the analysable figurative idiom such as *do a U-turn* or *a narrow shave* via expression in which not all speakers will be aware of a literal antecedent as in *stop the rot* or *bury the hatchet* to the most moribund expressions such as *kick the bucket* or *spill the beans*". While a diachronic perspective has greater explanatory potential than a strictly synchronic one, a description of conventional expressions in terms of metaphors of petrification and death again seems to reflect the underlying influence of a generative notion of linguistic creativity. Actually, far from being dead, such expressions are alive and kicking. Thus Moon (1998: 120 f.) has found in her corpus studies that variation of 'fixed expressions and idioms' is widespread and fairly consistent across types,[25] and she therefore suggests that the notion of fixedness should be replaced by newer models.

[25] Moon distinguishes the three macrocategories 'formulae', 'metaphors', and 'anomalous collocations'. The last group includes 'defective collocations', which correspond to 'restricted collocations' in Cowie's and Howarth's terminology.

As it turns out, there is a balance to be struck between restrictedness and fixedness on the one hand and variation on the other, and I will argue that a network model is needed to represent this balance and to capture the dynamics producing it. This line of argument, which was presented in Part 1 (cf. section 1.5), will be further pursued in Part 3, in the empirical part of this study.

2.1.5 A cognitive strand

Even if it has turned out that the continuum is not compatible with classical categories, phraseologists try to cope with the problems by relaxing the requirements of category membership rather than giving up the traditional model (see section 2.2.4 below). This does not mean that they are not concerned with functional and cognitive aspects; actually, because of the focus on the problems of the foreign learner striving to achieve native-like proficiency, there is a keen interest in the psychological status of conventional expressions as well as their social status, or 'institutionalization' (Howarth 1996: 48). In this section I will look at relevant aspects of the psycholinguistic evidence, whereas institutionalization, as one of the notions used by phraseologists to categorize collocations, will be discussed in section 2.2.2 below.

Psycholinguists ask questions about the way in which conventional expressions are stored and processed. Although studies normally focus on idioms and polymorphemic words rather than collocations, as remarked by Howarth (1996: 48 ff.), and in spite of the fact that experimental methods may be criticized - for example they mostly do not use natural language data, as noted by both Howarth (1996: 55) and Moon (1998: 32) - it is possible to make some inferences from the general direction (or directions) in which the findings seem to be pointing. In the context of the present study, it is interesting how these findings are presented and used by phraseologists, and particularly whether the system of categorization is consistent with conclusions drawn from the psycholinguistic evidence. Howarth's 1996 study will serve as a case in point, as it devotes a full chapter to the processing of conventional language with the purpose of

testing the hypothesis “that the role of prefabricated language in discourse is to facilitate the fluency and naturalness of an utterance” (Howarth 1996: 48).

Do conventional expressions constitute a problem or a solution to a problem? This must be the most basic question to ask a psycholinguist, but the answer depends on how the nature of the problem is perceived. From the viewpoint of linguists in the generative tradition, conventional expressions constitute a problem (see section 2.1.3.2 above), because they can only be described in terms of their idiosyncratic restrictions on substitution and their transformational deficiencies, which are conceived of as being stated separately in the mental lexicon, in the so-called ‘idiom list’ (Katz and Postal 1963, in Howarth 1996: 51). This notion has been said to reflect that “the generative research tradition was influenced in the past by the Bloomfieldian view of the lexicon as an appendix to the grammar, a list of basic irregularities” (Everaert et al., 1995: 2). It has led to the formulation of an **‘idiom list hypothesis’** also in psycholinguistics, according to which the strategy for dealing with idiomatic ambiguities as in *John gave Mary the slip* is supposed to be distinct from the strategy for dealing with ambiguities in literal statements like *Mary fed her dog biscuits* (Bobrow and Bell 1973, in Howarth 1996: 51). It is assumed that the literal meaning is normally processed first so that the idiom list is resorted to as a secondary option (Moon 1998: 33). Idioms are seen as problematic, because they complicate the operation of generative principles of creative competence. I would say that it is the traditional generative system of description that creates the problems, because it gives priority to the level of syntax and does not handle conventional expressions on their own premises. Jackendoff (1995: 138 f.), in response to this problem, proposes a system that integrates phonological, syntactic and conceptual formation rules, which are regarded as equal, autonomous levels. In this system, both individual lexical items and multi-word items are ‘triples’ including structures of all three kinds, and the lexicon is seen as part of a system of ‘correspondence’ rules. Phraseological categorization appears to be based on the traditional generative view of conventional expressions as ‘exceptions to rules’ and marginal to the language system, although most phraseologists would agree with Howarth’s conclusion (1996: 64) that conventional composite units do in fact have a significant communicative function:

> Sufficient evidence has been adduced that, far from being processed more slowly as a result of their semantic idiosyncrasy, idioms present no special problem for mental processes, and that the familiarity and fixedness of a sequence result in greater speed and ease of retrieval and comprehension.

The functionality of phraseology is underpinned by the work of Peters (1977) on **language acquisition**, which posits separate strategies for processing 'unitary' and 'structurally-formulated' expressions, just like the idiom list hypothesis. In this framework, however, the use of holistic, unanalysed chunks presents an efficient problem-solving strategy for children learning a language for the first time, and Peters suggests that it is also useful for adults (Peters 1983: 3-4, also cited in Howarth 1996: 60):

> For mature speakers of a language [...] formulaic speech may serve as a shortcutting device: It saves processing time and effort, allowing the speaker to focus attention elsewhere ... far from employing a minimal amount of storage space for our language, we keep on hand many representations of the "same" information, choosing in any given instance exactly that one which minimizes processing effort.

In this view, conventional expressions are not problems but solutions, and the problem is processing speed, not storage space. According to Peters, the processing of structurally-formulated expressions requires an **'analytic' strategy**, which she associates with the left hemisphere of the brain, whereas the strategy for processing 'unitary' expressions is a **'gestalt' strategy**, associated by Peters with the holistic, right hemisphere. The notion of gestalt perception, or perception of overall part-whole configuration, has been characterized as the most important determinant of the basic level of categorization, for example by Tversky and Hemenway (1984), and by Lakoff (1987: 47), and I see it as a link between conventional expressions and cognitive models based on prototypicality and basic level categorization (cf. section 1.5 above). In Part 3, I will relate gestalt perception to the abstract 'contexts of situation', in the Firthian sense, and to Fillmore's notion of 'frames' (cf. sections 1.5 and 2.1.2 above), suggesting that conventional expressions may reflect the gestalt perception of typical situations with characteristic configurations of elements.

If it is assumed that multiple storage is not a problem so that words are available both for the construction of new expressions and as part of more or less unitary, conventional

ones, the question remains how the two types are decoded. Will the reader/listener resort to a 'holistic' reading only after a structural one has failed, or will the holistic meaning be recognized first because of its conventional status? While the former hypothesis is more in line with generative thinking, most studies give conventional expressions a head start. Thus the **'direct access' hypothesis** (Gibbs 1980) contradicts the idiom list hypothesis by claiming that, if an expression has both a literal and an idiomatic meaning, the idiomatic meaning will be perceived first, and the literal meaning will be resorted to only if an idiomatic interpretation fails. Although "the ease of processing an expression depends on how conventional it is" (Howarth 1996: 54), it seems plausible that an expression is more likely to become conventional if it requires extra conceptual work the first time it is encountered. This seems consistent with results for unidiomatic versions of idioms, for which people show good short-term memory, reflecting the special effort that it takes to process them (Gibbs 1980: 155). Thus literal readings of an idiom will be perceived as stylistically marked, because they are unconventional. (See the examples in the context of the discussion of the 'configuration' and the 'decomposition' hypotheses later in this section.)

Also according to the **'lexical representation' hypothesis**, idiomatic meaning will be recognized faster than non-idiomatic meaning in the case of structurally identical expressions like *break the cup* and *break the ice*, but no special processing strategy is posited for idioms, which are assumed to be stored and accessed as 'long words' (Swinney and Cutler 1979). More recent studies reject this hypothesis in favour of models that focus on the progressive nature of interpretation. Thus the **'spreading activation model'** (Aitchison 1987, in Howarth 1996: 50) "integrates phonological and semantic/syntactic information in a process of narrowing down a range of phonologically similar possibilities as more data is received about the shape of the word and tested against knowledge of the wider context of meaning and structure". More particularly, according to Cacciari and Tabossi (1988) the meanings of idioms are associated with particular **'configurations of words'**, which are recognized as soon as enough input has been received. Specific **'key' words** will play an important role in this process. Approaches that allow for the dynamic nature of processing as well as the internal structure of expressions are better suited to account for the flexibility of natural

language, including lexical and syntactic variation and creative extension. As Howarth points out (1996: 57), variability in the strength of cohesion between the components of a composite unit can also be accommodated in such a model; it is therefore useful for the description not only of more unitary expressions like idioms, but also of collocations. However, I would not expect it to be compatible with categorization in terms of rigid criterial features, since patterns of variation are bound to be highly idiosyncratic. On the basis of findings by Swinney and Cutler (1979) that transformational deficiencies in composite units apply to individual units rather than to whole classes of expressions, Howarth (1996: 52) concludes "One cannot therefore postulate a single scale of **frozenness** as a correlate of degrees of unitary storage. This opens the door to a search for other features that may account for the mental organization of prefabricated units."

While, in the tradition of Russian phraseology, both pure and figurative idioms are defined as being units of meaning and therefore not 'compositional' in the sense of being predictable from the meaning of constituent items, several psycholinguistic studies focus on 'compositionality' in the sense of analysability and find a great deal of variation.[26] The '**decomposition hypothesis'** (Gibbs and Nayak 1989; Gibbs, Nayak and Cutting 1989; in Van de Voort and Vonk 1995: 284 f.) defines compositionality in terms of two criteria, which are based on Geeraerts (1990):

(1) motivatedness: the extent to which a plausible relation can be found between the figurative meaning of the whole phrase and the literal meaning of its parts

(2) isomorphicity: the extent to which its figurative meaning can be distributed over its parts

An idiom like *spill the beans* that meets both criteria is considered to be compositional, or decomposable, in this framework, while it is frequently quoted as an example of an unmotivated, or unanalysable idiom in the context of phraseological categorization. The decomposition hypothesis further distinguishes three classes of compositionality:

[26] Following Langacker, I will argue, in section 2.2.5 below, that compositionality and analysability are independent parameters.

1) ‘normally decomposable idioms’ like *pop the question*
(*pop* refers directly to ‘suddenly ask’ and *question* to ‘marriage proposal’)

2) ‘abnormally decomposable idioms’ like *carry a torch for someone*
(*torch* is a metaphor for ‘warm feelings’)

3) ‘noncompositional idioms’ like *kick the bucket*
(the individual parts refer neither directly nor indirectly to the individual parts of ‘the act of dying’)

The hypothesis predicts that the first two classes will be processed faster than both literal expressions and noncompositional idioms and furthermore, because their parts have independent meanings, they can be modified and moved just like parts of a literal expression. Internal modification of a compositional idiom as in *He left no legal stone untouched* will not prevent idiomatic processing, whereas modification of a noncompositional idiom as in *He kicked the empty bucket* will cause it to lose its idiomatic meaning. Van de Voort and Vonk find, however, that, unlike external adverb modification, internal adjective modification affects all classes of idioms as they “seem to postpone access of the idiomatic meaning to a later stage in processing or maybe even change the access route to idiomatic meaning” (1995: 292 f.). They also find that the processing of idioms involves accessing the literal meanings of individual idiom words and not just their idiomatic meanings. Consequently, they prefer a **modified version of the configuration hypothesis** mentioned above. To allow for the effect of internal modification, they assume an extra stage in processing during which “the system is still processing the literal meaning of the words, including the meaning of the key(s), although the overall idiomatic meaning of the phrase is already activated” (Van de Voort and Vonk 1995: 297). If the meaning of a further key is not compatible with the idiomatic meaning, as in the case of an internally modified idiom, the idiomatic meaning will initially be discarded in favour of a literal meaning, before the idiomatic meaning eventually prevails.

Such a dynamic and flexible strategy seems to be required to make sense of the following modified version of the idiom *to have a finger in every pie* from an article about Sainsbury’s in the *Guardian*, May 18, 1989, where *pies* refers to both real and figurative pies (Poulsen 1991: 7):

- "*So just in case we do some day reach the point where there are no more shops to build, it* [*Sainsbury's*] ***has a few fingers in pies other than in its meat counters.***"

This is exploited, for example in journalism and advertising, by forcing readers to decompose holistic conventional meanings and blend them with the subject matter at hand. In fact, sentence-length idioms may be so rarely used in their canonical form that it may be next to impossible to establish their specific wording (Moon 1998: 121). The exploitation of collocations for expressive purposes is not so widespread, but does occur, as in the following passage from Tolkien's *The Lord of the Rings* (1974: 49 f.), about the label on a present from Bilbo Baggins to a not-so-beloved relative:

> *Inside in the hall there was piled a large assortment of packages and parcels and small articles of furniture. On every item there was a label tied. There were several labels of this sort: [...]*
> *"For LOBELIA SACKVILLE-BAGGINS, as a PRESENT"; on a case of silver spoons. Bilbo believed that she had acquired a good many of his spoons, while he was away on his former journey. Lobelia knew that quite well. When she arrived later in the day,* ***she took the point at once, but she also took the spoons.*** [My emphasis].

Forcing a compositional interpretation of the conventional collocation *take the point* by juxtaposition with the non-conventional combination *take the spoon* is only successful if the reader feels that there is a pay-off, in this case a brief thrill of surprise at the unexpected pun. This requires the right setting, or 'situational context', which Tolkien indeed has been careful to provide in this example.

The narrowing down of the possible meaning of an expression involves disambiguation as part of the **'post-access decision process'**, which comes after initial word recognition. Since words are typically polysemous, this is a general process and does not just apply to words in conventional expressions: "For any word, some subset of all the information which is originally accessed […] may be selected for further processing and integration into "ongoing sentential analysis." (Swinney and Cutler 1979: 658). It follows that words are not monosemantic 'building blocks', and that the meaning of an expression does not follow from the juxtaposition of words in a simple way. This can be contrasted with the characterization of free combinations, which are excluded from the phraseological spectrum, exactly because the meaning of constituents is seen as being 'additive' (an example is Howarth 1996: 38). Polysemy is not a problem in a cognitive framework in which the component structures are schematic networks whose integration

motivates aspects of the composite structure (cf. section 1.5 above). Also the Firthian concept of meaning by collocation allows for the polysemous nature of words, which are disambiguated through the process of collocation (section 2.1.2).

The aspect of **context** is an important factor in the post-access disambiguation process and, as pointed out by Howarth (1996: 59), it operates at different levels. In psycholinguistic studies it is mostly used about what Howarth calls "the reference in discourse to the real world", typically in the form of "short pieces of preceding co-text constructed to provide semantic clues for the interpretation of pre-selected idiomatic strings" (Howarth 1996: 54). As an example, the expression *He's singing a different tune* is given in two different contexts, one of which involves the singer Jackson Browne, and one which involves former president of the US, Jimmy Carter (Gibbs 1980: 150). The creative examples given in this section (*have one's fingers in* (*meat*) *pies*; *take a point/spoon*) as well as *Emily-coloured primulas* (section 2.1.2) all rely on the context in this sense of **'co-text'.**[27] While extra prompting is required for processing an idiom the first time as well as for innovative extensions, the conventional use of idiomatic expressions is assumed to reduce dependence on co-text (Gibbs 1980). I suggest that this may be due to the gestalt perception associated with these expressions and may help to explain the faster processing.

Less frequently in psycholinguistics, 'context' is used about the **co-occurrence**, or collocation, of items in a conventional expression, corresponding to Langacker's notion of 'syntagmatic context' (section 1.5). Thus Cutler (1983: 45, in Howarth 1996: 59) has found that: "certain words carry implications about their surrounding sentence context as part of their intrinsic meaning. These implications involve restrictions on co-occurrence, which actually make processing faster rather than impeding it." In line with Cutler's findings is the contention in Cruse (1986, in Howarth 1996: 60) that combinations involving arbitrary collocational restrictions are semantically more cohesive than

[27] The meaning of 'Co-text' is explained in the OED as "The language which surrounds a particular word, phrase, or passage, and which can determine its meaning" and is illustrated with the following example [my emphasis]: **1966** A. MCINTOSH in C. E. Bazell *In Memory of J. R. Firth* 303. Within the confines of what is actually spoken or written, I have found it necessary to consider the influence of elements in the co-text (i.e. in the textual environment of the constructions under actual scrutiny).

those with generalizable selectional restrictions, the reason being that the former are more predictable. According to Cruse, not only idioms but also collocations show **semantic cohesion**: " … the constituent elements [of collocations] are, to varying degrees, mutually selective. The semantic integrity or cohesion of a collocation is the more marked if the meaning carried by one (or more) of its constituent elements is highly restricted contextually, and different from its meaning in more neutral contexts". The restrictedness of conventional expressions is thus seen as making them more cohesive, which in turn causes them to be more predictable and easier to process – at least for native speakers.

This notion of predictability is comparable to the Firthian notion of 'mutual expectancy' (cf. section 2.1.2). Howarth (1996: 60) points out that such conclusions are highly relevant for defining the characteristics of restricted collocations, the phraseological category which is the chief concern of his study. I would argue, however, that the actual categorization is influenced more by traditional generative ideas about linguistic competence and by the viewpoint of the foreign learner. Phraseologists associate predictability not with restricted collocations but with "the very large numbers of combinations that are entirely predictable and lacking in phraseological significance" (Howarth 1996: 8-9), and restrictedness is defined negatively as characterizing expressions that are not composed according to generative rules (Howarth 1996: 31).

In addition to the two types of context referred to in the psycholinguistic studies which have already been mentioned (syntagmatic context and co-text), three other types, or levels, have come up earlier in this study: the Firthian notion of an abstract 'context of situation', Langacker's notions of 'systemic context', and pragmatic 'situational context' (see sections 1.5 and 2.1.2). Table 5 below provides an overview.

	CONTEXT	Source	Definition	Example
1	**SYNTAGMATIC CONTEXT** Comparable to **COLLOCATION**	Langacker Firth	The other lexical items with which an item co-occurs in a complex expression	*take a point*
2	**SYSTEMIC CONTEXT**	Langacker	Position of lexical items in their respective schematic networks	*take a point* *take a break, a nap ..* *take a spoon, a fork ….*
3	**CONTEXT = CO-TEXT**	OED Howarth	The language which surrounds a particular word, phrase, or passage, and which can determine its meaning	*And the Emily-coloured primulas* *Bob in their pinafores on the grass*
4	**CONTEXT OF SITUATION** Comparable to **FRAMES**	Firth Fillmore	Abstract, typical situation with typical participants and role configurations	Commercial situation: buyer, seller, product, price; buying, selling
5	**SITUATIONAL CONTEXT**	Langacker	The pragmatic circumstances of a particular usage event	Bilbo's birthday party Particulars about host, guests, gifts and relations between people

Table 5: Levels of context

Howarth credits psycholinguists like Cutler (1983) and Gibbs and Gonzales (1985) with having recognized that "features of the context exert an influence on processing at an early stage, ruling out ambiguity and making the question of whether an idiom has a literal interpretation unnecessary or irrelevant" (Howarth 1996: 51). Nevertheless, this distinction between literal and figurative meaning is used in phraseological categorization to separate phraseological from non-phraseological expressions and to distinguish between phraseological categories (see section 2.2.5 below).

While many psycholinguists study how conventional expressions are processed (especially how they are decoded rather than encoded), few will ask themselves questions about their conceptual basis. Of those mentioned so far, Peters (1977, 1983) is an exception because of her notion of gestalt perception. Gibbs (1995: 104 f.), however, challenges traditional assumptions about the arbitrary nature of conventional expressions, positing that "the figurative meanings of idioms might well be motivated by people's conceptual knowledge that is itself constituted by metaphor". According to this view of motivation, which is rooted in a cognitive approach, conceptual metaphors do

not predict exactly what the form of a conventional expression will be - for example *spill the beans* rather than *spill the peas* - but they partially explain why a phrase like *spill the beans* is used to mean 'the revealing of secrets'. In a cognitive approach, conventional expressions are seen as reflecting general principles of conceptualization and categorization. Consequently, they are not studied in isolation, but as coherent systems of metaphorical concepts; thus ANGER IS HEATED FLUID IN A CONTAINER, which may be seen as motivating expressions like *blow your stack*, *flip your lid* and *get steamed up* (cf. Lakoff 1987: 380 ff.). Studies of people's mental images for idioms with similar meanings have provided evidence that these images are not merely generalizations of linguistic meaning, but are partly motivated by underlying conceptual metaphors (cf. section 3.1 below). Furthermore, it has been shown that people actually use their knowledge of such metaphors to make judgements on the appropriate use of specific idiomatic expressions, and that their understanding is facilitated by discourse contexts that are consistent with these expressions (Gibbs 1995: 108 f.). While Gibbs emphasizes that it is not predicted that conceptual metaphors will necessarily influence all aspects of idiom understanding, he argues that they are an essential part of a theory of idiomaticity, because "lexical meanings do not by themselves capture the complex inferences associated with idiomatic meanings" (Gibbs 1995: 108-9).

If such results from cognitive studies in psycholinguistics are to be applied to the categorization of collocations, I suggest that it will be necessary to abandon the classical model, which is based on the formal features of lexical items and cannot capture their conceptual basis. As already indicated, a model based on prototypes and basic level categories is compatible with the notion of gestalt perception, or perception of overall part-whole configuration, which may in turn reflect conceptual metaphors underlying conventional expressions. As mentioned earlier, cognitive models are beginning to be applied in phraseological studies (section 1.4 above). To a certain extent, the discussion within the traditional approach to phraseology that is being reviewed here also reflects ideas that are more compatible with a prototype model than with the classical one. The notion of a continuum discussed in 2.1.4, is the best example, because it reflects that phraseologists see no "watertight division between the various types of collocation and idiom" (Howarth 1996: 32), but also the mentioning of 'fuzzy zones' (Cowie 1998b: 6),

'fuzzy boundaries' (Howarth 1996: 33) and of more or less 'central' collocates (Howarth 1996: 31) conflicts with the idea of categories based on criterial features. Besides, results of psycholinguistic studies are used by phraseologists to confirm that conventional expressions have a significant communicative function and that a continuum model suitably captures the variability in the strength of cohesion of expressions. Further, the notion of frozenness has been questioned by phraseologists as a feature to be used in categorization, because it has been found to apply to individual expressions rather than to classes. On the other hand, findings questioning that notions like compositionality, predictability and the distinction between literal and figurative meaning can be used as criterial features are not heeded, just as the evidence of possible creative strategies and of a conceptual basis for conventional language are not yet being exploited. This may be because phraseologists focus on the struggles of the foreign learner while the psycholinguists are studying the fluent language of native speakers, and because the influence of traditional structural and generative thinking still prevails.

2.2 Collocations in a framework of phraseological categories

Having accounted for the theoretical influences and practical concerns that in my view determine the approach to categorization in phraseology, I will now turn to the resulting framework. I will first discuss the notions of grammatical welformedness and institutionalzation that are used to characterize phraseological expressions in a general way and which are applied also to expressions that are excluded from the phraseological spectrum. I will then discuss the notions of full compositionality, restrictedness and analysability, as well as the distinction between literal and figurative meaning, which are used to distinguish categories and to identify subcategories. The focus will be on the issues raised in Part 1 (section 1.3), and the notions will be evaluated in terms of their descriptive and explanatory potential as part of a framework for the categorization of collocations.

2.2.1. Grammatical well-formedness

While Firth argued that collocations should be treated as a separate level of meaning for the purpose of linguistic analysis, he also stressed the interdependence of levels (see section 2.1.2 above). Today two different approaches mirror the two perspectives: while corpus linguists such as Sinclair will examine collocations independently of syntax, phraseology studies only composite units that are "exponents of recognizable syntactic units (grammatically well-formed)" (Howarth 1996: 34).[28] The argument for treating lexical and grammatical patterns separately, as corpus linguists typically do, is that the lexical patterns of collocation and set are different from the grammatical patterns of structure and system and should therefore be described separately (section 2.1.2), whereas the argument for treating them together in phraseology is that, in an actual language event, they will always be integrated, and that a conventional expression, or composite unit, will be characterized not only by lexical, but also by lexicogrammatical restrictions. Structural variation is a normal feature, especially of collocations, but there are also idiosyncratic usage patterns, which have to be observed if a combination is to be 'lexicogrammatical' (cf. section 1.3). To phraseologists it seems artificial to separate the two types of restrictions because, for encoding purposes, it will always be necessary to combine knowledge of appropriate lexical choices with knowledge of possible structural variation (see for example Mitchell 1971: 53 f.). Consequently, an approach that examines lexical relations independently of syntax is felt to be inadequate. Thus Greenbaum (1970: 11) refers to such an approach as being 'item-oriented' and as "obscuring syntactic restrictions on collocations". Also it fails to show potential syntactic variation and how collocations and idioms can be flexibly integrated with other material. Cowie (1999: 76) used the idiom *to sell something* or *somebody short* to

[28] A definition that equates 'syntactic' with 'grammatically well-formed' does not include 'ill-formed collocations' and 'cranberry collocations', two types of FEI [fixed expressions and idioms] posited in Moon (1998: 21). The former are combinations like *at all, by and large, of course, stay put,* and *thank you,* which "break the conventional grammatical rules of English", while the latter "include items that are unique to the string and not found in other collocations" - just as the morpheme *cran-* was once unique to *cranberry* - and are exemplified by expressions like *kith and kin, short shrift* and *to and fro.*

Among sentence-length idioms, quotations which owe their status as conventional expressions to the fact that they are not grammatically well-formed, are also included:"*Curiouser and curiouser!" cried Alice.* Lewis Caroll [Charles Lutwidge Dodgson]. *Alice in Wonderland,* Part 1. In *The Concise Oxford Dictionary of Quotations.* London, Oxford, New York. Oxford University Press. 1964: 56.

illustrate how the *Oxford Dictionary of Current Idiomatic English. 2: English Idioms* indicates "that idioms could form collocational links outside their own strict limits with other sets of words":

> sell sth/sb short [V + O + A pass]... cheat sb in value or quantity; belittle oneself or sb/sth else O: ... country, economy; oneself, one's friends ...

The function label 'O' indicates that the listed items *country, economy, oneself, one's friends*, etc. can function as the direct object in a sentence containing *sell ... short*, while 'pass' indicates that the passive transformation is possible. Howarth (1996: 36) shows how collocations are often associated with 'grammatical fillers' such as articles: *carry a/the responsibility* or possessives: *summon up one's/my/his energy* as well as 'optional lexical items', (as in the example from the ODCIE 2 above): e.g. *carry the heavy/onerous/ ministerial/temporary*, etc. *responsibility*.

As mentioned in section 2.1.2 above, the approach that examines lexis and syntax separately has also produced evidence to the effect that the 'co-selection' not only of lexical items, but also of lexical items and grammatical constructions, is a general phenomenon. Consequently, I would argue that it is not useful to describe such patterns in terms of arbitrary transformational deficiencies, because this would imply that the phenomenon was exceptional. Assuming that there is truth in the claims made in Stubbs (1996: 40) that "Meaning is not constant across the inflected forms of a lemma" and that "Every sense or meaning of a word has its own grammar: each meaning is associated with a distinct formal patterning. Form and meaning are inseparable", it should be worthwhile to examine the syntactic forms of collocations and idioms for what they contribute to the meaning of specific expressions. Gibbs and Gonzales (1985) in a study of speakers' sensitivity to the frozenness of idioms, suggested that the restrictions on the possible forms of an idiom, its 'limited productivity', must be due to the relation between syntax, semantics, and pragmatics of the novel, metaphorical expression from which the idioms are derived, and that "people may still have knowledge of the internal semantics of idioms which constrains the possible syntactic forms these expressions may take" Gibbs and Gonzales (1985: 258). The findings of these different approaches point towards a need for researchers to go further than phraseologists today by trying to

account for the joint motivation of structural and lexical choices in conventional expressions, rather than just stating them in terms of restrictions. This would be in line with Firthian thinking that levels of meaning are 'integral in experience' although it may be useful to analyse them separately (section 2.1.2), and also with the notion in cognitive grammar of symbolic units with a semantic pole that includes lexical as well as grammatical structure.

I have suggested that Firth's notion of 'context of situation' may be seen as providing the functional framework that recruits and integrates the relevant lexicogrammatical patterns (section 2.1.2). The conceptualization of such a typical situation may be reflected in lexicogrammatical choices, which are in turn motivated by it. Thus, the use of the present and past participle forms of *negotiate* in the following examples is obviously motivated by the context of situation, a typical negotiation situation, and contributes to the meaning of the composite units:

negotiating table *negotiating partners* *negotiated settlement*

By contrast the equivalent Danish terms lack these structural meaning elements:

forhandlingsbord	*forhandlingspartner*	*forhandlingsløsning*
[negotiation table]	[negotiation partner]	[negotiation settlement]

While it is characteristic of the phraseological approach that composite lexical structures are studied as syntactic combinations, this feature is not used for categorization since it also applies to free combinations. What is thought to distinguish phraseological expressions is their lexicogrammatical usage restrictions, but only lexical restrictions are used for categorization, since phraseologists have come to the conclusion, on the basis of psycholinguistic evidence, that restrictions on syntactic variation are idiosyncratic and cannot be correlated with classes of expressions (cf. section 2.1.5 above).

2.2.2 Institutionalization

While phraseologists are interested in what psycholinguists have to say about the psychological status of conventional expressions, i.e. how they are stored and processed, the social status or 'institutionalization' of such expressions is used more directly as a criterion for categorization (Howarth 1996: 34; Moon 1998: 7). Language is an 'inter-organism' (Halliday 1978: 38 f.) phenomenon also according to Saussure (1966: 9) for whom 'language' (langue) represented the social aspect, being "both a social product of the faculty of speech and a collection of necessary conventions that have been adopted by a social body to permit individuals to exercise that faculty". He made it clear that language was the proper object of study and that psychological studies could not be expected to reveal anything worthwhile about the language system, belonging as they did to 'speech' (parole): "Taken as a whole, speech is many-sided and heterogeneous; straddling several areas simultaneously - physical, physiological, and psychological - it belongs both to the individual and to society; we cannot put it into any category of human facts, for we cannot discover its unity" (Saussure 1966: 9). Firth likewise believed that language should not be explained by referring to 'inner mental happenings', but by studying language in use, and his notion of 'context of situation' is to be understood as an abstraction, based on "typical 'repetitive' events in the social process" (section 2.1.2). Whereas Firth restricted himself to the social status of conventional expressions, Fillmore's notion of 'frames' includes their psychological status. I will pursue the idea that the two notions can be seen as complementary, with context of situation providing the socially based functional motivation of expressions and frame semantics accounting for their psychological or cognitive motivation (see section 3.1 below).

In his book on morphological productivity, Bauer (2001:46) distinguishes between two complementary types of established complex lexemes: those that are 'lexicalized' and those that are 'institutionalized'. The former, after becoming established as part of the language norm, have diverged from their original form or meaning to the extent that they could not be 'contemporary coinages', whereas the latter "still form part of a synchronically productive series, differing only from potential words in that, by being

used, they have come to have a specific reference". Bauer compares lexicalization to 'idiomatization', which is the term used by Lipka (1994) and which Bauer finds "nicely captures the idea that semantic lexicalization involves the loss of compositionality of meaning, just as the formation of idioms does" (Bauer 2001: 45). Bauer further points out that there are different types of lexicalization, e.g. semantic and phonological, and that words can be lexicalized to different extents (cf. the discussion of variations in analysability and compositionality in section 2.1.5 above).

It would seem that a borderline between what is phraseological and what is not might be drawn to coincide with the division between new coinages and institutionalized expressions, which according to Pawley (1985: 105, in Howarth 1996: 36) are typically "backed by customary ways of behaving which confirm and reinforce their status as a social institution". However, the primary concern of phraseologists is not the social or psychological status of expressions, but whether they are expected to pose problems for learners. Consequently institutionalized expressions which are not "restricted by semantic and collocational features" (Howarth 1996: 37 f.) are excluded from the phraseological spectrum and categorized as free collocations together with nonce-forms and 'run-of-the-mill combinations' like *affect world trade* "that are predictable and generated by the language system with nothing distinctive in their semantics or communicative function to make them institutionalized or memorable". I have pointed out (section 2.1.5) that the use of 'predictability' in this sense of 'rule based' conflicts with the use of 'predictability' about conventional expressions, indicating that the 'mutual expectancy' associated with them makes them easier to process, because certain key words will prepare the reader/hearer for what is coming next. I would infer from this that the predictability of conventional expressions facilitates encoding as well as decoding.

Howarth (1996: 37 f.) mentions *go to school/church/work* as institutionalized expressions that "would be considered by some as free collocations (since both elements are used in their literal senses)". On the other hand, he suggests that *take a vacation* would be classed as restricted, presumably because *take* is used in a figurative sense defined by *vacation*. Does this imply that while *take a vacation* is within the phraseological

spectrum, while *go on vacation,* in line with *go to school*, etc., is not? I will come back to these examples in a discussion of the relevance of the literal/figurative distinction (in section 2.2.5). Howarth's line of argument is similar to that found in Hausmann's discussion (1985: 118 ff.) of what word combinations to include in the example material of a dictionary, in which he relates combinations to Saussure's langue/parole dichotomy.

EXAMPLE	**LANGUE / PAROLE**	**COLLO-CATION**	**FEATURES** [my translations]	**EXPLANATION**
ein Buch aufschlagen	langue	yes	typical , specific, characteristic	The expression belongs to langue as norm.
ein fesselndes / spannendes / interessantes / Buch ...	langue	no	not typical non-specific	The adjectives also occur with many other nouns.
ein Buch kaufen	parole	no	non-specific banal	Generated by the language system without any usage restrictions.

Table 6: Collocations in the langue/parole dichotomy (based on Hausmann 1985)

As conventional expressions, collocations are said to belong to language (langue), which means that they are institutionalized, but only 'typical', 'specific' or 'characteristic' combinations ("typische, spezifische und characteristische Zweierkombinationen") like *ein Buch aufschlagen* ('open a book') are to be classified as collocations. The second type, exemplified by *ein fesselndes/ spannendes /interessantes Buch* (a captivating/absorbing/interesting book),[29] also belong to language, but are not classified as collocations, because the adjectives also occur with a wide range of other nouns. Hausmann's third type, 'non-specific' and 'banal' combinations, correspond to Howarth's 'run-of-the-mill combinations'; since they are generated by the language system without any usage restrictions, they are not considered to be of interest to phraseology. The general condition stated by Howarth - that only expressions which are restricted by semantic and collocational features should be included in the phraseological spectrum - is expressed more specifically here in terms of substitutability, which is the paradigmatic dimension of compositionality. Whether the dependent element of a

[29] The English equivalents were ascertained in the Cobuild Corpus Concordance Sampler, which draws on a corpus of about 56 million words of contemporary English.

combination, typically an adjective or verb combined with a noun, is to be considered collocationally restricted depends on the 'set' of other nouns with which it co-occurs in the same sense. This takes us back to the question inherited from Russian phraseology (cf. section 2.1.4 above): should the dependent element be determined by just one noun, or by a limited set of nouns, and in the latter case how should 'limited set' be defined?

As suggested in Part 1 (section 1.3), I would argue that such patterns of substitutability are extrinsic to word combinations and therefore not suitable as a basis for categorizing them. Neither do they seem helpful for foreign learners in an encoding situation. They may not be familiar with such patterns, or they may not want to rely on them without specific confirmation, for example that in the context of *book*, *captivate* is the English equivalent of German *fesseln* (literally 'fetter' or 'chain') and of Danish *fængsle* (direct translation: 'imprison'). Decoding is no problem here since all three languages draw on the same conceptual metaphor (cf. section 3.1 below). I find myself in agreement with Moon's suggestion (1998: 37) that it is "unlikely that the set of FEIs [fixed expressions and idioms] in English will be delimited until the syntagmatic and the paradigmatic properties of individual words and individual meanings have been properly explored". This would provide the knowledge about the component items of a composite structure that would allow statements about categorization to be based on intrinsic characteristics rather than on extrinsic distributional patterns. The purpose of this section has been to show how the decision to include only some institutionalized expressions in the phraseological spectrum has made phraseologists look for features that can be used to categorize combinations independently of their social and cognitive status. These features are based on the notions of full and restricted compositionality as well as analysability, to which I will turn next.

2.2.3 Full compositionality

Compositionality is the hub of the phraseological framework; it is used to distinguish combinations that are considered phraseological from combinations that are not, and it is used to distinguish between the main phraseological categories: pure idioms,

figurative idioms and restricted collocations. In this section I will discuss the underlying assumption that full compositionality is the norm from which phraseological expressions deviate whereas the syntagmatic and paradigmatic restrictions used to characterize collocations and idioms will be the subject of the following section. According to Howarth, whereas Russian lexicologists would focus on what is phraseological, English phraseologists like Aisenstadt and Cowie will ask themselves "at what point language users are manipulating expressions as wholes rather than composing them according to generative rules: in other words, what is not free?" (Howarth 1996: 31). This has involved dividing restricted collocations into subcategories and relaxing some of the criteria in order to be able to include in the phraseological spectrum more of the expressions formerly considered to be 'free' (cf. section 1).

The influence of the traditional generative approach is apparent here (cf. section 2.1.3) as is the concern about the needs of foreign learners, who are not supposed to experience any problems as long as combinations are "composed to standard rules of syntax and semantics" (Howarth 1996: 81). As mentioned in the previous section, this characterization leads Howarth (1996: 37) to identify three classes of **'free collocations'**:

(1) nonce-forms, which are coined to meet an immediate need
(2) 'run-of-the-mill combinations', which are predictable and generated by the language system
(3) institutionalized expressions that are not restricted by semantic and collocational features

Free collocations are characterized as "quite transparent, easily derivable from the juxtaposition of the elements in a recognizable syntactic pattern"; besides "both constituents are used in a primary literal sense" (Howarth: 38). This characterization of free collocations conflicts with the view in cognitive linguistics that component structures are not 'building blocks' from which the composite structures are assembled, but that, in Langacker's words, they serve to "motivate various aspects of it" (section 1.3). The same view was held by Saussure for whom the value of the whole term was never equal to the sum of the value of the parts (section 2.1.3.1) and by Firth who argued that words should not be treated as if they had 'isolate' meaning (section 2.1.2). The main argument against an analysis that treats word meanings as additive is based on the

polysemy of lexical items that is assumed in cognitive linguistics and which Geeraerts (1997: 2) associates with the "prototypical cognitive-semantic conception" of the lexicon.

Whereas prototypicality tends to multiply the meanings associated with a particular word, the structuralist conception of the lexicon is associated with the isomorphic principle of 'one form, one meaning'. Construing lexical items as having a unitary meaning is defended by Searle (1983: 135 ff., in Taylor 1992), who claims, for example, that *open* is not polysemous in combinations like *open the book/the door/a bottle/one's shirt/one's arms*. On the contrary, *open* is construed as having only one sense, which makes the same semantic contribution to all these expressions. To allow for the fact that the actions coded by *open* differ according to what is opened, Searle introduced a distinction between the understanding of an expression in terms of the unitary meanings of its component morphemes, and the manner in which they are combined, and its interpretation against a 'Background': "a set of beliefs, practices, assumptions, etc., that make it possible for a human being to interact with the world" (Taylor 1992: 136). A similar distinction is made by Bierwisch (1981, 1983, in Taylor 1992: 136 f.) between semantics as an exclusively linguistic phenomenon and conceptualization as an essentially non-linguistic phenomenon: "'Conceptual structures' (i.e. the meanings of words and sentences as used by a speaker in different kinds of contexts) arise through an "interpretation" of semantic structures relative to conceptual knowledge" (Taylor 1992: 138).

Approaches that posit linguistics as separate from pragmatics and general cognition are challenged by functional and cognitive linguists. Thus Taylor (1992: 140 f.) points to a number of problems for compositionality, if only the unitary, 'semantic' meanings of component items are taken into account. For one thing it presupposes that the unitary meanings of these items are adequately described and that people know what they are. Taylor claims that while English speakers "have not the slightest difficulty in explicating (e.g. by mime or ostentation) what is meant by *open a window, open a book, open one's arms*", they would be at a loss to explain "what it means 'to open X', where the value of 'X' ranges over the full set of nominals that can serve as the direct object of

open". Comparing languages highlights another problem: different unitary meanings would have to be postulated for English *open* and Italian *aprire*, because the collocational ranges of the two words are not identical. The English semantic representation for *open* would have to block **open the TV/the light/the hot water* while the Italian equivalents would be included in the representation for *aprire* (Taylor 1992: 145).

Langacker's 'network model' of category structure (Langacker 1988: 51 f.) is mentioned by Taylor (1992: 146) as an alternative that avoids these problems without leading to 'rampant polysemy'. This model is a synthesis between prototypicality, which focuses on individual instances of a category and the relations between them, and schematicity, which focuses on the network as an abstraction subsuming all the individual instances. The meanings of composite expressions would integrate relevant meanings from the two networks representing component items, and nonce-forms would find their place in the relevant networks by extension (cf. section 3.4.1.1 below).

Howarth does not attach any significance to the fact that 'open' is sometimes used instead of 'free' to characterize 'run-of-the-mill' combinations. However, he mentions that the reason why some phraseologists prefer to call them 'free (word-) combinations' or 'free constructions' rather than 'free collocations' may be that they want to reserve the term 'collocation' for expressions that are restricted, whereas his own use of 'free collocation' is meant to reinforce the continuum perspective (Howarth 1996: 34). This seems to imply that phraseology involves more than can be accounted for in terms of restrictedness. Both Cowie (1999: 79 f.) and Howarth (1996: 181) note that *The BBI Combinatory Dictionary of Collocations* (Benson et al.1986), like its successor *The BBI Dictionary of English Word Combinations* (Benson et al.1997), includes expressions like *close*/*shut* and *open the door*, although its authors claim in the introduction that the dictionary does not include free lexical combinations.[30] Howarth (1996: 181) concludes that the key to this apparent discrepancy must be 'the concept of recurrence and familiarity':

[30]"The Combinatory Dictionary does not include free lexical combinations. Free lexical combinations are those in which the two elements do not repeatedly co-occur; the elements are not bound specifically to each other; they occur with other lexical items freely."

> Taking *receive a letter* as an example, the question that needs to be asked is "What is it about a letter that makes it more likely to be the object of the verb *receive* than of many other verbs?" In other words, "What do we do with letters?" Typically we write, read, send and receive them. Tearing them up, throwing them away and losing them may also happen, but are arguably not acts intrinsic to the nature of letters.

Nevertheless, Howarth (1996: 181) insists that combinations like *writing a letter*, etc. should be excluded, because they "pose no problems for learners. Although recurrent and familiar, they are composed according to standard rules of syntax and semantics".

I have previously tried to make a case for 'open collocations' as a motivated term for combinations like *write, read, send* and *receive letters* in Howarth's example that cannot be subsumed under the definition of restricted collocation as defined in phraseology, but which are institutionalized (Poulsen, 1991: 41 f.). Combinations like *tearing up, throwing away* and *losing letters*, which are clearly not institutionalized, might then be referred to as 'free combinations' or 'free collocations'. The important criterion would be institutionalization, whereas the distinction between 'open' and 'restricted' collocations would be of secondary importance. A useful test of institutionalization might be to ask oneself whether there is a context of situation, or frame, to which the processes and participants would typically belong. This would make it possible to base categorization on patterns that are intrinsic to lexical categories..

foot	8	
pick up	4	
meet	12	*the bill*
settle	15	
pay	26	

The set of verbs shown above collocate with *bill* in the sense 'statement of money owed'. The figures indicate the number of instances found of each combination in a search of the BNC, which produced 1,452 concordances for *bill/bills.*[31] My point is that leaving out *pay the bill* from this range, because the combination is not 'restricted'

[31] The BNC, accessed at http://corp.hum.sdu.dk/corpustop.en.html, the English corpus page of VISL, which stands for "Visual Interactive Syntax Learning", a research and development project at the Institute of Language and Communication (ISK), University of Southern Denmark (SDU), Odense Campus. A search was conducted for the single and plural form of *bill*, which produced 3,158 hits, of which 1,452 were in the required sense: 'statement of money owed.

would lead to a misrepresentation of the collocational range of *bill*, as it would seem to imply that it is a less acceptable or idiomatic option than for example *foot the bill* and *pick up the bill*. Besides, the more 'open' collocations are also likely to be the more neutral ones stylistically, so that if they are left out, the 'unmarked' level of formality will be missing and the representation will be biased towards a more 'marked' level. (By the same token, and assuming that the numbers can be correlated with register, an approach based on frequency might be biased in favour of the unmarked level.) Also, as indicated several times before (cf. section 1.2), the question is where to draw the line. Should *settle a bill* be included in the phraseological spectrum or should it be considered to be free? And what about *meet the bill*? These are questions that phraseologists are trying to answer by categorizing word combinations in terms of restricted compositionality.

2.2.4 Restricted compositionality

While idioms are characterized as unmotivated (in the sense of not being 'analysable') and typically noncompositional, restricted collocations are described as motivated (in the sense of 'analysable') and partly compositional. They are categorized in terms of restrictedness, which is 'collocational' or 'lexical' rather than just 'semantic', meaning that it cannot be accounted for in terms of general selection restrictions. This reflects that phraseology takes the perspective of the foreign learner of English for whom conventional expressions represent encoding problems, while for the native speaker they are useful routines facilitating the language processes. This perspective also explains the preference for strict categorization in spite of the difficulties involved. Howarth recognizes that "there are items which could be considered central members of a category and others which straddle the rather fuzzy boundaries" and that this would seem to indicate that categories should "be presented […] as shaded and overlapping areas, rather than as discrete points". However, since this "does not represent a useful visualization of the framework for practical applications", he suggests that what is needed is greater differentiation of criteria (Howarth 1996: 33).

Restricted collocations are defined as partly compositional word combinations in the sense that "one element (here the verb) has a specialized meaning determined by the other element (here the noun)" (Cowie and Howarth 1996: 81). This implies that the noun has an independent sense and that determination is one-way. Alternatively, if it is assumed that words are typically polysemous (see the discussion in the previous section and also in section 3.4.4.2 below), it can be claimed that determination tends to be mutual, as in the following examples:

pick up the bill	(*pick up* = 'pay'; *bill* = 'bill of payment')
throw out the bill	(*throw out* = 'reject'; *bill* = 'bill of parliament')

(This ignores the way in which the co-text or pragmatic circumstances could lead to alternative, literal interpretations of 'bill' as a document). A definition that does not preclude polysemy of the base-word or mutual dependence is found in Cowie (1981: 224): "A collocation is by definition a composite unit which permits the substitutability of items for at least one of its constituent elements (the sense of the other element or elements remaining constant)".

While those who, following Firth, treat lexis as an independent level will consider collocation to be an equal and mutual relationship, phraseologists, who are only interested in syntactic combinations, see collocation as a hierarchical relationship between a collocate and a base that determines it. Hausmann (1985: 119 specified that the base would most typically be a noun, because "nouns express the things and phenomena in the world that people find something to say about," while "verbs and adjectives would be possible as base words only in combinations with adverbs":

> Die wichtigste Basiswortart ist das Substantiv, weil es die Substantive sind, welche die Dinge und Phänomene dieser Welt ausdrücken, über die es etwas zu sagen gibt. Adjektive und Verben kommen als Basiswörter nur insoweit in Frage, als sie durch Adverbien weiter determiniert werden können.

From an encoding point of view, it makes sense to focus on syntactic combinations and to look for a verb or an adjective to go with a noun and an adverb to go with a verb or adjective, rather than the other way round. The problem of polysemy of the noun will not exist for the speaker/writer since he or she will have chosen a sense to begin with.

Still, in many cases the verbal process is the focus of attention at least as much or even more than the nominal participant. An indication of the relative dependence/independence of the verb could be whether it alone would be able to evoke a typical context of situation, or frame, in which both verb and noun have their roles to play: if asked to state what they consider to be the most typical object of a given verb, say *ask*, how likely would people be to come up with a given noun, say *question*, *favour*, *price*, or *the neighbours*?

According to Langacker (1987: 349) "canonically the structures in a valence relation manifest substantial asymmetry, with one of them (on balance) clearly dependent, and the other autonomous." The dependent structure is construed as having a substructure that functions as an 'elaboration site', or 'e-site', which is elaborated by the autonomous component, and the degree of dependence correlates positively with the salience of the e-site within the dependent structure and the degree to which it is elaborated by the autonomous component (Langacker: 1987: 300 f.). I will return to this discussion in the empirical part of this study arguing that a noun may also be construed as having salient substructure, which may be elaborated by a verb to a greater or lesser extent (cf. section 3.4.4.2).

In phraseology, the verb in a verb + nominal object collocation has been characterized as typically having either (1) a weakened, grammaticalized or delexical meaning, as in *do business*, (2) a figurative meaning, as in *run a company*, or (3) a specialized or technical meaning, as in *conclude a contract* (cf. section 1.5). It is to be expected that verbs with a technical or specialized meaning will be less dependent on the noun than verbs with a delexical meaning, because they combine with relatively few nouns, which moreover will be semantically related, as in *conclude an agreement/an alliance/a contract/a deal/a treaty*. This has led some researchers to posit a separate category of 'specialized lexical combinations', arguing that combinations should not be classified as collocations if they can be defined in terms of more general selection restrictions (L'Homme and Bertrand 2000: 497 ff.; see also Heid 1994: 236 ff.). Another aspect is how narrowly 'technical' or 'specialized' should be defined. Should it be interpreted as 'belonging to a technical register' or could it also relate to 'domains of everyday life' as

in *write a letter* and *speak a language*? (Howarth 1996: 31). Howarth (1996: 92 f.) chooses the narrow definition for his own investigations, whereas he classifies the familiar, everyday combinations as fully compositional free collocations (section 2.2.3). Generally, it is a problem for the phraseological model how to treat institutionalized combinations in which both elements have a literal meaning, since basically the model equates idiomaticity with non-literal meaning (see section 2.5 below).

Phraseologists include both the syntagmatic and the paradigmatic dimension in their efforts to identify collocations on the basis of restricted compositionality. It is not enough that "an element of their meaning is indicated when their habitual word accompaniment is shown" (Firth 1968b: 106); it is a requirement that this element cannot be accounted for in terms of general selection restrictions. This has shifted the focus to the lexical set as the paradigmatic dimension of collocation, and to the phenomenon of 'substitutability'. As the paradigmatic set is defined by the syntagmatic relationship of collocation, it is taken as evidence of restricted collocability if sets cannot be freely extended to include synonyms. This line of argument is often supported by examples of 'overlapping lexical sets' showing how collocational restrictions make it necessary to discriminate between verbal expressions like *assume, acquire, take on, adopt* that are thought of as near synonyms, in order to form acceptable combinations with nouns like *importance, form, role* and *mantle*. The resulting overlaps (cf. Fig. 1, section 1.3) illustrate the nature of the problems facing foreign learners, but they also reinforce the impression that collocations are 'odd comings-together of words' as Palmer put it (cf. section 2.1.1), and it is implied that the normal state of language affairs would allow all (near) synonyms to replace each other. I would argue that the members of a set are rarely, if ever, fully synonymous, and that, as a matter of principle, the phraseological status of combinations like *adopt a role/a form* should not depend on whether **adopt importance* or **adopt a mantle* are established combinations. This is extrinsic evidence, whereas intrinsic evidence needs to be based on an analysis of the actual constituents and their mode of integration. As Howarth (1996: 107) points out, while some 'blockages' seem arbitrary, others are found to be semantically motivated by "slight, though significant differences of meanings between the nouns and consequently between verbs". What makes the difference between that which seems

arbitrary and that which seems motivated is presumably whether we can trace the origins of expressions in a way that allows us to see why in some cases lexical categories show overlap and in others they do not. It may not be entirely arbitrary, for example, that *adopt* (from *ad* + *optare*: 'choose') has become established in combinations with *role* as well as *form* rather than *importance.*[32]

Just as blockages of lexical sets are used to strengthen the case for collocational status in phraseology, relatively free substitutability is felt to weaken it, as noted in the discussion of 'technical' collocations above. According to Cowie (Cowie et al. 1983: xiii), "In expressions with more lexical variation like *cardinal error/sin/virtue/grace*, restricted collocations are seen as being more like 'open collocations' or free combinations". Amosova's 'phraseme' must therefore be considered the prototype of a restricted collocation, because it requires the meaning of a collocate to be determined by a single base-word, as in *grind one's teeth* or *jog somebody's memory* (cf. section 2.1.4). However, phraseologists today are casting the net wider than Amosova and even Vinogradov (section 2.1.4) in order to capture all the expressions that are felt to be a problem for learners although they are not so obviously restricted. Cowie (in Howarth 1996: 30 f.)[33] discusses a set of criteria for the definition of the 'phraseoloid', Amosova's term for combinations that allow contextual determination by a limited set of nouns and which she therefore did not consider to be phraseological. Cowie uses the example *pay one's respects/a compliment/court to someone*, which is characterized in terms of three criteria:

[32] Mel'čuk (1998: 31) uses the notion 'lexical function' to handle the problem of overlapping sets in a generative framework. A lexical function is defined as a very general and abstract meaning, coupled with a deep-syntactic role, which can be lexically expressed in a large variety of ways, depending on the lexical unit to which this meaning applies. This approach also defines collocations as not being constructed 'unrestrictedly' according to the' selection rules' of a language, but treats idiosyncracy as normal. The approach has shown that the number of 'general and abstract meanings' is likely to be very large and probably cannot be reduced to a finite set.

[33] Papers from the international symposium on phraseology. Leeds April 1994. These papers have subsequently been published in Cowie (1998b), but the above example is from the version cited by Howarth.

(a) a verb used in a figurative sense ('offer' or 'extend')

(b) contextual determination by an (arbitrarily) limited set of nouns (*one's respects/a compliment/court*)

(c) determination of no other verb in the same sense by the same limited context (*present*, if possible, would need a plural form of *compliment*)

Howarth (1996: 31) subsequently discusses ways in which these three criteria can be relaxed. Thus criterion (a) can be given a broad interpretation to include verbs used in 'technical' or 'specialized senses', and 'technical' can be interpreted as 'domains of everyday life', so that combinations like *write a letter* and *speak a language* could be included, although the senses of these verbs are literal. If criterion (b) is waived, it would allow "the inclusion in the phraseological band of the spectrum those combinations of a verb in a specialized sense and a potentially large number of collocating nouns" like the examples mentioned below with *run* + an object meaning 'human organization'. Finally, if criterion (c) is relaxed to allow a limited range of synonymous verbs, as suggested by Cowie, it would be possible to include as phraseological those combinations in which the figurative sense of a small number of synonymous verbs is determined by a limited set of nouns, as in *call/convene a meeting/session/gathering*. Cowie and Howarth's subcategorization (1996: 82 f.) reflects how the criteria have been relaxed:

(a) invariable collocation: *foot the bill, break a journey*

(b) collocation with limited choice at one point: *give/allow/permit access*

(c) collocation with limited choice at two points: *find/experience + trouble/difficulty*

(d) overlapping collocations

The question is just how limited 'limited choice' needs to be, and how much variation to allow before it can be argued that the meaning of the collocate is 'independent' rather than determined by its base. This is a crucial question since the framework makes phraseological status dependent on restrictedness, a choice which links categorization to factors which are extrinsic to individual word combinations and is perhaps not so helpful to foreign learners either. They would want to know specifically whether *run a company* or *run a factory* are acceptable collocations, irrespective of how many other nouns combine with *run* in the sense of 'manage' (*a business, a theatre, a school, a*

hospital, etc.). They are also likely to want information about possible alternatives in each case in order to be able to choose a combination that is not only acceptable, but optimal in terms of level of formality and degree of specificity. As Howarth (1996: 31) observes, even if - as in the case of *run* - "the restrictions on the object nouns can be stated in general semantic terms ('human organization') and do not form a limited set [...], these combinations could still be regarded as a phraseological problem for the learner". For example, Howarth says, *run* may be a more central choice with some types of 'human organization' than with others, and he concludes that "the significant phraseological focus is on the noun and on the fact that the selection of the verb is highly restricted" (1996: 31). Howarth's discussion shows how intrinsic evidence of the collocation of a given noun with a given verb may conflict with extrinsic evidence based on how many other nouns collocate with the same verb. It also shows that phraseological combinations cannot be reliably identified on the basis of extrinsic, distributional evidence of substitutability.

Collocations with a 'delexical' or 'light' verb like *make, break* and *take, have* are normally described as a special type of collocation, although they actually belong to the group in which the verbs are used in a figurative sense. What makes them seem different from combinations like *launch a campaign* and *lift a restriction* is that they are used with a wide range of nouns that cannot easily be subsumed under specific headings (*make a mistake, an improvement, a decision, a contribution*, etc.). They may consequently be construed as either extremely polysemous, changing their meaning in chameleon-like fashion to blend in with the nouns, or as having stable, but much weakened, 'grammaticalized' meanings. In the former case, they would fit the description of a restricted collocation very well, in that their meaning, in each case, would be almost entirely dependent on that of the noun. In the latter case, if they can be said to have a more abstract, grammatical or 'support verb' function, it makes no more sense to discuss them in terms of restricted compositionality than it would for, say, auxiliary verbs or affixes. These delexical verbs belong to the words that corpus studies have shown to be among the most frequent in language. Sinclair (1991: 113) makes the tentative generalization that such frequent words tend to have "less of a clear and independent meaning than less frequent words" and that consequently "we are reduced

to talking about uses rather than meanings". Sinclair refers to this process as 'progressive delexicalization'; whether it can also be described in terms of grammaticalization as implied by terms like 'support verb construction' and 'grammaticalized expression' is a question that I will come back to in the empirical part of this study in Part 3 (sections 3.4.4.3 and 3.4.4.4).

Recent phraseological studies (for example Howarth 1996; Cowie and Howarth 1996; Moon 1998) have demonstrated the pervasiveness of collocational patterns in language, and the attempt to refine the system of categorization inherited from the Russian lexicologists has revealed how idiosyncratic these patterns are, even if a few prototypes may serve to illustrate the range. Phraseologists have found it difficult to categorize combinations in terms of criterial attributes, even if they have tried to allow for the gradedness that they found by relaxing the criteria and subdividing the category of restricted collocations. But as Langacker has pointed out, if the defining criteria are loosened in a criterial attribute model, there is no arbitrary stopping point, and the model no longer serves its purpose (cf. section 1.5).

2.2.5 Analysability, compositionality and the literal/figurative distinction

Although they are not described as co-extensive, the parameters of compositionality, analysability and literalness go hand in hand in defining the phraseological continuum, sometimes referred to as a 'scale of idiomaticity' (Cowie 1999: 70-71). Thus free collocations are characterized as fully compositional, transparent and literal (cf. section 2.2.3) whereas, at the other end of the scale, pure idioms like *kick the bucket* or *shoot the breeze*, are defined as noncompositional, opaque and figurative. In contrast to 'pure' idioms, 'figurative' idioms like *change gear* and *make a U-turn* are found to be transparent, because they have a literal equivalent (Howarth 1996: 23). Restricted collocations are analysable like figurative idioms and free collocations, but occupy an intermediate position in being neither fully compositional nor unitary, because of the figurative or specialized meaning of one element and the literal meaning of the other. In this section I will discuss the difference between the two notions of compositionality

and analysability as well as the way in which the distinction between literal and figurative meaning is related to the other two notions.

Compositionality, as Langacker uses the term, is about "the regularity of compositional relationships, i.e. the degree to which the value of the whole is predictable from the value of the parts" (1987: 457). I have pursued the argument, in section 2.2.3 above, that full compositionality should not be seen as the norm from which conventional expressions deviate, because the integration of lexical items in a composite structure does not normally happen in an additive way. However, if it is normal for composite structures to be motivated by, if not predictable from, component items, as cognitive linguists would see it (cf. sections 1.3 and 2.2.3), how is it possible to account for the motivation of the unitary meaning of idioms, which cannot be motivated by individual items? From a structuralist viewpoint, idioms are basically arbitrary. Thus Weinreich (1969: 76), asked "why should the acceptance of what one has fought against be signified by *eating crow*, and not **eating dog* or **drinking mud* or **smelling rotten eggs*?" In response to Weinreich's comment (1969: 76) that the relation between idiomatic and literal meanings was "so unsystematic as to deserve no place in the theory", a cognitive linguist might point out that *eat crow*, like other expressions such as *swallow a bitter pill*, which happened to become established, could be seen as representing an underlying conceptual metaphor (something like: ACCEPTING SOMETHING AGAINST ONE'S WILL IS LIKE EATING SOMETHING NASTY) just as other idioms have been shown to reflect metaphorical concepts in a systematic way (cf. section 2.1.5 above). It should be noted that in a cognitive framework, which includes a diachronic view, 'motivation' relates to 'original motivation' and thus to compositionality, whereas in a structuralist framework, which focuses on synchronic facts, it relates to 'present motivation', or analysability (cf. section 2.1.3.1 above). It is assumed in this study that synchronic evidence in the form of the polysemy of words reflects their diachronic development from the perspective of a contemporary user (cf. section 3.2.3 above).

Analysability is an independent parameter that refers to the extent to which speakers are aware of "the contribution that individual component structures make to the composite

whole" (Langacker 1987: 457). Consequently, it is possible for somebody to be aware of the individual components of expressions like *blackboard* or *take a point* that are not related to the meaning of the whole in any obvious way, and to be unaware of the components of expressions like *computer* and *bed and breakfast*, which may seem to be more directly related to the meaning of the whole, although they do not exhaust it. Analysability is thus a relative matter depending as it does on the background of the individual person and on the particular context in which the expression is used, something that can be manipulated for stylistic effect (cf. section 2.1.5). While a freshly coined expression will be fully analysable, once an expression has gained unit status through institutionalization, it is more likely that "its composite structure [...] may be activated autonomously and a gradual loss of analysability may occur" (Langacker 1987: 457, 465). Psycholinguistic studies have found that idioms can be subdivided according to how likely people are to perceive them as compositional. It should be noted that when such studies discuss possible 'degrees of compositionality', they refer to the likelihood that people will 'decompose' an institutionalized expression, which corresponds to its analysability as defined here (cf. section 2.1.5).

How is the literal/figurative distinction used to define categories and how does it relate to compositionality and analysability? In phraseology, as illustrated by Howarth's definitions of verb + nominal object combinations (1996: 47), there is a clear correlation with degree of idiomaticity:[34] (1) in free collocations, outside the phraseological range, both elements are used in a literal sense, (2) restricted collocations have one element that is used in its literal sense, whereas the other is specialized (a delexical, technical or figurative verb), and (3) figurative idioms have figurative meanings as a whole, but retain a current literal interpretation in contrast to (4) pure idioms, which do not. The distinction between literal and figurative meaning is thus related to motivation, or analysability, rather than to compositionality: an expression can be motivated either in terms of its literal elements alone (in which case it will not count as phraseological), or in terms of figurative extensions of literal elements (restricted collocations and figurative idioms). Figurative expressions that have no current literal equivalents (pure

[34] For a discussion of the different meanings of the terms 'idiom' and 'idiomatic', see Moon (1998: 3 f.)

idioms) are not considered to be motivated at all but rank highest on the scale of idiomaticity, implying that idiomaticity is the opposite of literalness (cf. section 2.2.4).

As mentioned in connection with the discussion of the role of different aspects of context (section 2.1.5), psycholinguistic studies have found that the question of whether an idiom has a literal interpretation may not be all that important for analysability, because ambiguity will often be resolved by features of the context. Moreover, as pointed out in Howarth (1996: 53), "the idiomatic sense of an expression is not a subsidiary interpretation, but will usually dominate over the alternative literal interpretation, if one exists". Consequently, people will often use 'literal markers', such as *literally* to block a figurative reading.

Category	**Phraseological**	**Compo-sitional**	**Analysable / Motivated**	**Literal**	**Figurative**
free collocation	no	yes	yes	yes both elements	no
restricted collocation	yes	partly	yes	partly one element	partly one element
figurative idiom	yes	no	yes	no, but literal equivalent exists	yes both elements
pure idiom	yes	no	no	no, and no literal equivalent exists	yes both elements

Table 7: Phraseological categories

It appears from Table 7 above that the figurative meaning of a phraseological expression correlates negatively with compositionality. If instead, as in cognitive linguistics, elements are seen as dynamic schematic networks, or complex categories, the figurative extension of an element will be a normal compositional process. Collocations with a figurative element, such as *lift a restriction* can thus be accounted for in terms of a functional compositional strategy that makes it possible to express abstract, literal meaning in terms of figurative, concrete meaning. The same functionality operates when idioms are motivated by a conceptual metaphor that refers to complex real-life situations in terms of a concrete and familiar story in a way that makes it easier to grasp and to handle linguistically, and perhaps also psychologically. My point is that a category of collocations should integrate the literal/figurative distinction with the notions of compositionality and analysability in a way that brings

out the functionality of institutionalized expressions. Figurative meaning should not automatically be construed as restricting compositionality, and literal meaning should not automatically be equated with full compositionality, so that an expression with one literal element is considered to be less idiomatic than one without literal elements, and an expression with two literal elements is considered to be outside the scope of phraseology altogether.

Finally, a closer look at some examples from the literature may serve to illustrate that it is not always obvious to which of two neighbouring phraseological categories specific expressions should be assigned. Whether an expression should be classified as a pure idiom or a figurative idiom is made dependent on its analysability, or motivation, as mentioned above. While figurative idioms are characterized as clearly motivated in the sense of being fully analysable and transparent, pure idioms are said to be unmotivated and quite opaque (Howarth 1996: 24). The problem, as also noted by Howarth, is that analysability is relative; it depends on such factors as age and cultural background and on whether somebody is a native speaker or a foreign learner. To me as a non-native speaker of English, the following examples (from Cowie 1999: 71) have pretty much the same status:

pure idiom:	*spill the beans*
figurative idiom:	*call the shots*

I know the meaning of these expressions as wholes and they seem motivated in terms of the constituent items in the sense that I can associate the 'calling of shots' with the 'exertion of influence' (even if I was never in the army), and the 'spilling of beans' with the 'giving away of information'. Not that I have ever spilt any literal beans either, but the expression draws on the familiar container metaphor (Lakoff and Johnson 1980: 29 f.). It is used in other expressions with similar meanings like *let the cat out of the bag* and *have something up one's sleeve,* which, by the way, have Danish equivalents (*slippe katten ud af sækken, have noget i ærmet*). The point is that whether an expression seems opaque or transparent depends on real-life experience, which includes linguistic experience. A pure idiom like *kick the bucket,* whose origin is uncertain, does not draw on a general conceptual metaphor for dying in any obvious way. I believe, however, that

once people know the meaning of the idiom, they are likely to make up their own stories on the basis of the actual words, the transitive construction and the knowledge that such expressions will often be metonymic, in that they only represent a part of the picture. My own version is about someone hanging himself, which involves kicking away the bucket he is standing on. This rather morbid story may not be so far-fetched, since others have had a similar image of someone being hanged (Makkai 1993: 314). Another version (Makkai: 1977) is about the slaughtering of pigs.[35] To the extent that a metonymic link is perceived between an idiom and its symbolic meaning, I would argue that it does not appear to be entirely arbitrary even if it is not transparent and even if conclusive diachronic evidence is not available.

The borderline between idioms and restricted collocations, which is generally held to be the most clear-cut, is accounted for in terms of idioms having a unitary meaning whereas collocations are compositional with one element used in its 'normal' meaning, and a second element which is restricted by the former. Restricted collocations like *jog one's/somebody's memory* are sometimes described as 'idiom-like', or even 'semi-idioms' if the verb has a particular sense that occurs in no other context. (Cowie et al. 1983: xiii), but such an expression is clearly not unitary as *memory* still means 'memory', just as in *refresh somebody's memory*. Going from *jog one's/somebody's memory* to *refresh one's/somebody's memory* would be an example of 'substitution not amounting to commutation' (cf. section 2.1.2), whereas going from *call the shots* meaning 'give the order to shoot' as in *No team captain to call the shots?* to the sense 'make decisions' as in: *You call all the bloody shots round here!* exemplifies substitution that does amount to commutation (both examples are from the BNC):

figurative idiom:	*call the shots*
restricted collocation:	*call the shots*

The question is, however, whether the figurative idiom can be interpreted as having a compositional meaning in line with the literal expression:

[35] According to Makkai (1977), "[o]thers believe it comes from 16th century English farms where pigs were tied to beams, called *bucqet* (from French), and when the farmer slit their throats, their feet, that were tied to these beams, ´kicked´ in protest."

figurative idiom:	*make a U-turn*
restricted collocation:	*make a U-turn*

U-turn can be interpreted as having, in addition to a literal meaning (which includes the metaphor 'U-turn'), a metaphorical one (which does not involve any literal 'turning'). Also in its metaphorical sense, *U-turn* collocates with *make*, so that we have two collocations rather than an idiom and a collocation. Of the 56 instances of *U-turn* in the BNC, only twelve are literal, whereas the rest are figurative and document that this sense is well established and independent of *make*:

- *He needs to work furiously to rebuild his reputation and authority, shattered by a humiliating series of **U-turns** and political defeats which have all but torn asunder his party and even his Cabinet.*

By comparison, *shots* does not seem to have the meaning 'influence' except in combinations with *call*.[36] The point is that the figurative meaning of the noun is not enough to distinguish between figurative idioms and restricted collocations in such expressions.

The borderline between restricted collocations and free collocations is important, because it defines the limits of what is to be considered phraseological. The reason Howarth (1996: 37 f.) gives for classifying *go to school/church/work* as free collocations although they are institutionalized expressions is that both elements are used in their literal senses. *Take a vacation*, on the other hand, is classified as a restricted collocation (cf. section 2.2.2):

restricted collocation:	*take a vacation*
free collocation:	*go to school*

I would argue that firstly *go* is not more or less 'literal' in *go to school*, etc. than *take* in *take a vacation*; after all, the following example (from the BNC) does not imply that Steve Norris actually walked to school with the Beatles:

[36] A closely related sense is 'an influential person' (only in combinations with *big*): *It reminded him of the big shots in the trade union movement having sandwiches at Number 10 all those years ago*. (Example from BNC).

- *Mr Steve Norris, 46, a self-made millionaire representing the wealthy Essex constituency of Epping Forest, who went to school with two of the Beatles, becomes Parliamentary Under Secretary of Transport.*

Besides, as indicated above, the distinction between literal and figurative expressions is not a functional one for the purpose of separating what is phraseological from what is not, since literal expressions like *drill a hole* and *shift gear(s)* may be no less established than figurative ones. This also appears from the following examples, in which Cowie (1999: 56) uses literal meaning to identify a restricted collocation:

restricted collocation: Now *don't lose sight of the rabbit!*
([*lose + sight of*] + *the rabbit*)

figurative idiom: She *seems to have lost sight of the main purpose of the campaign.*
([*lose + sight* o*f*] + *purpose*)

The second example is described as a figurative extension of the literal, restricted collocation. Bearing in mind the 'decomposition hypothesis' (section 2.1.5), I would say that the restricted collocation *lose + sight of* and the figurative idiom *lose + sight of* are equally analysable in this case, since it is possible to find a plausible relation between the figurative meaning of the whole and the literal meaning of the parts, and it is no problem to distribute the figurative meaning over the parts of the figurative expression, which furthermore draws on a familiar conceptual metaphor: UNDERSTANDING IS SEEING (Lakoff and Johnson 1980: 48). I would consequently classify *lose sight of* as a conventional collocation in both cases. At the next level, the combination between *lose sight of* and *the rabbit* is not likely to be institutionalized, whereas the combination between *lose sight of* and *purpose* might be. This is not because of the literal/figurative distinction, but because I suspect that there are no typical contexts of situation (or frames) in which rabbits are being lost sight of, whereas this happens more routinely to purposes.

The last examples illustrate the three types of expression that are categorized as free collocations, which are not included in the phraseological spectrum. The first two expressions are used by Howarth (1996: 31) and have been referred to before (section 2.2.3), while I have added the last example to illustrate a nonce-formation (imagine a letter written on scented paper).

free collocation (institutionalized):	*receive a letter*
free collocation (not institutionalized):	*lose a letter*
free collocation (nonce formation):	*inhale a letter*

Howarth distinguishes between *tearing up*, throwing *away* and *losing letters*, which are "not acts intrinsic to the nature of letters" and "*writing, reading, sending* and *receiving letters*", which are the things we typically do with letters. Nevertheless he concludes that "Applying the main criterion [...] of semantic specialization of the verb, in each case the verb seems to be used in its primary, literal sense, as is the noun. None is delexical, nor technical; they do not therefore seem to qualify as restricted collocations" (Howarth 1996: 182). I have argued before (section 2.2.3) that all institutionalized expressions should be considered phraseological, irrespective of whether they are used in a more general literal sense or whether they are used in a more restricted sense. More specifically, while the literal/figurative distinction may serve to make a useful comparison between the literal and the figurative uses of specific expressions, it cannot be used as a criterial feature to identify collocations or other phraseological categories, or to distinguish between what is phraseological and what is not.

2.3 Summary and conclusions

The range of different uses of *collocation* as a technical term, which was introduced in Part 1 (section 1.2), can be said to form a schematic network of overlapping meanings. The growth of such a network was explained by Wittgenstein in functional terms: for example something would be referred to as a 'belief', because of its similarities with some of the things that were previously called beliefs. In this way the application of a term would be extended from previous cases to new cases "as in spinning a thread we twist fibre on fibre. And the strength of the thread does not reside in the fact that some one fibre runs through its whole length, but in the overlapping of many fibres" (Wittgenstein 1953: 32).

It may well be that it is normal for theoretical frameworks to be developed in the same way, not by discarding old theories altogether, but by taking elements from them and integrating them with new ideas for new purposes. On this assumption I have argued, in the first part of Part 2, that the phraseological approach to the categorization of collocations brings together a number of different notions and practical concerns. Thus I have claimed that the notion that conventional expressions are arbitrary has been taken from Saussure while the idea that they deviate from a standard of full compositionality is rooted in generative linguistics. Whereas the technical notion of collocation comes from Firth, it was preceded by an interest in the phenomenon from the perspective of the foreign learner of English. As a consequence, the focus has been on the problems created by restrictions on substitutability (also referred to as commutability or recombinability), and the solution has been sought in a strict system of categorization based on the model developed by Russian phraseologists.

In the second part of Part 2, I have critically examined the resulting framework, claiming, for instance, that it tries to combine the ideal of classical categories based on criterial features with a continuum model, which implies gradedness. This framework only includes collocations in syntactic combinations (cf. definition c in section 1.2 above), and only those institutionalized expressions that are found to deviate from an assumed standard of full compositionality. This means that institutionalized expressions that are not found to deviate from this standard are not included. In this view, restricted compositionality is what defines phraseological combinations in contrast to free collocations. Restricted compositionality is defined syntagmatically in terms of the semantic dependence of the collocate on the base and paradigmatically in terms of arbitrarily restricted substitutability. Predictability, analysability, and literalness are used as evidence of full compositionality, whereas restricted compositionality is characterized by the absence of these features.

I have questioned the notion of full compositionality, arguing that lexical items are generally polysemous. I have further argued that evidence of restricted substitutability is extrinsic to a given word combination, since it is based on the extent to which constituent items also co-occur with other items. In order to account for the complex

patterns of collocability that phraseologists have found, I suggest that it is necessary to give up the notion of paradigmatic restrictedness in terms of arbitrarily limited substitutability and to reassess the notion of syntagmatic restrictedness as deviation from a standard of full compositionality.

Assuming that combinations are partly motivated by constituent items, I consider them to be neither arbitrary nor fully predictable. I see analysability as relative and independent of the factors motivating a word combination and its psychological and social status, just as the literal or figurative meaning of a combination does not determine its phraseological status or category. I have suggested that all institutionalized expressions should be included in the phraseological spectrum and that the idiosyncrasy of entrenched combinations should be seen as a normal characteristic, which can be understood as being motivated by typical contexts of situation and frames. In this view, entrenched linguistic structures reflect typical situations of use as well as the cognitive models associated with them and are available for routine use as well as for the creative exploitation.

3. **Collocations in a functional and cognitive framework**

The general purpose of this part is to test the assumption that a functionally and cognitively based framework is descriptively more adequate and has greater explanatory potential than the framework of phraseology, which was discussed in Part 2. As far as methodology is concerned, I will try to make good the claim that using corpus data on frequency as just one input for qualitative analysis is more suitable than a quantitative corpus linguistic approach which relies mainly on statistical observations of frequency of co-occurrence. Before presenting my research questions and the data that I will use to test them, I will discuss the notions 'functional' and 'cognitive' in somewhat greater detail (in section 1.4), in order to explain what I see as the implications of using them in a study that is based on linguistic data. I will also discuss some methodological issues related to the nature of my data. My research aims, which spring from the general questions raised in Part 1 (section 1.5), will be specified in section 3.3, where I will also describe the design of my case study of the collocation *break an appointment* and give a general presentation of the data. Section 3.4 will include the analyses based on the case study and structured according to my research questions. In section 3.5, I will summarize my findings and evaluate the suitability of the methodology used.

3.1 What is 'cognitive' and what is 'functional' about language?

Basic notions in cognitive linguistics like 'schema', 'gestalt' and 'prototype', have their roots in a tradition that antedates the period of Behaviourism when what was going on in the mind was not to be the concern of linguists. According to Sinha (2001: 1), cognitive

linguistics is the “rightful inheritor” of this older tradition, which is described as “centred in psychology but drawing heavily on biology, linguistics, philosophy, anthropology and sociology”. The swing of the pendulum in linguistics, after the ‘Chomskyan revolution’ (cf. section 2.1.3.2 above) introduced the period of Classical Cognitive Science. Because its focus is on innate mental properties as the basis for formal principles of syntax, it does not end the segregation between body and mind and consequently has no use for notions that try to capture the continuity between them. Second Generation Cognitive Science, of which Cognitive Linguistics is a part, tries to do just that: “The higher mental processes are considered [...] to occupy the problematic and indeterminate zone at which biologically-based psychological processes shared by human organisms with other mammals, interface with, and are perhaps transformed by, the processes of social life, symbolization, and cultural tradition.” (Sinha 2001: 3). For linguists this means that they share their general subject matter with neuroscientists, psychologists, philosophers and others and have to take their work into account when drawing inferences from their own, linguistic, data. The question is what criteria they have to go by in a framework which does not define itself as narrowly linguistic.

Functional and cognitive linguistics are normally defined in terms of what they have in common and in contrast to generative linguistics (cf. section 1.4 above). Again the focus is normally on their agreement to abolish distinctions, e.g. between syntax, semantics and pragmatics, rather than on differences between them. Any differences are typically characterized as a matter of emphasis, in that functionalists will attach greater importance to the interactive role of language as communication whereas cognitive linguists will stress that linguistic knowledge is part of general cognition. The question is whether the differences are more substantial and, if so, what practical consequences they should have for linguists defining the object of their research. This study agrees with the approach taken in Harder (1996a, 1999), according to which both ‘cognition’ and ‘function’ should be defined more precisely than it is generally done. Harder suggests that the domain of a cognitive semantics should only include the core of higher-level cognitive competence

constituted by 'conceptual, intentional representations'.[37] This implies that 'world states' are conceived as existing independently of conceptualization, but may be linked to it by 'intentional relation' - meaning that people are aware of the fact that they are conceptualizing something outside themselves. Moreover, perception skills and motor routines are conceived as pre-conceptual, but linked with conceptualization "as part of the overall processing-and-action competence of the subject" (Harder 1996a: 75). Finally, it is argued that actual structure of experience should not be confused with secondary representations of experience, as when "metaphors are understood as operating in the domain of primary experience" (Harder 1996a: 72, 75).

In contrast to this, Lakoff and Johnson (1999: 20) argue, on the basis of ongoing research[38] in cognitive science, that there is no "fully autonomous faculty of reason separate from and independent of bodily capacities such as perception and movement". The question they are researching is whether conceptual inference uses the same brain structures as perceptual motor inference and that consequently much of conceptual inference can be said to *be* sensorimotor inference. In the process of language acquisition, after a period in which experience in the sensorimotor domain, like 'seeing', is conflated with evaluation in the epistemic domain, like 'knowing', the child manages to distinguish between the two, which is when 'conceptual metaphor' emerges (Lakoff and Johnson 1999: 48), accounting for statements like "*You see what I'm saying?*" Harder recognizes it as a key achievement of cognitive linguistics that it demonstrates "the many ways in which the process of articulating and encoding secondary representations must be understood by reference to the experiential basis of conceptualization" (Harder 1996a: 75). From a linguist's point of

[37] Searle, from whom the concept 'intentionality' is taken, explains it as follows: "Most important, conscious states typically have "intentionality," that property of mental states by which they are directed at or about objects and states of affairs in the world. Philosophers use the word intentionality not just for "intending" in the ordinary sense but for any mental phenomena at all that have referential content. According to this usage, beliefs, hopes, intentions, fears, desires and perceptions all are intentional. So if I have a belief, I must have a belief about something. If I have a normal visual experience, it must seem to me that I am actually seeing something, etc.", (John R. Searle: *Consciousness*).
http://philosophy.berkeley.edu/jsearle/Consciousness1.rtf

[38]The Neural Theory of Language (NTL) project at the University of California, Berkeley, is an interdisciplinary research effort to answer the question: How does the brain compute the mind? http://www.icsi.berkeley.edu/NTL/overview.html

view, I agree with Harder that it is necessary to try to tell things apart in order to understand the connection between them; we need to "refer to objects, experiences, words, and concepts separately, so as to be able to describe how they are related." (Harder 1996a: 75)[39].

By the same token, 'function' should be defined in greater detail to be able to appreciate how functions 'recruit conceptualizations' (cf. section 1.4 above). The definition used in Harder (1996a) comes close to the most general dictionary definition: 'the special kind of activity proper to anything; the mode of action by which it fulfils its purpose' (OED). More special meanings are then derived from different kinds of purposes dependent on special contexts, as when we talk about the 'function' of a word in contrast to its 'form', or about the possibility of a lexical item developing a more 'functional' role. A function thus involves action for the purpose of achieving an effect defined by the context: "when we describe something in terms of function, we change the emphasis from the thing itself to its contextual role, which determines the norm in terms of which its function is defined" (Harder 1996a: 88). The norm that Harder posits for linguistic expressions is derived by analogy from biology: just as wings have the function of contributing to the survival of birds, and are therefore themselves preserved, so it is argued, linguistic constructions persist if they "further the reproduction of tokens of which they form part" (Harder 1996a: 91). In the case of entrenched collocations, I would argue that their function, or contextual role, is the reproduction, through renewal of connection, of the contexts of situation or frames which they evoke (cf. sections 1.5 and 2.1.2 above).

If this line of argumentation is accepted, it follows that the primary function of language is communicative interaction, since talking to oneself does not improve the chances of

[39] That experiences and concepts are related seems to be confirmed by brain scan research at the Neurobiology Research Unit of Rigshospitalet, Copenhagen University Hospital, by neuropsychologist Christian Gerlach: "If you look at the areas of the brain that are activated when we understand different concepts, you may say that they more or less reflect what we are doing as human beings" [my translation, based on the web source mentioned below]. The results were presented in the science programme 'Viden om' on the Danish Radio's TV-station DR2, on January 27, 2004.
http://www.dr.dk/Videnskab/viden_om/Programmer/139tale/Ordbilleder+på+hjernen

'survival' of linguistic structures. What it is interesting to study, in this view, is therefore not so much the language of individuals, but that of a speech community during a particular period. More specifically, the focus would be on the way in which linguistic expressions contribute to 'communicative events': "The (linguistic) meaning of a linguistic expression is its (canonical, proper) communicative function, i.e. its potential contribution to the communicative function of utterances of which it forms part" (Harder 1996a: 101). This definition emphasizes the interactive function of conceptualization and subsumes the more general definition of 'meaning as conceptualization' in cognitive linguistics, as it is assumed that, in order to make its contribution, a linguistic expression "calls on all the skills that are necessary for the success of the act to which it contributes - including most saliently the cognitive capacity of the speaker and hearer" (Harder 1996a: 101). Such a definition offers the possibility of distinguishing between *what* the effect of a linguistic expression is, viz. its contribution to a communicative event, and *how* its ability to achieve this effect can be accounted for by means of assuming certain underlying cognitive skills and models that are separately supported by the work of neurolinguists, psycholinguists and others.

Both walking and talking can be approached with questions of *what* and *how*, both are means to an end, and both rely extensively on routines which we only become conscious of if, for some reason, they are disrupted. However, the interactive nature of language involves the use of symbols encoding the speakers' conceptualization of a 'referential situation' and instructing hearers to generate their own representation. Unlike a signal, which may be described as a (coded) instruction to behave in a certain way, a linguistic expression, or 'symbol', "directs and guides, not the behaviour of the organism receiving the signal, but their intentional stance, or minimally, their attention" (Sinha 2001: 20). Sinha uses a modified variant of Karl Bühler's Organon Model (cf. Bühler 1990: 30 ff.) to illustrate what he refers to as the three 'meta-functions' of the symbolic sign.[40] In this model, the symbolic sign reflects the speaker's conceptualization of the referential

[40] The term 'organon' comes from Greek 'organum', which means 'tool'. Sinha's version is a somewhat simplified version of Bühler's model.

situation. Consequently, it is said to represent, or 'code', the referential situation as conceived by the speaker, which is its first meta-function. As the second meta-function, it expresses the speaker's communicative intention, and as the third, it appeals to the hearer "to direct his own intentional processes towards the referential situation represented by the symbolic sign". It is emphasized as the special feature of Bühler's model that it places symbolic representation in the context of communication, which makes it both cognitive and functional.

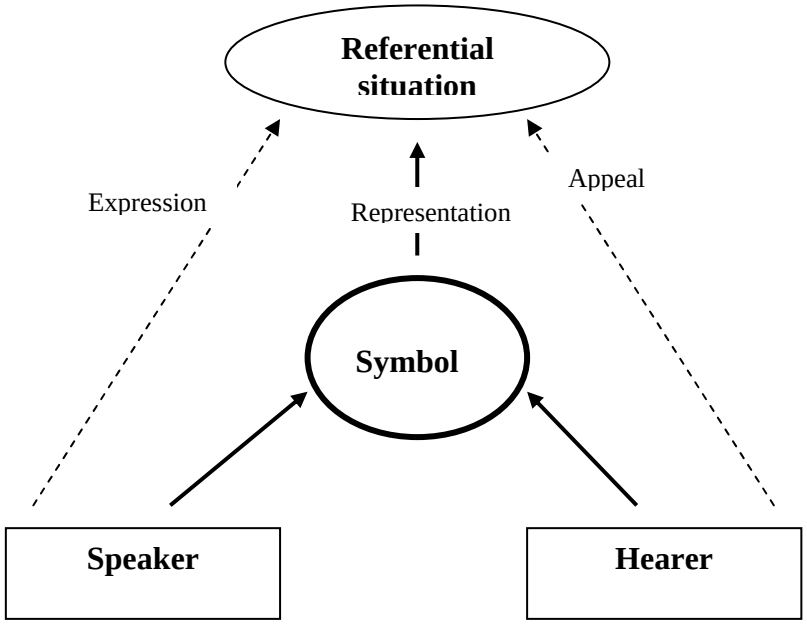


Fig. 4: A modified variant of Bühler's Organon Model (Sinha 2001: 21)
(Broken lines represent joint attention).

In the model, the relationship of representation is between the symbol and the referential situation, which is "linguistically conceptualized" by the symbol. This relationship is not likely to be exactly the same for the hearer, even if the element of joint attention directed towards a referential situation will tend to make his conceptualization converge with that of the speaker. This seems most obviously to be the case when we think of the referential situation as part of the situational context (cf. sections 1.5 and 2.1.5 above), a 'world state'

existing independently of conceptualization, but linked to it by ‘intentional relation’. Often, however, not much situational context is provided, and the efficiency of linguistic expressions, or symbols, depends entirely on their ability to ‘call on’ the cognitive capacity of the hearer. This they do by setting up actual or hypothetical frames, and “[a]pparent core cases - where frames apply directly to real-world referents - are merely a subset of a more inclusive phenomenon” (Coulson 2000: 20). The conventional status of individual words will typically allow the hearer to generate a range of different representations, and the integration with other symbols in the linguistic context is therefore an important factor.

Harder (1996a: 86, 96) describes conventions as resulting from social coordination, which “prestructures hearer expectations”, and as “aspects of life which the members of the community ‘rely on’ just as much as they rely on features of the physical environment like the solidity of the ground.” Also Firth found an ‘order of mutual expectancy’ in language, which he saw as a basic principle of syntagmatic organisation at both the syntactic and the collocational levels of meaning (cf. section 2.1.2 above), and psycholinguistic studies have found that conventional expressions seem to add to predictability because of this factor of expectation (cf. section 2.1.5 above). Harder argues that although it is not fixed once and for all what is the “coded, proper function of a word", the definition of ‘function’ should specifically exclude occasion-specific aspects. He therefore posits a functional definition of meaning that differs from Wittgenstein’s definition of meaning as usage-based: “the meaning of words lies in their *use*” (cf. 2.1.2), both by “avoiding the anti-mentalist stance” (Harder 1996a: 105), and by excluding occasional uses. I would argue, however, that a definition that does not allow for the creative exploitation of conventional expressions ignores a functional aspect which is important in communication. Just as reliance on the solidity of the ground allows people not only to walk on it, but also to skip and dance, the reliance on the common ground of linguistic convention makes it possible to achieve a special effect by sometimes *not* meeting expectations but deviating from the routine in order to surprise hearers, or simply get their attention. While it is important to distinguish between convention and ad hoc creative exploitation, the latter should be included in the

definition of what is functional about language. This would seem to be in agreement with the following line of argument (Harder 1996a: 102):

> The understanding whereby meanings are (aspects of) interactive routines may also have a behaviourist flavour in that it emphasizes the role of habit. What was wrong about behaviourism, however, was not the claim that a competent speaker has automatized linguistic routines at his disposal - the problem is whether the routines are always mechanically triggered, as assumed in behaviourism, or are at the service of a human subject who uses those routines in a creative manner.

I would say that an important function of having routines is the possibility of exploiting them creatively for communicative effect.

3.2. Methodology

First of all, I want to explain why I intend to use evidence of statistical frequency of co-occurrence as just one kind of input for qualitative analysis and not as evidence in its own right. Secondly, I will comment on how I propose to deal with the challenge of analysing linguistic data in terms of cognitive models. Finally, I will discuss how synchronic data can be used to say something about the diachronic processes of entrenchment and innovation.

3.2.1 Corpus studies and frequency

I base my analyses on corpus data, which raises the question of how the notion of 'frequency' relates to the 'tendency of lexical items to co-occur', or their 'recurrence', which is, by definition, a central feature of conventional collocations (cf. section 1.2 above). One interesting aspect is that there is often a discrepancy between people's intuitions about frequency and the frequencies recorded by means of statistical corpus tools. In the introduction to his *Dictionary of English Collocations*, Kjellmer (1994: xiv) noted that, intuitively, an essential characteristic of collocations was their recurring, repetitive quality:

> We recognize the [word] clusters as clusters, because we have heard or seen them many times. It is natural, then, that recurrence should be one of the defining factors of a collocation. At the same time, it is often striking that even collocations that are well established in the language may occur relatively seldom in sizable samples of writing.

For the purposes of the dictionary, he therefore concluded that repetitiveness as a criterion for collocationhood, in the context of a one-million-word corpus, would have to be ”repetitiveness at the lowest possible level, i.e. simple recurrence” (Kjellmer 1994: xv). Pedersen (1995: 64 f.) concluded that, for the purposes of specialized lexicography, exclusion should not be based on frequency, and that ”lexically determined co-occurrence should therefore be understood as actual occurrence in texts, i.e. frequency $\geq$ 1”. Halliday (1966: 159) pointed out that "since lexical patterns are of low generality, they appear only as properties of very large samples”. More than 30 years later larger corpora have not solved the problem, and some linguists now think it is a matter of improving statistical corpus tools. Thus Krenn (2000: 359) finds that her empirical study confirms “the weakness of the statistical measures with respect to identifying collocations from data with a high proportion of low frequency data” and concludes that ”methods need to be developed which allow collocations to be identified from low frequency data, because the vast majority of word combinations in corpora is infrequent” (Krenn 2000: 369)[41].

Does this mean that, because of the insufficient size of corpora or the lack in sophistication of corpus tools, we should rely on intuition to tell us what is frequent or recurrent? Rosch (1973) warns us not to do this, as in her studies of prototypicality she has found that people’s intuitions about frequency are not reliable at all since “the impression of a higher frequency of occurrence of prototypical members may well be a symptom of prototypicality, and not its cause”. She found degree of category membership, which was based on judgements of prototypicality, to be independent of actual frequency. This does not imply that prototypical members of a category do not recur, that they are necessarily infrequent and can never be the most frequent, nor does it imply that linguists should rely on their intuition to tell them what are the most prototypical examples of collocations,

[41] ‘Collostruction’ analysis, which investigates the interaction of collocations and constructions, is now being used by corpus linguists as a method by which corpus data may be put to use within a cognitive approach (see for example Stefanowich and Gries: 2003).

forgetting about the rest. If they do so, they are likely to construe categories in terms of criterial features, based on the prototype, an approach which leaves the rich patterning and variability of language unaccounted for.

Even if native speakers can easily activate just the appropriate sense of a polysemous noun in the context of use, they have problems when they are to account for all the senses of a word and the finer distinctions between them. This is not something that human beings naturally have to do, because it is not functional in a normal context of communication, and even lexicographers, who are the experts, cannot do it. Thus Fillmore and Atkins (1994) in a critical analysis of the entries for 'risk' in ten current one-volume monolingual dictionaries found them incomplete and inconsistent. The alternative of 'frame semantics' takes into account the cognitive models underlying word meaning (Fillmore and Atkins 1994: 370):

> [it]begins with the effort to discover and describe the conceptual framework underlying the meaning of a word, and ends with an explanation of the relationships between elements of the conceptual frame and their realizations within the linguistic structures that are grammatically built up around the word.

It is hard to see how such an enterprise could be possible without access to comprehensive corpus data.[42] In a framework that wants to be usage based and also take into account the cognitive structures of language, corpus data provide useful evidence for qualitative analysis. Such data may include evidence of the relative frequencies of co-occurrence of combinations.

3.2.2 Linguistic evidence of cognitive routines

In functional cognitive linguistics, symbols are understood as conceptualizations (1) by somebody, (2) of something, (3) for somebody. In order to be able to explain what colloca-

[42] Fillmore comments on the usefulness and limitations of corpus data in the article: "Corpus linguistics" or "computer-aided armchair linguistics" (Fillmore 1992).

tions do and how they do it, it is necessary to make assumptions about the nature of the cognitive models that the symbols 'call on'. These models draw on research in psychology, neuroscience and philosophy, which apply their own methodologies to aspects of language (cf. section 3.1 above). They are not rules to be imposed on linguistic data, but, like theoretical notions generally, they are tools that can be used to make sense also of linguistic data. As such they are subject to constant reappraisal and revision. Thus the models are symbols themselves, used for conceptualizing the areas of cognition that Lakoff and Johnson (1999) refer to as 'the cognitive unconscious', because they are not directly accessible and have to be 'teased out'. Like other abstract domains, 'the cognitive unconscious' is typically conceptualized in metaphorical terms like 'gestalt', 'schema', 'frame', and 'space'. This makes it possible to discuss aspects of the cognitive domain in terms of the structure of more concrete domains, which can be used to draw inferences about the data that the models are applied to (cf. section 3.4.1.1 below).

Even if the metaphors are apt, not everything from the more concrete 'source domains', from which the image-schematic structure is recruited, maps onto the more abstract 'target domains' that it is used to discuss (cf. sections 3.4.1.1 and 3.4.2.4 below). The trap to be avoided is taking these metaphorical conceptualizations too literally, for instance by referring to schematic networks as if they were reified static structures using the brain as a storehouse. Instead, schemas should be seen as representing dynamic processing activity (Rumelhart, McClelland and the PDP Research Group; 1986, in Sinha 2001: 7):

> There is no representational object which is a schema. Rather, schemata emerge at the moment that they are needed from the interaction of large numbers of much simpler elements working in concert with one another. Schemata are not explicit entities, but rather are implicit in our knowledge and are created by the very environment that they are trying to interpret - as it is interpreting them.

On the one hand, researchers using purely linguistic data should be careful not to draw conclusions that would require psycholinguistic or neurolinguistic methodologies. On the other hand, it is not possible to make sense of linguistic data without assuming that cognition plays a central role, so the solution would seem to be for the linguist to make a

clear distinction between what can be said on the basis of the linguistic evidence and what needs to be assumed for the purpose of interpreting the data.

3.2.3 Synchronic evidence of diachronic processes

The continuity not only between body and mind, but also between language use and language system is central to cognitive linguistics, which implies a basically diachronic approach. In this study I will, however, confine myself to the synchronic evidence of entrenchment and innovation. Since earlier meanings of lexical items coexist with later extensions, current language variation will reflect diachronic processes, even if it is not likely to disclose their etymology, as some meanings will have gone out of use. Creative collocations used on a specific occasion for expressive effect or for the solution of a problem of conceptualization, may end up being entrenched. I see them as related to entrenched collocations, because they can only fulfil their communicative function if they are understood as exploitations of that category. An example is *rancid trousers*, which is likely to evoke a conventional collocation like *rancid butter*. All in all, I would thus like to claim that since synchronic corpus evidence reflects past language development as well as ongoing change, it is compatible with a diachronic viewpoint .

3.3 Introduction to the empirical part

A critical review of the framework used to categorize collocations in studies of phraseology was carried out in Part 2. I will now present and motivate my research questions. After that, I will briefly describe the source and nature of the data to be analysed in section 3.4.

3.3.1 Research questions and motivation

The questions address the issues claimed in Part 1 (section 1.5) to be central aspects of what a theory of collocations should account for: the internal structure of component items and their mode of integration as well as patterns of entrenchment and innovation. Since evidence from one study cannot be expected to be conclusive, the focus is on the suitability of the methodology. The aim is to show that the questions are relevant and make for a plausible hypothesis about what elements are required in a functionally and cognitively based framework for the study of entrenched verb + nominal object collocations. The general assumption is that entrenched collocations are speech routines allowing speakers to guide hearers by evoking cognitive routines associated with familiar semantic frames. As linguistic expressions, their function is to further the reproduction, through renewal of connection, of the contexts of situation and the underlying cognitive models, or semantic frames, to which they belong. The following four claims will be tested:

1) The contexts of situation to which entrenched collocations contribute and their underlying frames can be identified by analysing the internal structure of component items and their mode of integration. In this entrenched collocations do not differ from other composite structures.

2) In entrenched collocations consisting of a verb and a nominal object, the noun evokes the dominant frame while the verb profiles a specific aspect of the frame. In collocations that are not entrenched, it is the verb that evokes the dominant frame.

3) Entrenched collocations can be characterized in terms of prototypicality that varies with the schematicity of the verb and its salience in the frame evoked by the noun.

4) The verb in entrenched collocations has a functional, grammaticalized, role as support verb.

3.3.2 Design of case study

To test the claims stated above, I propose to carry out a corpus study of *break an appointment* as an example of an entrenched verb + nominal object collocation. More specifically, *appointment* will serve as an example of the group of 'deverbal' nouns that are derived from verbs by suffixation. Other examples are *agreement, regulation, lighting,* and *recep-*

tion. Typically deverbal nouns, which also include nouns like *grip, record,* and *strike*, can denote action as well as the product or result of an action. *Break* has been chosen as an example of the group of verbs that are used both in the sensorimotor and in the nonsensorimotor domain. *Break* is one of the very frequent words in English, which Sinclair has indicated may have less of a clear and independent meaning than less frequent ones (cf. section 2.2.5 above), and it is one of the verbs which characterize action at the basic level, which is said to be cognitively and linguistically the most salient (cf. section 1.5 above).

The first claim will be tested by first analysing *break* and *appointment* separately, as complex, polysemous categories and by subsequently analysing the composite structure *break an appointment*. The composite structure will be analysed in terms of the autonomy/dependence alignment of verb and noun with a view to testing the second claim, that in entrenched collocations it is the noun that evokes the frame while the verb profiles a certain aspect of it. The third claim, that entrenched collocations can be characterized in terms of prototypicality, will be tested by an analysis that relates schematicity to basic level categorization at the place of the verb. Finally, the findings relating to entrenchment will form the basis for a discussion of the fourth claim about a possible role for *break* as a support verb.

3.3.3 A general presentation of the data

The source of the data on the two lexical items *break* and *appointment* is the British National Corpus, which includes over 100 million (100,106,008) words of contemporary English and consists of 90% written and 10% spoken text. I have used a version made available online for research at the University of Southern Denmark.[43] This version, which

[43] The search engine, at http://corp.hum.sdu.dk/corpustop.en.html, was developed by Eckhard Bick, the project leader of VISL, which stands for "Visual Interactive Syntax Learning". This is a research and development project at the Institute of Language and Communication, University of Southern Denmark, Odense Campus.

is not tagged,[44] recognizes regular expressions (Dienhart and Kasch 2000), it is sentence based, and concordances show the target in boldface with 80 letters of context and with a source identification code (text segment, text type, paragraph, and sentence number). There are options to view either full sentences or more running context (approximately 175 words) in a special 'context window'.

In the case of *break* a search was conducted for 1000 concordances, targeting all the forms of the verb: *break, breaks, breaking, broke,* and *broken*.[45] This represents approximately 9.5% of the total of 10,494 occurrences in the BNC. In the case of *appointment*, the search included the singular and plural forms, and all of the 908 occurrences found are included in the data. In both cases, the context was increased to full sentences, which were then imported into two separate data bases,[46] one for *break* and one for *appointment*. The possibility of viewing more running context has been used when required for purposes of analysis. Halliday's claim that "lexical patterns are of low generality" is confirmed also by this study. Thus, in a corpus with over 100 million words, only five examples were found of forms of *break* in combinations with *appointment* - not much to go by in terms of numbers.

The purpose of the following analyses, then, is to look for the patterns in which this entrenched collocation has its place and for principles that may account for the patterns. What is required of the data in this perspective is that they should make it possible to detect such patterns and principles.

---

[44] After I collected my data, at tagged concordancer , also developed by Eckhard Bick, has been made available at the VISL site.

[45] The search string, using regular expressions, was: _break._|_broke_|_broken_|_breaking_ , where _ stands for a space, the full stop stands for any character and exactly one, and the vertical bar separates expressions, allowing a search for alternative patterns (Dienhart and Kasch 2000: 7).

[46] The programme used is File Maker Pro 4.1.

3.4 Case study: *break an appointment*

This section contains the analyses outlined in section 3.3.2 above, presenting the corpus evidence and interpreting it according to the principles outlined in section 3.4.1. Like all other composite items, collocations need to be understood as integrations of their component items. The first two analyses therefore target *break* and *appointment* separately as **complex categories** with the aim of accounting for their internal, polysemous structure (sections 3.4.2 and 3.4.3). The position of each item is examined in its systemic context, which in turn will make it possible to account for the way in which they are integrated syntagmatically (cf. sections 1.5 and 2.1.5, Table 5). The composite structure is subsequently analysed in section 3.4.4.

3.4.1 How to approach the analysis of a complex category

Before launching into analysis, I will discuss how I propose to go about it. In section 3.4.1.1, I will motivate an approach that combines a cognitive understanding of polysemy with the schematic network model, basic level categorization and semantic frames, and in section 3.4.1.2, I will outline how analyses in terms of domain, image schematic structure and construction type will be used to identify the readings that can plausibly be posited on the basis of the data.

3.4.1.1 How many meanings does a word have?[47]

At one extreme, it could be claimed that words have innumerable meanings, since no two situations of use are ever quite the same. At the other extreme, the one form/one meaning hypothesis claims that a lexical item has only one unitary sense that can be called

[47] This is the title of Taylor's 1992 article, which has been one of the sources of inspiration for this section.

'semantic' and that conceptualization, which involves interpretation against background knowledge of the world, is not part of linguistics proper (cf. section 2.2.3). As a middle way between rampant polysemy and isomorphism, Taylor (1992) proposes Langacker's **network model** (Langacker 1988: 51 f.) in which all nodes represent "acts of categorisation by the language user" based on his judgements of similarity. Such networks grow 'from the bottom up' as language users apply a term to new real world items or situations. To the extent that specific meanings are perceived as 'similar', a more abstract sense may emerge, which is said to be 'schematic' for the more specific senses, which in turn elaborate the schema. New meanings that conflict with the schema in certain respects are seen as extensions from it, which may become a permanent part of the network through conventionalization. Which instantiations of a schema and which extensions from it are "conventionally sanctioned" by the language is not predictable; it is something language users have to learn (Taylor 1992: 150)[48].

Judgements about similarity decide whether a given lexical category is seen as monosemous or polysemous. If the lexical category *bird* is often categorized as **monosemous**, or '**monocentric'**, it is because all members are perceived as similar to a single prototype, and if the category *school* is categorized as **polysemous**, it is because it is judged to be '**polycentric'** with members being similar to one of several related prototypes (such as a 'school for educating children', a 'school which is a division of a university', and 'school as an intellectual trend'), which are related through family resemblances (Taylor 1995: 99). In practice, the line can be hard to draw, and a number of tests have been suggested to help decide whether a lexical item is '**ambiguous'**, which indicates polysemy, or merely '**vague'**, in which case it is said to be monosemous. If a statement such as *I don't want a pig in the house,* has two readings which are not compatible ('pig as a farm animal' and 'pig as a person with messy eating habits') so that one must be discarded, we have a case of

[48] I suggest that 'schematicity' as used in this context may be characterized as 'semasiological', because it refers to the general meaning abstracted from the range of entities and situations that may be named by a given linguistic expression (cf. Taylor 1995: 261-62). By contrast, schematicity within lexical sets may be said to be 'onomasiological', because it refers to the level of generality of a range of expressions used to evoke specific substructure of a semantic frame (cf. section 3.4.3.2 below).

ambiguity caused by the polysemy of *pig*. A sentence like *There is a bird in the garden* is vague with respect to what kind of bird was observed, as long as we remain in the same **domain**, and *bird* is therefore sometimes given as an example of a monosemous word (Taylor 1995: 101). However, it is ambiguous in contexts where it might refer either to a person or a bird as, again, one meaning has to be discarded. If different uses of a lexical item, such as *pig* or *bird* belong to different domains, this is a strong indication of polysemy, but an item can also be polysemous if it is structured by different schemas within the same domain. Thus *a high building* and *a high ceiling* both belong to the domain of vertical space, but the former is about the vertical extent of an entity and the latter about the position of an entity in vertical space (Taylor 1995: 100). It should be noted that the notion of domain is a flexible concept, which can be given a very broad interpretation, as when it is used about the sensorimotor domain, or a very narrow one, as when the word *knuckle* is characterized as belonging to the domain of 'finger'. The definition given by Langacker (1987: 147) is a "context for the characterization of a semantic unit" and domains are characterized as being "necessarily cognitive entities".

As a general test, it is assumed that if an **anaphoric expression** like "do so too" requires that a choice be made between different senses of an item, we have a case of ambiguity, and the item is polysemous. Langacker (1999b: 126) points out, however, that judgements of ambiguity and vagueness are often graded: if only the specific senses are entrenched, we have a clear-cut case of ambiguity, and if only the schematic meaning is entrenched, we have a clear case of vagueness. In between there may be cases in which two meanings are sufficiently similar for a neutral value to emerge that can be used for anaphoric purposes: *Bill has been painting, and so has Jane* would normally be taken to mean that both had been doing either a utilitarian or an artistic job (i.e. *paint* is ambiguous), but in some usage situations the two types of work might be so closely related that the statement could be used even if Bill and Jane were engaged in activities that, strictly speaking, do not belong to the same category; i.e. *paint* becomes vague rather than ambiguous with regard to the specific nature of the activity (Langacker 1999b: 126).

A related type of test involves the rhetorical figure '**zeugma**' (or 'syllepsis'), which the OED defines as "a figure by which a single word is made to refer to two or more words in a sentence; esp. when properly applying in sense to only one of them, or applying to them in different senses […]". An example from section 2.1.5 above can be used to illustrate this:

- […] *she took the point, but she also took the spoons.*

Although the verb is repeated, co-ordination with *also* creates a zeugmatic effect. Using zeugma for expressive purposes is generally referred to as 'punning', which the OED defines as "The use of a word in such a way as to suggest two or more meanings or different associations, or the use of two or more words of the same or nearly the same sound with different meanings, so as to produce a humorous effect; a play on words". Furthermore, punning provides evidence of polysemy since it shows that the sender of a message considers a word to be ambiguous and expects the receiver to do so, too. In terms of frame semantics, one would say that the word is used to evoke two different frames at the same time; this creates an expressive effect as the receiver has to adjust his expectations. In many cases it seems that word meaning varies according to syntagmatic context, whereas in other cases it remains constant. By means of a **commutation** test (cf. section 2.1.2 above) in which one component of a composite expression is held constant while the other is changed, it is possible to elicit judgements about polysemy (Taylor 1992: 133). The question is whether any change in reference should be construed as evidence of polysemy. Lakoff (1987: 426) offers *window* as an example of polysemy, because it can refer "either to an opening in the wall or to the glass-filled frame in that opening". Following this approach, it could even be argued that *window* is four ways ambiguous as it can also refer specifically to the window-pane or the window frame.

1. *to brick up a window*	('opening in the wall')
2. *to put in a window*	('glass-filled frame')
3. *to paint a window*	('window frame')
4. *to break a window*	('window-pane')

Fig. 5: Different meanings or different active zones

An alternative construal would be that the four examples are cases of substitution that do not amount to commutation, but represent different '**active zones**' of the same sense of *window*. According to Langacker (1999b: 62 f.), we often find a kind of metonymy where the term for the whole (*window*) is used instead of the more precise term for the part (e.g. *window-pane*). What is said to be involved is "a conflict between two competing desiderata: that of being precise and accurate in regard to which entities actually participate in the profiled interaction; and that of focusing attention on entities that are inherently salient or of primary interest". Often it is the more salient 'whole' that is profiled as landmark or trajector rather than the subpart which is directly involved: the active zone. In such cases it seems plausible to say that the meaning of *window* is vague rather than ambiguous, also because all expressions belong to the same semantic frame. By contrast, cases of ambiguity would involve the choice between a literal and a metaphorical interpretation in examples like *to throw something out of the window*, which involves different domains.

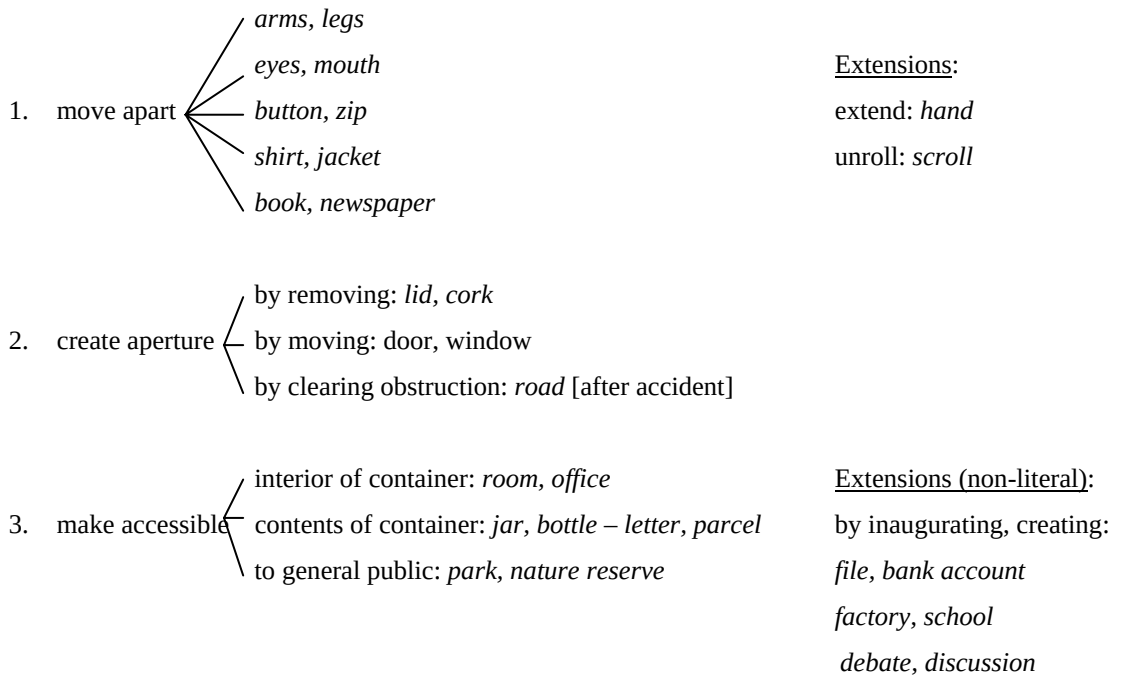


Fig. 6: A partial network for *open* (based on Taylor 1992: 152)

What if it is the noun that is substituted? Taylor (1992: 152) uses the example of *open* in a range of different combinations that he construes as instantiations of three 'high-level' schemas. They are not construed as parts of a 'superschema' but as linked by relations of extension, each being elaborated by a range of combinations, and there are overlaps between them; thus *open one's mouth* can be seen as an instantiation of all three. Taylor's network (cf. Fig. 6 above) is offered as a hypothesis based on his intuition and specifically not as a definite account, but it illustrates how distinct meanings can be seen as instantiations of more abstract meanings on the basis of perceived commonality and how these abstract meanings may in turn be linked through 'similarity of association'. Thus *open the door* and *open the office* may refer to the same situation, but they focus on different aspects and may be said to be related through metonymy (Taylor 1992: 150). The question is whether we should posit

1. one meaning for the entire network
2. a different meaning for each of the three 'high-level' schemas (plus extra meanings for extensions)
3. a different meaning for each of the subschemas (e.g. *open one's arms/legs*; *open one's eyes/mouth*; *open a book/newspaper*)
4. a different meaning for each individual combination.

In other words, where should we draw the line dividing the intensional level, the level of senses, from the referential level, the range of application of a word in a given sense (cf. section 3.4.1.2 below)? The commutation test would help us discard the **first alternative** in many cases, as substituting a member of one schema for an item from one of the others would seem to change the meaning of *open*: e.g. *open one's arms* vs. *open the door* vs. *open the room*. However, it would not work so well where items fit into more schemas. As pointed out by Taylor, *mouth* would fit into all three, and it seems that *road* would fit into the 'make accessible' schema as well as the 'create aperture' schema so that going from *open the road* to *open the park* might not be seen as changing the meaning of *open*. As regards the **second alternative**, positing one meaning for each high-level schema, it also

seems that substituting members of subschemas for each other would generally involve commutation. However, this is more clearly the case in some cases than in others: going from *opening one's arms* to *opening a shirt* in the 'move-apart' schema seems to change the meaning of *open*, whereas going from *opening a room* to *opening a park* in the 'make accessible' schema does not obviously do so. Also in the case of the **third alternative**, there is some vacillation, as substitution within a subschema sometimes does seem to change the meaning of *open*, as in *opening a jar* vs. *opening a letter* in the 'make accessible' schema, while in other cases it seems not to do so, as in *opening a door/window*, or in *opening a lid/cork* in the 'create aperture' schema. I will come back to the **fourth alternative** in connection with the discussion of basic level categorization and semantic frames below.

"Other linguists, working on their intuitions, might possibly come up with different proposals", as pointed out by Taylor (1992: 153), who goes on to mention various kinds of evidence that can be used to test the hypothesized structure of a network, including diachronic data and cross-language data, as well as data elicited from speakers submitting their subjective judgements of meaning similarity. This would produce a more broadly based hypothesis, but any typology imposed on a lexical category is bound to remain hypothetical, and the problems involved in the struggles of categorization merely confirm that such a category, which has developed over time based on people's associations and judgements of similarity, cannot be reduced to a classical category. The various tests mentioned above only confirm this. In the last analysis, "the tests rely for their success, on the very fine judgements which they were designed to replace" (Taylor 1995: 102). The advantage of using the network as an explanatory model is that it can accommodate the fact that it is not possible to draw clear boundaries between different senses of a word, and that new senses may develop over time as "a non-central member of a monosemous category increases in salience to the point where it constitutes a secondary conceptual centre of the category" (Taylor 1995: 103).

Rather than worrying about the fuzzy boundaries of meanings, cognitive linguists will focus on **the basic level of categorization** in the middle of taxonomic hierarchies (cf. sections 1.5 and 2.1.5 above). Taylor (1992: 154 ff.) proposes that this is the most salient level and the level at which meanings of lexical items are accessed – at least "in default cases", that is if no specific contextual information is available. Examples from the lexical category *open* are *open a window*; *open a book*; *open a bottle*; *open a parcel* and *open a discussion*, which occupy an intermediate level of abstraction ignoring distinctions, for example between the ways in which different kinds of window and different kinds of bottle are opened. According to Taylor, this is the highest level of abstraction at which people can form mental images of a category, while more schematic senses "lack salience, and are not readily activated in speech production and reception, nor are they easily available to introspection" (Taylor 1992: 156). This seems to imply that we are to posit a different meaning for each combination that relational predications like *open* and *break* enter into (the fourth alternative listed above), although it does not prevent us from seeing meanings as related. "Indeed, the very essence of polysemy (as opposed to chance homonymy) is traditionally said to reside in the relatedness of the separate meanings." (Taylor 1992: 146).

One reason for the cognitive salience of basic level categories is said to be that they are perceived in terms of their **gestalt** or overall part-whole configuration (cf. section 2.1.5 above). What this means becomes clear if word meaning is understood against the background of semantic frames, which I have argued are compatible with the Firthian notion of context of situation (section 2.2.2 above). Fillmore's explanation of **semantic frames** (1985: 224) also includes the notion of gestalt:

> Borrowing from the language of gestalt psychology we could say that the assumed background of knowledge and practices - the complex frame behind this vocabulary domain [weekday names] - stands as a common ground to the figure representable by any of the individual words. Such a frame represents the particular organization of knowledge which stands as a prerequisite to our ability to understand the meanings of the associated words.

Focusing on the basic level of a lexical category thus offers the possibility of linking the question of how many readings a word should be construed as having to the number of

semantic frames in which that word participates. However, for basic level verbs like *open* and *break*, this would mean that a very large number of readings would have to be posited. For practical reasons, it is therefore preferable to posit readings at a more schematic level, which is compatible with a network model of lexical categorization (cf. Fig. 6 above). For nouns the situation is different. They typically participate in a limited number of frames, and readings can therefore be directly related to the semantic frames that they evoke. Fillmore and his associates in the FrameNet Project have found that frames "can be evoked by words in any of the major lexical categories: noun, verb, adjective or preposition" (Johnson et al 2003, section 2.2).[49] In the category of nouns, it has been found that **'event nouns'** like *withdrawal* and *replacement* and **'relational nouns'** like *brother* and *girlfriend* are "most clearly frame-bearing" (Johnson et al 2003, section 2.3). In the expression *take revenge*, the **'dominant frame'** is evoked by the event noun rather than the verb, which is therefore characterized as having a support function. Such verbs, which are about the same event or state as the noun, are treated as **'support verbs'**, which

- turn a target event or state noun into a verb phrase-like predicate
- allow for the expression of a Frame Element as their subject
- are semantically neutral (to the degree that that is possible)

I would argue that this category should include not only verbs that are 'semantically neutral', but also other verbs belonging to frames evoked by event nouns; that is not only *make a promise*, but also *break a promise*, not only *have/hold a debate*, but also *open/close a debate*. Arguably, it might even be extended to include verbs in frames evoked by nouns denoting artefacts and natural things such as *drill* in *drill a hole*, *play* in *play the piano*, and *hold* in *hold hands*, which can also be said to turn a target noun into a verb phrase-like predicate. Distinguishing between frames evoked by the verb and frames evoked by the noun would moreover be a means of distinguishing between different readings of a verb.

[49] In an earlier edition (June 24, 2003), the formulation gave stronger emphasis to the role of the verb: "In principle, members of all three major lexical semantic categories (verb, noun, adjective) can be frame bearing, that is to evoke a semantic frame. However, the most prominent semantic frame evoked in a particular sentence is usually one evoked by a verb."

- *she **held Patrick** against her shoulder* (frame evoked by the verb: *hold*)
- *he wants to **hold meetings** with opposition figures* (frame evoked by the noun: *meeting*)
- *Peru **held elections** for a new Congress* (frame evoked by the noun: *election*)
- *he no longer **holds office*** (frame evoked by the noun: *office*)
- *and **held hands** for a circle dance* (frame evoked by the noun: *hands*)

Fig. 7: Full-verb vs. support-verb readings of *hold*[50]

As a minimum, one full-verb reading (the first example in Fig. 7) and one support-verb reading (all the others) can be distinguished. As an alternative, a support-verb reading would be taken to apply only in combinations with event nouns (*meeting* and *election*) and separate full-verb readings would have to be posited for *hold office* and *hold hands.*

This discussion will be continued in section 3.4.4.4 below. For now I will conclude that an approach combining the network model with the semantic frame model seems to be suitable for the analysis of lexical categories, since it takes into consideration the cognitive salience of the basic level of categorization and, at the same time, allows for the more general statements of meaning of higher-level schemas.

3.4.1.2 Domains, image schemas and construction types

One way of accessing the wide range of meanings expressed by a lexical item like *break* is to ask what conceptual content is expressed by individual instantiations and to group them according to abstract **'domains'**, which means to record the referential situations (with real-world situations as a subset) that are coded by these examples of usage (cf. sections 3.1 and 3.4.1.1 above). As pointed out by Harder (1996: 49), "information is always information-about and thus involves a two-level element: a representation and something else which it

[50] The examples are from the Cobuild Corpus Concordance Sampler.

represents". 'Aboutness' denotes the pre-linguistic 'substance'[51] from which we "carve out linguistic content elements" (Harder 1996b: 439), and as pointed out by Langacker (1999b: 5), "… we are able to construe the same content in alternate ways, resulting in substantially different meanings". Even at the pre-linguistic stage, the content or 'substance' is cognitive: it refers to real world situations (situational context) as perceived by somebody, or to semantic frames (cognitive context) abstracted from typical contexts of situation (cf. sections 1.5 and 2.1.5, Table 5 above).

An important element of construal involves **image-schematic structure**,[52] which is grounded in physical experience and is projected metaphorically to structure areas of less tangible physical experience as well as areas of abstract experience that do not have any gestalt of their own (as in *the sun broke the clouds*; *he broke her heart* - cf. section 3.4.2 below). I find that distinguishing only between sensorimotor domains as source domains, on the one hand, and domains of 'subjective experience' as target domains (as in Lakoff and Johnson 1999: 58), on the other hand, downplays the fact that image-schematic

	LITERAL CONCEPTUALIZATION	**METAPHORICAL CONCEPTUALIZATION**
SENSORIMOTOR DOMAIN	*similar colours* *start sweating*	*close colours* *break into a sweat*
NONSENSORIMOTOR DOMAIN	*understand something*	*grasp/ see something*

Table 8. Literal and metaphorical conceptualization in different domains[53]

[51] The notion of 'substance' for the content side is taken from Hjelmslev (1953: 52), who used the notion of 'form' for the expression side. Harder prefers 'structure' or 'content-form' as less ambiguous terms for the expression side (Harder 1996b: 439).

[52] Mark Turner 1993: 304 refers to 'image-schema' as "Mark Johnson's term", and renders his definition as "a recurring, dynamic pattern of our perceptual interactions and motor-programs that give coherence and structure to our experience" (Johnson, Mark, 1987: *The body in the mind*. Chicago: University of Chicago Press.

[53] Lakoff and Johnson (1999: 58) give the example contrasting *"These colors are similar"* and *"These colors are close"*.

structure also plays an important role within the sensorimotor domain (as in the expression *break into a sweat*).[54] As pointed out in Lakoff and Johnson (1999: 58), 'basic sensorimotor concepts', exemplified by *grasp* in the sense of 'holding' rather than 'understanding', tend to be literal, but I find it important to stress that the dividing line between literal and metaphorical conceptualization does not coincide with the division between the sensorimotor and the nonsensorimotor domains. Often experience in both domains can be construed literally[55] as well as metaphorically, as shown in Table 8 above. As a general principle, I therefore suggest that, for purposes of analysis, domain and image-schematic structure should be treated as separate but related aspects.

As a further aspect of construal, I will include **construction type,**[56] including (1) active and (2) passive transitive constructions, (3) intransitive constructions, (4) constructions with the past participle as premodifier and (5) postmodifier, as well as (6) constructions in which *break* is nominalized. Like image-schematic structure, the choice of construction type provides clues as to *how* content is conceptualized, for instance in terms of 'figure/ground organization' (Langacker 1987: 120 f.) and in terms of 'sequential scanning', which follows the temporal evolution of a process (as if it was a film), versus 'summary scanning' which construes it in a holistic fashion (as if it was a photo) (Langacker 1987: 144; 1999b: 309). Including construction type in the analysis is also motivated by the finding by corpus linguists that grammatical and lexical choices tend to be closely related (cf. section 2.1.2 above).

[54] The expression exemplifies the 'Location Event-structure Metaphor' which draws on the 'primary metaphors' STATES ARE LOCATIONS and CHANGES ARE MOVEMENTS (Lakoff and Johnson 1999: 179); cf. section 3.4.2.2 below.

[55] The distinction is not absolute - if we construe an expression as literal, it may simply be because we are not aware of its metaphorical origin.

[56] In Langacker's cognitive grammar, a grammatical construction is defined as "[a] symbolic structure involving the syntagmatic combination of morphemes and/or larger expressions. The construction consists of a set of component structures, their mode of integration, and the composite structure resulting from this integration" (Langacker 1987: 489).

The approach is **semasiological** in that it starts from 1000 examples of the lexical item *break* and asks what domains they represent. It may also be seen as **onomasiological** to the extent that it goes on to ask how image-schematic structure and construction type are used for conceptualizing content (Geeraerts 1997: 17). Following Geeraerts, I understand senses to be '**readings',** which do not imply a strict dichotomy between semantic and encyclopedic information or a strict separation between the **'intensional level',** the level of senses or the '**definitional'** level, and the **'extensional'** or **'referential level',** the range of application of a word as used in a specific sense (1997: 17 f.). Identifying a given number of meanings for a word is in the last instance a matter of construal. As indicated above, I see the analysis of the internal structure of lexical categories as eventually involving the identification of the **semantic frames** (cf. sections 1.5; 2.2.2 and 3.4.1.1) against which individual concepts are understood. However, it is beyond the scope of this study to describe all the frames in which *break* participates; rather the readings to be identified will be abstractions from the meaning of *break* in related frames. In the case of *appointment*, the number of frames is expected to be more manageable, so that it will be possible to describe each reading as a semantic frame.

3.4.2 The internal structure of BREAK

Taking on this basic-level and multi-purpose verb means to engage the container metaphor as well as the 'correspondence theory of truth' according to which a statement is held to be true "when it fits the way things are in the world" (Lakoff and Johnson 1999: 98). Breaking objects and breaking arms and legs may seem very 'literal' and easy to distinguish from each other and from the breaking of hearts and promises, but when less central readings (at the intensional as well as the referential level) are taken into account, a more complex picture emerges, and it becomes clear that literal meanings are based on conceptualization as much as metaphorical ones, and that the use of metaphor is also found in the sensorimotor domain.

As mentioned in section 3.3.3 above, the data for *break* are the result of a search specifying 1000 concordances of the forms *break, breaks, broke, broken* and *breaking*. The concordances have been expanded to full sentences, which have subsequently been imported into a data base. When necessary, the facility to view more running context has been used to resolve ambiguities. Sentences have been deleted from the data in a limited number of cases, if they were found to occur a second time in identical form; on the other hand, sentences including two constructions with *break*, have been duplicated. After revisions, the data base contains 993 records. Each record has been coded with the following information on *break*: form class (transitive/intransitive verb, phrasal verb, noun or adjective), verb form (infinitive, finite/infinite, active/passive, present/past participle) and function (predicator, premodifier/ postmodifier; head/dependent). In addition the nominal object or subject has been recorded, and coded corresponding to the reading which the example is construed as exemplifying.

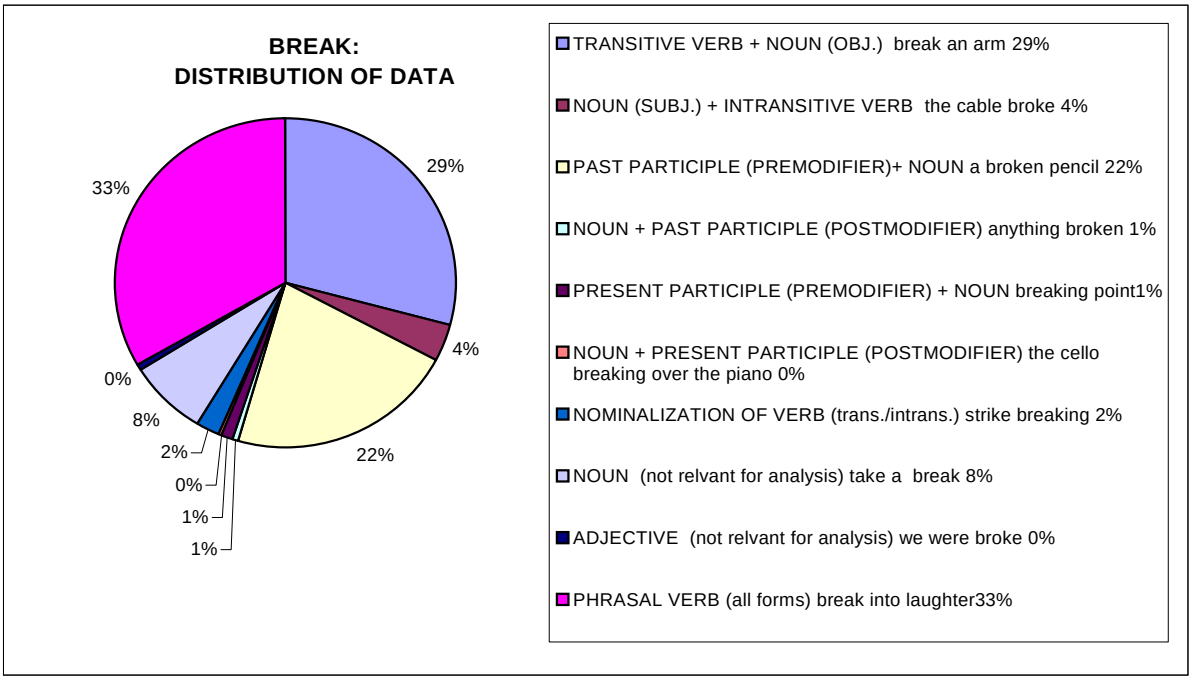


Fig. 8: Distribution of the data by construction type

The construction transitive verb + nominal object accounts for 29% of the 993 examples, whereas *break* as an intransitive verb makes up only 4%. The construction in which the past participle *broken* is in adjectival position premodifying the noun is found in 22% of the examples, while *broken* as postmodifier accounts for only 1%. Phrasal verbs[57] account for 33%, which is the largest share. Present participle forms functioning as pre- and postmodifiers make up just over 1% in all. Nominal forms, *break* and *breaking*, account for roughly 10%, but only those which are nominalizations of transitive or intransitive constructions (2%) are considered relevant for this study. This includes those examples that mention the entity that is 'breaking' or 'broken' after the present participle form as in *breaking the weak link*, *the breaking of the day*; *the breaking of a bond*, as well as those examples in which the entity precedes the nominal forms in compounds like *law breaking* and *cable breaks*. Most of the examples of *break* or *breaking* as a noun are not considered relevant for this study (8%) as they cannot be directly related to the transitive or intransitive uses. They include examples like *breaking and entering*, *maternity break*, *tea break*, and *autumn break*. Finally, the material includes four examples of the adjective *broke*, as in *to be* or *go broke*, which are not included in the analysis. Altogether 83 examples out of the 993 in the data base have been excluded, leaving 910 to be used for purposes of analysis.

3.4.2.1 Abstract domains and referential range

The data show that *break* is used to categorize a wide range of experience in the sensorimotor domain as well as in other domains, which will be referred to collectively as

[57] In addition to verb + adverb combinations (*break up/down*, etc.), the term 'phrasal verbs' here includes a few combinations of verbs and preposition (*break into*), and a few verb + adjective combinations (*break even/free/open*).

the 'nonsensorimotor domain'.[58] The network of domains and subdomains charted in Fig. 9 below is a first hypothesis about the internal structure of the lexical category BREAK, based on the construal of the data in terms of 'aboutness'. The aim is to describe the network in terms of image-schematic structure and construction type before discussing when two nodes in a network are so far apart that it is plausible to construe them as different readings (in section 3.4.2.4 below). Such a hypothesis can never be turned into a definite account, because it is bound to rely partly on intuitive judgements and is moreover based on a given amount of linguistic data. Eventually, it should be tested by a psycho-linguistic study to see if these judgements are in line with the intuitive judgements of native speakers.

Both the sensorimotor domain and the nonsensorimotor domain show substantial internal complexity, and it is not possible to define subdomains in terms of criterial features. Rather, the internal structure of break shows a recurring pattern of prototypicality at all levels as well as overlaps between the subdomains. The rather clumsy terms used to refer to sub-domains are a result of the frustrating struggle to label categories without clear boundaries. Thus *breaking a branch* has been categorized as belonging to the domain of 'artefacts and natural things', whereas *breaking an arm* has been characterized as belonging to the domain of 'body parts', which also consists of 'natural things', and both these domains can be said to be part of 'the physical environment', which also includes experience of 'physical activity'. Such overlaps are to be expected, since it is assumed that the network of meanings categorized by *break* has gradually grown up from below by processes of extension and elaboration motivated by judgements of similarity with familiar uses.

Fig. 9 below shows the domains of sensorimotor as well as nonsensorimotor experience categorized by the examples of *break* that occur in the data. The intention is to illustrate the internal complexity of conceptualizations in each domain and how the internal structure

[58] Lakoff and Johnson (1999: 46) contrast domains of sensorimotor experience with 'domains of subjective experience', but since the sensorimotor domain is also subjectively construed, I use the expression 'nonsensorimotor domain' instead.

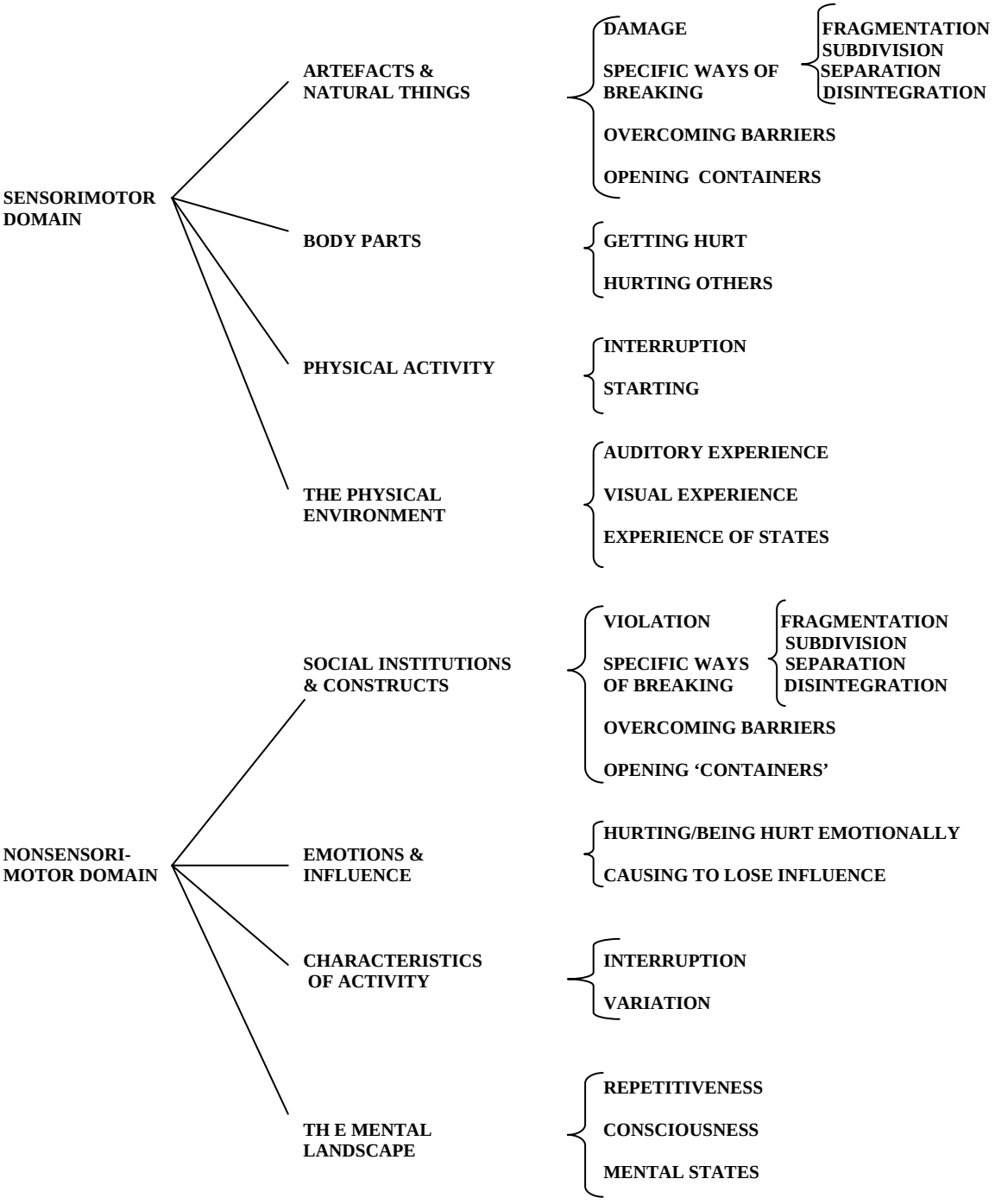


Fig. 9: BREAK domains[59]

[59] Examples to illustrate the subdomains are given in Tables 9 and 10 below.

of the nonsensorimotor domain mirrors that of the sensorimotor domain. In the **sensorimotor domain** the lexical item *break* is used to categorize aspects of our experience of moving about in the world, using our body and our senses, observing and interacting with natural things and artefacts, and with our physical environment. In the **nonsensorimotor domain**, *break*, by metaphorical extension, categorizes experience that does not involve our senses directly, but seems to be conceived of in terms of experience that does.

Within the sensorimotor domain, I would claim that the subdomain of **'artefacts and natural things'** is more 'basic'[60] than the other subdomains, since it can be seen as the source domain not only for nonsensorimotor domains, but also for less central sensorimotor domains ('body parts' and 'physical activity', as well as 'auditory and visual experience' of the physical environment).

- **'oblong things' that break like sticks of wood:**
  branches, chair-legs, frames, masts, pencils, pipes, rifles, rigs, stems, stakes, twigs, wickets
- **'chains or rope-like things' that break through strain or may be severed:**
  cables, chains, connections, elastic, links, ropes, straps, washing lines
- **'brittle things' that splinter or crack like glass or pottery:**
  bottles, cups, eggs, glass, glasses, glass doors, light bulbs, mirrors, pots, records, windows, a WC
- **'stones or stone-like things' that crack or split:**
  brick, churches, flintstone, pavements, pediments, rock, slabs of marble, stone, tile, the wings of a stone dragon
- **'simple functional things' that may break in different ways, related to materials and design:**
  boots, boxes, chairs, doors, drains, fences, furniture, gates, handles, pictures, rails
- **'complex functional things' whose breaking involves the mechanism as a whole:**
  cameras, cars, differentials, fans, gearboxes, gliders, lighting, machine, microscopes, silencers, trains
- **'the outcome of breaking' (only in the construction with broken as premodifier):**
  bits, crocks, ends, fragments, rubbish

Fig. 10: Subdomain: artefacts and natural things (damage)

[60] This would seem to be in accordance with Langacker's definition of 'basic domain' (1987: 486) as "A domain that is primitive and not characterized in terms of more fundamental domains" although it is used here more specifically to mean 'not characterized metaphorically in terms of a source domain'.

Its suitability as a source domain would seem to reflect the great variety in the types of event involving physical objects that have come to be categorized by *break*. Fig. 10 above shows the subset in which 'artefacts and natural things' are 'damaged' so as to be no longer functional,[61] with examples being grouped according to the kind of process denoted by *break*. It seems plausible to assume that uses of *break* in the less basic sensorimotor domains as well as in nonsensorimotor domains are not motivated by a general 'literal' meaning of *break*, but by the relevant substructure of the basic domain of 'artefacts and natural things'. Thus *breaking relations* (example b in Table 9 below), may be motivated by the breaking of 'chains and rope-like things', which would seem to offer suitable structure for drawing inferences about connections between people (for example it is possible to talk about *strained relations*).

This does not mean that it is predictable what substructure will be recruited for metaphorical extension; typically there are alternative possibilities reflecting different conceptualizations of the same situation. Thus *negotiations* as well as *relations* may be *broken off*, in the same way that you may break off the end of a stick, implying that they are left in an incomplete state, and that somebody is responsible; or they may be said to have *broken down*, like a machine, backgrounding human agency. What matters, then, does not seem to be whether experiences in different domains are comparable as such, but whether image-schemas used in one domain are found to be useful for conceptualizing and drawing inferences about experience in another domain.

Also in the subdomain of **'body parts'**, *break* is used to categorize a range of experience according to what kind of body part is conceived of as 'breaking' or 'being broken'. It is not unlikely that the fracturing of bones is conceptualized in terms of breaking things ('oblong things' that break like sticks). After all, most of us have no direct experience of what a broken human bone looks like; we may have seen X-rays, but it normally takes a

[61] Depending on the context, 'breaking' may be conceived of as purposeful or harmful, so that 'damage' here should be understood as a technical term meaning 'no longer fit for its original function'. There is no clear border line between purposeful and harmful 'breaking'.

doctor to make a judgement as to whether a bone is actually broken or not (cf. example 17 in Table 10 below).

- **fracturing a bone, or a 'bone-like' body part:**
  arms, bones, fingers, jaws, knees, legs, limbs, necks, ribs, shoulders, skulls, thumbs, tibias, toes, wrists, finger-nails, teeth
- **bursting due to internal pressure; rupture:**
  blood vessels, veins
- **cutting the surface:**
  skin
- **undifferentiated:**
  bodies

Fig. 11: Subdomain: body parts

While the use of *break* in the domain of body terms may be motivated by its use in the domain of artefacts and natural things, it is in turn extended to conceptualizations in the nonsensorimotor domain of 'human emotions and influence'. In the case of some expressions (e.g. examples p and q in Table 10), it is obvious that the extension is from (relevant substructure of) the subdomain of 'body terms' since the actual name of the body part (*back; backbone*) is directly involved in the mapping from source to target domain (*the back of the rebellion; backbone industries)*. It is not predictable, however, that a given target domain will draw on a given source domain. Different construal is always a possibility as illustrated by the following example where the mapping is from the domain of 'artefacts and natural things', as shown by the addition *to pieces*:

- *a* [...] ***bow*** [...] *which would by now no doubt* *be broken into unidentifiable pieces* ..
- *The* ***adversaries*** *of the Lord shall* *be broken to pieces* ...

Moreover, the data show that experience in the less basic sensorimotor domain involving **'physical activity'** and observations about **'the physical environment'** can also be conceptualized in terms of 'breaking' (examples 18-23 in Table 10), whereas examples r-z may be construed as extensions from these sensorimotor domains to nonsensorimotor domains. Is it possible that the metaphorical extension could even go from an original

sensorimotor source domain via a nonsensorimotor domain and back to a target in the sensorimotor domain? On the basis of the synchronic evidence made up of the following examples of *break in,* this seems plausible:

- *Shallow tunnels* […] *are best dealt with by* <u>*breaking in*</u> ***the roof*** *of the tunnels and filling in the trench created*. [Sensorimotor domain: artefacts and natural things]

- *The young* ***conscript*** <u>*was broken in*</u> *slowly.* [Nonsensorimotor domain: emotions and influence]

- *[…]a* ***horse*** *that is galled by the saddle when it* <u>*is initially broken in*</u> *may always have a "cold back" and buck,* .[Nonsensorimotor domain: emotions and influence]

- *My first outing in these tough looking* ***boots*** *was three miles in good weather. For the first mile they were brilliant, very comfortable, but then the Big Bs struck. Dreaded blisters on the back of both my heels. Undaunted I was sure it would be worth the effort of* <u>*breaking*</u> ***them*** <u>*in*</u> *and it was.* [Sensorimotor domain: artefacts and natural things]

Fig. 12: Sensorimotor domain as target domain: *breaking in boots*

There is of course a possibility that the 'breaking in of boots' is just another literal use of *breaking in*, but then again it might be the case that boots, being made of leather, are somehow conceptualized as having a spirit that needs to be tamed like that of a wild horse (or an unruly conscript). A similar example in the OED Online that refers to bringing (virgin) land under control clearly involves such personification:

- *In many parts of the North Island,* ***farm country*** *still offers a stubborn resistance to* <u>*breaking-in*</u>*.*

Such examples show that metaphorical extension need not always be from a sensorimotor source domain, but that the source can also be found in the nonsensorimotor domain. As already indicated, literalness is restricted to the most basic domains of sensorimotor experience, which in turn cover a range of conceptualizations of real-world phenomena. It is evident that the correspondence theory of truth does not hold, not only from the metaphorical use of a term like *backbone*, but also from the use of this term in the domain of body parts to refer to the human spine, which is not '*a* bone', but 'a row of bones' that tend to get 'disconnected' rather than 'broken'.

	SENSORIMOTOR DOMAINS	NONSENSORIMOTOR DOMAINS
	ARTEFACTS & NATURAL THINGS	**SOCIAL INSTITUTIONS & CONSTRUCTS**
to damage sth so that it is no longer functional/ to violate sth (cancellation or failure to observe)	1) *With the weak-link strength recommended by the glider manufacturer, breaking the weak **link** is now comparatively rare ...* 2) *"Their minds are like trains going along a track which here and there has a broken **connection**," said the nurse.*	a) *But within two days the RPR had broken the initial **agreement**.* b) *.. Saudi Arabia signalled Iran's return to international respectability by resuming the diplomatic **relations** that had been broken in 1987.*
to fragment sth so that it is no longer 'intact'	3) *The main good grains […]cannot be readily digested until the tough outer **shells** are broken up and, sometimes, removed.* 4) ***Barges** had been [...]left by their previous owners to be broken up by ice ...*	*c) … the existing danger of skilled house-building **teams** being broken up and not replaced ..* *d) The electricity supply **industry** is to be broken up and sold.* *e) "Ari, the world, our **world**, is breaking up around us.*
to subdivide sth into pieces of a serviceable size	5) *Then Mr. Spurgeon rose, and all present rose with him, and, after breaking the **bread,** spoke the accustomed words: "This is my body --."* 6) *200g/7oz can **tuna fish**, drained and broken into large pieces.*	*f) Psychologists have broken down "real" **jobs** into a series of "simple" components ..* *f) .. if **IBM** had been broken up into five pieces 20 years ago,* […]
to separate sth from sth. else	7) *Not only has he again fallen flat on the ground, but his **head** and his **hands** have broken off and lie at the entrance of the building.*	*g) .. a coalition between one major party and a **fragment** from another which has broken off because it disagrees with one of its parent party's central tenets.*
to cause sth to disinte-grate	8) *.. great quantities of **waste** can be broken down by the bacteria in the water.*	*h) White is one of the most stylish throwers in a **game** which only too easily breaks down into brutishness.*
to overcome barriers	9) *Once the **blockade** of the river leading into the city was broken by English ships, James and his besiegers lost heart and abandoned the siege.* 10) *The **tape** was broken at the tunnel's exit by Isambard Kingdom Brunel in green with GWR-type plates and name.*	*i) BRITAIN has broken a vital **barrier** in developing computers that recognise continuous speech …* *j) These guys seem to think all speed **limits** are there to be broken and they just don't believe they will ever be caught."*
to open 'containers' and get things out of them	11) *It keeps the tension going until the Tsarevich breaks the **egg** and thus signifies the death of Kostchei.* 12) *The spear breaks the **pack** and allows water to surge out of nozzles above the spear-end to rinse it.* 13) *Madge's **house** was broken into last night.* 14) *Would "Flash" be breaking out the **champagne**?*	*k) He might have been intending to go back to his family and appear utterly horrified while breaking the **news** of my death.* *l) Meiko has broken into the American and Japanese supercomputer **markets**, and now has offices across the US and in Europe.* *m) The don in him could be satisfied and its conscience allayed by having intellectually broken into the **truth**.*

Table 9: BREAK domains I

SENSORIMOTOR DOMAINS	NONSENSORIMOTOR DOMAINS
BODY PARTS	**EMOTIONS & INFLUENCE**
getting hurt 15) *A fussy neighbour broke his **collar bone** building barricades against the hippies.* **hurting others** *16) Edmund-Davies(dissenting) gave the example of breaking someone's **arm**: that is a really serious injury, but one which is unlikely to endanger the victim's life.* *17) Some broken **bones** have no more than a hairline crack in them, but this is enough to cause your withdrawal from competition, regardless of the stage you are at.*	**getting hurt/hurting others** *n) Hum," said Jay, at least that way you don't get your **heart** broken ..* *o) Table Tennis: Mason breaks Soviet **spirit.*** **causing to lose influence** *p) The Colombo government believes it has broken the **back of the rebellion** in the south ..* q) *The miners' strikes that have paralysed about one-third of Soviet pits are slowly breaking **backbone industries**, such as steel, gas and chemicals.*
PHYSICAL ACTIVITY	**CHARACTERISTICS OF ACTIVITY**
interruption *18) So I walked away, breaking my **journey** home to leave the key with the solicitors.* **starting to do something** *19) .. he had no sooner broken into a **sprint** along the alley than the sound of his own footsteps stopped him in his tracks.*	**interruption** *r) It breaks the **continuity** of the dance.* *s) He talks of responsibility, wrote Goldberg, half-tearing the page of his pad in his hurry to turn it over without breaking the **flow** of his thought.* **variation** *t) The whole cast muster in the market place before breaking into extravagant **variations** of the tarantella.*
THE PHYSICAL ENVIRONMENT	**THE MENTAL LANDSCAPE**
auditory experience *20) He glanced at the clock and nodded his approval without breaking the rhythm created by his fingers on the strings.* **visual experience** *21) Ahead and to our right a head breaks the **surface**.* **experience of states** *22) Insomnia is a common complaint in which the sufferer has poor **sleep** that is often broken many times during the course of the night.* *23) With reference to the article in June on breaking **dormancy** in sweet peas: experiments conducted over a period of several years at New College, Pontefract certainly seem to bear out Anne Swithinbank's remarks.*	**repetitiveness** *u) The **pattern**, that this was the way archbishops happened, was broken by the steady increase in lay headmasters during the earlier twentieth century.* *v) Only by breaking the **habit** of what seems a lifetime will England this evening jeopardise their hopes of reaching the World Cup in Italy next summer.* **consciousness** *w) But a convoluted new thought is breaking **surface*** **mental states** x) *The **deadlock** over farm-trade reform between Europe and America and its allies was broken when the EC agreed to negotiate cuts ..* *y) Fancy a game of darts, lad? Jos said to Mungo breaking what was left of the **ice**.* *z) It is 1900 hours when the **peace** of the evening is broken by the cough of the Challenger engine ..*

Table 10: BREAK domains II

In Table 9 above, examples 1-14 illustrate the use of *break* in the subdomain of artefacts and natural things to refer to 'damage' (examples 1, 2, 4, 7, and 13) as well as to various purposeful activities such as subdividing things, overcoming barriers, and opening containers and getting things out of them (examples 3, 5, 6, 8, 9-12, and 14). Examples a-m show how structure from the subdomain of artefacts and natural things is systematically used outside the sensorimotor domain for conceptualizations about what I have called 'social institutions & constructs'. The other subdomains of sensorimotor experience - body parts, physical activity and the physical environment - also serve as source domains for experience outside the sensorimotor domain, as shown in Table 10 above.

A range of specific ways of breaking are conceptualized in terms of phrasal verbs like *break up, break down, break off* and *break out*, a group in which I have included prepositional verbs like *break into*, or combinations with adjectives like *break open*. This group of verbs, which accounts for roughly one third of the data, can be construed as a 'low-level schema', which Langacker characterizes as being "extracted from the conception of specific instances" (1987: 382).[62] In a taxonomic hierarchy in which *break* is the basic level term, such verbs can be construed as making up the subordinate level of categorization (cf. section 1.5 above). The image-schematic structure underlying the more specialized conceptualizations of 'breaking' will be discussed in the following section, in terms of the general conceptual metaphors that they can be construed as examples of.

3.4.2.2 Image schemas and event structure

Approaching the lexical category BREAK by asking what the lexical item *break* categorizes, as has been done in the previous section, reveals that many uses can be accounted for as motivated extensions from more basic literal uses. Cross-domain metaphorical mappings seem to be an important constitutive feature of such a lexical

[62] Thus the category *tree* may come to include pine trees and palm trees by 'extension', and specific instances like apple trees, peach trees and cherry trees by 'specialization' (Langacker 1987: 382).

category, which has grown into a comprehensive network as new domains have been conceptualized in terms of those which are already an established part of the network. Trying to answer the question *What?* thus automatically leads to the question *How?* In this section I will look at the way in which the image-schematic structure of *break* can be related to more general patterns of metaphor underlying the conceptualization of causation and events.

Causation, according to Lakoff and Johnson (1999: 177 f.) is a radial category with "conscious volitional human agency acting via direct physical force" at the centre. The category also includes less prototypical literal cases, which vary in degree of directness, as well as metaphorical forms of causation. The way we conceptualize causes and the way we conceptualize events are seen as closely related. Both causes and events are understood by means of basic event-structure metaphors, which in turn are based on so-called **'primary' metaphors**.[63] The notion of primary metaphor (Grady 1997, in Lakoff and Johnson 1999: 49), is based on the hypothesis that early childhood experience conflates different domains, like 'knowing' and 'seeing'. The 'conflation stage' in which connections are established between the domains, which are not experienced as separate, is followed by a 'differentiation stage' when metaphorical source and target domains are distinguished, and **'conceptual metaphor'** emerges based on mappings between them, as in the expression *I see what you are saying* (Lakoff and Johnson 1999: 48 f.). I agree with the criticisms levelled at this strong version of the 'embodiment thesis' for not making sufficient allowance for the sociocultural aspect of language acquisition.[64] However, if we take into account that we learn about language not just from our individual bodily experience, but

[63] The notion 'metaphor' in this framework refers to a cross-domain mapping in the conceptual system, whereas the linguistic expression that results from such a mapping is referred to as a 'metaphorical expression' (Lakoff 1991: 203).

[64] For a critical discussion of the 'embodiment' thesis, see Sinha and Jensen de López (2000: 25, 36). They refer to findings showing that "when Zapotec children learn to use body-part terms, they learn to use them appropriately and consistently with the speech practices of the surrounding linguistic community, rather than assimilating them to an overriding basic meaning derived from their experience of their own bodies" and conclude that "the embodiment thesis needs to be extended to take account of the role of cultural meaning in motivating linguistic structure, and more widely, the sociocultural grounding of language."

also from experience that is socially and culturally mediated, I find that notions like primary and conceptual metaphor have substantial explanatory potential.

Our understanding of causation as well as events, according to Lakoff and Johnson (1999: 178 f.), draws on primary metaphors like CAUSES ARE FORCES and CHANGES ARE MOVEMENTS. Events can be subdivided into two main categories depending on whether they are conceptualized in terms of locations or objects:

1. **the 'location event-structure' metaphor**: *Harry got into trouble*
2. **the 'object event-structure' metaphor:** *Harry has trouble*

In the first example, which draws on the primary metaphor STATES ARE LOCATIONS, 'trouble' is conceived as a location, and movement to that location causes it to affect Harry. In the second example, which draws on the primary metaphor ATTRIBUTES ARE POSSESSIONS, 'trouble' is conceived as an object that one can possess and that can be 'given' to somebody. A purpose in a location-event construction would be conceived as a destination, whereas in an object-event construction it would be seen as a desired object (Lakoff and Johnson 1999: 196), as in *come to a conclusion* and *receive an answer*, respectively.

In the analysis of domains (section 3.4.2.1 above), the data were construed as showing metaphorical mappings both between subdomains within the sensorimotor domain, and from subdomains in the sensorimotor domain to subdomains in the nonsensorimotor domain. In the following I will relate the metaphorical uses of *break* emerging from the analysis in terms of domains to the notions of location and object event-structure. I thus see metaphorical mappings as operating on two dimensions: between domains in a network constituting a specific lexical category like BREAK and between underlying, general event-structure patterns and their instantiations across many different lexical categories.

The **object event-structure metaphor** has the following entailments (Lakoff and Johnson 1999: 196):

- attributes are possessions
- changes are movements of possessions (acquisitions or losses)
- causation is transfer of possessions (giving or taking)
- purposes are desired objects

The object event-structure metaphor underlies much of the use of *break* in the nonsensorimotor domain (Table 9: examples a, b, i, j, k; Table 10: examples n-z), although there is also specific motivation from corresponding uses in the sensorimotor domain. Also, the object event-structure metaphor is the general conceptual metaphor underlying the less 'tangible' subdomains of the sensorimotor domain: 'physical activity' and 'the physical environment' as well as the corresponding nonsensorimotor subdomains 'characteristics of activity' and 'the mental landscape'. Here the distinction between sensorimotor and nonsensorimotor domains begins to get somewhat blurred (Table 10, examples r-z). However, it is still possible in some cases to see the nonsensorimotor target domain as specifically motivated by the corresponding sensorimotor source domain, especially when the same term is used (e.g. *surface* in Table 10: examples 21 + w). In other cases, when the terms used in the nonsensorimotor domain have no direct equivalent in the sensorimotor domain (e.g. *rhythm* and *habit* in Table 10: examples 20 and v), expressions in both domains may be construed as drawing on the basic domain of artefacts and natural things, at the same time instantiating an underlying object event-structure metaphor.

Often specific ways of 'handling' are entrenched as collocations using basic level verbs. Such collocations can be seen as 'polysemy evidence' (Lakoff and Johnson 1999: 180) of the object event-structure metaphor in that all the verbs can be used to refer to the handling of objects in a literal as well as a metaphorical sense (cf. Fig. 13 below). The fact that the uses of a verb like *break* in the sensorimotor domain are systematically exploited for metaphorical extension, (cf. Tables 9 and 10 above) is a further source of polysemy evidence of an underlying conceptual event-structure metaphor

.

- *make a promise*
- *have an agreement*
- *give a lecture*
- *take a point*
- *hold an office*
- *keep an appointment*
- *put a question*
- *set a limit*
- *lift a restriction*
- *break the law*

Fig. 13: Entrenched collocations using object event-structure

As mentioned at the end of the previous section (3.4.2.1), about a third of the examples in the data specify the outcome of the 'breaking' by adding an adverb and/or a prepositional phrase, or an adjective. Most clearly in the case of examples where *break* is followed by *into* + nominal, the underlying event structure changes from object event-structure to **location event-structure**, with the following entailments (Lakoff and Johnson 1999: 196):

- states are locations
- changes are movements (to or from locations)
- causation is forced movement (to or from locations)
- purposes are desired locations (destinations)

Again, conventional collocations provide polysemy evidence in the form of short, basic level verbs referring to literal as well as metaphorical motion:

- *go into/out of business*
- *run out of money, into trouble*
- *fall in love, into disrepair*
- *jump to conclusions, to somebody's defence*
- *fly into a rage*

Fig. 14: Entrenched collocations using location event-structure

In Table 11 below, examples of *break into* show how the location event-structure metaphor is used in the sensorimotor domain to denote change of state (in contrast to the literal

	SENSORI-MOTOR DOMAIN	LOCATION EVENT-STRUCTURE	MOVING OBJECT (FIGURE)	CHANGE = MOVEMENT	STATE = LOCATION (GROUND)
1.	**ARTEFACTS AND NATURAL THINGS**	*He must have constructed a pretty good strong bow [...]which would by now no doubt be broken into **unidentifiable pieces** in distant undergrowth.*	bow	be broken into	unidenti-fiable pieces
2.	**PHYSICAL ACTIVITY**	*I remember screaming and my scream breaking into **ecstatic laughter and relief***	my scream	breaking into	ecstatic laughter and relief
3.	**PHYSICAL ACTIVITY**	*The silhouette materialises into a colonel in uniform, standing at ease, his face breaking into **a tentative smile**.*	face	breaking into	a tentative smile
4.	**PHYSICAL ACTIVITY**	*I offer a flask of Featherwheel's famous body lotion [...] to anyone who can describe, without breaking into **a light sweat**, what was going on.*	anyone	without breaking into	a light sweat
5.	**PHYSICAL ACTIVITY**	*"I thought he had come to take you away, Jim," she said, breaking into **tears**.*	she	breaking into	tears
6.	**PHYSICAL ACTIVITY**	*He shouts " Beats International!" with the kind of enthusiasm most sane people save for a winning goal or a multiple birth, and then he breaks into **rhyme**.*	he	breaks into	rhyme
7.	**PHYSICAL ACTIVITY**	*Thoughts of butchers' knives and bloodied axes galvanized the boy into frantic action, but he had no sooner broken into **a sprint** along the alley than the sound of his own footsteps stopped him in his tracks.*	he	had broken into	a sprint
8.	**PHYSICAL ACTIVITY**	*[...]North would produce his spiral notebook or pieces of yellow legal paper, asking" how in God's name we can expect these young men and women to fight against Communism [...]", and often breaking into **tears**.*	North	breaking into	tears
9.	**PHYSICAL ACTIVITY**	*To his name one could add that of Jelly Roll Morton -- who had a splendid, faded baritone, suggesting Humphrey Bogart suddenly breaking into **song** [...].*	Humphrey Bogart	breaking into	song
10.	**THE PHYSICAL ENVIRON-MENT**	*Long after even the latest apple tree has finally broken into **leaf**, the mulberry's branches remain stubbornly bare.*	apple tree	has finally broken into	leaf
11.	**THE PHYSICAL ENVIRON-MENT**	*Katya went crazy over it before it had broken into **leaf**, just as she did with cat-mint.*	it (plant)	had broken into	leaf
12.	**THE PHYSICAL ENVIRON-MENT**	*A rolling like the banging of many drums, distant at first, then breaking into **a staccato crackling**, announced the return of the storm*	a rolling	breaking into	a staccato crackling

Table 11: Location event-structure in the sensorimotor domain (*break into)*

	NON-SENSORI-MOTOR DOMAIN	LOCATION EVENT-STRUCTURE	MOVING OBJECT (FIGURE)	CHANGE = MOVEMENT	STATE = LOCATION (GROUND)
1.	**EMOTIONS & INFLUENCE**	*Over and over again initiatives towards "Imperial union" had broken into* ***spray*** *against the rock.*	initiatives	had broken into	spray
2.	**EMOTIONS & INFLUENCE**	[...] *he said that Warnie and Mrs Moore liked each other, "and, I hope, as W. gets broken into* ***domestic life****, they may come to do so still more*	Warnie	gets broken into	domestic life
3.	**SOCIAL INSTITU-TIONS &CON-STRUCTS**	*The acquisitive sales promotion company FKB Group has broken into* ***the lucrative US medical promotions market*** [...]	company	has broken into	the lucrative [...]market
4.	**SOCIAL INSTITU-TIONS &CON-STRUCTS**	*Edwards, 21*[...]*had 18 months with Rangers without breaking into* ***the first team****.*	Edwards	without breaking into	the first team
5.	**THE MENTAL LANDSCAPE**	*It is about a […]friendship that will take us through life, facing the black moments as well as the good, it is about God's time breaking into* ***our lives*** […]	God's time	breaking into	our lives
6.	**THE MENTAL LANDSCAPE**	*The other word, kairos, is a word that is used a great deal to describe God's breaking into* ***our time and history*** *to bring his salvation.*	God	breaking into	our time and history
7.	**THE MENTAL LANDSCAPE**	*Of course, that is what Advent is all about -- God's adventure in breaking into* ***our history*** *and making it "his-story".*	God's adventure	breaking into	our history
8.	**THE MENTAL LANDSCAPE**	*The don in him could be satisfied and its conscience allayed by having intellect-ually broken into* ***the truth****.*	conscience	having broken into	the truth

Table 12: Location event-structure in the nonsensorimotor domain (*break into)*

meaning of e.g. example 13, Table 9: *Madge's house was broken into last night*, which involves a literal change of location of the burglar doing the breaking in). Table 12 shows examples of *break into* from the nonsensorimotor domain. The claim that states are conceptualized as locations and attributes as objects or possessions, can also be supported by means of 'inferential evidence' and 'poetic evidence' according to Lakoff and Johnson (1999: 180 f.). Inferences that apply to a given source domain in space are mapped to the corresponding target domain both in conventional and novel expressions. Thus someone who is *in business* cannot at the same time be *out of business*, and it seems natural to talk about somebody who has *gone into business* eventually being *driven out of business*; although the last expression is probably not entrenched, it seems a logical extension.

DOMAIN	EXAMPLES OF *break* IN PHRASAL VERB AND RELATED CONSTRUCTIONS (with adverbs, prepositions, and adjectives)	MOVING OBJECT (FIGURE)	CHANGE = (SELF-PRO-PELLED) MOVEMENT	STATE = LOCATION (GROUND)
1. NSM	*We sell a great many long-haul holidays, often breaking apart the **packages** then putting them together again at lower cost.*	packages	break, vt	apart
2. NSM	*[...] a pattern emerged through Europe of societies of **artists** breaking away from official organisations,*	artists	break, vi	away from off. org.
3. SM	*The increased speed [...] has led to more crowded play so that linesmen are apt to flag when a **forward** suddenly breaks clear to receive a pass.*	forward	break, vi	clear
4. SM	*lo and behold the **kitchen door** was broken right down!*	kitchen door	break, vt	down
5. NSM	*By educating children that a large part of society would like to see behind bars, we are breaking down important **taboos**," said Alda Marco Antonio, [...].*	taboos	break, vt	down
6. SM	*Shallow tunnels [...]are a frequent cause of concern, [...] are best dealt with by breaking in the **roof** of the tunnels and filling in the trench created.*	roof	break, vt	in
7. NSM	*Secret **talks** with the government, [...] were also broken off.*	talks	break, vt	off
8. SM	*In the case of an eighth-century Pre-Khmer bronze figure [...] the **head** had once been broken off […].*	head	break, vt	off
9. SM	*Parents know what their children are longing for and will give them what they want, even if every **piggy bank** in the house has to be broken open.*	piggy bank	break, vt	open
10. SM	*However, while returning home with the Cup might not quite be an occasion for breaking out **the ticker tape**, [...] it would be nice to see England do well,*	ticker tape	break, vt	out
11. SM	*When hooligans themselves are asked to define the term they usually say "a right little hardnut", someone willing to get "stuck in" when a **fight** breaks out.*	fight	break, vi	out
12. SM	*Outside there are scudding clouds, but they are high up and **the sun** occasionally breaks through.*	the sun	break, vi	through
13. SM	*Projection of another dimension or dimensions as when **the sun** breaks through a cloud*	the sun	break, vi	through a cloud
14. NSM	*[...]a statutory commission to find out why **women** have not broken through the "glass ceiling" that keeps them from top management jobs.*	women	break, vi	through the "glass ceiling"
15. NSM	*The lyrics deal with teenage frustrations and breaking through to the other side.*	-	break, vi	through to the other side
16. NSM	*[…] they went on to do other things, [...],a "beautiful" working **relationship** was thereby broken up,*	relation-ship	break, vt	up
17. NSM	*The consultants recommended that the **Plastics Division** be broken up into four groups, […].*	Plastics Division	break ,vt	up into four groups
18. NSM	***Mrs Thatcher** was anxious not to give many hostages to fortune by breaking too openly with the one-nation traditions of Butler and Macleod.*	Mrs Thatcher	break, vi	with traditions

Table 13: Phrasal verbs and related constructions analysed as location event-structure
(SM: Sensorimotor domain; NSM: Nonsensorimotor domain; vi: verb intransitive; vt: verb transitive)

The question is whether the location event-structure metaphor can be said to account for all the phrasal verb constructions (combinations with adverbs as well as prepositions, and adjectives). As shown in Table 13 above, this would involve conceptualizing the change of state categorized by *break* in terms of forced or self-propelled movement to or from a location, although in some cases the state may be less clearly associated with a location (*break clear; break off; break open*) than in others (*break away; break up; break down; break out; break in; break through*). In most of these expressions, the adverb, preposition or adjective denotes the resulting state or the location towards which the figure is moving; only in two cases does it denote the original state or the location the figure is moving away from (*break away from; break with*). As implied by the term 'phrasal verb', it is common to analyse the adverb, preposition or adjective as part of a verb group, which supports the idea that the change of location is actually conceptualized as an integral part of the event. With the reservations made, phrasal verbs may be said to provide evidence for the location event-structure metaphor. The fact that they account for such a large share of the examples of *break* in the data can be related to the important function they have as the precision tools of the BREAK category, allowing the detailed conceptualization of events in the sensorimotor as well as the nonsensorimotor domain.

In the previous section it was found that it is not possible to study *break* in terms of domains without becoming aware of the metaphorical mappings that provide the internal coherence of this lexical category. The purpose of this section has been to relate lexically specific image-schematic structure to conceptual metaphors underlying events in general. This means that the discussion has already turned from lexical to grammatical structure, which in cognitive grammar is taken to differ from lexical structure only by being more schematic. It is claimed "that grammar itself serves an 'imagic' function and that much of it has a figurative character" (Langacker 1987: 38 f.). So far, the focus has been on the way in which domains share image-schematic structure, but in the next section I will look at patterns of construction for which the data show substantial differences between domains.

3.4.2.3 Construction types

Grammatical organisation contributes to the meaning of linguistic expressions, because it "structures a scene in a particular way" so that "roughly synonymous sentences with the same content words but different grammatical structures [...] are claimed [...] to be semantically distinct by virtue of their different grammatical organization per se" (Langacker 1987: 39). By the same token, the construction types[65] found in the data for *break* can be seen as alternative profiles on the same base (see below). In this section I will compare construction types in the sensorimotor domain of 'artefacts & natural things' with those in the nonsensorimotor domain of 'social institutions & constructs' and discuss the patterns found.

In Langacker's cognitive grammar, **'predication'** refers to the 'semantic pole' of a linguistic expression, the other pole being phonological (Langacker 1987: 97, 491). Basic classes of predications are **'nominal predications'** (or 'things'**)** and **'relational predications'**. The latter are in turn subdivided into **'atemporal relations'** and '**processual predications**', (or simply 'processes'), depending on whether they have a 'positive temporal profile', meaning that "the evolution of a relation is followed through time" (Langacker 1987: 249). The mode of cognitive processing used for nominal predications and atemporal relations is compared to how we see a still photograph and characterized as **'summary scanning'**: all aspects of an event are simultaneously available and constitute a coherent gestalt. By contrast, the mode used for processes is characterized as **'sequential scanning**', which involves viewing a scene as a succession of transformations and can be compared to what we do when we watch a motion picture (Langacker 1987: 144 f, 248 f.).

Predications are seen as always having a certain **'scope'** or **'base'**, defined as "those aspects of a scene that are specifically included in a particular predication" (Langacker

[65] In Langacker's cognitive grammar, 'grammatical construction' refers to: the component structures, their mode of integration, and the resulting composite structure" (Langacker 1987: 277) (cf. section 1.5 above.

1987: 493), and the substructure selected for 'designation'[66] is referred to as the **'profile'**, which is perceived as somehow "standing out" against the base (Langacker 1987: 183). Thus the base for the past participle *broken* used as a premodifier, as in *a broken cup*, is the processual predication *break*, which designates "a continuous series of states distributed through time". The past participle form *broken*, however, only profiles the final state, and is therefore given a stative, adjectival construal as an atemporal relation, although "the conception of a process evolving through time provides the necessary context" (Langacker 1987: 221).

The profile/base distinction is part of the more general notion of **'figure/ground'** organization, which tries to capture the importance in grammatical structure of the "perspective taken on a scene" (Langacker 1987: 120). Another example of figure/ground organization is the **'trajector/landmark asymmetry'**, exemplified by the subject/object distinction in clause-level syntax (Langacker 1987: 232). In a processual predication like *Peter broke the record*, *Peter* and *record* are said to be, or more correctly, *correspond to*, respectively the trajector (tr) and landmark (lm) of *break*, which are construed as belonging to the internal structure of relational predications. In an atemporal relation like *a broken vase*, *vase* corresponds to the trajector of *broken*. Prepositions also profile atemporal relations; in the case of a noun group such as *the book on the table*, in which the noun is modified by a prepositional phrase, *the book* would thus be said to correspond to the trajector of *on*, while *table* corresponds to its landmark (cf. section 3.4.3.2 below).

Table 14 and Fig. 15 below give an overview of the distribution of construction types in the sensorimotor domain of 'artefacts & natural things' as well as in the nonsensorimotor domain of 'social institutions & constructs'. In this section, I will refer to the two domains as the ANT–domain and the SIC-domain, respectively. In the former domain, *break* is used to categorize 'damage' (intentional or accidental) and in the latter it categorizes what has been construed as the equivalent of damage in the nonsensorimotor domain: the 'violation'

[66] 'Designation' is defined as "The relation within a semantic structure between the base as a whole and some substructure selected as profile." (Langacker 1987: 488). The relation between a linguistic unit and a usage event is referred to as 'coding' (Langacker 1987: 487).

	SENSORIMOTOR DOMAIN			NONSENSORIMOTOR DOMAIN		
	ARTEFACTS & NATURAL THINGS:			SOCIAL INSTITUTIONS & CONSTRUCTS:		
	'DAMAGE'			'VIOLATION'		
PREDICATION TYPE:	**CONSTRUCTION TYPE:**	**NO.**	**%**	**CONSTRUCTION TYPE:**	**NO.**	**%**
	1. TRANS. VERB + NOUN (OBJ.), ACTIVE	12	7.59	**1. TRANS. VERB + NOUN (OBJ.), ACTIVE**	65	74.71
PROCESSES	*Somebody had broken the glass door of the shop, but Michael wasn't worried..*			*Or suppose A has broken his contract to sell land to B*		
	2. TRANS. VERB + NOUN (SUBJ.) PASSIVE	9	5.70	**2. TRANS. VERB + NOUN (SUBJ.), PASSIVE**	11	12.64
	The wooden gate was broken and the hedge overgrown.			*If a law is not being broken, you cannot do anything.*		
	3. NOUN (SUBJ.) + INTRANS. VERB	10	6.33	**3. NOUN (SUBJ.) + INTRANS. VERB**	0	0
	The cable has broken.					
	4. PAST PARTICIPLE + NOUN	117	74.05	**4. PAST PARTICIPLE + NOUN**	5	5.75
ATEMPORAL RELATIONS	*A park strewn with dog-shit and broken bottles …*			*The reverberations of a broken marriage ...*		
	5. NOUN + PAST PARTICIPLE	2	1.27	**5. NOUN + PAST PARTICIPLE**	1	1.15
	Then I would be ashamed, seeing his little body broken on the rocks below.			*Legends full of taboos broken with dire results ...*		
	6. NOMINALIZATION OF VERB	8	5.06	**6. NOMINALIZATION OF VERB**	5	5.75
NOMINAL PREDICATIONS	*However, with fewer breaks pilots tend to take it for granted that they will not get a cable break.*			*Such potential law breaking would threaten the deal.*		
	Breaking the weak link is now comparatively rare.			*The breaking of rules and the taking of liberties ...*		
	TOTAL	158	100		87	100

Table 14: Construction types in two BREAK domains

of agreements, rules, laws, etc. (cf. Table 9, section 3.4.2.1). Six different construction types are included: (1) transitive verb + noun (obj.), active form, (2) transitive verb + noun (subj.), passive form, (3) noun (subj.) + intransitive verb, (4) past participle of verb + noun, (5) noun + past participle of verb, and (6) constructions with a nominalized verb; either preceded by a noun in compounds like *rope breaks*, or followed by a nominal group as in *breaking the weak link* (or *the breaking of the weak link*). Only the nominalizations in (6) are nominal predications or things; the other five types are relational predications of which the first three are processes and the other two are atemporal relations. Construction (5), in which the noun, or pronoun, is followed by the past participle, as in *Anything broken had to be hidden*, can be construed as a **'complex atemporal relation'**, which is like a process in that it profiles all the component states of its base, but like an atemporal relation in that it is scanned in summary fashion (Langacker 1987: 249). The table shows the number of occurrences of each type as well as its percentage share of the total number of occurrences. Percentages are given with two decimals, but in my comments I will round off to the nearest whole number.

The first three types of relational predications are processes. As indicated above, they are characterized as profiling a series of component states that can be scanned sequentially, not just the final state like atemporal relations (types 4 and 5 in Table 14). Unlike nominal predications (6), which may profile component states collectively, a process "must have the profile of a relation throughout its temporal extension" (Langacker 1987: 247). In the ANT-domain, the three types of processes between them account for a relatively moderate share of roughly 20% of total occurrences, whereas in the SIC-domain, processes account for the vast majority of occurrences with a share of almost 88%.

It is especially striking that the percentage share of **the first construction type, transitive verb + noun (subj.), active form**, is almost ten times larger than in the ANT-domain, while in the case of **the second type, transitive verb + noun, passive form**, the ratio is a little over two to one. This indicates that an important function of *break* in the SIC-domain is to combine with nominal predications to form processual predications, whereas in the

ANT-domain this function is relatively less important. If this should turn out to be a general pattern, it might support a construal of *break* in the SIC-domain as having, in addition to its role as a 'full verb', a kind of 'support verb' function.

It is also interesting in this respect that there are no examples of **the third construction type, noun (subj.) + intransitive verb** in the data for the SIC-domain. Actually, there are only two examples in the data where *break* alone (that is not in phrasal verb constructions) occurs as an intransitive verb in the nonsensorimotor domain, viz. *the crisis broke* and *the scandal broke* where the meaning is 'break out' (subdomain: opening 'containers'). However, in the ANT-domain the construction is found to occur quite naturally with subjects like *cable*, *film*, *glass*, *pencil*, *rope*, and *shaft*. The difference between the two domains is illustrated by the following abbreviated examples:

➢ *he (tr) broke the glass* (*lm*)	*the glass* (*tr*) *was broken*	*the glass* (*tr*) *has broken*
➢ *he (tr) broke the law* (*lm*)	*the law* (*tr*) *was broken*	* *the law* (*tr*) *has broken*

I think the difference between the two domains can be explained in terms of ergativity and thematic relationships. According to Langacker (1990: 84 f.), ergativity "reflects an aspect of the structure of event conceptions whereby a transitive object or an intransitive subject has a greater degree of intrinsicness than do other participants." Both are said to encode the 'theme', which functions as the 'conceptually autonomous core' of the event. In the nonsensorimotor SIC-domain, it seems that the object, while still encoding the theme, is not conceptually autonomous to the same extent. If *break* is seen as possibly having a support verb function in the SIC-domain, the composite structure [*break the law*] can be construed as integrating the theme so closely with the verb that it becomes part of the process and is no longer sufficiently independent to function as subject, or trajector, in an active, ergative construction. However, the periphrastic form of the construction still allows a change of perspective by means of passivization, so that the landmark in the active sentence now becomes the trajector, without the noun becoming a fully independent nominal participant.

Turning now from processes to atemporal relations, we find **the fourth construction type, premodifying past participle + noun**, accounting for 74% of occurrences in the ANT-domain. The question is why this construction, which profiles only the final state of the process, should be so much more frequent in the sensorimotor domain than transitive constructions profiling the process of an agent doing damage to an object (6%). Since *break* designates a perfective process involving a change through time, it seems quite plausible that the result of that change should often be perceived as the most salient of the component states, but why, then, does this construction account for only 6% of occurrences in the SIC-domain? Again, if *break* is assumed to have a full-verb function in the ANT-domain and a support-verb function in the SIC-domain, this might explain the predominance of processual relations over atemporal relations in the latter domain. If the comparison is extended to include all subdomains (excluding phrasal verbs), the imbalance is less pronounced, but still quite clear with this construction accounting for 55% and 13% in the sensorimotor and nonsensorimotor domain, respectively.[67] Provided that this distribution reflects a general trend, it may serve as an example of the way in which corpus data may reveal information that is not accessible through introspection (cf. section 3.2.1. above).

The fifth construction type, noun + past participle, has been construed above as a complex atemporal relation, intermediate between a process and an atemporal relation (Langacker 1987: 249). In the first of the following two examples, it is followed by an adverbial group, which highlights its processual side, whereas its stative, adjectival side comes out in the second example, where it occurs in parallel with the adjective *untidy*:

- *Far below us pine forest bearded the foothills, broken here and there by the scars of red roofed villages*
- *"I could mend that for you," he had said, hating anything broken or untidy only for want of a firm nail.*

[67] 483 examples have been recorded as belonging to a sensorimotor domain, of which 144 are accounted for by phrasal verbs and related constructions (with prepositions and adjectives), leaving 339 examples. 187 of these, or 55.16%, are of the construction type past participle + noun. 424 examples have been recorded as belonging to a nonsensorimotor domain, of which 178 are accounted for by phrasal verbs and related constructions. This leaves 246 examples of which 33, or 13.41%, are of the past participle + noun construction.

In the ANT-domain there are 2 occurrences (1.19%), out of 9 (1.88%) in the sensorimotor domain as a whole. In the SIC-domain there is one occurrence (1.16%) out of six (1.4%) for the whole nonsensorimotor domain. This is not much to go by, but in line with what has been said above, I would argue that examples in which *break* is followed by an adverbial phrase, a construction which can be seen as a reduced passive relative sentence with the relative pronoun and the finite auxiliary verb missing, would be compatible with a support verb construal. Actually, the only example of this construction in the SIC-domain is of this kind:

- *Cuchulain legends, written down by monks in the early Middle Ages, are full of* <u>*taboos*</u> [which have been] <u>*broken with dire results*</u>*, totem animals and other Apache touches.*

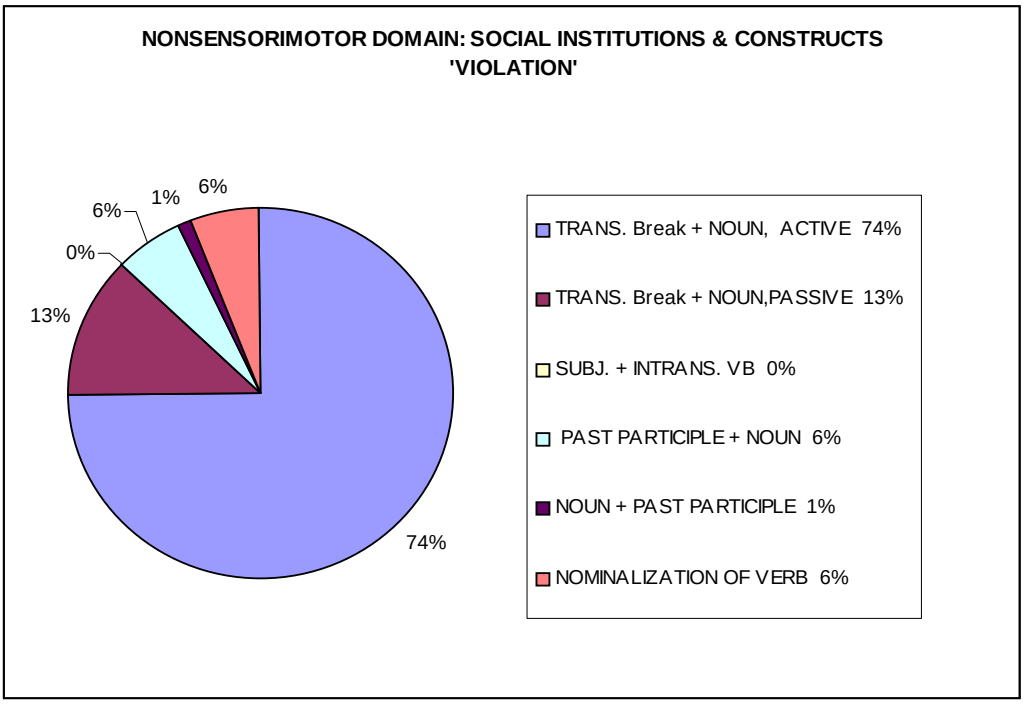
In the nonsensorimotor domain as a whole, four of the six examples are phrasal verbs. In the final example of this construction (also from the nonsensorimotor domain), *broken* is followed by another past participle:

- [...] *I was off once more into the land of longing,* <u>*my heart at once broken and exalted*</u> *as it had never been since the old days at Bookham."*

The data furthermore include two occurrences of a noun followed by the present participle, which can be seen as a reduced active relative sentence:

- […] *the swell of Brahms cello* [which is] <u>*breaking over the piano*</u> ..
- *With Skerrett and* […] *John Hamer* [who is] *now* <u>*breaking almost unopposed*</u>

Both of these examples are from the sensorimotor domain (subdomains: 'overcoming barriers' and 'the physical environment') and *break* is used as an intransitive verb. In accordance with what was said about the third construction type, nominal subject + intransitive verb, I would not expect this type to occur in the nonsensorimotor domain, and I do not see it as compatible with a construal of *break* as a support verb.



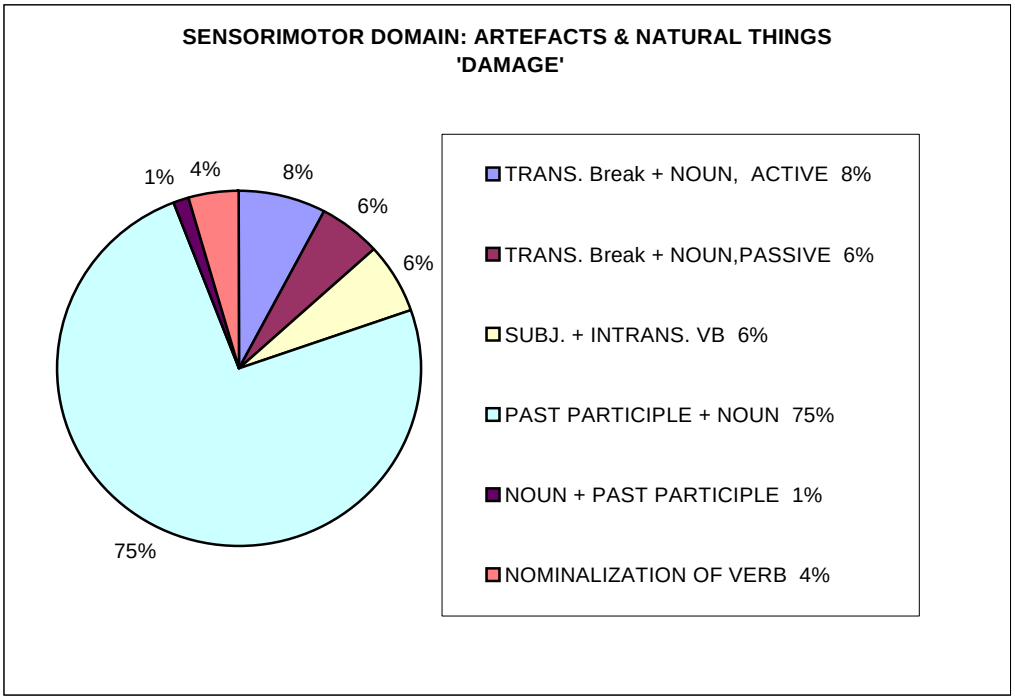


Fig. 15: Construction types in two domains

The sixth type is nominalizations of the verb *break* and includes four different constructions. In the sensorimotor domain, 7 of the 8 occurrences are (1) compounds like *cable break* and *rope break* consisting of a noun coding the breaking object + the nominal predication break. The last occurrence,(2) *breaking the weak link* consists of the nominal predication *breaking* + a noun group coding the physical object that is broken. In the SIC-domain, 3 of the 5 occurrences take the form (3) *a/the breaking of* + noun or nominal group coding the 'social institution or construct' that is 'violated'. The last two are compounds: (4) *law breaking*, in which the violated entity is followed by *breaking* (present participle as deverbal nominalization). Table 15 below gives an overview of the distribution of types of nominalization according to domain.

DEVERBAL NOMINALIZATION	TYPE	ARTEFACTS & NATURAL THINGS	SOCIAL INSTITUTIONS & CONSTRUCTS	EXAMPLE
mass noun	1. *breaking the* + noun	x		*Breaking the weak link is now comparatively rare.*
mass noun bounding imposed by other predications	2. *a/the breaking of* + noun		x	*The breaking of rules and the taking of liberties ...* […] *the threatened self-annihilation of the women is also a breaking of the circuit*
mass noun	3. noun + *breaking*		x	*Such potential law breaking would threaten the deal*
count noun	4. noun + *break*	x		[…] *with fewer breaks pilots tend to take it for granted that they will not get a cable break.*

Table 15: Deverbal nominalizations of *break*

In these constructions, both *break* and *breaking* are categorized as nominal predications, which are cognitive events characterized abstractly as 'bounded regions in some domain' (Langacker 1987: 183, 189). They are defined as symbolic structures whose semantic pole designates a '**thing**'. A '**region'** is further characterized as a set of interconnected entities which are not in profile, but only presupposed as part of the base. The difference between

the nominal predications, or nouns, *break* and *breaking*, and the processual predication, or verb, *break* is that the former profile the component states collectively as a thing whereas the latter profiles them individually as relations (Langacker 1987: 246 f.). Moreover, it is suggested (ibid. 203 f.) that '**bounding**' can be used to distinguish between '**count nouns'** and '**mass nouns**'. The former designate a single instance of the process, "a region that is specifically construed as being bounded within the scope of predication in a primary domain", in contrast to the latter, which "refers to it in a generalized, even generic fashion". Deverbal nominalizations like *jump, dance* and *throw* that designate a single episode of the process are categorized as count nouns, whereas nominalizations like *jumping, destruction* and *love* are categorized as mass nouns (Langacker 1987: 207 f.).

According to this distinction, the first three types of nominalization are mass nouns whereas the fourth type is count nouns. In the second type, bounding is imposed on the mass noun by syntagmatic combination with the indefinite or definite article; it is not a part of the nominal predication itself (cf. Langacker 1987: 204). Even if there are very few examples, it may be of some interest that examples of the fourth type, in which *break* is construed as a count noun, are not found in the data for the SIC-domain. In the data as a whole, only one example is found in the nonsensorimotor domain: *a break with the past*, which is a phrasal (prepositional) verb construction. A supplementary search of the BNC produced only one example, viz. *career break*, which turns out to be not the nominalization of *to break a career* (as in *they can make or break political careers*), but to be the equivalent of the expression *a break in somebody's career*. In an attempt to build a case for a support-verb construal of *break* in the nonsensorimotor domain, it makes sense that the verb is nominalized as a mass noun designating "an indefinite sequence of internal states construed atemporally" (Langacker 1987: 188) rather than as a count noun which designates only a single episode. However, the count noun is not a general type of nominalization for full verbs either, so even if it holds, the 'non-use' of this construction is not all that useful as evidence of the possible status of *break* as a support verb.

It is generally assumed that differences in form imply differences in function. I have tentatively suggested that some of the differences between the two domains that have been compared might be symptomatic of the use of *break* as either a full or a support verb, an issue that I will return to in sections 3.4.3.3 and 3.4.4.3 below.

3.4.2.4 How many meanings does *break* have?

As a verb, *break* is a relational predication, which means that it is a 'dependent structure' (Langacker 1987: 488) understood as part of a syntagmatic combination with a nominal predication: the object that is conceived of as 'breaking' or 'being broken'. Expressions like *break one's leg*, *break somebody's heart*, *break a gramophone record*, and *break a sports record* evoke different semantic frames. However, in the case of a high-frequency, multi-purpose verb like *break*, positing a reading for every frame that it evokes would not have much descriptive or explanatory value. The network model offers the possibility of moving to a higher level of abstraction reflecting the extent to which expressions that evoke different frames can be construed as extensions from the same prototype. The construal will draw on the notions of domain, image-schematic structure and construction type, as well as on notions that may help distinguish between ambiguity and vagueness. As mentioned before, the aim is to produce a plausible hypothesis of the nature of the internal structure of this lexical category which is consistent with the data, while it is not the ambition to produce a definite account.

Ten possible readings will be discussed for the sensorimotor domain. The discussion of each will be followed by a discussion of its extension to the nonsensorimotor domain, adding up to a total of twenty possible readings, as listed in Fig. 16 below (cf. Fig. 9 as well as Tables 9 and 10 in section 3.4.2.1). Fig. 17 below is a blueprint for figures 18-28, which will be used to describe these pairs of readings. BREAK represents the lexical category and the arrow indicates its extension to accommodate the examples shown of the sensorimotor domain reading in question, which is itself a complex category. The broken arrow in turn

SENSORIMOTOR DOMAIN	**NONSENSORIMOTOR DOMAIN**
(source domain)	**(target domain)**
1. Damaging physical objects	11. Violation of social institutions & constructs
2. Specific ways of 'breaking' physical objects	12. Specific ways of 'breaking' social institutions & constructs
3. Overcoming physical barriers	13. Overcoming psychological barriers
4. Opening physical containers	14. Opening metaphorical containers
5. Damaging body parts	15. Causing psychological damage
6. Interrupting an activity	16. Changing from a way of doing something
7. Starting an activity	17. Changing to a new way of doing something
8. Interrupting auditory and visual experience	18. Changing a pattern
9. Becoming visible	19. Entering consciousness
10. Interrupting a physical state	20. Interrupting a psychological state

Fig. 16: Proposed readings of *break*

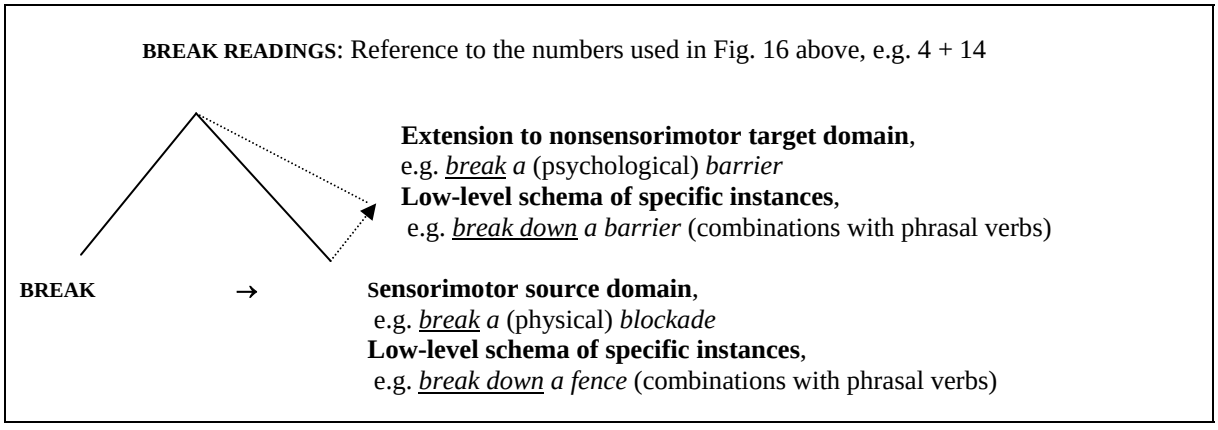


Fig. 17: Blueprint for figures 18-28

indicates the metaphorical extension to the nonsensorimotor domain reading, also a complex category represented by the examples from which it is abstracted. The examples in figures 18-28 are of *break* as a transitive verb with a nominal object and in the past participle form as pre- or postmodifier of a noun. Following each reading, the phrasal-verb combinations construed as the low-level schema of specific instances of that reading are shown. Finally, the 'pyramid' is meant to show that, at a higher level of abstraction, a

schema can be extracted which is in turn instantiated by readings in the two domains. The motivation for conceptualizing a given target domain in terms of a given source domain is assumed to be that the image-schematic structure of the source can be used to draw inferences about the target. However, according to 'the invariance hypothesis' (Turner 1993), target domains have image-schematic structure of their own, and mappings are constrained by the need not to violate this structure, unless such violation has a special significance. This implies that "[m]any components of image-schematic structure in the source are simply not involved in the mapping" and that "components of the source that are indeed involved in the mapping often have image-schematic structure that is not mapped onto the target" (Turner 1993: 301 f.). Each reading is a complex category with more or less prototypical members, and the different readings can be seen as forming a 'radial category', which is described by Lakoff (1987: 91 ff.) as having a cluster of converging cognitive models as a central subcategory, which is surrounded by a number of conventional extensions that are not predictable from, but motivated by the centre.[68]

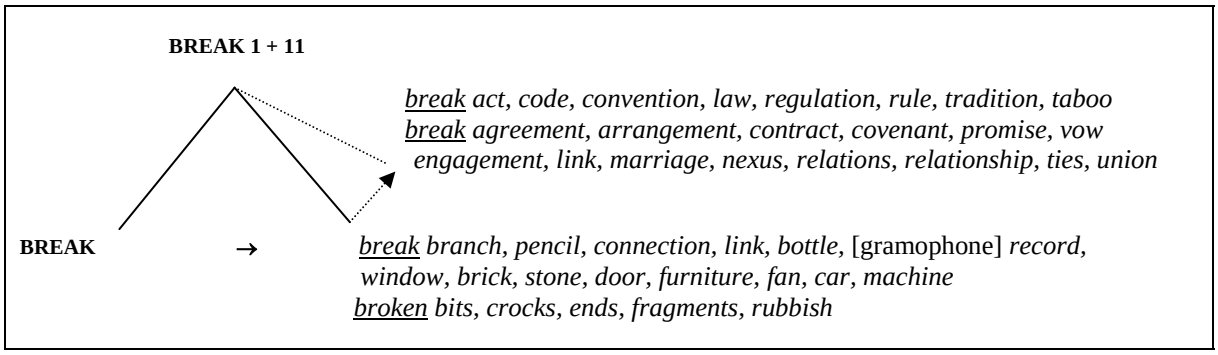


Fig. 18: Damaging physical objects & Violation of social institutions and constructs

(1) 'Damaging physical objects' is posited as one of the central meanings of BREAK conceived of as a radial category. It is a high-level schema which is abstracted from a wide range of expressions at the basic level of categorization, exemplified in Fig. 18 above (cf. Fig. 10 in section 3.4.2.1). However, it seems that the differences in meaning can be

[68] Lakoff uses the category *mother* as an example, with the birth model and the nurturance model as part of the centre and with adoptive mothers and marital models (*mother* refers to 'the woman who married the father') among the non-central extensions.

attributed to vagueness rather than ambiguity in that substituting nouns for each other does not lead to commutation, and using *break* to refer to different types of object does not produce zeugmatic effects (cf. section 3.4.1.1 above). This intuitive judgement is supported by the following example, in which *broken* profiles a composite structure with two nominal predications:

- *The small room looked like Darnley Tip, piled high with broken* ***furniture and cardboard boxes****, with newspapers strewn all over the floor and across various chairs.* (1)

Both *furniture* and *cardboard boxes* have been characterized as coding 'simple functional things' (Fig. 10, section 3.4.2.1), and the example shows that the meaning of *break* can be vague as regards the difference between the breaking of quite different objects. An alternative analysis would be that *broken* modifies *furniture* only, in which case the argument fails. One type of combination seems to stand out from the rest, as the noun, which occurs only with the past participle form *broken*, denotes the result of the breaking (*bits, ends, fragments, rubbish*) rather than the object that is broken. Talking about *broken bits and chains* would seem zeugmatic (unless *bit* referred to a broken object) just as substituting *bits* for *chains* would seem to involve commutation of meaning. This might justify positing a separate but related meaning for *break* when referring to objects resulting from the process rather than to the objects exposed to it. As an alternative, which I prefer, such examples, which only occur in the construction type in which the past participle premodifies the noun, may be assimilated to the category as peripheral members that differ from more prototypical examples in emphasizing the stative and adjectival profile imposed by the construction type. The advantage of prototype categories is that they can accommodate such deviation from the centre.

Reading **(11) 'Violation of social institutions & constructs'** in the nonsensorimotor domain is construed as an extension from 'damaging physical objects' and, at the same time, as motivated by the general object event-structure metaphor (cf. section 3.4.2.2 above). 'Damage' in the sensorimotor domain corresponds to 'violation' in the nonsensorimotor domain; in both domains causation is a central factor and *break* denotes a perfective

process of sudden change. In line with the invariance hypothesis, it can be argued that mappings are constrained by the image-schematic structure of the target domain in various ways. In the sensorimotor domain, the object is a fully independent participant which is clearly impacted by the process, whereas in the nonsensorimotor domain, the object can be seen as elaborating the process, and in many cases it is not impacted by it at all (cf. sections 2.2.4 and 3.4.2.3 above). Thus *acts, codes, conventions, laws, regulations, rules, traditions*, or *taboos* will still exist after the breaking, like *agreements*, *contracts*, *vows* and *promises* whereas breaking mutual commitments such as *appointments*, *engagements*, *marriages*, *links*, *relations*, *relationships* and *unions* will typically amount to cancelling them.

In terms of frame semantics, the difference between readings (1) and (11) would seem to be that in reading (1) it is the verb that evokes the dominant frame, so that *break a pencil, a rope, a stone, a camera*, etc., can all be seen as instantiating a frame evoked by the verb rather than a frame evoked by the noun. In reading (11), i.e. 'violation of social institutions & constructs', it seems to be the case that the frame is typically evoked by the noun; if this is so, it opens the possibility of construing *break* as a support verb as discussed at the end of the previous section (3.4.1.1)[69]. Alternatively, two different readings may be posited: One in which *break* means 'failure to observe', as in *breaking the law*, which relates to social institutions, another in which it means 'cancellation', as in *breaking an engagement*, which relates to the termination of mutual commitments between individuals. There are three examples in the data for this reading in which *break* takes two objects:

- *Even more exciting changes of éaulement can be found in Ashton's Birthday Offering where each soloist dances an old step at a new angle, without* <u>breaking</u> ***the rules or older conventions*** *of nineteenth-century ballet.* (11)

- *In effect the insider who questions the gross systems of classification which define police practice seems set to join those deviant criminals who contest the system of law and order by* <u>breaking</u> ***its rules and regulations****.* (11)

[69] It is not assumed, however, that it is only in the nonsensorimotor domain that the noun can evoke a frame. In the literal expressions *break a link* or *break a connection*, it would seem that it is also the noun that evokes the frame. However, this intuitive judgement needs to be tested by corpus studies of the respective lexical categories as well as by psycholinguistic studies.

- *Most athletes first encountered him as a voice, bellowing in multi-lingual fury at officials who had broken* ***rules or arrangements*** *designed to make racing safer or more fair.* (11)

In the two first examples, the nouns belong to the first reading suggested above, so that no zeugmatic effect is produced. In the third example, the effect does not seem to be zeugmatic either as the coordination of *arrangements* with *rules* invites an interpretation of *arrangements* as a general social institution rather than a mutual commitment between individuals, so that *breaking* in both cases means 'failure to observe'. An alternative explanation might be that *break* is vague rather than ambiguous as far as this distinction is concerned. Thus it does not seem zeugmatic (or not obviously so) to talk about *breaking rules and contracts.* Would talking about *breaking rules and relationships* create a zeugmatic effect – and would substituting these nouns for each other involve commutation? According to my subjective judgement, it would, and consequently two readings of *break* are posited as extensions from the source domain of 'damage to physical objects', namely (1) 'failure to observe' and (2) 'cancellation'. Another possibility, as suggested above, is a more general construal of both readings as exemplifying a support verb function of *break.*

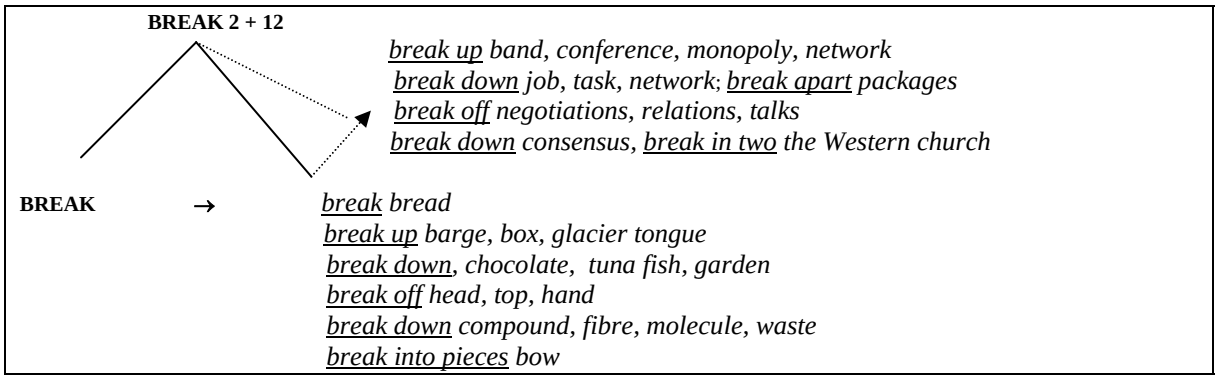


Fig. 19: Specific ways of breaking physical objects and Specific ways of breaking social institutions & constructs

Reading **(2) 'Specific ways of 'breaking' physical objects'** in the domain of 'artefacts and natural things' shows overlaps with reading (1), in which some of the examples also involve 'damage'. However, I take the prototype for this reading to be a process that is purposeful or neutral with regard to its effect. What characterizes this category is the level

of specificity at which the process is coded, evidenced by the predominance of phrasal verbs: *break up/down/into pieces/off* (cf. examples 3-8 in Table 9, section 3.4.2.1 above). In fact, only one example has been recorded which does not have a phrasal verb:

- *Then Mr. Spurgeon rose, and all present rose with him, and, after breaking* ***the bread****, spoke the accustomed words: "This is my body --."* (12)

This example of a literal use of *break* in the sensorimotor domain is rather special in that it refers to the breaking of the sacramental bread in the Communion of the Lord's Supper. Normally no actual breaking of bread takes place; instead the words refer to the handing out of a special kind of thin wafer. The example has been recorded as an instance of the sense 'subdivide into pieces of a serviceable size', but probably the subdivision has happened already at the manufacturing stage while the idea of sharing is preserved by the expression *to break the bread*. The same idea is also expressed by the idiom *to break bread with somebody*, which means 'associate with somebody'. In addition to the symbolic and metaphorical uses, *break bread* still has the literal meaning of 'subdivision'.

The examples of specific ways of breaking have been characterized as denoting fragmentation, subdivision, separation or disintegration, but there is no one-to-one correspondence between these subtypes and specific phrasal verbs. Thus *break into* + noun may mean to 'fragment something so that it is no longer whole' (a sense which overlaps with 'damage') or to 'subdivide something into pieces of a serviceable size':

- *He must have constructed* ***a pretty good strong bow*** *too (according to my detailed instructions) which would by now no doubt be broken into unidentifiable pieces in distant undergrowth.* (2)
- *1 × 200g/7oz can* ***tuna fish,*** *drained and broken into large pieces ...* (2)

Likewise, *break down* may mean to 'subdivide something into pieces of a serviceable size' or to 'cause something to disintegrate':

- *In fact,* ***gardens*** *that are long and narrow are among the more simple to design because they can be broken down very easily into separate garden areas or "rooms", each of them having a different function or theme.* (2)

- *The basic biochemical process is one by which **glycogen** in the muscle (a store of glucose) <u>is broken down</u> to release energy which can then be used to drive the muscles.* (2)

I have argued above that phrasal verbs can be seen as making up a subordinate level of categorization elaborating the various readings of *break* as a basic level term (section 3.4.2.1). Actually, reading (2) might be construed as being related to (1) 'damaging physical objects' as a special low-level schema of specific instances. However, as argued above, I have chosen to construe it as a separate reading in which the basic level (represented by *breaking bread*) has been made more or less redundant by the more specific conceptualizations. The category construed as reading **(12) 'specific ways of 'breaking' social institutions & constructs'** includes examples that seem to be motivated by the sensorimotor source domain just described. It draws on the same underlying location event-structure metaphor (section 3.4.2.2), and extensions of all the subtypes are found (cf. examples c-h in Table 9, section 3.4.2.1 above). The rich inferential structure of the sensorimotor domain is utilized, including the potential for alternative construal.

Reading **(3) 'Overcoming physical barriers'** overlaps with reading (1) in that damage to physical objects is involved. What sets this reading apart is a special image-schema in which a physical object like *a blockade, boom, chain, door, fence, gate,* or *tape* is conceived as a 'barrier' that somebody is trying to 'overcome'. Its extension to the nonsen-sorimotor domain, which uses the same image-schematic structure, has been construed as reading **(13) 'Overcoming psychological barriers',** as shown in Fig. 20:

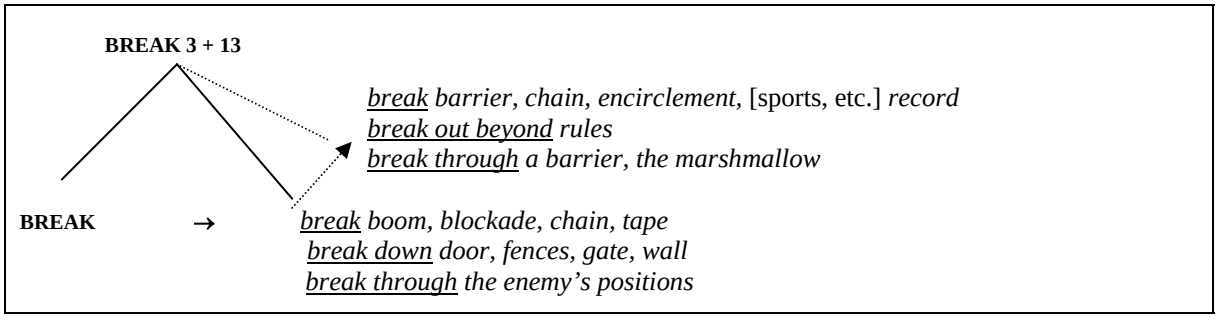


Fig. 20: Overcoming physical barriers & Overcoming psychological barriers

The following is a prototypical example:

- *Britain has broken a vital **barrier** in developing computers that recognise continuous speech rather than needing a keyboard to type in information.* (13)

Combinations with phrasal verbs like *break through* and *break out beyond* code specific instances of this category, as does *break down.* The latter can also be construed as involving 'damage', showing the overlap (and family resemblance) with reading (1).

- *One politician argues that "Arizonans have broken through a **shame barrier**, they now just find it all funny rather than sad."* (13)

- *The idea of the New has always been a crucial part of pop's rhetoric: the idea that the future is going to be an improvement on the past (that sixties' feeling of being on the brink of a whole new order, the beginning of an endless breaking down of **barriers and limits**).* (13)

While it seems plausible to construe *breaking barriers* in the metaphorical sense as an extension from *breaking barriers* in the literal sense, it is less obvious that *breaking limits* and *breaking records* should be so. However, the co-ordination of *barriers* and *limits* in the example above, which does not seem to produce any zeugmatic effect, shows that such less prototypical members can be accommodated, arguably because they are conceived of as similar to more prototypical members.

- *By common consent, Ravenscraig has been kept alive since 1982 because workers have broken all production **records**.* (13)

Along the same lines, I would argue that the expression *breaking a record* can be assumed to evoke a frame in which *record*, meaning 'the best recorded achievement', is conceptualized as a psychological barrier and can therefore also be 'broken'. To the extent that the combination *break* + object evokes the frame of the noun, as seems to be the case for this reading of *record*, it is a possibility to characterize *break* more generally as a support verb.

I agree with the claim made in Cowie (1998a) that collocations mostly do not attract attention to themselves, but examples of creative exploitation of conventional collocations

do occur, even if they are few and far between. In the data for *break*, there is one example of a creative collocation in which two different and incongruous meanings of *break* are evoked at the same time, with a zeugmatic effect:

- *PETER SCUDAMORE, who breaks **records** as easily as a psychotic disc jockey was at it again yesterday, shattering his own mark for the fastest 50 winners by a National Hunt jockey when he rode In-Keeping to an easy victory at Wincanton.* (1 + 13)

As mentioned above (in section 1.5), such an example can be analysed as a blend which integrates input from two different mental spaces. In this example one input space is in the domain of sports, which is reflected in the frame elements and relations which make up the internal structure of this space: a jockey and his horse who break a record in a race. This is the space which refers to the actual topic of the article in question. The other input space is in the domain of discotheques, and the frame elements and relations are a disc jockey who breaks gramophone records. Further details are found in both spaces, for example we learn the horse jockey's name and that he breaks records 'easily' as well as the name of the race, and the disc jockey is characterized as 'psychotic'. Mappings between the elements and relations in the two spaces and selective projections from the two inputs to the blended space are likely to produce a conceptual integration in the mind of the reader of the breaking of sports records and of gramophone records. It thus becomes possible to refer to the former as 'shattering a mark', as if it was a physical object.

The intended effect is presumably to impress on the reader that this was quite an extraordinary performance, and the journalist expresses this by evoking frames that are basically incongruous, not only because they belong to different domains, but also because *to break a* (sports) *record* is a conventional collocation while *to break a* (gramophone) *record* is not. The blend thus forces a compositional, literal interpretation on the conventional expression. The fact that, in spite of the incongruities, readers will have no problem 'running the blend', seems to confirm that it is cognitively plausible to conceptualize the meanings of *break* as being linked together in a network. Prompted by the context, each meaning can be accessed from any other in the process of meaning construction, no matter how great the distance between them.

Reading **(4) 'Opening physical containers'** and reading **(14) 'Opening metaphorical containers'** share flexible image-schematic structure abstracted from the sensorimotor experience of 'containers' such as *houses, eggs* and *piggy banks*. While the basic-level expression *break the pack* can be considered prototypical for reading (4), expressions using the phrasal verbs *break open, break into* and *break out* can be seen as a low-level schema of

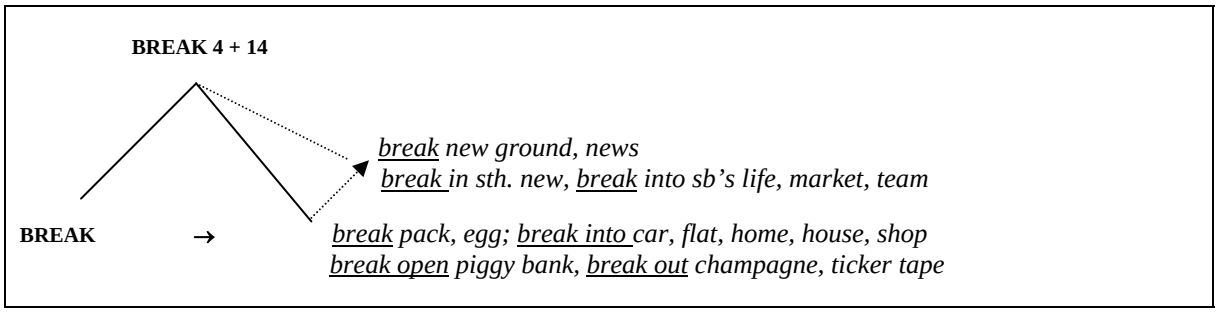


Fig. 21: Opening physical containers & Opening metaphorical containers

specific instances (cf. examples 11-14 in Table 9, section 3.4.2.1 above). In the prototypical example using object event-structure, *break* is followed directly by the object that denotes the container. The phrasal verb *break open* also focuses on the container, specifying that the focus is on getting access to something inside it, while *break into* and *break out* offer the contrasting perspectives of 'entering a container from the outside' and of 'taking something out from the inside of a container', respectively. To *break eggs* refers not only to the breaking of a container, the egg shell, but also implies getting at the contents. In the metaphorical extension *break news*, there also seems to be a merging of container and contents:

- She did not want to have the **news** broken to her carefully. (14)

This is also the case in the idiom *to break new ground*, in the sense of trying out new ideas. It is a lexicalized extension of the institutionalized collocation meaning 'to plough up new land' (cf. sections 2.1.5 and 2.2.2 above) of which no examples are found in the material. Examples of *breaking into* are also found in the nonsensorimotor domain (cf. examples in

Table 9, section 3.4.2.1 above), while there are no metaphorical examples of *breaking something out*; the idea of getting something out of a container so that it is visible and people may know about it[70] is expressed by *break* followed directly by its object as in *break the news* (compare also the intransitive construction *the scandal broke*). Finally, there is one example of *breaking in* used metaphorically, of which there is no equivalent example in the sensorimotor domain:

- It is a breaking in of **something wholly new**: it is an act of creation which parallels the story of creation in Genesis. (14)

It is not quite clear whether *break in* is transitive or intransitive in this example, but the expression "*an act of creation*" seems to tip the scale in favour of the former, since it implies an agent. Readings (4) and (14) can be distinguished from readings (3) and (13) in terms of image-schematic structure. This becomes evident at the subordinate level of categorization where *break in/into, break open* and *break out* are typically associated with the notion of a container whereas *break through* and *break down* are associated with the notion of a barrier. However, overlaps occur reflecting flexible conceptualization, as in the example below, where *break out beyond* can be construed as a blend of the two image-schemas:

- *In the fifties, when the grey-suited corporation man, his corsetted wife right beside him, was so very anxious to conform to every social nicety, all thought of breaking out beyond* ***highly defined rules of social behaviour*** *was severely repressed.* (13)

Here, 'rules' are conceptualized both as a container to escape from and as a barrier to be overcome, not as an object which may be broken, as in reading (11) 'social institutions & constructs'.

I have argued (in section 3.4.2.1) that it is not unlikely that the breaking of bones is conceptualized in terms of objects like sticks or branches. At the same time, we often conceptualize artefacts such as furniture in body terms:

[70] Cf. the conceptual metaphor UNDERSTANDING IS SEEING, in Lakoff and Johnson (1980: 48).

- *The upper rooms were silent when he finally went to his own room, but before getting into bed he pushed his wedge of broken **chair-leg** under the door to keep the Bogeyman out.* (1)

Considering that all artefacts start as ideas, it is not so strange that body terms are used to name the parts of objects, especially items designed to meet the needs of the human body, such as tables, chairs and beds. On the other hand, we have much more experience of

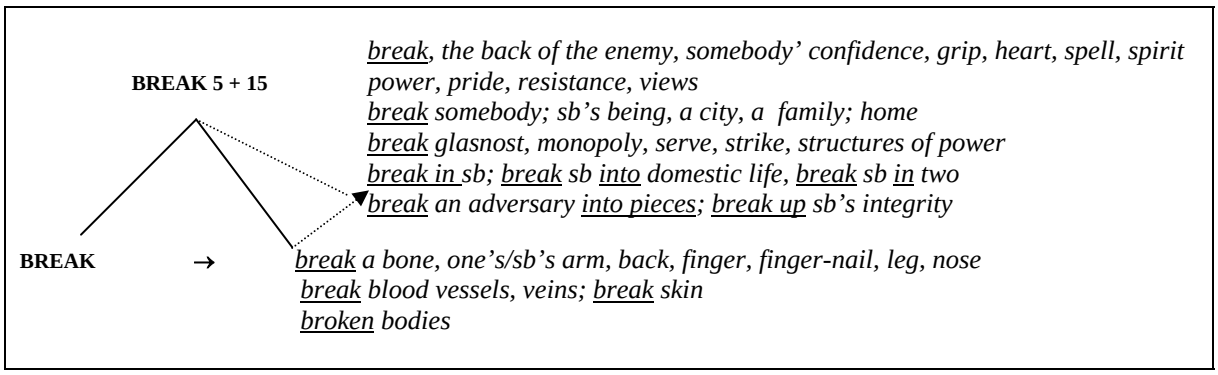


Fig. 22: Damaging body parts & Causing psychological damage

handling things, so it would seem plausible that bodily damage should be conceptualized in terms of breaking things. The reason for positing **(5) 'damaging body parts'** as a separate, reading from artefacts and natural things, in spite of this conceptual interrelationship, is that it belongs to a different domain, and that terms like *arms* and *legs* are ambiguous rather than vague with regard to the two readings. To make sense of the expression *broken arms and legs*, we have to choose one meaning and discard the other, and it would seem odd if, in the same sentence, one term referred to body parts and the other to parts of a piece of furniture.

- *Looking through the metal bars, hypnotized by the revolving wheel, he could see his own body, broken **arms and legs** flopping grotesquely, bouncing from girder to girder on its way down.* (5)

In the domain of body parts, *break* is used not only about bones, but also about *skin*, referring to an injury from the outside by cutting or penetration, and about *veins* and *blood vessels*, referring to bursting because of pressure from the inside:

- A wound has been defined as an injury which breaks both the outer and inner **skin** - a bruise or a burst blood-vessel in an eye would not amount to a wound. (5)

As implied by the example, *broken* does not mean the same in the expression *broken skin* as in *broken blood vessels* (or, by analogy, *broken veins*), and in neither expression does it mean the same as in *a broken arm*. Coordinating the nouns in paratactic constructions would have a zeugmatic effect, and substituting nouns for one another would amount to commutation, so that three different readings need to be posited for 'damaging body parts': (a) 'fracturing', (b) 'cutting/penetrating', and (c) 'causing to burst'. The expression *broken body* is vague with regard to the type of injury suffered, although it is hard to imagine a broken body without any fractured bones.

Reading **(15) 'Causing psychological damage'** is construed as an extension mainly from (5), especially (a) 'fracturing', which I consider to be the prototype of (5). What is psychologically 'broken' is categorized by means of body part terms like *back, backbone* and *heart* as well as terms related to functions of the human body (*grip, views*), the human mind (*spirit, resistance*), and human influence (*monopoly, strike*). Although real hearts are not very much like bones, metaphorical hearts may still be conceived of as being fractured, rather than bursting or being cut. (This seems to be supported by the way people normally draw a broken heart as a heart that is broken in two, while the piercing of a heart by an arrow normally symbolizes 'falling in love'). The lexical evidence of mappings can be supported by parallels in construction type, as both readings show extensive use of possessive pronouns and genitive constructions (see also examples 15-16 and n-p in Table 10, section 3.4.2.1 above):

- *She had broken **her neck***. (5)
- *And you - you're sitting there breaking **my** fucking **heart**.* (15)
- *The fourth-seeded American looked sluggish as **his serve** was broken twice in the first set ...* (15)
- *... low sales would indicate that **consumer's confidence** is broken ...* (15)

In 28 of the 39 examples of reading (5), or 72%, the object is preceded by a genitive or a possessive pronoun, while the corresponding numbers for reading (15) are 17 out of 53 examples, or 32%. Besides, 7 of the examples of reading (15) have of-constructions and in 2 further examples the object is premodified by an adjective or an adjectival noun. These two kinds of constructions, which are exemplified below, also relate the entity that is conceptualized as being 'broken' to the 'body' that it is part of. Adding them in brings the total for reading (15) to 26 or 49%:

- *Michael Caine plays a cockney Don Juan who "don't believe in making anyone unhappy, not if you don't have to do it" but is forever breaking* ***the hearts of young girls*** *...* (15)
- *Eliminating tax breaks on corporate entertaining would probably do more to break* ***the power of keiretsu*** *than beefing up Japan's feeble antitrust laws...* (15)
- *... our work will not end when we have broken* ***the grip of the six major killer diseases****.* (15)
- *The advent of satellite television and the imminent breaking* ***of what has been a BBC monopoly*** *of cricket offers opportunities of expanding TCCB income.* (15)
- *Table Tennis: Mason breaks* ***Soviet spirit*** (15)

By comparison, readings (1) and (11) have only 9% and 13% of the object types mentioned (*the car had* ***its windows*** *broken*, *A has broken* ***his contract***), while objects are mostly in the definite or indefinite form: *somebody has broken* ***the glass door***, *I have never broken* ***the law***), which I construe as showing a difference between the prototypes for the two sets of readings as well as overlap between them.

The decision to construe terms like *glasnost*, *monopoly*, *strike*, and *structures of power* as belonging to reading (15) rather than (11) 'social institutions & constructs' is partly based on the syntactic evidence and partly based on the intuitive judgement that, in combinations with the nouns listed, *break* is found to express the thwarting of human endeavours rather than failure to observe a rule or to meet a commitment. *Breaking* does not seem to mean the same in *breaking a strike* as in *breaking a rule*, or in *breaking an agreement*.

In the evidence of metaphorical extension from reading (5) to (15), a further parallel is found where the 'physical breaking' refers to the whole body and the 'psychological breaking' refers to the whole person:

- *Then I would be ashamed, seeing **his little body** broken on the rocks below.* (5)
- *After having been broken by torture, **the poor Nepomuk** was bound and thrown into the Vltava from Charles Bridge.* (15)
- ***He** was broken with compassion as he watched her crashing like a falling star, pulled by gravity and centripetal forces towards the consummation of the sun ...* (15)

At the subordinate level of categorization, no transitive phrasal verb constructions were recorded for reading (5)[71], while (15) includes examples of *break in/into/in two* and *break up*:

- *The young **conscript** was broken in slowly.* (15)
- *Ernie had a handsome black **mare** for sale that he had bred and broken in himself.* (15)
- *... he said that Warnie and Mrs Moore liked each other, "and, I hope, as **W**. gets broken into domestic life, they may come to do so still more ..."* (15)
- *she felt as if **she** had been broken in two and glued back together again all wrong.* (15)
- *Sex, which breaks up **our integrity**, **our single inviolability**, **our deep silence** ...* (15)

For specific ways of conceptualizing 'psychological damage', it seems that inferential structure has been found not in the source domain of 'damaging body parts', but in subdomains of 'artefacts and natural things'. Thus *break up* and *break in two* have also been recorded for reading (12) 'Specific ways of breaking social institutions & constructs' and reading (14) 'Opening metaphorical containers', which were construed as extensions from the corresponding readings in the sensorimotor domain, (2) and (4). This may be taken as a reminder that it is easy to underestimate the dynamic nature of a network model. Although, as a general principle, readings in the nonsensorimotor domain can be seen as

[71] Intransitive constructions include *break down* (people, immune system) and *break out* (*... she was eating fruit and breaking out all the time as a result ...*).

extensions from specific readings in the sensorimotor domain, alternative conceptualizations are always possible.

In the sensorimotor domain, *break* is used about interrupting an activity, **reading (6)**, and starting an activity, **reading (7)**. **Readings (16)** and **(17)** are construed as categorizing interruption and change not of physical action itself, but of the way it is performed.

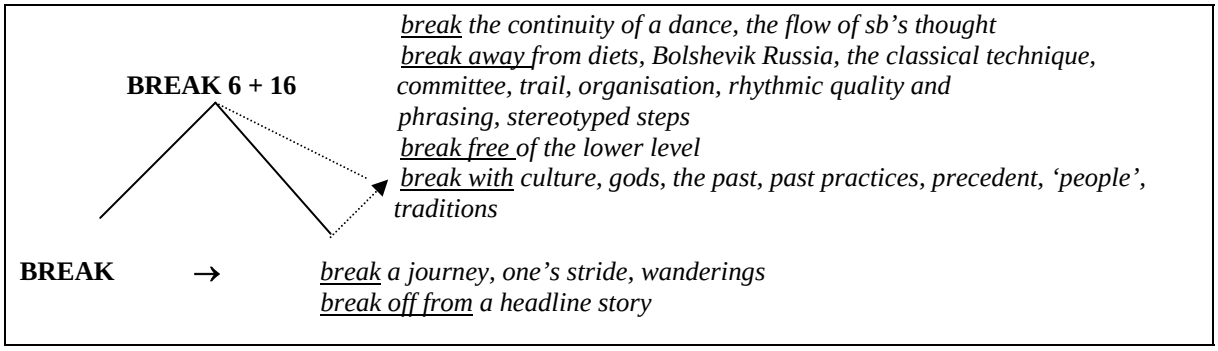


Fig. 23: Interrupting an activity & Changing from a way of doing something

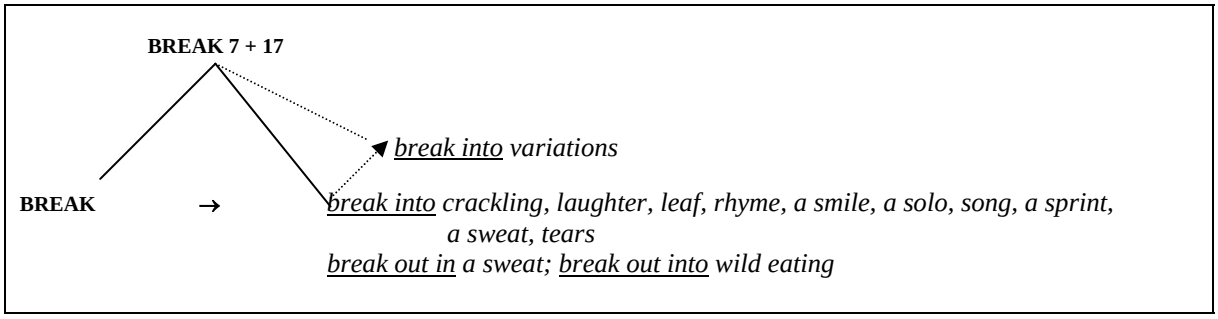


Fig. 24: Starting an activity & Changing to a new way of doing something

Readings (6/16) and (7/17) differ in that the former refer to the interruption of an ongoing activity and the latter to the starting of an activity. The difference between sensorimotor and nonsensorimotor domain readings is quite subtle and hard to perceive. It seems clear that *break into a sweat* is a physical activity, but is *break into rhyme* really more 'physical' than *break into variations (of the tarantella)*? And is *break a journey* more 'physical' than

break the continuity (*of the dance*)? Probably the differences are not cognitively salient, and it does not seem plausible to posit different readings for the two domains.

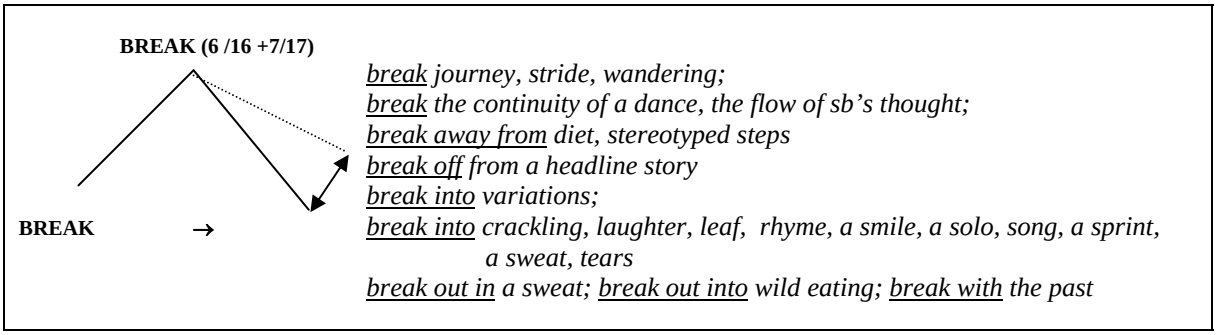


Fig. 25: Conflation of readings (6/16) Interrupting an activity / Changing from a way of doing sth. with (7/17) Starting an activity / Changing to a new way of doing something

Moreover, if the meaning of *break* in these examples is held to be vague as regards the difference between 'changing to' and 'changing from' a way of doing things, a further conflation of readings is possible as shown in Fig. 25 above, in which the conflation is indicated by means of the solid, double-headed arrow. Also, since *break* alone can only mean 'interrupt' and not 'start' an activity, it does not seem plausible, from a semasiological perspective, to posit the latter as a separate reading. An onomasiological perspective, brings out the differences in image schematic structure at the subordinate level of categorization. While location event-structure is found in both (6/17) and (7/17), the underlying conceptual metaphors are different. In the case of readings (6/16), 'Interrupting an activity' or 'Changing from a way of doing something' is typically conceptualized as deviating from a path, based on the underlying metaphor LIFE IS A JOURNEY:

- *So there is little chance of our news readers breaking off, US-style,* ***from a headline story*** *to extol the virtues of a brand of paint, or even a national newspaper.* (6)
- *But as the Tory wagon-train breaks away* ***from Mrs Thatcher's trail****, Mr Major's gentler style inevitably means slacker reins.* (16)

By contrast, in all the examples of (7/17), 'Starting an activity' or 'Changing to a new way of doing something' the image-schema of a path was combined with that of a container conceptualized as the 'location' of the 'activity' or 'new way':

- *"I thought he had come to take you away, Jim;" she said, <u>breaking into</u>* ***tears****.* (7)
- *The whole cast muster in the market place before <u>breaking into</u> extravagant* ***variations*** *of the tarantella..* (17)

The last three pairs of readings code sensorimotor experience in 'the physical environment' while extensions code nonsensorimotor experience in 'the mental landscape' (cf. examples 20-23 and u-z in table 10, section 3.4.2.1 above). Thus reading **(18) 'Changing a pattern'** is construed as an extension of **(8) 'Interrupting auditory and visual experience'** with *break the rhythm* as a prototypical example.

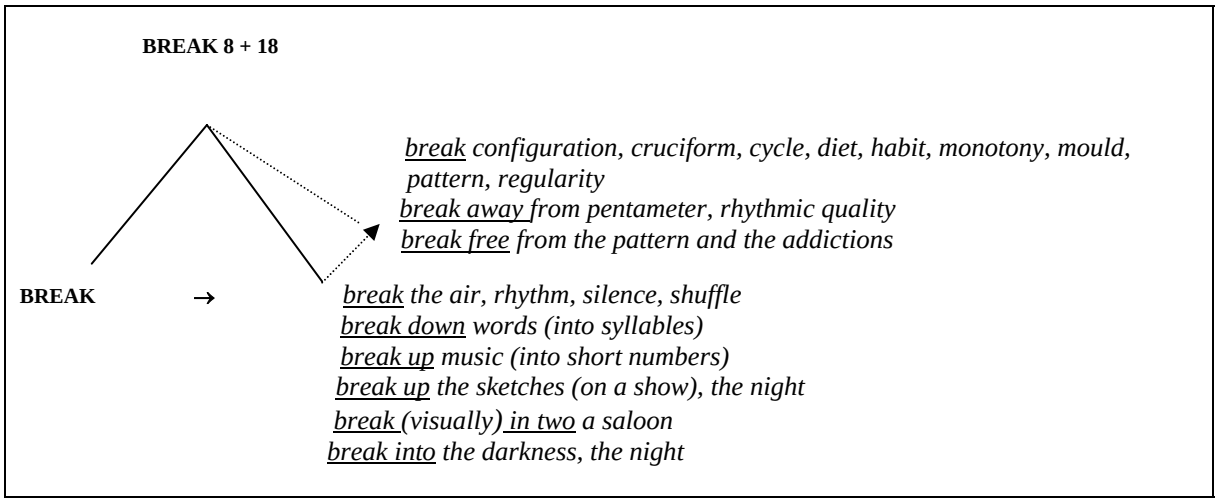


Fig. 26: Interrupting auditory or visual experience & Changing a pattern

The auditory experience may be associated with experience of making a noise by drumming one's fingers, stamping one's feet, singing, or playing an instrument, or it may be associated with the experience of observing similar behaviour in others. The use of *break* about the interruption of a pattern also in 'the mental landscape', in combinations with *habit, pattern, regularity*, and *monotony* seems to be motivated by its use in the realm of physical experience, even if the auditory element is not mapped (in line with what was said about the invariance hypothesis at the beginning of this section) or is not very salient.

That examples of sensorimotor experience also use image-schematic structure is apparent from the following example of visual experience:

- *The first-class **saloon** is visually broken in two, less tube-like, the gangway's position changing half way through.* (9)

Motivation for expressions coding interruption of a pattern of behaviour also seems to come from the basic domain of 'artefacts and natural things' as in combinations with *break up* and *break down*, and in the expression *break the mould*:

- *But if Chapman's death is to provoke any analysis it might be to ask whether the Pythons achieved their original aim of breaking **the mould** of English television comedy, doing away with the studio applause, the guest star, the musical, even on occasion the punchline.* (18)

Mould is a physical object with the function of producing identical copies by the process of casting and therefore has come to be associated with a pattern as in the example above. Here elements follow each other in a certain 'rhythm', which becomes apparent if the sentence is read aloud. Moreover, interrupting a pattern may be conceptualized as interrupting physical movement along a given path (compare reading (17) above).

- *.. when modernist poetry* [...] *breaks away from **the pentameter**.* (18)

To account for the variability that alternative conceptualizations exemplify, it is necessary to combine a semasiological approach asking what domains are categorized by *break* with an onomasiological approach asking how meaning in these domains is expressed by means of *break*.

The conceptual metaphor underlying the extension from **(9) 'Becoming visible'** to **(19) 'Entering consciousness'** is SEEING IS UNDERSTANDING (or SEEING IS KNOWING) with *breaking surface* as an example of how a physical barrier to vision can be conceptualized as a psychological barrier to understanding (cf. examples 21 and w in Table 10, section 3.4.2.1).

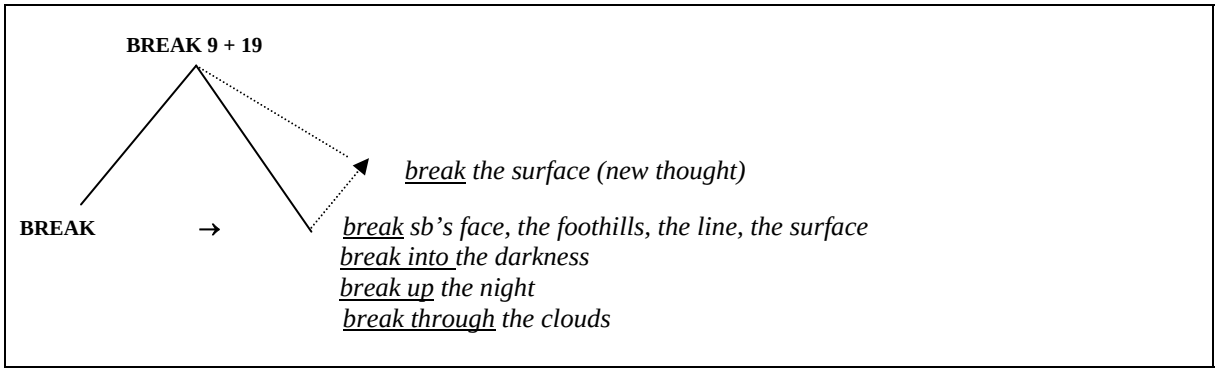


Fig. 27: Becoming visible & Entering consciousness

The low-level schema of specific instances shows how image-schematic event structure links reading (9) to subdomains of 'artefacts and natural things'. Thus *break up* can be linked to reading (2) 'breaking physical objects in specific ways', while *break through* and *break into* can be linked to readings (3) 'overcoming physical barriers' and (4) 'opening physical containers' (cf. Table 9 in section 3.4.2.1 above). There are examples in which both trajector and landmark are visual elements, as in *the sun breaks* (*through*) *the clouds*, as well as examples in which visual and auditory experience are mixed in that the trajector is a sound, or combines visual and sound elements (*noise; explosions and flashes*) while the landmark is visual:

- *It must be nearly dawn, for there are more traffic noises breaking into* ***the darkness*** *outside* (9)
- *... explosions and flashes were breaking up* ***the Baghdad night****.* (9)

Finally, readings **(10) 'Interrupting a physical state'** and its extension **(20) 'Interrupting a psychological state'** refer to situations that are conceptualized as lasting for a certain period of time, until they are interrupted or 'broken', and which are typically associated with a characteristic kind of 'inactivity' (cf. examples 22-23 and x-z in Table 10, section 3.4.2.1).

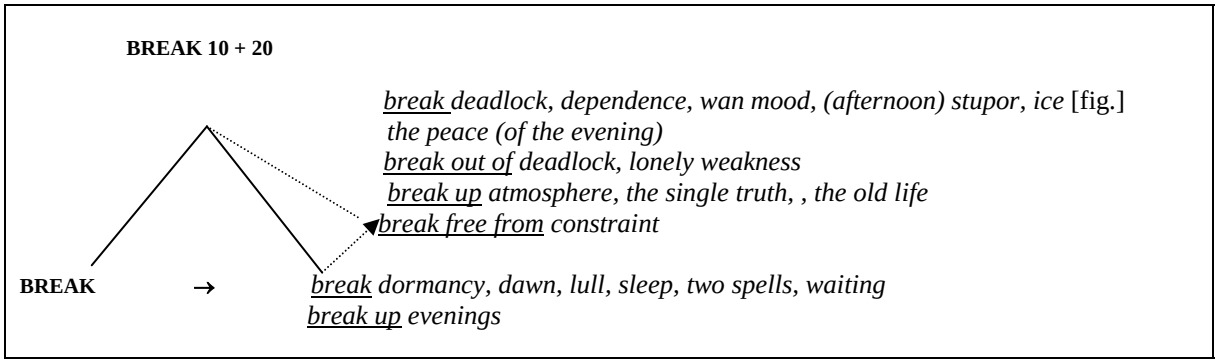


Fig. 28: Interrupting a physical state & Interrupting a psychological state

In the examples construed as reading (10), objects like *dawn* and *evening* are conventional terms for times of the day characterized in relation to the rising and setting of the sun (therefore they might also be associated with 'visual experience'; cf. reading (8)), *dormancy* and *sleep* refer to a physical state of plants and mammals respectively, and *lull* and *waiting* define a period in relation to some other state or activity (*a lull in the storm; waiting for his arrival*), whereas *spell* simply means 'period of time during which something lasts'[72] (*a spell of bad weather*). In reading (20), objects have positive or negative connotations so that interruption is conceptualized as either harmful or beneficial:

- *Erika laughed and Omi smiled, breaking* ***her wan mood****.* (20)
- *The only sound that breaks* ***the afternoon stupor*** *is the gentle pounding of the pastry roller ...* (20)
- *It is 1900 hours when* ***the peace of the evening*** *is broken by the cough of the Challenger engine ...* (20)
-

States of mind are conceptualized in terms of periods of time, which makes it possible to draw inferences about duration and interruption. In addition to the image-schematic structure inherited specifically from reading (10), the low-level schema of specific combinations shows that structure is recruited from other parts in the network for alternative conceptualizations in both domains. Thus *break up the evenings* (10) and *break up the old*

[72] The definition is from Cowie1989.

life (20) is construed in terms of 'specific ways of breaking physical objects' (2), and *break out of deadlock* (20) is construed in terms of 'opening physical containers' (4).

Also in the idiom *break the ice*, inference structure comes from the domain of 'artefacts and natural things', as appears from this variation:

- *"Fancy a game of darts, lad?" Jos said to Mungo, <u>breaking</u> what was left of **the ice**.* (20)

It is plausible, if not predictable, that *ice* should be used for drawing inferences about a tense situation, which is often associated with a feeling of coldness and stagnation,[73] and that relieving the tension should be conceptualized in terms of 'breaking' it. Apart from the image-schematic structure involved in mappings within the network, general object and location event-structure can be seen as providing underlying motivation.

On the basis of the discussion above, the list of readings proposed in Fig. 16 at the beginning of this section has been adjusted as indicated in the **revised proposal** in Fig. 29 below. For reading (11) 'Violation of social institutions & constructs', which was construed as an extension of reading (1) 'Damaging physical objects', two readings have been posited: (a) 'Failure to observe' and (b) 'Cancellation'. For reading (5), 'Damaging body parts', three readings have been posited: (a) 'Fracturing', (b) 'Cutting/penetrating', and (c) 'Causing to burst'. It was further suggested that readings (6) 'Interrupting an activity', (16) 'Changing from a way of doing something', (7) 'Starting an activity', and (17) 'Changing to a new way of doing something', should be conflated to one reading, as shown in Fig. 25 above.

After the revision, there are still 20 readings, which are high-level schemas abstracted from expressions at the basic level of categorization, with low-level schemas of specific instances. A hypothesis of how the posited readings of *break* relate to each other will conclude this part of the study of BREAK. The hypothesis assumes the network model, but

[73] This may be related to an underlying metaphor that allows relations between people to be conceptualized in terms of temperature, e.g. a relationship may be warm or cold, and a person may be *frozen out*.

SENSORIMOTOR DOMAIN	**NONSENSORIMOTOR DOMAIN**
(source domain)	**(target domain)**
1. Damaging physical objects	11. Violation of social institutions & constructs (a) Failure to observe (b) Cancellation
2. Specific ways of breaking physical objects	12. Specific ways of breaking social institutions & constructs
3. Overcoming physical barriers	13. Overcoming psychological barriers
4. Opening physical containers	14. Opening metaphorical containers
5. Damaging body parts (a) Fracturing (b) Cutting; penetrating (c) Causing to burst	15. Causing psychological damage
6. Interrupting an activity 7. Starting an activity 8. Interrupting auditory and visual experience	16. Changing from a way of doing something 17. Changing to a new way of doing something 18. Changing a pattern
9. Becoming visible	19. Entering consciousness
10. Interrupting a physical state	20. Interrupting a psychological stat

Fig. 29: Readings of *break*: revised proposal

Fig. 30 below is not meant to represent the network as such, but to illustrate the principles of extension that are seen as linking the different readings. The judgments underlying the hypothesis are based on the analysis of the linguistic data, and evidence from a psycholinguistic study might reveal a somewhat different pattern. One principle of extension is seen as operating within the sensorimotor domain where two central readings are construed as source domains for the others. The former, (1) 'Damaging physical objects', is seen as motivating extensions to readings (3-5), in which the 'objects' are conceptualized as 'barriers', 'containers', and 'body parts'. The latter, (6/7) 'interrupting/starting an activity', is seen as motivating extensions to readings (8-10), in which the 'activity' is conceptualized as auditory and visual experience, and as the experience of states. Reading (2), which is also construed as an extension of (1), mainly codes 'specific

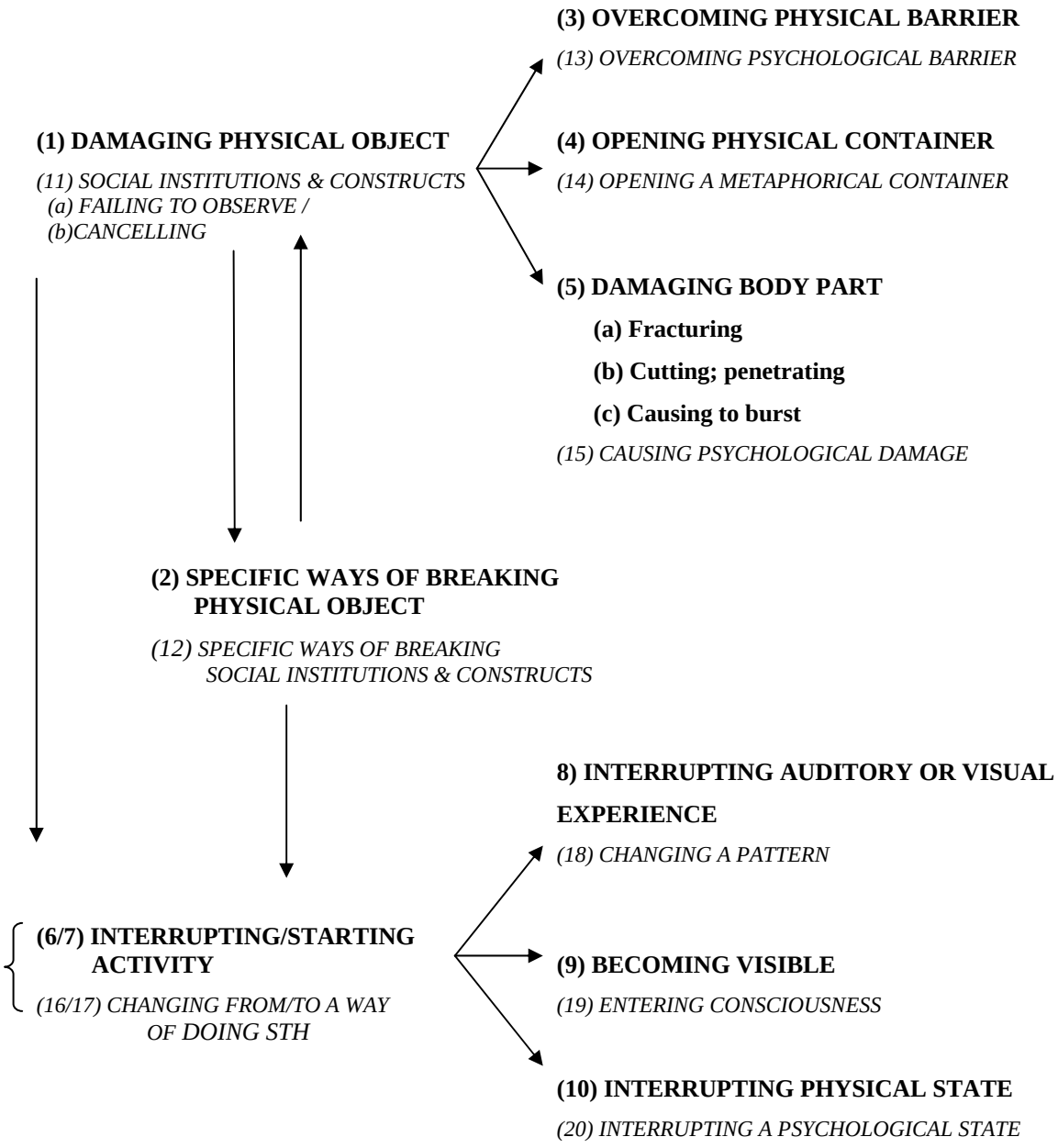


Fig. 30: BREAK as a complex category

ways of breaking physical objects', which are mostly conceived of as purposeful or neutral as regards their effect. As indicated by the arrows, the phrasal verbs that dominate this reading are found in the low-level schemas of specific instances throughout the network, although some phrasal verbs are related to specific readings.

As a second principle, readings in the sensorimotor domain (printed in bold type in Fig. 30) are systematically extended to readings in the nonsensorimotor domain (shown in italics below them). Between them, these two principles allow image-schematic, inferential structure from the basic sensorimotor domain readings to be used for less basic sensorimotor readings as well as for nonsensorimotor readings. While these two principles may account for relations between prototypical examples of the readings proposed, the overlaps between them can be explained by taking a diachronic perspective of the synchronic evidence. Since lexical categories grow up gradually based on judgements of similarity with prior uses, overlaps, or 'family resemblances', are to be expected.

However, a further factor seems to be at work, which can be called the principle of 'alternative conceptualization', which allows image-schematic structure to be recruited from all over the network, as long as the invariance hypothesis is respected – or failure to do so is justified by situational context and communicative purpose. This phenomenon, which is documented by phrasal verb expressions at the subordinate level of categorization, is taken as evidence that the network model is suitable for characterizing the complex and dynamic nature of a lexical category like BREAK. Also creative exploitation of conventional collocations can be explained by appealing to the network model.

The analysis started by asking what domains are categorized by *break* (a semasiological approach) and found that a first hypothesis could be made according to which a range of readings in the sensorimotor domain were systematically extended to readings in the nonsensorimotor domain using the same image-schematic structure, which was furthermore seen as reflecting general event structure. In the final part of the analysis a revised hypothesis was offered and examples were assigned to the readings now stipulated. This

involved including an onomasiological perspective to account for the variability in image-schematic structure found within each reading.

3.4.3 The internal structure of APPOINTMENT

In this section, I will turn to the analysis of the APPOINTMENT category following the same principle as in the analysis of BREAK, before the composite structure integrating the two categories is analysed in section 3.4.4.

The Collins Cobuild English Dictionary (Sinclair 1995) divides words into 'frequency bands' based on their occurrence in The Bank of English, a corpus of over 200 million words of written and spoken English. It puts *break* in the highest band, indicated iconically by means of five black 'diamonds' (p. xiii of the introductory matter), which means that it is in the same league as "common grammar words" such as *the, of,* and *to,* and other very frequent vocabulary items, e.g. *like, go, paper, return.* This band includes about 700 words, which are said to account for 75% of all English usage. By comparison, *appointment* has only been allocated three diamonds as one of about 1,500 words the knowledge of which "extends the range of topics which you can talk about".

As mentioned above (in section 3.3.2), the BNC includes over 10,000 examples of *break,* of which approximately 9.5% have been analysed in this study. A search for singular and plural forms of *appointment* produced 908 concordances, all of which are included in the data. At the editing stage, a number of concordances have been deleted, because they only have a code but no content, or because they occur twice in identical form; on the other hand some have been duplicated if two co-ordinated verb forms occur with *appointment(s)* as the object, as in the example *Several appointments were made and broken before he and his wife finally arrived.* After editing, the file contains 931 records. As appears from Fig. 31 below, the construction transitive verb + nominal object accounts for 46% of the 931 examples. The verbs in this construction type are processual predications, which are the

base for the alternative profiles found in constructions with a nominalized verb (2%), or with the past participle as premodifier (1%) or postmodifier (2%) of *appointment* (cf. 3.4.2.3 above). Together these four construction types account for 51% of the examples. Of the 17% of examples in which *appointment* is recorded as subject or subject complement, 42% have a copula verb. Other intransitive constructions and transitive constructions each make up about 20%, while 18% of the examples in this group consist of a nominal group without a verbal predicate. Furthermore, the group in which *appointment* is recorded as the premodifier in compounds accounts for 4% of the examples, while the group in which the noun is the complement in a prepositional phrase makes up 28% of the total.

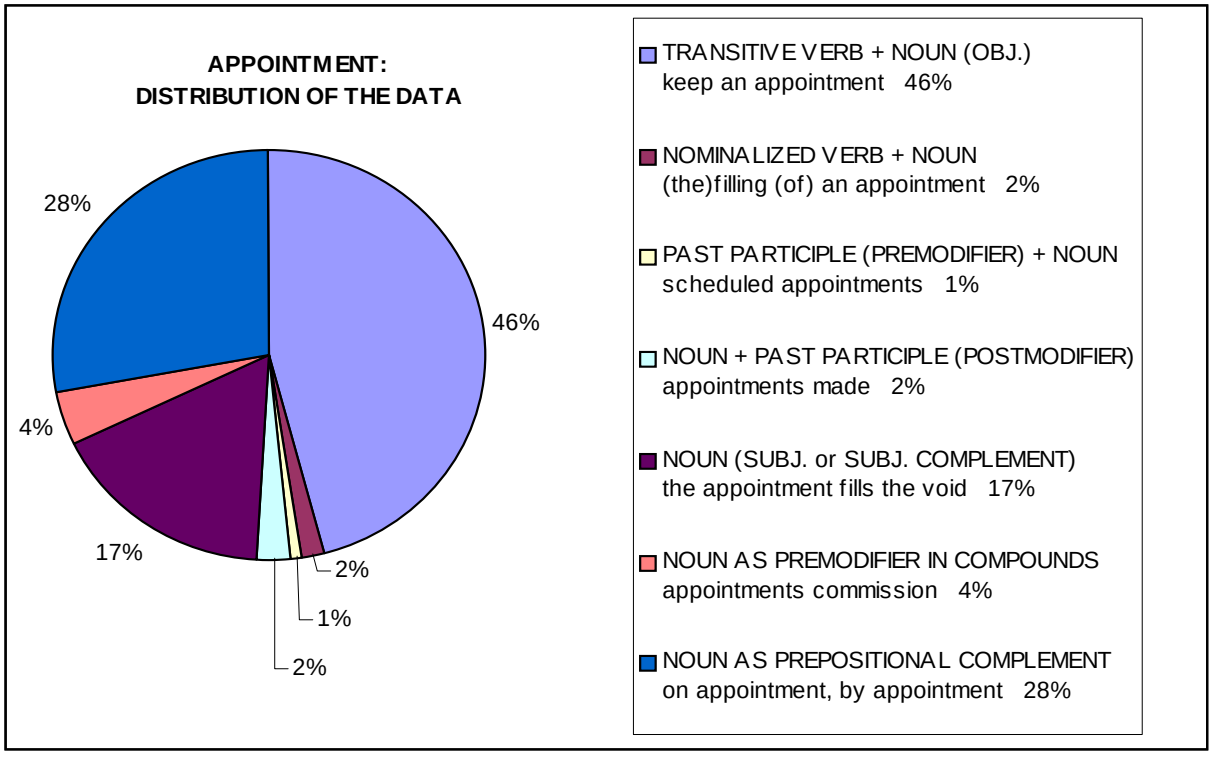


Fig. 31: Distribution of data by construction type

As in the case of BREAK, the analysis of APPOINTMENT as a complex lexical category will include the aspects of domain (and subdomain), image-schematic structure, and construction type of combinations with forms of *break*. How many readings can be posited

for *appointment* (as a linguistic form) will be linked to the number of frames that can be identified. In constructions profiling processual relations, the verb (or the nominalized verb) will be the basis of categorization; in addition prepositional phrases with *appoint-ment* as a complement will be analysed, as they are expected to be a good hunting-ground for image-schematic structure (cf. section 3.4.3.2 below).

The data are seen as 'synchronic evidence of diachronic processes' (cf. section 2.3.2 above), which implies that the evidence is *not* diachronic. Diachronic knowledge may be included in the discussion, such as the fact that *appointment* is a deverbal noun derived from the verb *appoint*, which is in turn derived from the prepositional phrase: *at/to a point* (cf. footnote 75 below), but it is not the purpose of the study to trace the etymology of verb and noun, and it ignores meanings that have gone out of use. It is therefore understood that a diachronic study would be likely to find motivation that is not apparent from the synchronic data, just as a synchronic study may assume motivation where historically there was none. However, seeing current readings as resulting from diachronic developments makes it possible to construe them as a coherent network which determines the meaning potential of the category for current users.

3.4.3.1 Abstract domains and referential range

As a deverbal noun, *appointment* is used both as a noun of action and as a noun coding the result of action in the nonsensorimotor domain, and in addition it is used to refer to artefacts in the sensorimotor domain. While *break* as a relational predication and a delexical verb is semantically dependent on the noun that it profiles, *appointment* as a nominal predication, is relatively independent. However, as noted before, (in section 2.2.4) autonomy/dependence asymmetry is relative, thus collocation with verbal and often also other co-textual elements are required to distinguish between the different APPOINT-MENT frames that the noun evokes (cf. section 3.4.3.4 below).

Fig. 32 below is a first hypothesis about the internal structure of APPOINTMENT based on 'aboutness' or domain. A comparison with Fig. 9 (section 3.4.2.1) shows that *appointment* is used to categorize a much narrower range of experience than *break*, viz. 'social institutions & constructs' in the nonsensorimotor domain, accounting for 99% of examples, and 'artefacts' in the sensorimotor domain, accounting for just 1% (only plural forms).

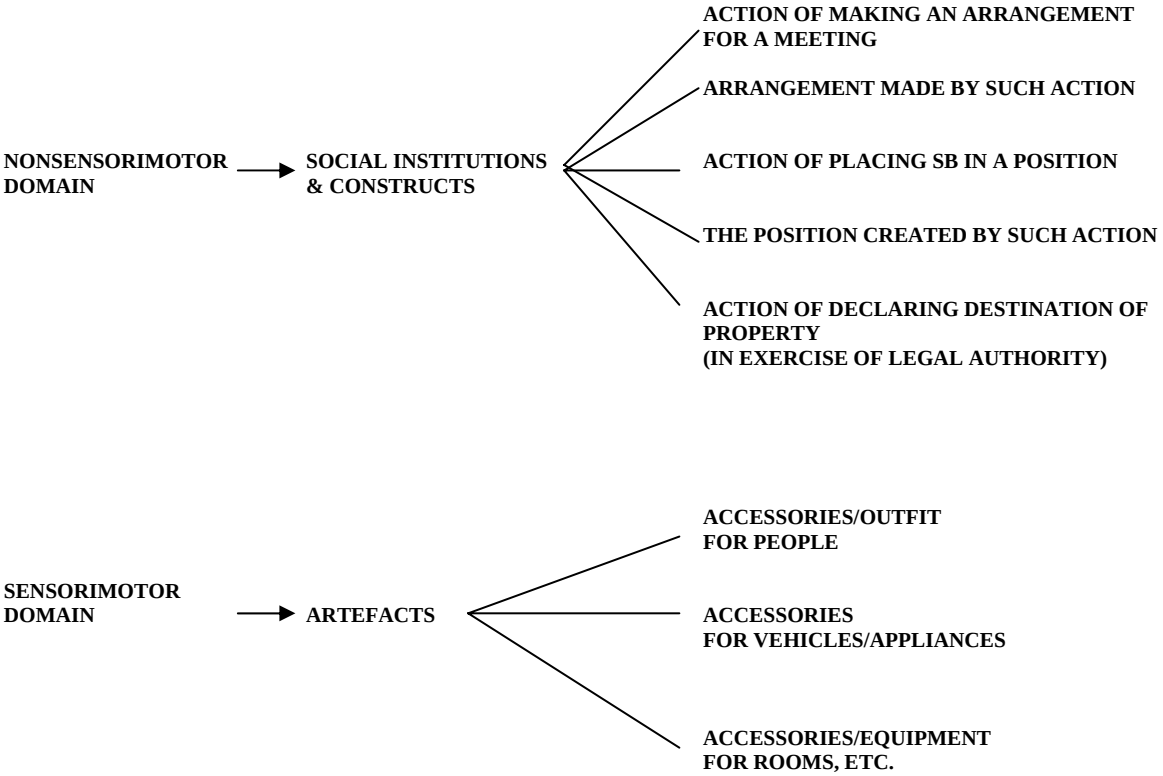


Fig. 32: APPOINTMENT domains

Table 16 below shows how, according to the data, the potential of *appointment* for categorizing both action and the result of action is exploited in the nonsensorimotor domain, while in the sensorimotor domain *appointment* categorizes equipment with specific functions.

NONSENSORIMOTOR DOMAIN		SENSORIMOTOR DOMAIN
SOCIAL INSTITUTIONS & CONSTRUCTS		ARTEFACTS
ACTION	RESULT OF ACTION	EQUIPMENT
arranging a meeting 1) *She <u>made</u> **appointments** to meet me at different places: restaurants, art galleries.*	**arrangement for a meeting** 2) *He's already <u>broken</u> three **appointments** to see me.*	
placing sb in a position *3) On March 23 she <u>made</u> the first **appointments** to a new government.*	**position** *4) As he never <u>held</u> a hospital or university **appointment**, his work was slow to gain recognition.*	**accessories/outfit for people in a position** 5) *Then you had to <u>produce</u> your **appointments** -- which was your baton and brass keys for the police telephone boxes, hold them in your hands, and then you were issued with the Beat Book on parade.*
action of declaring destination of property (in exercise of legal authority)[74] *8) Her father's will also made provision for the destination of this life interest should she fail to <u>make</u> any such **appointment**: the money would be payable to a number of her distant relatives.*		**accessories for vehicles/appliances** *6) In June 1964; The Autocar's testers remarked; "In its **appointments** the Austin-Healey 3000 MkIII **is** now more of a touring car than a sports car.*
		accessories/equipment for rooms 7) *Cleo <u>exclaimed over</u> the rooms and their **appointments** with delight.*

Table 16: APPOINTMENT domains

Whereas in the BREAK category, nonsensorimotor uses were seen as extensions from sensorimotor uses, motivation in the APPOINTMENT category seems to flow in the opposite direction. The meaning 'placing somebody in a position' (example 3 in Table 16) may be understood as having first been extended to mean 'the position' itself (example 4), which may have motivated a further extension to denote the 'accessories' or 'outfit' needed for a position (example 5), and even further to denote 'accessories' not only for

[74] This definition is adapted from the OED.

people but also for cars, homes, etc. (examples 6-7). Alternatively, uses in the two domains might be seen as unrelated, an example of homonymy rather than polysemy.[75]

Subdomains categorized by *appointment* cover a range of experience, apart from 'action'. In the sense 'arrangement for a meeting', *appointment* may thus refer more specifically to a point of time, a period of time, the place, record or purpose of a meting (cf. Fig. 33 below). Besides, the 'arrangement' itself may be conceptualized as a kind of 'instrument' or 'tool', as when you are allowed to see somebody *by* or *with an appointment* (cf. Table 20 below).

- **'meeting as a point of time '**
  "But all of it," the Dean says firmly, checking his watch for his next ***appointment,*** *" all of it is to the glory of God.*

- **'meeting as a period of time'**
  His nurse had brought the paper back from her lunch break and left it lying on the low coffee table during his ***appointments*** *with his two previous patients.*

- **'meeting as a place'**
  .. after ten minutes of chat we leave for our ***appointment****, promising to rejoin them later.*
  During the week, I often shuttle easily to ***appointments*** *in central London by car.*

- **'meeting as a record'**
  One of the ***appointments*** *we found in Miss Morgan's diary was with a Miss Huntley.*

- **'meeting as a purpose'**
  ... attendance for ***appointments*** *(including tutorials and examinations)*

Fig. 33: Subdomain: social institutions & constructs ('arranging/arrangement for a meeting')

These specific meanings can be understood to be like the different active zones of *window* discussed above (cf. Fig. 5 in section 3.4.1.1). I will return to the notion of active zones in the discussion of how many readings (or frames) should be posited for *appointment*, in section 3.4.3.4 below.

[75] According to the OED, the English verb *appoint* originates from the Old French verb *apointe-r*, *-ier* from *à point*: 'to the point', 'into condition', but one of the old uses of the verb, 'to put into proper state or condition', is related to an old sense of the noun *appoint*: 'array', 'equipment' and French *en bon point*.

Also in the case of the subdomain 'position', *appointment* includes a range of specific aspects apart from the action, viz. the position as such, the person appointed to it, job content and period of employment, as well as source of income:

- **'the position as such'**
  .. the present Government can barely find enough Members from Scottish constituencies to fill the ***ministerial appointments*** *in the Scottish Office.*

- **'person appointed to the position'**
  .. now, new ***appointments*** [...] *no longer have the same job security.*

- **'content of the position'**
  Medical ***appointments*** *in military units were believed to provide useful experience .*

- **'period of employment'**
  It is hoped that it will be possible to continue the post after ***the*** *first* ***appointment*** *so that there can be a series of three year* ***appointments***

- **'position as source of income'**
  the fortunate few who enjoyed a measure of affluence were chiefly to be found in towns where lucrative guild ***appointments*** *were available.*

Fig. 34: Subdomain: social institutions & constructs ('position')

In many cases it is not possible to say for certain whether an example of usage refers to the action of placing somebody in a position or to the position itself, and sometimes both seem to be involved at the same time:

- *Failure to secure an* ***appointment*** […] *would produce lasting enmity more certainly than the gift of patronage could ensure permanent friendship.*

- *I would have to lose 10 teaching* ***posts****,* ***and*** *several ancillary* ***appointments*** *we've made to allow teachers to concentrate on teaching.*

In the second example above, *lose*, as well as co-ordination with *post*, implies that the item *appointments* refers to the positions as such, whereas *we've made* implies that it refers to the action of placing people in these positions. In the example below, *appointment* also codes action and result of action at the same time, as *made* implies the arranging of a meeting and *broken* refers to the resulting arrangement for a meeting:

- *Several* ***appointments*** *were made and broken before he and his wife finally arrived.*

Finally, the following example could belong to any of the domains or subdomains categorized by *appointment*, depending on whether the co-text or situational context is about appointments for cars, people, meetings, jobs, or about the lawful right to receive a sum of money:

- .. *don't* <u>*make*</u> *the* ***appointments****, you see*.[76]

While I construe *shuttle to an appointment* and *leave for an appointment* (examples given in Fig. 33 above) as the specific instances making up the low-level schemas of the APPOINTMENT category, combinations with verbs like *make, break, keep, lose, hold, give, take* and *fill* are seen as the basic level of categorization. Some of these verbs are used in more than one domain, and they provide important clues to the image-schematic structure of the APPOINTMENT category, which will be discussed next.

3.4.3.2 Image schemas and event structure

In this section I will focus on the two subdomains 'arranging/arrangement for a meeting' and '(placing sb in) a position', which both include action and the result of action. The other domains 'action of declaring destination of property' and 'accessories', each accounting for about 1% of the examples, will be briefly commented on at the end of the section (cf. Table 16 in section 3.4.3.1). Tables 17 and 18 below give an overview of transitive verb + nominal object constructions in the two subdomains, based on a distinction between the basic level of categorization and a subordinate level.

At the **basic level**, we find *make, keep, break* and *hold* and other highly schematic verbs that are expected to fill the slots in semantic frames 'by default' (cf. section 3.4.2.3 above and 3.4.3.4 below). Combinations between these basic level verbs and *appointment* are construed as prototypical examples of entrenched collocations (cf. section 3.4.4.2 below).

[76] In this case, the BNC, in the version I used, does not provide the full sentence (cf. section 3.3.3 above).

It should be noted, however, that it is not automatically assumed that combinations with basic level verbs are entrenched collocations. At the **subordinate level**, the less schematic codings are construed as low-level schemas of specific instances. In some cases, no basic level verb has been recorded, but only examples coding more specific experience, i.e. *terminate* and *create appointments* in Table 18.

No examples have been recorded as belonging to the **superordinate level**. Although verbs like *arrange, attend, cancel,* and *obtain* are quite general in meaning, I have construed them as belonging to the subordinate level, because I see them as more specific than the basic level verbs. Thus *to arrange an appointment* is seen as elaborating *make an appointment* by adding an element of 'detailed planning' and 'organization', and *to attend an appointment* is seen as elaborating *keep an appointment* by adding an element of 'regular activity' or of 'presence throughout an arrangement'. Likewise, *to cancel an appointment* is seen as elaborating *to break an appointment*, by adding an element of 'formal correctness'.[77] Moreover, *to obtain an appointment* may be seen as adding the information that a certain 'effort' is required in order to 'get' the appointment, whereas *to attain an appointment* implies 'succeed in getting', and *to secure an appointment* implies 'obtain with difficulty'. Finally, *to receive an appointment* may be the expression closest to *to get an appointment*, but seems to stress that the appointment has been received 'from somebody', which 'get' does not seem to do.[78]

[77] At the same time, these verbs have been considered to be sufficiently general in meaning to be used as an ad hoc technical term to characterize the meaning of basic level verbs; thus *make an appointment* has been defined as 'arranging a meeting' in this subdomain of APPOINTMENT.

[78] The attempt to account for the elaborations are partly based on definitions in the Oxford Advanced Learners Dictionary (Cowie 1989), except in the case of *cancel*, where I have relied exclusively on my own intuitive interpretation of the data.

LEVEL	APPOINTMENT: 'ARRANGING/ARRANGEMENT FOR A MEETING'					
BASIC prototype, default case most schematic	***make***	***give***	***get***	***have***	***keep***	***break***
SUBORDINATE not prototypical less schematic more specific	*arrange, book, fix*	*offer*			*attend*	*cancel* *miss*
least schematic most specific	*do, plan, rearrange, schedule, time, juggle around, confirm (in writing), note down, put in a diary, record*		*accept, ask for, call for, telephone for, claim, need*	*cope with, go through, find, look at, sort out, show, print out*	*leave for, get to, get there, arrive for, cycle to, shuttle to, cover, rush through*	*arrive late for, forget*

Table 17: Levels of categorization: APPOINTMENT I

LEVEL	APPOINTMENT: '(PLACING SB IN) A POSITION'									
BASIC prototype, default case most schematic	***make***	***stop***			***fill***	***give***	***get***	***take***	***hold***	***lose***
SUBORDINATE not prototypical less schematic more specific	*handle*	*veto* *block*	*create*	*terminate*		*offer*	*attain, make* (= *get*), *obtain, receive, secure*	*accept, assume, take up, take on, undertake*	*have got, occupy, fulfil, serve in*	
least schematic most specific	*advertise, announce, propose, advise on, affect, agree on, comment on, react to concur in, control, deal with, discuss, finance, focus on, initiate, insist on, go ahead with, complete, work on, review, sit in on, oversee, supervise, take action on, take over, keep in sb's hands, politicize, ring-fence, accept, allow, confirm, ratify, recommend*	*be opposed to, complain about, condemn, botch, refuse, suspend*				*fax out, shower on sb*	*seek, claim, covet, interest oneself in*			

Table 18: Levels of categorization: APPOINTMENT II

A distinction has been attempted between 'more' and 'most specific' to indicate that we have a continuum covering a range from most schematic to most specific reflecting the variability of conceptualization. While it is not predictable which, or how many, combinations will be entrenched, it is argued in this study that combinations with basic level verbs are most likely to be so, while the most specific expressions are less likely to be entrenched. Mostly, they are ad hoc combinations designed to fit the needs of a particular usage situation. The following examples illustrate the continuum (cf. Table 17 above):

- *She made **appointments** to meet me at different places: restaurants, art galleries.*
- *"They asked for an **appointment**, so I arranged **it** for this afternoon.*
- *However, people cancelled meetings and re-arranged **appointments** with remarkable good humour.*

While a psycholinguistic study will be needed to test these ideas about entrenchment, relative frequencies of combinations provide linguistic evidence in some cases. Thus the 141 occurrences of transitive verb + object constructions[79] for *appointment* in the domain 'arranging/arrangement for a meeting', include 43 different verbs, of which only one third (14), which account for almost four fifths of all occurrences, were recorded more than once, and two thirds (29), which account for the last one fifth of occurrences, were recorded only once. Those which occur only once include some that are not all that specific such as *attend, confirm, miss*, and *receive*. The pattern is similar for the 257 occurrences of transitive verb + object constructions for *appointment* in the domain' (placing sb in) a position': a little less than a third of the 88 different verbs (25), which account for almost three fourths of all occurrences, were recorded more than once, while over two thirds (69), which account for a little less than on quarter of occurrences, were recorded only once. For the most frequent verbs, such as *make, hold*, *have*, and *keep*, frequency can probably be seen as reflecting institutionalization or entrenchment, but it does not seem safe to conclude that frequencies below a certain rate imply the opposite (cf. section 3.2.1 above).

[79] Constructions in which the past participle form is pre- or postmodifier are not included in these figures.

Subdomain of APPOINTMENT	**'arranging/ arrangement for a meeting'**	**occ.**	**'(placing sb in) a position'**	**occ.**
	14 verbs **33%**	**112** **79%**	**25 verbs** **28%**	**188** **73%**
Verbs that occur more than once	*make* *have* *keep* *cancel* *arrange* *book* *call for* ***break*** *get* *ask for* *give* *accept* *fix* *offer*	41 15 11 10 6 5 4 **4** 4 3 3 2 2 2	*make* *hold* *confirm* *give* *announce* *receive* *secure* *comment on* *obtain* *offer* *control* *lose* *accept* *advertise* *attain* *approve* *deal with* *fill* *fulfil* *get* *influence* *lead to* *suspend* *vet* *veto*	91 20 11 7 6 5 5 4 4 3 3 3 2 2 2 2 2 2 2 2 2 2 2 2 2
	29 verbs **67%**	**29** **21%**	**63 verbs** **72%**	**69** **27%**
Verbs that occur once	*attend, claim, come back to, confirm, cope with, cover, do, find, follow, forget, go through, include, juggle around, look at, miss, need, note down, plan, print out, put in one's diary, put in writing, re-arrange, receive, reschedule, rush through, send, show, sort out, time*		*add, advertise, agree, agree on, allow, base on, be attributed to, be in the hands of, be opposed to, block, cause, claim, complete, condemn, congratulate – on, consider preferable to, cover, discuss, fax, fill, finance, focus on, follow, give rise to, go ahead with, handle, initiate, insist on, interest oneself in, introduce, issue, keep in the hands of, know, limit, list, make* [= 'get'], *make – attractive, monopolise, need, note, obstruct, offer, omit, outnumber, oversee, perceive, place, precede, process, propose, provide for, ratify, recommend, record, refuse, regard – as, reserve, review, ring-fence, seek, shower on sb, sit in on, stop, subject – to vetting, supervise, surround, take over, use, work on*	
TOTAL	**43 verbs**	**141**	**88**	**257**

Table 19: Frequencies of co-occurrence in two APPOINTMENT domains

It has been argued (in section 3.4.2.2 above) that metaphorical mappings can be seen as operating both between and within domains in the network of a specific lexical category and between underlying, general event-structure patterns and their instantiations. As an example of the former, the use of *break* in the sensorimotor domain to refer to 'artefacts' was seen as a motivation for its use in the nonsensorimotor domain to refer to 'social institutions & constructs'. Although no examples of *appointment* were found in the data for *break*, it can be categorized together with other nouns coding the 'violation of social institutions & constructs' (cf. Fig. 18, section 3.5.1.2.5). The fact that a range of other basic level verbs from the sensorimotor domain of 'artefacts', e.g. *make, give, get, take, have, keep, fill,* and *hold,* are also used in the APPOINTMENT category is seen as evidence of the claim that this category uses inferential structure from that particular source domain.

The systematic use of such verbs broadens the perspective, so that *break an appointment* can be seen as an example of general object event-structure cutting across many different lexical categories (cf. Fig. 13, section 3.4.2.2 above). However, as predicted by the invariance hypothesis (cf. section 3.4.2.4 above), mappings from source to target domain need to respect the image-schematic structure inherent in the target domain (cf. Figs 36 and 37, this section, above). Fig. 35 below gives an overview of mappings from the sensorimotor domain of 'artefacts and natural things' to the two subdomains 'arranging/arrangement for a meeting' and '(placing somebody in) a position'.[80] While some of the basic level verbs are associated with only one of the domains (*break, fill, hold*), most occur in both (*make, give, get, take, have*), so that the context needs to be taken into account to determine which domain is referred to. In four cases, no mapping has been recorded, which does not mean that it could not occur in a given context. Thus *to keep an* appointment referring to a position would mean 'not lose it' while *to hold an appointment* referring to an appointment for a meeting could refer to the literal holding of a document.

[80] The convention of indicating mappings from source to target domain by means of an arrow is taken from Lakoff and Johnson 1999. As an alternative to the arrow notation, the target domain meaning can be stated first, before indicating the source domain that it is expressed in terms of, e.g. 'Journey → Purposeful Life', or 'A Purposeful Life is a Journey' (Lakoff and Johnson 1999: 61-62).

SOURCE DOMAIN: 'artefacts and' natural things' (A.)		TARGET DOMAIN I: 'arranging/arrangement for a meeting' (M.)	TARGET DOMAIN II: '(placing sb in) a position' (P.)
MAKE A.	→	ARRANGE M.	PLACE SB IN P.
KEEP A.	→	COME TO M.	-
BREAK A.	→	NOT COME TO M.	-
FILL A.	→	-	PUT SB IN P.
HOLD A.	→	-	HAVE/BE IN P.
GIVE A.	→	AGREE TO MEET	ACCEPT FOR P.
GET A.	→	OBTAIN PERMISSION TO MEET	BE ACCEPTED FOR P.
TAKE A.	→	-	ACCEPT P.
HAVE A.	→	HAVING OBTAINED PERMISSION TO MEET	HAVING BEEN ACCEPTED FOR P.

Fig. 35: Object event-structure in two APPOINTMENT domains

The specific instances at the subordinate level of categorization give an impression of the range of experience that basic level verbs are seen as being schematic for, in the two subdomains (cf. *make* in Tables 17 and 18 above). The verbal element may not suffice to determine which domain a given example belongs to, so that it is necessary to take the situational context, or more co-text, into account.

- *Phone call from a parent saying she cannot make next week's appointment. Could we postpone it until the following week? Juggle the **appointments** around and manage to fit her in.*

The findings so far are in line with the invariance hypothesis (cf. section 3.4.2.4 above), which predicts that, although image-schemas come with specific inferential structure from a source domain, this structure is adjusted to the target domain. It is in the target domain that words evoke the frames which enable us to understand what they mean (cf. section 3.4.3.4 below). Also prepositional phrases with *appointment* as the complement show that the APPOINTMENT category has image-schematic structure of its own. As mentioned above (in section 3.4.3), constructions in which *appointment* is the complement of a

preposition were found in almost one third of the examples in the data. Furthermore, constructions in which *appointment* is postmodified by a prepositional phrase were found in almost 10% of the examples. In terms of figure/ground organisation, *appointment* corresponds to the landmark (ground) of the preposition in the former type and to the trajector (figure) of the preposition in the latter (cf. section 3.4.2.3 above). The distinction between count nouns and mass nouns (in section 3.4.2.3) is also involved.

	SUBJECT	**PREDI-CATOR**	**ADVERBIAL PHRASE** (with ***appointment*** as **dependent)**					**ADVERBIAL PHRASE** (postmodifying ***appointment***)	
	noun phrase trajector➔	**verbal**	**prep.**	**pos.**	**art.**	**adj.**	**nominal ←landmark trajector➔**	**prep.**	**noun phrase ←landmark**
1.	*We*	*leave*	*for*		*our*		*appointment*	*with*	*the dentist*
2.	*I*	*shuttle*	*to*				*appointments*	*in*	*central London*
3.	*Information*	*will be given*	*at*		*the*	*first*	*appointment*		
4.	*Our client's wife*	*was unable to* ***attend***	*in*			*two*	*appointments*	*at*	*the hospital*
5.	*The paper*	*was left on the*	*during*	*his*			*appointments*	*with*	*his two patients*
6.	*She*	*lingered in the waiting room*	*after*	*her*			*appointment*		
7.	*Consultations*	*are*	*by*				*appointment*		
8.	*Visitors*	*are welcome*	*with*		*an*		*appointment*		
9.	*Visitors*	*may not call*	*without*				*appointments*		
10.	*They*	*had to* ***keep***		*their*			*appointments*	*for*	*treatment sessions*

Table 20: Prepositional phrases in the APPOINTMENT domain 'arranging/arrangement for a meeting'

The examples in Table 20 above and in Table 21 below illustrate how prepositional phrases reflect image-schematic structure in the two subdomains of APPOINTMENT. Table 20 shows how in the subdomain 'arranging/arrangement for a meeting', both location and object-event-structure are used in prepositional phrases (cf. section 3.4.2.2 above). In examples 1-4 and 6, *appointment* corresponds to the landmark of the prepositions *for* (20 occ.), *to* (7 occ.), *at* (1 occ.), *in* (1 occ.), and *after* (1 occ.), which indicates that it is conceptualized as a 'desired location, or destination'. This conceptualization partly conflates TIME and SPACE (cf. Lakoff and Johnson 1999: 137 ff.) as *for*, *at*, *in* and *after* may refer to both, whereas *to* refers only to a location in SPACE, and *during* (1 occ.)

refers only to a period of TIME. It should be noted that only one of the 20 examples of *for*, in the combination *leave for an appointment*, construes *appointment* as a 'destination'; most typically *for* occurs in combinations like *ask/call/telephone for an appointment*, in which *appointment* is construed as a 'desired object' (cf. section 3.4.2.2. above).

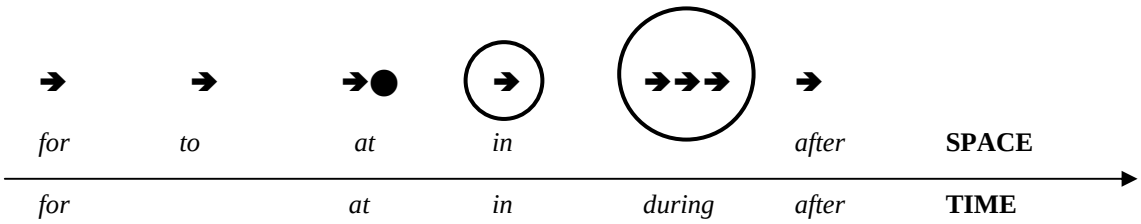


Fig. 36: Location event-structure in APPOINTMENT subdomain ('arranging/arrangement for a meeting')

This is illustrated in Fig. 36 above, where arrows symbolize the trajector (the person conceptualized as moving towards the place of appointment) and circles symbolize *appointment* as corresponding to the landmark of the prepositions. The different size of circles is meant to indicate that the prepositions place the trajector at varying distances from the landmark, so that there is a kind of 'zooming in' effect. The conceptualization of *appointment* as a location is emphasized in cases where it is related to specific locations, corresponding to the landmarks of the prepositions *with*, *in*, and *at* (examples 1-2 and 4-5 in Table 20). These postmodifying expressions may also follow expressions in which *appointment* is conceptualized as an object, as in *to get/have/keep an appointment with somebody*.

In examples 7-9 of Table 20, *appointment* corresponds to the landmark of the prepositions *by* (8 occ.), *with* (2 occ.), and *without* (2 occ.), indicating that object event-structure is being used to conceptualize *appointment* as a sort of 'instrument'. In combinations in which *by* is followed directly by *appointment* (example 7), I construe *appointment* as a mass noun, which profiles the component states of the process of *appointing* in a summary and general fashion (cf. section 3.4.2.3 above). This is supported by its uncountable form.

Whereas in such combinations with *by* the process of 'appointing' is still clearly there as part of the base, the process is completely backgrounded in combinations using *with* or *without* (examples 8-9), in which *appointment* is conceptualized as an object.[81] It occurs

	SUBJECT	**PREDI-CATOR**	**ADVERBIAL PHRASE (with *appointment* as dependent)**					**ADVERBIAL PHRASE** (postmodifying ***appointment***)	
	noun phrase trajector➔	**verbal**	**prep.**	**pos.**	**art.**	**adj.**	**nominal ←landmark trajector➔**	**prep.**	**noun phrase ←landmark**
11.	*They*	*got the job*	*by*				*appointment*		
12.	*His salary*	*was increased*	*on*				*appointment*		
13.	*The first meeting*	*was held*	*after*	*their*			*appointment*		
14.	*They*	*had **made***				*senior*	*appointments*	*from*	*external sources*
15.	*The president*	*makes*					*appointments*	*to*	*the Federal government*
16.	*The general*	*was*	*among*	*his*			*appointments*		
17.	*British officers*		*in*			*ex-change*	*appointments*	*at*	*US military establishments*
18.	*She*	*has **held***				*teaching*	*appointments*	*at*	*several universities*
19.	*The government*	*changed*			*the*	*top*	*appointments*	*in*	*broadcasting*
20.	*They*	*secured*					*appointments*	*as*	*veterinary surgeons*

Table 21: Prepositional phrases in the APPOINTMENT domain '(placing sb in) a position'

only in the countable form with the indefinite article or plural ending, and can also be construed as a count noun, because it refers to the result of action only, without any reference to component states.[82]

Also in the subdomain '(placing sb in) a position', prepositional phrases reflect event structure, as shown in Table 21 above. Here, we find *appointment* as the landmark of *by* (4 occ.), *on* (6 occ.) and *after* (6 occ.) coding the action of 'being appointed' in a summary

[81] In Swan (1995: 117), the difference between *by* and *with* is explained like this: "We use *by* when we talk about an action - what do we **do** to get a result. We use *with* when we talk about a tool or other object - what we **use** to get a result."

[82] This differs from the examples of count nouns given in section 3.4.2.3 above, in which, following Langacker 1987, for example *cable break* was characterized as a count noun, because it designates only a single episode of the process.

fashion (examples 11, 12 and 13). In these combinations no plural forms are found, and the preposition is either followed directly by the noun, or the noun is preceded by a genitive, or a possessive pronoun, which corresponds to the subject or object in a processual predication. However, *after* also occurs in constructions in which *appointment* is a count noun denoting the result of action:

- *It is hoped that it will be possible to continue the post after the first* ***appointment*** *so that there can be a series of three year appointments.*
- *After* ***appointments*** *in community service, she had joined the Navy.*

Whereas combinations with *by* cast the noun in an instrumental role ('with the action of') and exemplify object event-structure, combinations with *on* and *after* give it a temporal profile. Since, as shown in Fig. 36 above, the conceptualization of TIME and SPACE may be conflated, these examples (12-13) are construed as location event-structure. In examples 14-20 *appointment* corresponds to the trajector of *from* (3 occ.), *to* (49 occ.), *at* (5 occ.), *in* (11 occ.), and *as* (12 occ.). In combinations with *to*, which account for over 60% of the examples, as well as with *from*, the noun is typically preceded by a possessive pronoun or genitive, or succeeded by an of-construction:

- *The earliest sources give no clue to the date of Molla Fenari's* ***appointment*** *to the Muftilik*
- *The Great War also caused the* ***appointment****, for the first time because of the absence of men in the Forces, of lady teachers to the staff.*

These examples show one of the entailments of location event-structure, namely that causation can be conceptualized as 'forced movement (to or from) locations' (cf. section 3.4.2.2 above). In this case, *appointment* expresses the exertion of the moving force, and the object being moved is expressed as the possessive pronoun, genitive form, or complement in the of-construction, while the complement of *to* is the destination. In the corresponding processual predication with *appoint* as a finite verb, it is the verb that expresses the exertion of the moving force, while the object moved is a proper noun, and the destination is still expressed as the complement of the preposition *to*:

- *Ralph Neville **was appointed** to that office, and several of his descendants secured a like **appointment** ..*

In other words, the location event-structure reflects the fact that the noun is derived from the transitive verb *appoint.* All the examples are in categories for which *make* is construed as the prototypical basic level verb (cf. the two first columns of Table 18 above).

In combinations with *as,* the noun is also often preceded by a possessive pronoun or genitive, or succeeded by an of-construction, and although the preposition does not indicate motion, an element of causation can still be perceived, as the adverbial phrase *as* + noun emphasises the deverbal nature of *appointment,* especially if there is an adverbial of time to emphasize that the statement describes an event:

- *But **his appointment** as steward of the royal household in November 1318 was both a snub to Lancaster, who claimed the right of appointment, and ..*
- *The clampdown was continued by **the appointments of Sir John Nott-Boyer** as Commissioner of the Metropolitan Police* [...] ***and a new Home Secretary**, **Sir David Maxwell-Fyfe*** [...]

As the landmark in combinations with *among* and *in*, *appointment* is conceptualized *as* a location (examples 16 and 17), and as the trajector in combinations with *at* and *in* (examples 17, 18, and 19) it is conceptualized as being *in* a location. In examples 18 and 19, the use of the verbs *hold* and *change* supports a construal in terms of object event-structure: *appointment* as an object in a location.

Summing up the position as far as image-schematic structure is concerned, I would claim that the use of verbs like *make, break, keep* and *hold* show that the two domains 'arranging/arrangement for a meeting' and '(placing sb in) a position' use image-schematic structure from the sensorimotor domain of 'artefacts and natural things'. As the extension of such basic level verbs to categorize experience in the nonsensorimotor domain is pervasive, their use in the APPOINTMENT domain can be seen as an example of object event-structure in general. In accordance with the invariance hypothesis, actual mappings from source to specific target domains are restricted by the internal structure of each

domain, as shown in Fig. 35. Prepositional phrases are also seen as reflecting inherent image-schematic structure as illustrated in Tables 19 and 20. *Appointment* is thus conceptualized both *as* a location and as an object *in* a location, and as a force moving an 'object' *to* and *from* locations. In other words, the APPOINTMENT category can be characterized as having inherited as well as inherent event structure.

The two subdomains which have not been discussed so far exemplify *appointment* as action and result of action respectively (cf. Table 16 in section 3.4.3.1 above). In a legal context, *appointment* is used about the action of 'declaring destination of property', and its image-schematic structure is location event-structure with the entailment that causation is forced movement to a location as in the case of 'placing sb in a position'. Finally, the use of *appointment* in the sense 'accessories' refers to actual physical objects. Still, a trace of location-event structure may be preserved to the extent that the object symbolizes that somebody has moved, or has been moved to, a location, for example in order to take up a post.

3.4.3.3 Construction types

In this section the analysis of construction types carried out for the BREAK category (in section 3.4.2.3 above) will be repeated for the two largest subdomains of the APPOINTMENT category, 'arranging/arrangement for a meeting' and '(placing sb in) a position', which both code nonsensorimotor experience involving 'social institutions and constructs'. The purpose is to test whether the pattern appearing from the data for *break* is also found in the data on *appointment*, which include combinations with 155 different verbs.

The results of the analysis of construction types in the two APPOINTMENT domains are shown in Table 22 below and can be compared with the results for 'violation' in the non-sensorimotor domain of 'social institutions & constructs' in the BREAK analysis (Table 14 in section 3.4.2.3 above). **The first two types of construction,** active and passive forms of

processual predications taking *appointment* as object, account for over 80% of examples in both APPOINTMENT domains: 86% in the case of' arranging/arangement for a meeting' and 83% in '(placing sb in) a position'. These shares are on the same level as the 87% for the same construction types in the BREAK domain of 'violation'. In the latter domain no examples were found of **the third type of processual predication**, in which the object of *break* as a transitive verb appears as the subject of *break* as an intransitive verb. This construction was only found in the sensorimotor domain of 'damage', as in *The cable has broken.* In the two APPOINTMENT domains, *appointment* is the subject of an intransitive verb in almost 7% of examples (23 occ.), but only in one example, in which the verb is *increase*, can the intransitive construction be seen as a figure/ground reversal of a transitive construction taking *appointment* as object:

- Media-only **appointments** increased at a faster rate than those involving creative and full service during the year.

In this example it is understood that it is the *number* of appointments that increases. There are no examples of intransitive uses of basic level transitive verbs like *keep*, *hold*, and *break*, such as **the appointment held*, or **the appointment broke*.

Turning now to atemporal relations, **the fourth construction type**, past participle + noun, occurs in close to 6% of the examples in the 'meeting' domain and in less than 1% of examples in the 'position' domain'. In the BREAK analysis the corresponding figure for the 'violation' domain was close to 6%. By comparison, 74% of examples in the 'damage' domain were of this type. For **the fifth construction type**, noun + past participle, percentages in the APPOINTMENT domains are higher than in both BREAK domains and

NONSENSORI-MOTOR DOMAIN	SOCIAL INSTITUTIONS & CONSTRUCTS:			SOCIAL INSTITUTIONS & CONSTRUCTS:		
	APPOINTMENT: 'ARRANGING/ARRANGEMENT FOR A MEETING'			APPOINTMENT: '(PLACING SB IN) A POSITION'		
PREDICATION TYPE:	**CONSTRUCTION TYPE:**	**NO.**	**%**	**CONSTRUCTION TYPE:**	**NO.**	**%**
	1. TRANS. VERB + NOUN (OBJ.), ACTIVE	83	70.33	**1. TRANS. VERB + NOUN (OBJ.), ACTIVE**	189	57.27
PROCESSES	*We will keep all appointments on the day notified to you.*			*.. don't make the appointments you see.*		
	2. TRANS. VERB + NOUN (SUBJ.) PASSIVE	19	16.10	**2. TRANS. VERB + NOUN (SUBJ.), PASSIVE**	84	25.45
	Several appointments were made and broken before he and his wife finally arrived.			*A number of blacks were given government appointments.*		
	3. NOUN (SUBJ.) + INTRANS. VERB	4	3.38	**3. NOUN (SUBJ.) + INTRANS. VERB**	23	6.96
	Hospital appointments seemed to go on for ever..			*.. sub-cabinet appointments* [..] *went; as far as possible; to candidates who were both competent and ideologically sound.*		
	4. PAST PARTICIPLE + NOUN	7	5.93	**4. PAST PARTICIPLE + NOUN**	2	0.60
ATEMPORAL RELATIONS	*The resented traffic jam, the heated political argument, the broken appointment, can cause strong feelings of resentment.*			*The nomenklatura system of party-controlled appointments is being dismantled.*		
	5. NOUN + PAST PARTICIPLE	3	2.54	**5. NOUN + PAST PARTICIPLE**	19	5.75
	Rose and Dora were full of their "projects" and the appointments arranged for the following morning.			*.. this led to appointments made for political rather than academic reasons.* *The appointment held until 1670 associated him with an energetic department.*		
	6. NOMINALIZATION OF VERB	2	1.69	**6. NOMINALIZATION OF VERB**	13	3.93
NOMINAL PREDICATIONS	*The timing of all subsequent follow up appointments and endoscopies in the healing phase was discretionary.*			*.. the holding of outside appointments such as consultancies must be approved by faculty boards.*		
	TOTAL	118	99.97		330	99.93

Table 22: Construction types in two APPOINTMENT domains

higher for the 'position' domain (at almost 6%) than for the 'meeting' domain (at a little less than 3%). As pointed out in connection with the BREAK analysis, this construction, which is construed as a complex atemporal relation, can be seen as a reduced relative clause with a transitive verb. Finally in the case of nominal predications, **construction type 6** in both Table 22 and Table 14, the percentages found in the APPOINTMENT domains, at under 2% in the 'meeting' domain and close to 4% in the 'position' domain, were slightly lower than in the BREAK domain of 'violation' which was almost 6%, about the same as in the 'damage' domain. What I find more interesting than the actual percentages, however, is that in the APPOINTMENT domains, as in the 'violation' domain of BREAK, the only type of nominalization found were mass nouns preceded by the definite article and followed by an of-construction, e.g. *the creation, holding, making,* or *timing of an appointment* (cf. Table 15, section 3.4.2.3 above).

The analysis shown in Table 22 above, which includes all the 155 different verbs recorded, thus produces results (referred to in rounded figures) that are comparable to those for the nonsensorimotor domain of 'violation' in the BREAK analysis. Table 23 below shows the results of the same analysis applied to the ten verbs construed as belonging to the basic level of categorization. For these verbs the domination of processual relations is even more pronounced. Between them, active and passive forms of the transitive verb + object construction account for 212 of the 232 occurrences, or 91%, while there are no examples in which *appointment* as subject is combined with an intransitive form of one of these verbs. Atemporal relations account for 18 occurrences, or almost 8%, including 2, a little less than 1% in which the past participle of the verb premodifies the noun. It should be noted that in the first of these examples the verb has the prefix: 'un-' (cf. Fig. 37 below):

- .. A&R is a common term of abuse, synonymous with unanswered telephone calls, unkept **appointments** and broken promises.

- The resented traffic jam, the heated political argument, the broken **appointment**, can cause strong feelings of resentment, exasperation and frustration which have a physical effect.

		PROCESSES							ATEMPORAL RELATIONS				NOMINAL PRED	
BASIC LEVEL VERB	**total occ.**	**% of total occ.**	**1. TRANSITIVE VERB + NOUN (OBJ.) ACTIVE**		**2. TRANSITIVE VERB + NOUN (OBJ.) PASSIVE**		**3. NOUN (SUBJ.) + INTR. VERB**		**4. PAST PART. + NOUN**		**5.NOUN + PAST PART.**		**6. NOMINAL-IZATION OF VERB + NOUN**	
	Occ.	**%**	**Occ.**	**%**	**Occ.**	**%**	**Occ.**	**%**	**Occ.**	**%**	**Occ.**	**%**	**Occ.**	**%**
make	145	62.5	85	58.6	48	33.1	0	0	0	0	12	8.3	0	0
hold	26	11.2	16	61.5	4	15.4	0	0	0	0	4	15.4	2	7.7
have	17	7.3	17	100.0	0	0	0	0	0	0	0	0	0	0
*keep**	13	5.6	11	84.6	1	7.7	0	0	1	7.7	0	0	0	0
give	9	3.9	1	11.1	8	88.9	0	0	0	0	0	0	0	0
get	6	2.6	6	100.0	0	0	0	0	0	0	0	0	0	0
*take***	5	2.1	5	100.0	0	0	0	0	0	0	0	0	0	0
break	5	2.1	2	40.0	2	40.0	0	0	1	20.0	0	0	0	0
fill	3	1.3	2	66.7	1	33.3	0	0	0	0	0	0	0	0
lose	3	1.3	2	66.7	1	33.3	0	0	0	0	0	0	0	0
	232	99.9	147	63.4	65	28.0	0	0	2	0.9	12	6.9	2	0.4

Table 23: Basic level predications types in two APPOINTMENT domains

(Percentages rounded to one decimal)

* The example of construction type 4 is *<u>un</u>kept appointments*.

** The five examples of *take* include the phrasal verbs *take up* (2), *take on* (1), and *take over* (1).

Construction type 5, in which a past participle form follows the noun, accounts for most examples of atemporal relations: 16 occurrences, or 7%. Finally, there are 2 examples of a nominalized verb form, which is close to 1%. The nominalized forms are mass nouns as defined above (in section 3.4.2.3):

- .. the incident affords a clear illustration of the close links between parliamentary and municipal politics and the holding of official **appointments**

Table 23 also shows that one verb, *make*, accounts for 145 out of the 232 occurrences of basic level verbs, or 63%, while *hold* accounts for 26 occurrences, or just over 11%, and *have* with 17 occurrences has a share of 7%, followed by *keep* with 13 occurrences and 6%. The other six verbs: *give, get, take* (including *take up, take on*, and *take over*), *break, fill* and *lose* each occur less than 10 times, and between them they account for only 31 occurrences or 13.4%. Excepting *make*, which is in a league of its own as far as frequency of occurrence is concerned, and perhaps *hold*, *have* and *keep*, it seems that it is not possible to claim a special status for these verbs in the APPOINTMENT domains on the basis of sheer numbers. The point of view pursued in this study is that their special status is based on their schematicity and their role as default verbs in the semantic frames that correspond to the readings that can be posited for *appointment* (cf. section 3.4.3.4 below).

The diagram in Fig. 38 below shows the great variation in the frequency of occurrence of the basic level verbs and illustrates the predominance of processual relations. Lining up the verbs in this way also draws attention to the fact that most of the verbs simply do not normally occur in this form.[83]

** a held camera*	*a hand-held camera; an enemy-held coast*
** a kept appointment*	*an unkept appointment; a well-kept garden*
** a made cake*	*a home-made cake; ready-made desserts*
**a filled container*	*trout-filled streams*

Fig. 37: Past participles of basic level verbs as premodifiers [84]

[83] According to Swan (1995: 405), it is not yet possible to give rules for which participles can be used as adjectives before nouns. Why, for instance, is it perfectly acceptable to speak of *a lost dog*, while *a found dog* seem decidedly odd.

[84] Most examples are from the Cobuild Corpus Concordance Sampler. http://www.collins.co.uk/Corpus/CorpusSearch.aspx

While *break* and *lose* can be given a stative, adjectival construal, profiling the final state of a process, it seems that the others are not used in this way unless they have a prefix that adds more specific information, as in the case of *unkept* (cf. example in this section above). In the case of *give*, the past participle seems only to be used as a premodifier in the sense 'the particular kind that is being referred to', about a specific time or place, etc. Bache and Davidsen-Nielsen (1997: 450) describe the difference between past participles used as premodifier and postmodifier in terms of the 'temporary vs. permanent' distinction. Just as *the only stars visible* are stars that are 'visible at the moment', and *the only visible stars* are those that are 'normally within sight', past participles of a verb express a temporary meaning as postmodifiers, as in *the statement written* and a more permanent quality as premodifiers, as in *the written statement.* This is in line with the distinction made above (in section 3.4.2.3) between premodifying past participles as atemporal relations, as in *a broken cup* and postmodifying past participles as complex atemporal relations, as *in plain chocolate, broken into pieces.* The former construction was characterized as 'stative', which is equivalent to saying that it expresses a more permanent quality, whereas the latter was like a process scanned in a summary fashion, which means that the 'temporary' aspect is still present although it is not profiled as clearly as if it was scanned sequentially in a processual relation: *the cup was broken* or *she broke the cup*.

What is worth noticing is that most of the basic level verbs recruited for these nonsensorimotor domains typically code processual relations. In the case of those that also have the potential to code atemporal relations, like *break*, the forms that do so occur much less frequently in the nonsensorimotor domain than they do in the sensorimotor domain. I have argued that this would be in line with a construal of such verbs as having a support verb function (cf. section 3.4.2.3). The question remains why past participle forms of some verbs do not tend to be used as premodifiers. As indicated by the examples in Fig. 37 above, adding a prefix changes the situation in the case of verbs like *hold, keep, make* and *fill*, and an adverb often has the same effect: *a carefully kept secret; a freshly filled grave*, and *fortuitously gotten lottery gains* .[85] However, the past participles of *have* and *take* do not seem to

[85] The examples are from the Cobuild Corpus Concordance Sampler.

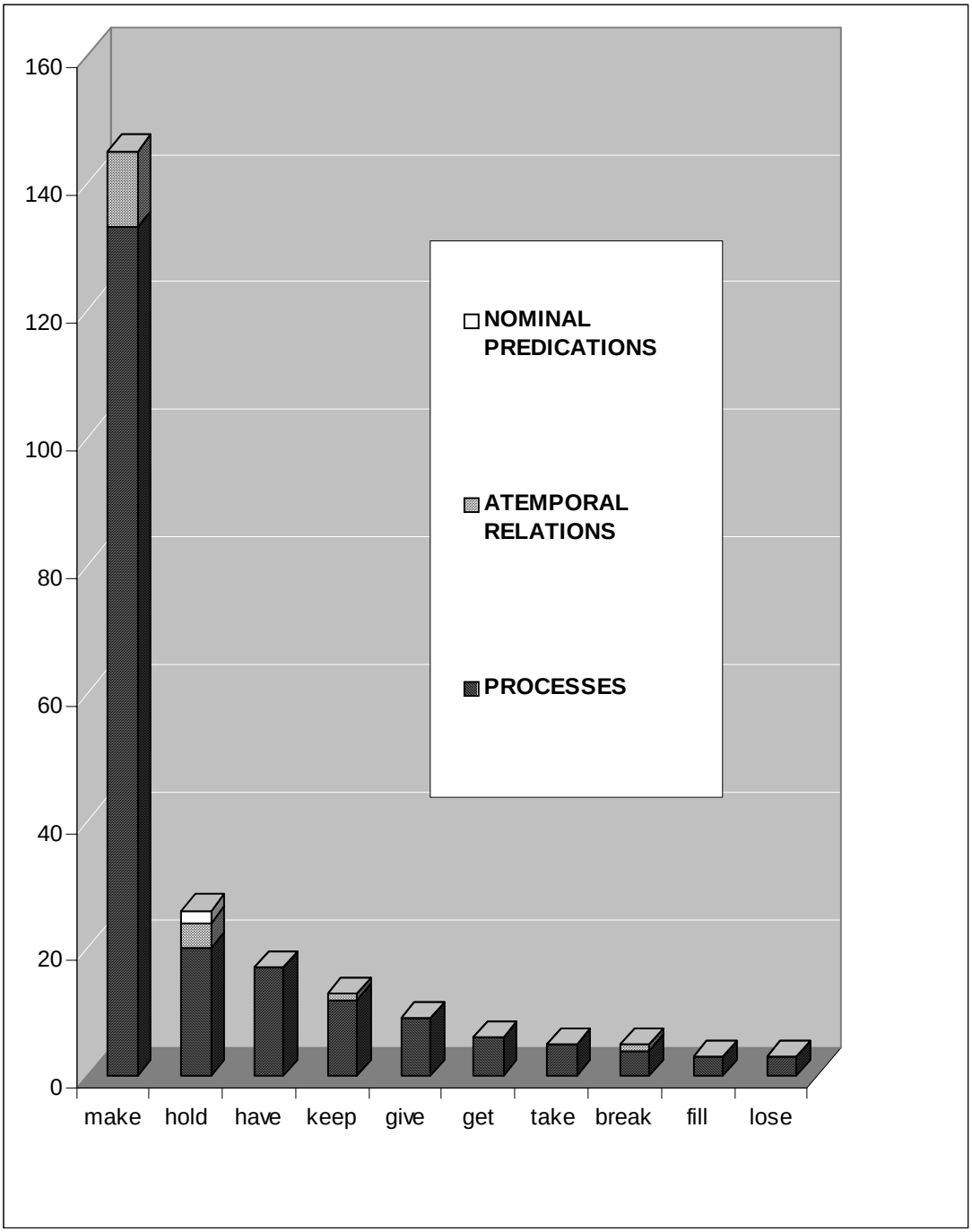


Fig. 38: Basic level predications in two APPOINTMENT domains (by number of occurrence)

be used as premodifiers, and in the case of *get* the construction seems to be rare. My hypothesis is that basic-level transitive verbs are mainly simply too schematic to denote a more permanent quality, something which, on the other hand, makes them perfectly suited for the function as support verbs denoting processes associated with deverbal nouns of action and their semantic frames. The discussion of the role of these verbs will be continued in the next section.

3.4.3.4 How many meanings does *appointment* have?

The APPOINTMENT category offers examples of commutation without substitution as well as of substitution without commutation. Thus *making an appointment* may turn out to be about a job, a meeting, a sum of money, or some accessory, depending on the context, whereas *keeping* and *breaking* may both be about an arrangement for a meeting, and *filling* and *holding* both refer to a job or 'position'. Talking about *holding* rather than *keeping an appointment*, on the other hand, will often involve a different conceptualization of what type of experience it is that *appointment* is coding, so that we have a case of commutation, i.e. ambiguity rather than vagueness. In this section, I will relate the question of how many meanings, or readings, can be posited for *appointment* to the number of semantic frames that can be identified on the basis of the data for the category APPOINTMENT. One type of frame that is considered relevant for this purpose is the one referred to as 'schematic' in the annotation principles of the FrameNet Project at the University of California (Johnson et al 2003: section 5.1.4). Furthermore, the discussion will include the role of 'specific frames', which provide each their perspective on the schematic frame, with which, in the FrameNet terminology, they are said to have a 'use relationship'.

What is cognitive and what is linguistic about semantic frames? In this study, frames have been described as cognitive structures which form the background against which linguistic expressions are understood. This is what sets apart semantic frames from the Firthian 'contexts of situation', a notion which is otherwise quite compatible with semantic frames (cf. section 2.2.1 above). Semantic frames can be traced back to Fillmore's case grammar (1968, 1977a), which was based on the generative distinction between deep-structure cases and surface transformations. In his article *Scenes-and-frames*

semantics (1977b: 60), however, Fillmore said that he had "pulled back" from case grammar for "the same reason I get dissatisfied with a filing system for my notes when I suddenly become aware that the box labelled "MISCELLANEOUS" contains more than all the rest. There were just too many things I could not account for." The same article (1977b: 63) referred to frames as 'linguistic' and associated with 'scenes':

> I want to say that people, in learning a language, come to associate certain scenes with certain linguistic frames. [...] I intend to use the word frame for referring to any system of linguistic choices (the easiest cases being collections of words, but also including choices of grammatical rules, or grammatical categories) - that can get associated with prototypical instances of scenes.

Over time, the word 'frame' itself has come to be used as a unifying term for a range of notions that refer to the cognitive structures underlying word meaning, including Fillmore's own former use of 'scene', (1985: 223).[86] In 1985, Fillmore still speaks of "the set of interpretive frames provided by language" (1985: 229), but makes the reservation that it is not really the language that provides the frames: "The language provides the mappings between linguistic choices and the interpretive frames, but while some of them are 'created by the language', most of them can be said to exist independently of the language". The current use equates 'semantic frames' with conceptual structures as appears from the website of the FrameNet Project, which describes frame semantics as an "approach to the understanding and description of the meanings of lexical items and grammatical constructions" and is based on the assumption that " ... in order to understand the meanings of the words in a language we must first have knowledge of *the conceptual structures*, or *semantic frames* [my emphasis], which provide the background and motivation for their existence in the language and for their use in discourse" (Johnson et al 2001). The motivation for frame semantic research, according to Fillmore (1985: 234) is "the effort to understand what reason a speech community might have found for creating the category represented by the word, and to explain the word's meaning by presenting and clarifying that reason." Frame semantics has thus developed into a cognitively based as well as a functionally motivated approach to understanding and explaining the meaning of words and grammatical constructions.

[86] These notions include 'script', 'schema', 'scenario', 'idealized cognitive model', and 'folk theory' (Coulson 2000: 20). See also Fillmore (1985: 223).

The question in the context of this study is how many schematic frames can be identified for the APPOINTMENT category and how many readings should be posited as a consequence. It follows that the amount of text that needs to be included in the analysis should be sufficient to allow the frame, or 'context of situation', to be recognized in each individual case (cf. the discussion of 'significant proximity' in section 2.1.2 above). As a point of departure, I will use the subdomains already discussed. As in section 3.4.3.2, I will begin with 'arranging/arrangement for a meeting' and '(placing sb in) a position' and then move on to the domains 'action of declaring destination of property' and 'accessories', of which there are but a few occurrences in the data. It is not the ambition to draw up a complete inventory of frame elements and construction types for each specific frame as would be required in a complete study for lexicographical purposes, but to discuss whether it is possible to identify a schematic frame for each subdomain and to posit a reading for each frame.

The subdomain **'arranging/arrangement for a meeting'** is construed as a schematic frame whose specific frames are evoked by the following composite structures, in which the basic level verbs represent the default case:

1. *make an appointment*
2. *give an appointment*
3. *get an appointment*
4. *have an appointment*
5. *keep an appointment*
6. *break an appointment*

Within the schematic frame, *appointment* combines with the different basic level verbs to evoke the specific frames. This is illustrated in Fig. 39 below, which can be compared with Table 17 (in section 3.4.3.2). The idea is to show that, in each frame, we find a continuum of specificity without any clear boundaries, ranging from the most schematic default verb to the most specific ad-hoc instantiations. The basic level verbs in the inner circle are seen as the prototypical default fillers of the 'verb slot' in the specific frames, which combine with *appointment* to provide each their perspective on the schematic frame. The verbs in the outer circle are construed as belonging to the subordinate level of categorization; they are less prototypical and schematic, and I consider that they are also quite likely to be entrenched, while those outside the circle are given as examples of quite specific ad hoc instantiations, which are not considered likely to achieve conventional status, or become

'institutionalized' (cf. section 2.2.2 above). The two-headed arrows are meant to illustrate that combinations range from the most schematic and entrenched to the most specific ad hoc expressions. This variability may be seen as evidence of entrenchment and innovation.

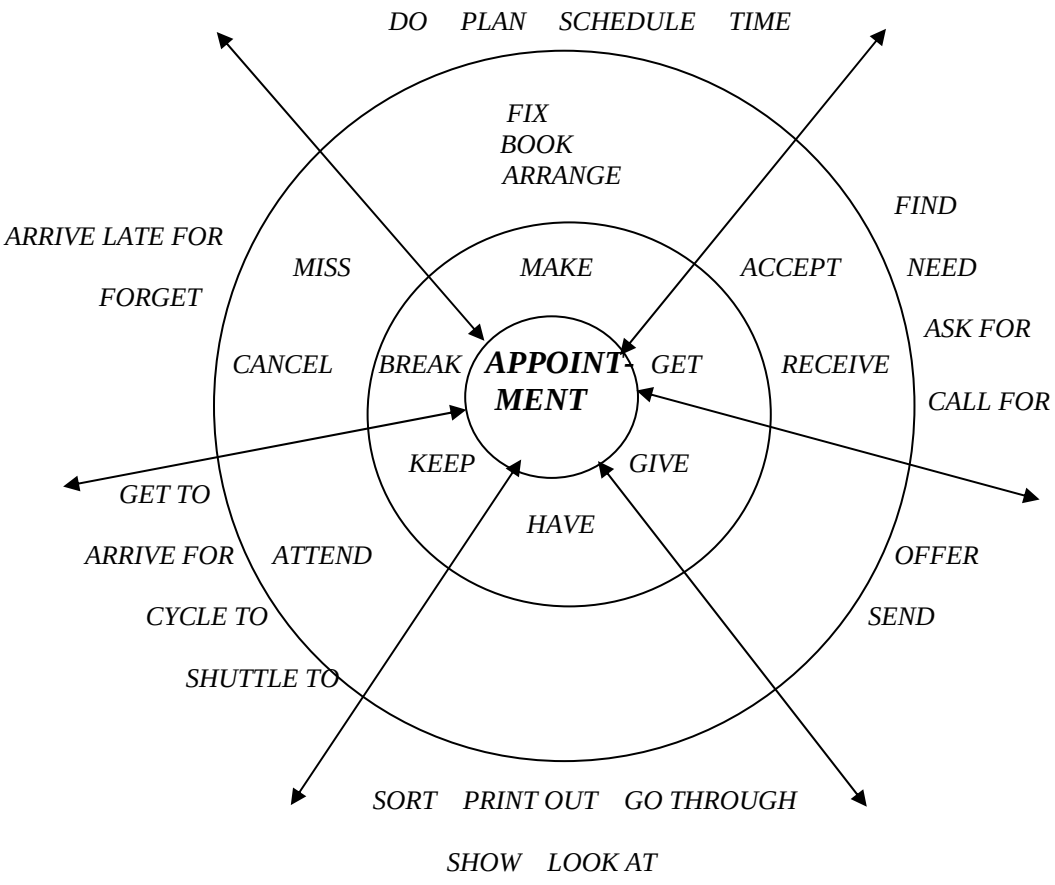


Fig. 39: Schematic frame APPOINTMENT I: 'arranging/arrangement for a meeting'

The following elements could be identified in this schematic APPOINTMENT frame:

1. the meeting itself
2. the party (or parties) arranging or cancelling the meeting
3. the party (or parties) the meeting is 'with'
4. the institution the meeting is 'at'
5. the location the meeting is 'at' or 'in'
6. the purpose the meeting is 'for'
   (also expressed as an infinitive)
7. the time the meeting is arranged 'for'
8. (or the day it is 'on')

An example including all seven elements is not found in the data, but can easily be made up:

> ***He*** (2) *made an* ***appointment*** (1) *with* ***his doctor / to see his doctor*** (3) *at* ***the hospital*** (4) *in* ***London*** (5) *for a* ***health check*** (6) *on /* for **next Monday** (7).

In this example, replacing *make* by the other basic level verbs: *get*, *have*, *keep* and *break* does not lead to a commutation in meaning. Rather, it involves a change of perspective, so that if for example the verb is give, we now find 'the party (or parties) the meeting is with' as the trajector of the verb rather than 'the party arranging the meeting'(cf. section 3.4.2.3 above) .

> ***His doctor*** (3) *gave him an appointment* (1) *for a* ***health check*** (6) *at* ***the hospital*** (4) *in* ***London*** (5) *on /* for **next Monday** (7).

As noted previously, substitution does not even seem to lead to commutation in the case of the two basic level verbs that seem to be most far apart in meaning, viz. *make* and *break*. The example given (in section 3.4.3.1) was

- *Several* ***appointments*** *were made and broken before he and his wife finally arrived.*

Although *appointment* is a noun of action in the combination with *make*, whereas it codes the result of action in the combination with *break*, no zeugmatic effect is created, rather two 'zones' of the schematic frame are activated in the same sentence (cf. section 3.4.1.1 above). Consequently, only one reading is posited for the schematic frame corresponding to the APPOINTMENT subdomain 'arranging/arrangement for a meeting'.

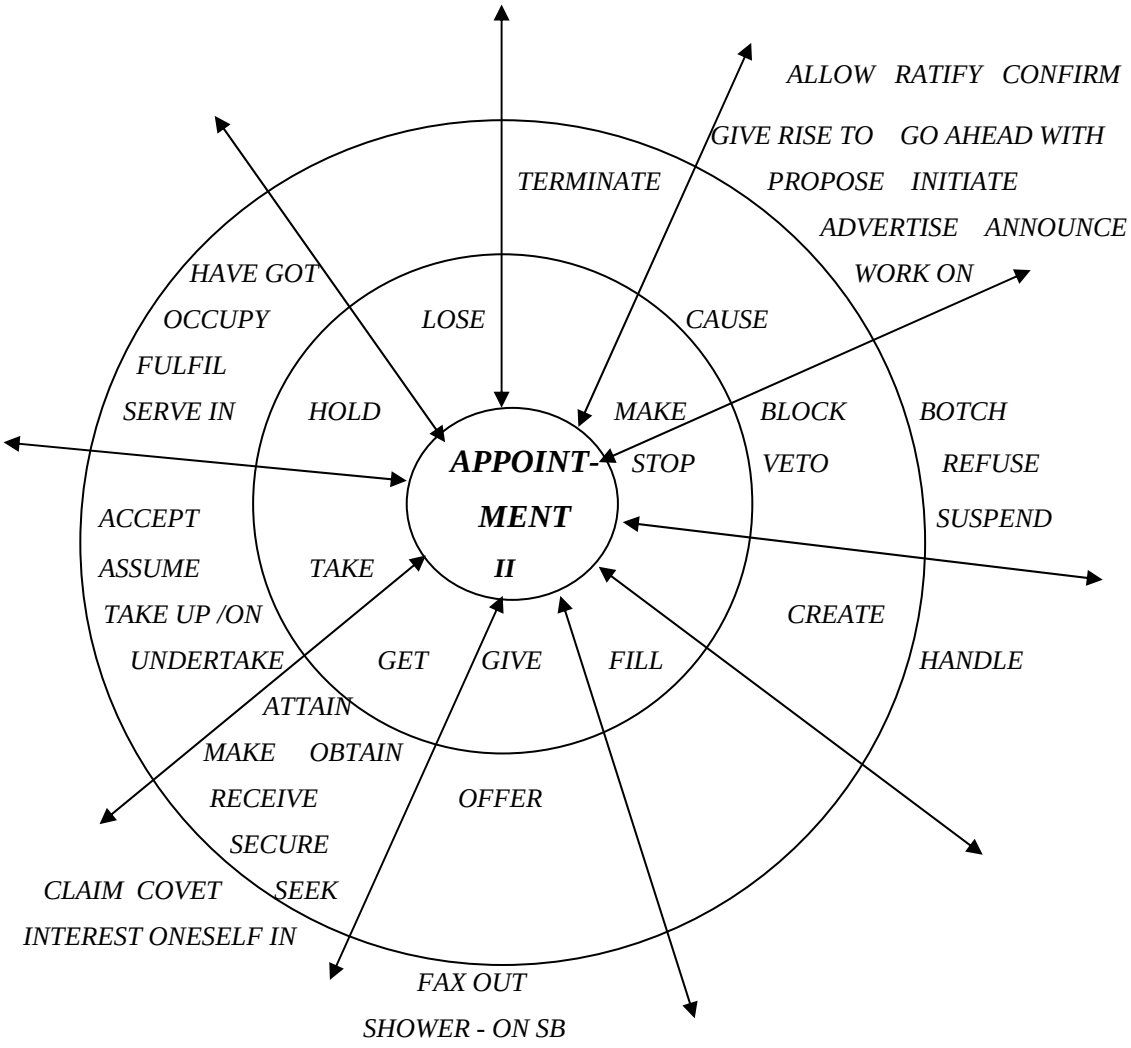


Fig. 40: Schematic frame APPOINTMENT II: '(placing sb in) a position'

In the schematic frame corresponding to the subdomain **'(placing sb in) a position'**, the following ten specific frames were identified:

1. *make an appointment*
2. *stop an appointment*
3. *create an appointment*
4. *fill an appointment*
5. *give an appointment*
6 . *get an appointment*
7. *take an appointment*
8. *hold an appointment*
9. *lose an appointment*
10. *terminate an appointment*

In some frames no basic level verb was found (*create* and *terminate*) and in others *only* a basic level verb was found (*fill* and *lose*)[87]. As in APPOINTMENT frame I, *appointment* here combines with the basic level verbs to evoke specific frames. This is illustrated by Fig. 40 below, which can be compared with Table 18 (section 3.4.3.2). The basic level verbs in the inner circle are seen as the prototypical default fillers of the 'verb slot' in the specific frames, combining with *appointment* to provide each their perspective on the schematic frame. As in Fig. 39 above, a distinction is made between more and less schematic verbs at the subordinate level of categorization.

An analysis of the elements occurring in this schematic APPOINTMENT frame shows that they belong to two sets, one of which is centred on the act of 'placing sb in a position' (location event structure):

1. the act of placing sb in a position
2. the party placing sb in a position
3. the party preventing sb from being placed in a position
4. the party being placed in the position
5. the location sb is moved 'to' or 'from'
   (or 'at'/'in' which sb is placed in a position)
6. the time at which sb is placed in the position

- [...] ***early in 1988*** (6) ***they*** (2) <u>*had made*</u> *a number of senior financial* ***appointments*** (1) ***from external sources*** (5).
- *One thing* ***she*** (2) *can do is to* <u>*make*</u> ***appointments*** (1) ***to the state's numerous policy-making commissions and boards*** (5).
- *Following* ***your*** (4) *recent* ***appointment*** (1), *I confirm that this Firm will be pleased to act on your behalf in connection with your claim for damages arising out of the incident.*

The second set of elements is centred on the act of 'creating a position' and 'handling' it (object event structure):

[87] It is assumed that combinations with basic level verbs are more likely to be entrenched, but it is not assumed that they always are. See the discussion of *fill* and *lose an appointment* in section 3.4.4.2 below.

a. the position itself
b. the party creating/ filling/giving the position
c. the party taking/getting/holding/losing the position
d. the location at or the circumstances in which the position is found
e. the time at which the position is created, taken, etc.

- *In order to further develop our services to the Agrochemical Industry* ***we*** (b) *wish to* *create* *two* ***appointments*** (a) ***at Section Manager level*** (d).

- ***they*** (c) *hold* ***their*** (c) ***appointments*** (a) *at the pleasure of the Crown* (d).

Because of the difference in event structure, an example like the following (also shown in section 3.4.3.1) is ambiguous between the two readings of *appointment* and one of them has to be discarded:

- *I would have* *to lose* *10 teaching* ***posts*** (a) ***and*** *several ancillary* ***appointments*** (1 or a) ***we*** (2 or b) *'ve made* *to allow teachers to concentrate on teaching.*

The coordination of *appointments* with *posts* as the object of *lose* invites a reading in terms of object event-structure ('create a position')[88], but it may also be the case that, in mid-sentence, the writer switches to a construal in terms of location-event structure ('place sb in a position'). At any rate, it seems to me that the reader has to make a choice, if the ambiguity is to be resolved. In the case of the schematic frame '(placing sb in) a position', I therefore find it plausible to posit two different readings: one for the specific frames 'make an appointment' and 'stop an appointment', which are based on location event structure,[89] and one for all the other specific frames, which are based on object event structure. In line with what has been argued above (in section 3.4.1.1), differences in meaning between specific frames can be accounted for in terms of different active zones. That overlaps should occur between the readings, in the form of examples that are ambiguous between them, is to be expected and may be taken as evidence for the

[88] There is no example in the data in which *make an appointment* clearly means 'create a position', but there are a few examples in which this expression is ambiguous between 'create a position' and 'place sb in a position'.

[89] Actually, *to make an appointment to a position* is a 'caused-motion' construction like *to push the crumbs off the table* or *to sneeze the napkin off the table* (Goldberg 1995)

descriptive and explanatory relevance of the network model, which is seen as linking the different meanings of a complex lexical category.

The schematic frame corresponding to the nonsensorimotor subdomains **'action of declaring destination of property'** (cf. Table 16 in section 3.4.3.1) is similar to '(placing sb in) a position' in its location event structure. Only ten examples were found, two of which may not belong to this frame at all. On the basis of the examples, two specific frames can be posited:

1. *make an appointment*
2. *stop an appointment*

The following frame elements were identified:

1. the act of declaring destination of property
2. the source of authority of this act
3. the party declaring destination of property
4. the act of stopping property from going to a destination
5. the party stopping property from going to a destination
6. the property
7. destination of property

- *Her father's will also made provision for the destination of this life interest should **she** (3) fail to make any such **appointment** (1): **the money** (6) would be payable to a number of **her distant relatives** (7).*
- *This would revive **the 1979 will** (2) with its power of **appointment** (1), and he thought that that more closely approximated to those intentions than the **1989 will** (2) without any such power.*
- *It is likely that the court would take a similar attitude to attempts to stop **appointments** (1) by other **bodies** (5).*

Finally, a small number of examples (10 occ.) were recorded as belonging to the sensorimotor subdomain **'accessories'** (cf. Table 16 in section 3.4.3.1 above). They belong to two subdomains construed as corresponding to the two schematic frames: 'accessories for people' and 'accessories for cars, rooms, etc.'. Consequently, two readings are posited for APPOINTMENT in the sensorimotor domain: 'accessories for people' and 'accessories for cars, rooms, etc.'. The following specific frames can be identified as belonging to the former:

1. *have appointments*
2. *produce appointments*

The following elements could be identified for this frame:

1. the accessories
2. the person having the accessories
3. the authority demanding that the accessories be produced

- *Then **you*** (2) *had to <u>produce</u> **your*** (2) ***appointments*** (1) *which **was your*** (2) ***baton and brass keys for the police telephone boxes*** (1), *<u>hold</u> **them*** (1) *in your hands, and then you were issued with the Beat Book on parade.*

- ***He*** (3) *stated himself pleased with the general appearance and **appointments*** (1) *of the **unit*** (2).

For the second schematic frame, just one specific frame was documented:

1. *have appointments*

The following elements were found:

1. the accessories
2. the entity equipped with the accessories (expressed as a possessive pronoun)
3. the party benefiting from the accessories

- ***Cleo*** (3) *exclaimed over the **rooms*** (2) *and **their*** (2) ***appointments*** (1) *with delight.*

- *In **its*** (2) ***appointments*** (1) ***the Austin-Healey 3000 MkIII*** (2) *is now more of a touring car than a sports car.*

The discussion is summarized in Table 24 below, from which it appears that six readings are posited for *appointment* on the basis of the schematic frames identified. In the case of the schematic frame 'arranging/ arrangement for a meeting', it was found that it was enough to posit one reading as the specific frames evoked by the basic level verbs could be accounted for in terms of active zones. In the case of the schematic frame, '(placing sb in) a position', two readings were posited because of the difference in event structure between one specific frame, *making an appointment*, using location event structure and the other specific frames using object event structure. The schematic frame 'action of declaring

	SCHEMATIC FRAME with elements	SPECIFIC FRAME as active zone	READING
1.	**arranging/arrangement for a meeting** 1. the meeting itself 2. the party arranging or cancelling the meeting 3. the party/parties the meeting is 'with' 4. the institution the meeting is 'at' 5. the location the meeting is 'at' or 'in' 6. the purpose the meeting is 'for' (or the day it is 'on')	*make an appointment* *give an appointment* *get an appointment* *have an appointment* *keep an appointment* *break an appointment*	arranging/arrangement for a meeting
2.	**(<u>placing sb in</u>) a position** **a. location event structure** 1. the act of placing sb in a position 2. the party placing sb in a position 3. the party being placed in a position 4. the location sb is moved 'to' or 'from' (or 'at'/'in' which sb is place in a position) 5.the time at which sb is placed in the position	*make an appointment*	placing sb in a position
3.	**(placing sb in) <u>a position</u>** **b. object event structure** a. the position itself b. the party creating/filling/giving/terminating the position c. the party getting/taking/holding/losing the position d. the location 'at' or 'in' which the position is found e. the time at which the position is created, etc.	*create an appointment* *fill an appointment* *give an appointment* *terminate an appointment* *get an appointment* *take an appointment* *hold an appointment* *lose an appointment*	position
4.	**accessories for cars, rooms, etc.** 1. the accessories 2. the entity equipped with the accessories 3. the party benefiting from the accessories	*have appointments*	accessories for cars, rooms, etc.
5.	**accessories for people** 1. the accessories 2. the person having the accessories 3. the authority demanding that the accessories be produced	*have appointments* *produce appointments*	accessories for people
6.	**action of declaring destination of property** 1. the act of declaring destination of property 2. the source of authority of this act 3. the party declaring destination of property 4. the act of stopping property from going to a destination 5. the party stopping property from going to a destination 6. the property 7. destination of property	*make an appointment* *stop an appointment*	declaring destination of property

Table 24: Readings of *appointment*

destination of property', was construed as corresponding to a further reading in the nonsensorimotor domain. For the sensorimotor domain, finally, two readings were posited on the basis of the frame elements and specific frames identified: 'accessories for people' and 'accessories for cars, rooms, etc. In Fig. 41 below an attempt has been made to illustrate how, in principle, the different APPOINTMENT frames could be related in the mind of a contemporary user. It is meant to illustrate how somebody might gradually become familiar with different uses of a word as it occurs in actual contexts of situation, so that eventually the knowledge of that word will grow into a complex category, which can

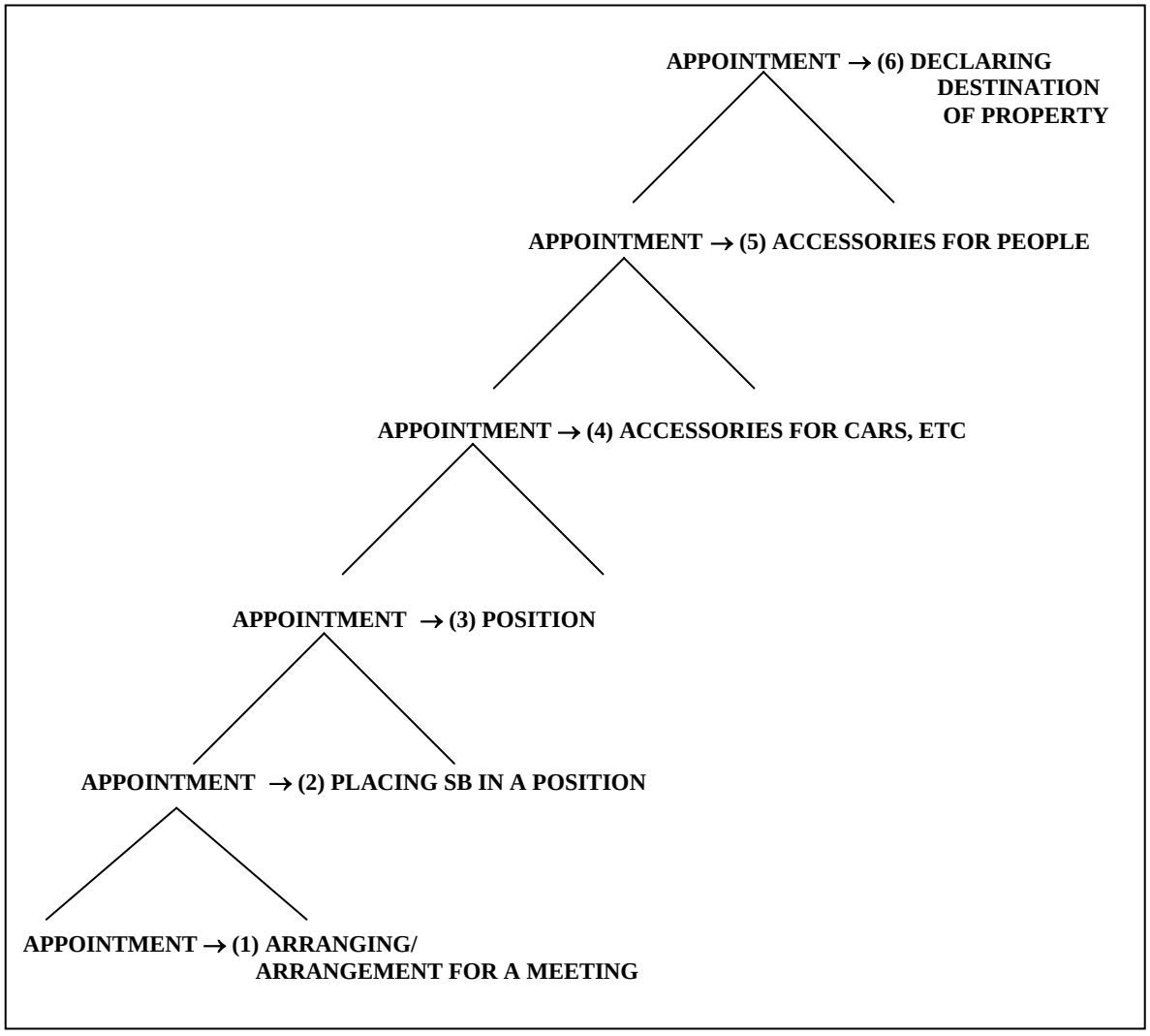


Fig. 41: APPOINTMENT as a complex category and a network of meanings

be used flexibly and dynamically for meaning construction. Which meanings people become familiar with and in what order is bound to be idiosyncratic, for example some might never come across readings (5) and (6).

What seems certain from the linguistic evidence is that the APPOINTMENT category is a complex one, although much less complex than the BREAK category analysed in section 3.4.2. I have tried to show how the internal structure of the category can be accounted for at the referential (or 'extensional') level, in terms of polysemy relating readings to schematic frames, and how specific combinations with basic level verbs can be accounted for at the intensional level, as active zones imposing different perspectives on these frames.

3.4.4 The integration of BREAK and APPOINTMENT

After accounting for the complex internal structure of the component items BREAK and APPOINTMENT, I will now turn to an analysis of their mode of integration. A **composite structure** like *break an appointment* is characterized as an assembly of symbolic structures (Langacker 1999b: 13) which is motivated by its **components** but not predictable from them (cf. section 1.3). I will elaborate on theoretical aspects of composition involving the integration of a relational and a nominal predication (cf. section 1.5 above), of which the composite item *break an-appointment* is an example.

Like each of the components, the composite structure is treated as a complex category, and the five examples in the data are seen as instantiations of this category or, in the terms of Firth, as the occasions on which it finds "application in renewal of connection with the sources of the abstractions" (Firth 1968a: 200, quoted in section 2.1.2 above). This evidence of integration at the schematic level, as well as its instantiations, will be supplemented with an analysis of lexical sets representing the range of variability at the place of the verb. The purpose is to identify characteristics of entrenched collocations as a

prototype category and to test the claim that *break* may be construed as having a grammaticalized role as support verb.

3.4.4.1 *Break an-appointment* as a composite structure

The account is given in terms of Langacker's cognitive grammar, according to the principles outlined in the introduction to this study (section 1.5 above). The evidence presented about the internal structure of BREAK and APPOINTMENT respectively is in line with this approach, which sees component items not as building blocks that can be assembled in an additive way, but as complex categories, or dynamic networks, whose integration is not predictable from, but motivated by them. To see how the composite structure *break an-appointment* is motivated by the two component items, it is necessary to account for the combinatorial potential of each item in terms of its polysemy, as attempted in section 3.4. As pointed out earlier in this study (cf. Table 5 in section 2.1.5), composition involves several levels of context, not only syntagmatic context, or collocation, but also systemic context: the position of lexical items in their respective schematic networks. This in turn makes it possible to identify the valence relations between them, including correspondences, or overlaps, between substructures as well as autonomy and dependence relations. The diagrammatical presentation in Fig. 42 below is based on similar presentations by Langacker (especially 1992: 489: *under (the) tree*, in which *under* is the relational predication in a prepositional phrase).

At the schematic level, *break an-appointment* is an example of the integration of the relational predication *break* and the nominal predication *appointment.* Since *appointment* is a countable noun in these combinations, there will have been a previous level, or levels, of constituency integrating *appointment* minimally with an article, as indicated schematically by *an-*. As a relational predication, the verb profiles a relation between a subject corresponding to its trajector (*tr*), to be specified at the next level of constituency, and the nominal object, corresponding to its landmark (*lm*). The overall profile of the composite

structure is determined by the relational predication, which is said to be 'fully schematic' for the composite structure (Langacker 1987: 492). This is indicated in Fig. 42 by means of the heavy-line box enclosing *break*. The nominal predication, displays only 'partial schematicity', as it profiles only the landmark of the composite expression.

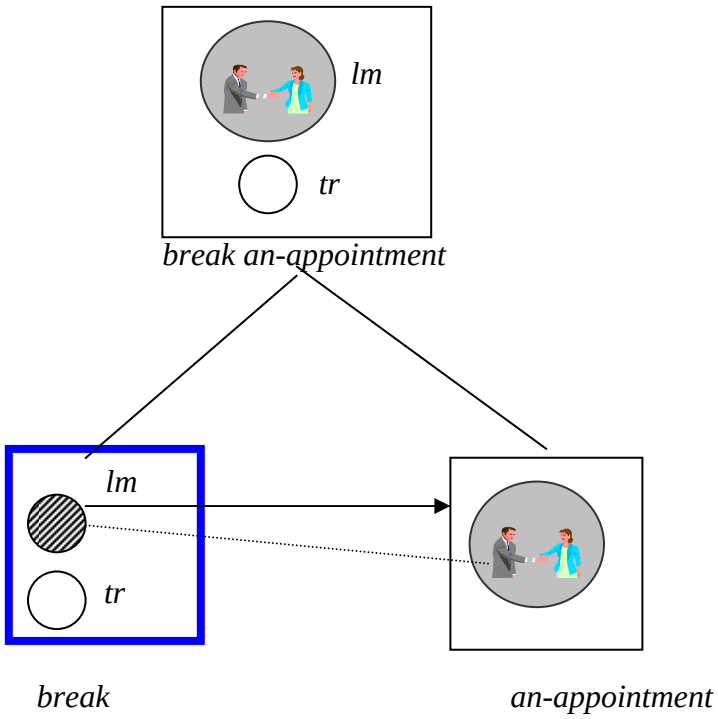


Fig. 42: The composite structure *break an-appointment*
(Based on Langacker, 1992: 489: *under (the) tree*)

In spite of the status of the verb as 'profile determinant', a relationship of (relative) asymmetry is posited between verb and noun in which the verb is seen as (relatively) dependent, because it requires elaboration of a salient substructure by the noun, which is seen as (relatively) autonomous (cf. section 3.4.4.2 below). The substructure of the relational predication *break* that is elaborated by *an-appointment* is its schematic landmark, which is also referred to as its e-site (elaboration site), and is marked diagrammatically in Fig. 42 by cross-hatching. The elaboratory relationship between the landmark of *break* and the nominal predication *an-appointment* is shown in Fig. 42 as a solid arrow. The dotted

line between the landmark of *break* and the symbol for *appointment* indicates a correspondence, not between the lexical categories BREAK and APPOINTMENT as such, but between subparts of the two complex categories that correspond, or show overlaps, when they are merged in the process of integration. The mnemonic presentation of *appointment* as two people shaking hands is meant to indicate that the schematic APPOINTMENT frame evoked by the composite structure is 'arranging/arrangement for a meeting', (cf. table 24 in section 3.4.3.4). This is actually anticipating the discussion of the five examples found in the data, since *break an-appointment* could also refer to the damaging of a physical object in the sensorimotor domain (reading 6, Table 24 in section 3.4.3.4 above).

NONSENSORIMOTOR DOMAIN →	**VIOLATION OF SOCIAL INSTITUTIONS & CONSTRUCTS**		
SCHEMATIC FRAME →	**APPOINTMENT (READING 1): 'ARRANGING/ARRANGEMENT FOR A MEETING'**		
SPECIFIC FRAME →	**BREAK (READING 11b) + APPOINTMENT**		
PREDICATION TYPES ↓	**CONSTRUCTION TYPE:**	**NO.**	**%**
PROCESSES	**1. TRANS. VERB + NOUN (OBJ.), ACTIVE**	2	40
	*1) He's already broken three **appointments** to see me.*		
	2) *She made some weak excuse when she telephoned to break the **appointment**, but she knew perfectly well that she was apprehensive about going out in case the weather changed and it began to rain.*		
	2. TRANS. VERB + NOUN (SUBJ.) PASSIVE	2	40
	*3) In the earlier stages, social and other **appointments** (particularly for the time to return home) are repeatedly broken because "someone turned up."*		
	4) *Several **appointments** were made and broken before he and his wife finally arrived*		
	3. NOUN (SUBJ.) + INTRANS. VERB	0	0
ATEMPORAL RELATIONS	**4. PAST PARTICIPLE + NOUN**	1	20
	*5) The resented traffic jam, the heated political argument, the broken **appointment**, can cause strong feelings of resentment.*		
	5. NOUN + PAST PARTICIPLE	0	0
NOMINAL PREDICATIONS	**6. NOMINALIZATION OF VERB**	0	0
	TOTAL	5	100

Table 25: Five instantiations of *break (an) appointment*

Each of the examples shown in Table 25 above, represents an instantiation: an actual usage situation in which only a subpart of the complex category BREAK is involved in correspondence relations with a subpart of the complex category APPOINTMENT. The data on *break* analysed in section 3.4.2 did not include any examples of combinations with *appointment*, but Reading 11 b, in which the 'violation of social institutions & constructs' was construed as 'cancellation', included combinations with a range of nouns coding mutual commitments such as *engagements*, *marriages, links, relations*, *relationships* and *unions*, to which I have argued that *appointment* may be added as an extension (cf. Table 22 in section 3.4.2.4 above). The five examples shown in Table 25, which were all found in the data on *appointment*, were analysed as belonging to this reading of *break*. They were furthermore construed as evoking a specific frame, or profiling an active zone, of the schematic APPOINTMENT frame that corresponds to Reading 1 of *appointment*: 'arranging/arrangement for a meeting', (cf. Table 24 in section 3.4.3.4). As indicated in Table 25 above, the integration of the two composite items is construed as involving a correspondence, or an overlap, between Reading 11b of *break* and Reading 1 of *appointment*, and as taking place within the nonsensorimotor domain VIOLATION OF SOCIAL INSTITUTIONS & CONSTRUCTS.

As *break* participates in three different APPOINTMENT frames, it is necessary to include the context, or co-text, to determine which one is being evoked by *break an-appointment* in a particular usage situation. In the discussion of Firthian linguistics (section 2.1.2 above), I suggested, "as a principled way of determining cut-off points, that the amount of actual text included for the purposes of analysis at the collocational level should be enough to identify the abstract context of situation involved." This flexible, qualitative approach, which was contrasted (in section 2.1.2) with the definition of 'significant proximity' in corpus linguistics as "a maximum of four words intervening" (Sinclair 1991: 170), follows Firth, according to whom the analysis should include complete sentences and might even be extended to include the utterances of preceding and following speakers (Firth 1968b: 106). Along the same lines, it was assumed in the frame analysis of *appointment* (in section 3.4.3.4) that the amount of text included should be sufficient to allow the cognitive frame

underlying the context of situation to be recognized. In three out of the five examples in Table 25, the collocate is only separated from the base word by a single word: the numeral *three* in the first example, the definite article in the second, and the indefinite pronoun *several* in the fourth; and in example 5, the base word directly follows the past participle of the verb premodifying it. In the passive construction in example 3, however, the base is separated from the collocate by an eight-word-long apposition, which goes to show that cut-off points that are rigidly defined in terms of the number of intervening words are arbitrary and need to be rejected for reasons of principle. Based on the iconic tendency for syntax and semantics to go hand in hand (cf. Langacker 1987: 361), I would say that what could be called ‘syntactic’ or ‘functional’ proximity, such as that between subject and object, or verb and object, is a better indicator of the significance of collocation, but that it needs to be supplemented with an analysis of the cognitive models underlying composite structures (cf. the discussion of lexicogrammaticalness in section 2.2.1).

In the discussion of the way in which the human mind processes conventional expressions (in section 2.1.5), it was argued that psycholinguistic theories which explain the understanding of idioms as a dynamic and flexible process are helpful also in explaining how we process collocations. Thus, according to the 'spreading activation model' (Aitchison 1987, in Howarth 1996: 50) and the 'configuration hypothesis' (Cacciari and Tabossi 1988; Van de Voort and Vonk 1995: 292 f.), we gradually narrow down our range of possible understandings of a word as we take in more information, especially in the form of key words, and we flexibly adjust our interpretations even after we have 'recognized' one meaning. I find it plausible that what has been found to apply to idioms also applies in the case of semantic frames. We take in clues in the form of elements and processes until we have enough information to be able to recognize a frame. If additional key words are perceived that do not fit that frame, we discard our first analysis as a different frame is evoked; alternatively we may find that two or more frames are evoked at the same time to produce an expressive effect (cf. section 3.4.3 below). In Table 26 below, it is shown

Key words (frame elements and processes/atemporal relations) evoking the APPOINTMENT frame 'ARRANGING/ARRANGEMENT FOR A MEETING'								
	the party arranging or cancel-ling the meeting	**process or atemporal relation**	**the meeting itself**	**the party (or parties) the meeting is 'with'**	**the institution the meeting is 'at'**	**the location the meeting is 'at' or 'in'**	**the purpose the meeting is 'for' (also expressed as an infinitive)**	**the time the meeting is 'for'(or the day it is 'on')**
1.	*He*	*'s already* ***broken***	*three* ***appointments***				*to see me*	
2.	*she*	*telephoned* ***to break***	*the* ***appointment***					
3.		***are*** *repeatedly* ***broken***	*social and other* ***appointments***				*for the time to return home*	
4.		***were made and broken***	*Several* ***appointments***					
5.		***broken***	*the -* ***appointment***					

Table 26: Key words evoking an APPOINTMENT frame

what key words coding the process (or atemporal relation) and salient frame elements may be assumed to prompt readers to identify the five sentences as belonging to the APPOINTMENT ‘arranging/arrangement for a meeting’. The items printed in bold type are lexically specific and I consider them to be especially important for the recognition of the frame. They are the noun *appointment* itself and a form of the verb *break* coding the process or an atemporal relation which is the result of the process. While the noun narrows down the range of possible choices to one of the six frames that I have argued may be evoked by *appointment*, the verb *break* limits the range to three, two of which are in the sensorimotor domain. The remaining ambiguity is resolved by the presence of one or several of the other frame elements, which are not lexically specific, but tend to take a certain form, such as 'the purpose of the meeting' which is typically expressed as a prepositional phrase with *for*, or as an infinitive.

Expectations concerning the frame may also be confirmed by details concerning the process, as in the first example where we learn that *she telephoned to break the appointment*, in the third example where *appointments are repeatedly broken*, and in the fourth example where it is said that the breaking of appointments had happened *before he and his wife finally arrived*, which implies that a time and a place for a meeting is involved. As shown in Fig. 39 (in section 3.4.3.4 above), the process of 'breaking' may also be coded by means of less schematic verbal expressions such as *miss* or *arrive late for*, which may or may not be entrenched. Also adjectives premodifying a frame element may serve to clarify what domain we are in, as in the third example where *social and other appointments* is most likely to evoke the frame ‘arranging/arrangement for a meeting’ in connections with *break*. In the fifth example, it is hard to pin down exactly which parts confirm the expectation that *broken appointment* is about the meeting frame; rather it is the whole sentence including the other frames that it evokes. By analogy with what we know about *traffic jam* and *political argument*, we are instructed to interpret *broken appointment* as a scenario involving a potential conflict between people, which points us towards the nonsensorimotor domain and the frame ‘arranging/arrangement for a meeting’. This interpretation is supported by the fact that definite singular forms are used to indicate that the situations are typical and familiar.

The five examples in Table 25 are subdivided according to construction type, in the same way as in the analyses of *break* (Table 14, section 3.4.2.3) and of *appointment* (Table 22, section 3.4.3.3). The distribution of predication types for the composite structure is much in line with what was found in

combinations with *appointment* and in combinations with *break* belonging to the nonsensorimotor domain. Thus active and passive processes between them amounted to roughly 87% and 83% for *break* and *appointment* respectively and for 80%, or four out of five, in the case of *break an-appointment*. As regards atemporal relations, the past participle + noun construction accounted for a relatively much smaller percentage in all three analyses, although there was more variation with almost 6% (five occurrences) in the case of *break*, less than 1% (two occurrences) in the case of *appointment*, and 20% (one occurrence) in the case of the composite structure. It is of course easy to exaggerate the importance of these low figures. Nevertheless, the predominance of processual relations in the nonsensorimotor domain seems to hold up. The higher figures for the past participle + noun construction in the case of *break* and *break an-appointment* than for *appointment* alone, may have something to do with the fact pointed out above (in section 3.4.3.3) that *break* is the only one of the basic level verbs that can be given a stative, adjectival construal, profiling the final state of a process, unlike others that can only be used in this way if they have a prefix that adds more specific information, or are preceded by an adverb. If *break* is understood as equalling 'un-make', or 'un-keep', the use of *broken* as premodifier can be seen as semantically parallel to the use of *unanswered*, from 'un + answer', and *unkept*, from 'un + keep', in the following example, also quoted (in section 3.4.3.3 above):

- *A&R is a common term of abuse, synonymous with unanswered telephone **calls**, unkept **appointments** and broken **promises**.*

In other words, *break* includes an element of negation that needs to be added to the other verbs as a prefix, and I would argue that it is the salience of this element across a wide range of frames and the highly schematic way in which it is expressed by *break* which has motivated its use in many entrenched collocations and which may justify positing it as a grammaticalized support verb (cf. section 3.4.4.3 below).

3.4.4.2 Evidence of entrenchment

The general claim to be tested by this study is that, as linguistic expressions, entrenched collocations contribute to the communicative utterances of which they form part by evoking semantic frames, which are the cognitive models underlying typical contexts of situation abstracted from use. I have claimed, as part of my third research question (cf. section 3.3.1) that salience, as

reflected in the autonomy/dependence alignment of the component structures, and schematicity, which is related to basic level categorization, are useful notions for this purpose, and I will now test these claims against my data.

I will start by pursuing the point that **salience** is reflected in the relative conceptual autonomy and dependence relations between the two components of a composite structure, also referred to as their 'A/D asymmetry' or 'A/D alignment' (Langacker 1987: 306; 356). As already mentioned, the verb, which is a relational predication, will typically be the more dependent component in a verb + noun combination, but this does not imply that dependence is 'unidirectional' (Langacker 1987: 358). Both component items have substructure that is elaborated to a certain extent by the other component, and the balance of dependence between the items will vary with the salience of the respective e-sites and the degree to which they are elaborated by the other item. The e-site of the relational predication *break* is its landmark (*lm*). The landmark is a salient substructure, which is moreover highly schematic and must therefore be said to be elaborated to a considerable degree by the specific nominal predication *appointment*. However, the noun also has substructure that is in its turn elaborated by the verb, and I would like to pursue the argument that, in an entrenched collocation, the salience of the substructure which functions as the e-site of the noun is greater than in a combination that is not entrenched.

As part of the analysis of APPOINTMENT as a complex category, ‘arranging/arrangement for a meeting’ was identified as one of the meanings of *appointment.* Using the FrameNet notions, the cognitive model on which it calls was referred to as a schematic frame, whereas the verbs *make, give, get, have, keep* and *break* were said to each profile a particular specific frame within the schematic frame, as the active zone in a typical usage situation (cf. section 3.4.3.4 above). These claims are based on my construal of the linguistic data. They should eventually be submitted to a psycholinguistic test, but for the purposes of the present argument I will assume that they have cognitive validity. In Fig. 43 below, specific frames are enclosed in solid circles to indicate that they are construed as salient substructure as well as potential e-sites in this particular schematic frame. Whereas elaboration of the verb is obligatory, elaboration of the noun is not, but still there is

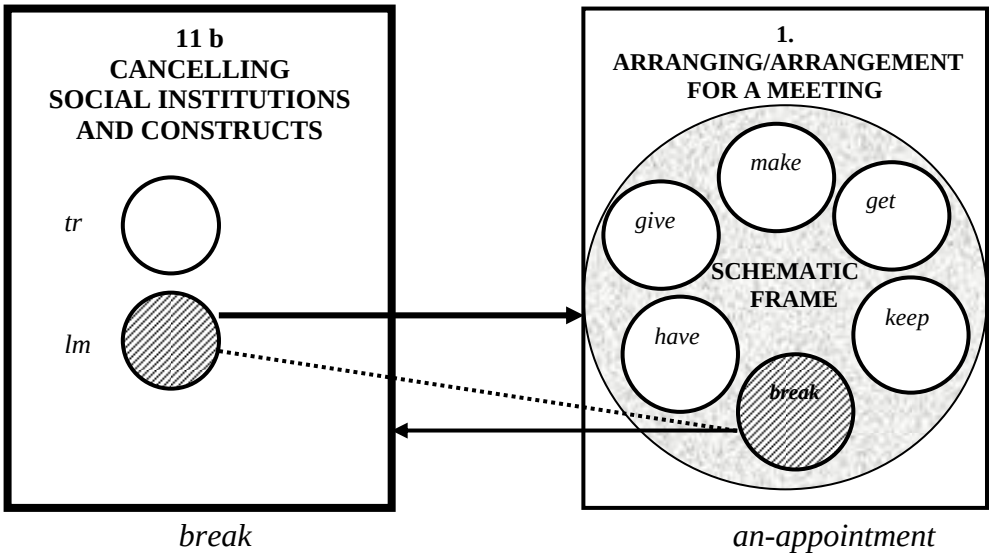


Fig. 43: A/D alignment of *break an-appointment* construed as an entrenched collocation

an 'elaboratory distance'[90] between the schematic frame and the active zone evoked by each of the specific frames. That the relationship between *break* and *appointment* in Fig. 43 is still seen as reflecting a certain asymmetry, with *break* being, on balance, the more dependent item, is indicated by the heavy solid arrow going from the landmark of *break* to *an-appointment*. However, my claim is that there is also substantial dependence going the other way, because of the elaboratory relationship between *break* and the corresponding specific frame of the noun. This is indicated by means of a somewhat thinner solid arrow. The cross-hatching of the circle symbolizing the specific frame indicates that it is construed as the e-site, whose elaboration by *break* both evokes a particular APPOINTMENT frame and imposes a specific perspective on it. In my view, the resulting A/D-alignment between the two items can be characterized as reflecting a substantial degree of **interdependence**, which I construe as being symptomatic of the entrenchment of this combination.

In the analysis of the composite structure *break an-appointment* (in section 3.4.4.1 above), all five examples found in the data were construed as belonging to APPOINTMENT frame 1, ‘arranging/ arrangement for a meeting’, but it is also possible to think of examples in which *break an-appointment* could evoke APPOINTMENT frame 5, 'accessories for people', or 4, 'accessories for cars, rooms, etc.', in which *appointment* refers to objects, such as a rear-view mirror, which can be

[90] ‘Elaboratory distance’ is explained by Langacker (1987: 301, note 21) as the amount of precision and detail that one component adds to another by elaborating it.

physically 'broken'. In such an example, I see the nominal predication as being less dependent on elaboration by the verb, resulting in the A/D alignment shown in Fig. 44 below.

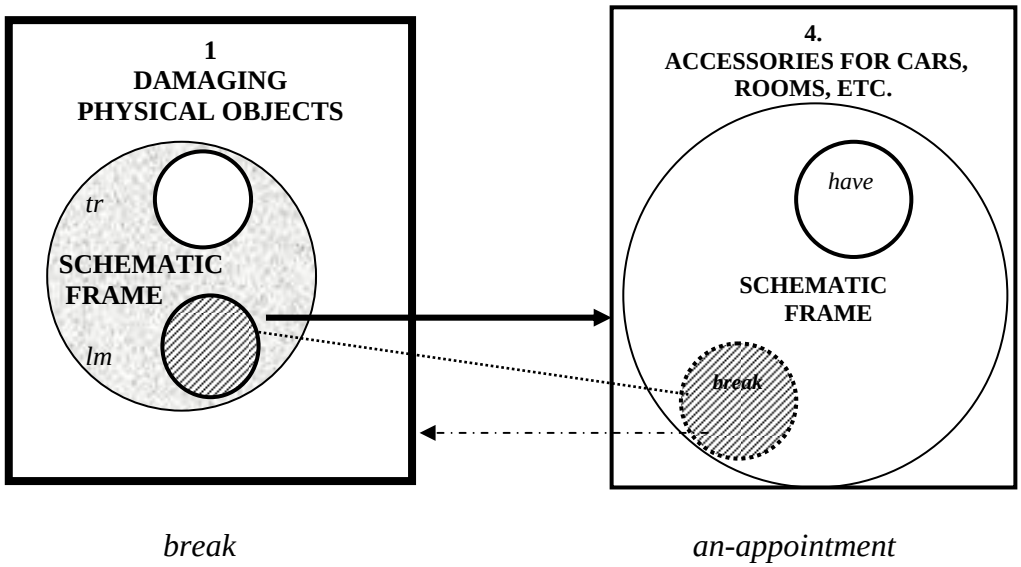


Fig. 44: A/D alignment of *break an-appointment* construed as a free collocation

Although the verb would help to identify the frame, I do not find that it elaborates any salient substructure of the frame, which is indicated by enclosing the e-site in a dotted circle alongside the more salient *have* frame, which was the only specific frame that could be documented by the data. The noun is therefore construed as being more autonomous in this schematic frame, and, by the same token, the A/D asymmetry is taken to be greater. The cognitive model that the composite structure calls on is understood to be the 'damaging of physical objects', which is the basic schematic frame evoked by *break* in the sensorimotor domain. The relatively high autonomy of the noun, which does not depend on the verb for elaboration of a salient substructure, is seen as evidence that *break an-appointment* is not entrenched in this frame.

The analysis of the schematic APPOINTMENT frame 'arranging/arrangement for a meeting', identified a number of lexical sets including the verbs found in specific frames. Moreover, the verbs belonging to each set were construed as showing a continuum of **schematicity**, or specificity, at the place of the verb (cf. Table 17 in section 3.4.3.2 and Fig. 39 in section 3.4.3.4). At the most schematic end of the continuum, the basic level verbs *make, get, give, have, keep* and *break* were seen as the default fillers of the verb slots in the corresponding specific frames. It was assumed that composite structures including these verbs were most likely to be entrenched while those using a

more specific verb might be so and those including the most specific expressions were considered to be ad hoc combinations and not very likely to be entrenched.

LEVEL OF CATEGORIZATION	**VERB EVOKING SPECIFIC 'BREAK FRAME'**	**SALIENCE**	**LIKELIHOOD OF ENTRENCHMENT**
Basic level default case, most schematic	***break*** *an-appointment*	high	high
Subordinate level less schematic, more specific	*cancel/* *miss an-appointment*	fairly high	fairly high
Subordinate level least schematic, most specific	*arrive late for/* *forget an-appointment*	fairly low	fairly low

Table 27: Schematicity and salience related to likelihood of entrenchment

As has been mentioned several times before in this study (cf. sections 2.1.5 and 3.4.1.1 above), basic level verbs are assumed to have a privileged cognitive status. They belong to the most frequent words in the language and, as pointed out by Taylor (1995: 49), they are "structurally simple', (i.e. monomorphemic)". They structure events as locations or objects according to general conceptual metaphors, or image-schemas, (cf. section 3.4.2.2), and they are associated with gestalt perception of the overall part-whole configuration of the background situation, or frame, which allows us to understand the meanings of words. Consequently, I would claim that basic level verbs are especially salient and therefore effective in evoking specific schematic frames, which I see as the main functional motivation for their participation in entrenched collocations.

Table 27 above shows the range of examples that were used as evidence for a continuum with *break* as the prototype (cf. Table 17 in section 3.4.3.2) and rates them for salience within the corresponding specific frame of the schematic APPOINTMENT frame ‘arranging/arrangement for a meeting’, which is related to their likelihood of entrenchment. In Fig. 45 below, this variation in salience is indicated by enclosing *break* in a heavier solid circle than *cancel* and *miss*, and by enclosing *arrive late for* and *forget* in 'dotted circles'. As in Fig. 44, a link is assumed between salience and the likelihood of entrenchment. Within the circle indicating the specific 'BREAK' frame, the solid circles enclosing *break, cancel*, and *miss* indicate that it is highly, or fairly likely, that *break, cancel*, and *miss an-appointment* are entrenched collocations. The dotted circles enclosing *forget* and *arrive late for* indicate that they may not be salient enough for the composite

structures *forget an-appointment* and *arrive late for an-appointment* to be entrenched. However, I still construe them as alternatives to the more schematic verbs. On the one hand, these more specific expressions elaborate the frame in more detail, which is reflected iconically in their greater length and complexity. On the other hand, they lack the schematicity of the basic level verbs which are the prototypical default fillers in the respective specific frames that, according to my analysis, constitute the salient substructure of the APPOINTMENT frame 'arranging/arrangement for a meeting'.

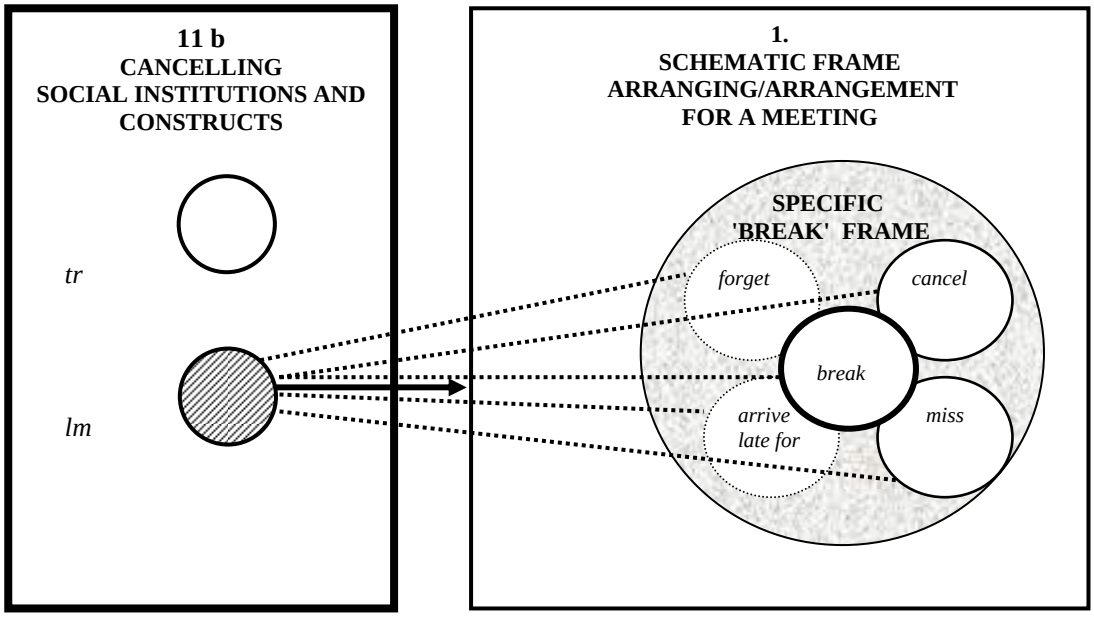


Fig 45: Range of collocational variability in the specific frame 'break an-appointment'

In contrast to entrenched collocations, which elaborate salient substructure of a frame in a more schematic way, I propose that such more specific ad hoc combinations could be categorized as '**open collocations**'. Referring to these expressions as ‘collocations’ is meant to emphasize the continuity with the more schematic entrenched collocations and to capture variability as an important dimension of convention. Referring to them as ‘open’ challenges Saussure's principle of 'the stacked deck' according to which individuals do not have much choice once a language convention has been established (cf. section 2.1.3.1 above). In a given usage situation, it would seem that there is typically freedom both to choose from a range of entrenched collocations and to choose a more specific expression which may not be conventional, but still perfectly 'normal'. The borderline between those combinations that are entrenched and those that are not is likely to be

fuzzy. For example, the difference in specificity between *fix* and *schedule an-appointment* does not seem all that great, yet the *Oxford Collocations dictionary for students of English* (Lea 2002), which is based on the BNC, includes only *fix* (two occurrences in the BNC), while *schedule* (one occurrence) is not included. We are dealing with prototype categories characterized by gradience, and no definite cut-off point can be posited for what is entrenched, although the data do reveal certain characteristics of entrenched combinations.

To make a distinction between entrenched and open collocations that evoke the frame of the noun and those expressions in which it is clearly the verb that evokes the dominant frame, I suggest that the latter may be categorized as **'free collocations'**. The use of 'free' implies that the choice of verb is only subject to general selection preferences, while the use of 'collocation' implies continuity between free collocations on the one hand and entrenched and open collocations on the other. Collocations that are perceived as free at one time may be perceived as entrenched at some other time simply because of a change in the contexts of situation leading to a change in the underlying cognitive models. Also, sociolinguistic differences of various kinds are likely to make a difference as to what expressions are understood to be conventional. If you are a disc jockey in a certain type of discotheque, *breaking* or *scratching* may be part of a familiar scenario associated with what you do with gramophone records; if you are a devoted lover of classical music, this is less likely to be the case.

In connection with my discussion of Halliday's definition of the 'lexical set' (in section 2.1.2 above), I suggested that rather than describing membership of the set as 'probabilistic', the set could be construed as a prototype category that mirrors habitual variations in use. In the light of my empirical studies, I see entrenched combinations with basic level verbs as typically forming the prototypical centre with more specific alternatives that may or may not be entrenched as natural extensions. The members of the set may be more or less synonymous, as in Halliday's example *strong*/*powerful tea*, but typically they represent a range of variability allowing for different conceptualizations of a situation, as in *break/cancel/miss an-appointment*, *make/arrange/rearrange* or *keep/attend/cycle to an-appointment*. Where Halliday (1966: 152) talks about members of the set having 'like privilege of occurrence in collocation', I would stress that the members of the set are typically not fully synonymous, and that their 'privilege of occurrence' depends very much on the actual usage situation. In my data, the combination *cancel an-appointment* occurred ten times,

while *break an-appointment* was found only five times. An explanation may be that the most schematic expressions are found to be too blunt to categorize a social situation that is inherently problematic; *cancel* may solve the problem by implying that the proper procedure is followed, while *miss* implies that the failure to turn up was not intentional. Defined in this way, as a prototype category, the set is seen as cognitively and functionally motivated and as cutting across the fuzzy borderline between combinations that are entrenched and those that are not.

Although it is assumed that combinations with basic level verbs are more likely to be entrenched than combinations with subordinate level verbs, it is not predictable that they will be so, as illustrated by the example in which *break an-appointment* refers to the damaging of a physical object (cf. Fig. 44 above in this section). Another example is *keep an-appointment*. Out of the 13 occurrences of the combination found in the data, 12 were recorded as belonging to the schematic frame 'arranging/arrangement for a meeting'. They were all prototypical entrenched collocations as defined above, in which *keep an-appointment* means 'come to', but it is also possible to find a context in which the verb means 'not change', or 'not cancel' as in the following example, which was found by means of a search in Google for ‘keep my appointment for’:

- *I talked to his nurse; she said that the results were "atypical". What does this mean. I asked her if she thought I should get in sooner for the colposcopy; she told me to keep my **appointment** for when it was scheduled, that way if anything was there it would give it a chance to grow more. I don't like the sounds of that.*[91]

Here it seems to me that the dominant frame is that evoked by the verb, i.e. 'to keep *something*' rather than elaborating any conceptually salient substructure of the APPOINTMENT frame. I therefore suggest that it is classified as a **free collocation**. The distinction between the two uses of *keep*, as an entrenched and a free collocation respectively, is in line with Bauer's (2001: 45) distinction between institutionalized expressions and potential words. The former are said to differ from the latter only in that "by being used they have come to have specific reference". In the last occurrence of the combination *keep* + *appointment*, it appears from the context that *appointment* means '(placing sb in) a position':

- *but the Duke of Montrose kept such **appointments** firmly in the hands of Graham gentlemen who looked upon him as their chief.*

[91] The example was found in the document titled *Abnormal paps and abnormal biopsies*, at http://forums.obgyn.net/womens-health/WHF.0301/1241.html

However, the dominant frame is that of the verb, i.e. 'to keep *something* in sb's hands', and it is therefore classified as a **free collocation**.

It follows from what has been said so far that the notion of schematicity associated with basic level categorization needs to be combined with a judgement about the salience of the substructure of a frame that is elaborated by a given verb. This is a matter of construal, and although linguistic data may be of help, a good deal of intuition will often be needed. An example of a combination that I find it quite hard to classify on the basis of the linguistic evidence is *lose an-appointment*, which was recorded as evoking the schematic APPOINTMENT frame '(placing sb in) a position' (cf. Fig. 40 and Table 24 in section 3.4.3.4). It fits into the object event-structure found in this frame, where the other basic level verbs are: *make, fill, give, get, take* and *hold*, but does it elaborate a salient substructure? My own intuition says probably not, especially as, in two of the three examples in the data, the participant 'losing' the appointment is not the one 'holding' it, but rather an institution (school, college). Also, I do not find that it lives up to Bauer's definition of institutionalized combinations that "by being used they have come to have specific reference", whereas *fill*, in the same frame, does. My intuitive judgement would be that *fill*, but not *lose*, elaborates a salient substructure of this APPOINTMENT frame and that the former is therefore an entrenched collocation evoking this frame, whereas the latter is a free collocation evoking the frame of the verb *lose*.

Again, it should be emphasized that we are not dealing with classical categories that can be defined in terms of criterial features. Rather it is assumed that salience is a gradable cognitive phenomenon and that it would be possible, by means of a psycholinguistic study, to place these composite expressions on a continuum according to the salience ratings of subjects. Just as the continuum of schematicity mentioned above was expected to cut across the borderline between entrenched collocations and open collocations, such a continuum of salience should be expected to cut across the borderline between entrenched collocations and free collocations. This results in the hypothesis for a typology of collocations, based on examples from the APPOINTMENT FRAME 'arranging/arrangement for a meeting' which is illustrated in Fig. 46 below. According to this typology, which is based on continuity and gradeability, the main distinction is between expressions that are cognitively salient and those that are not. Thus entrenched collocations (1), can be characterised in terms of high salience combined with high schematicity, whereas open collocations (2) combine high salience with low schematicity. High schematicity is also found in free

collocations (3), which are characterized as 'free', because they are construed as evoking the frame of the verb rather than that of the noun.

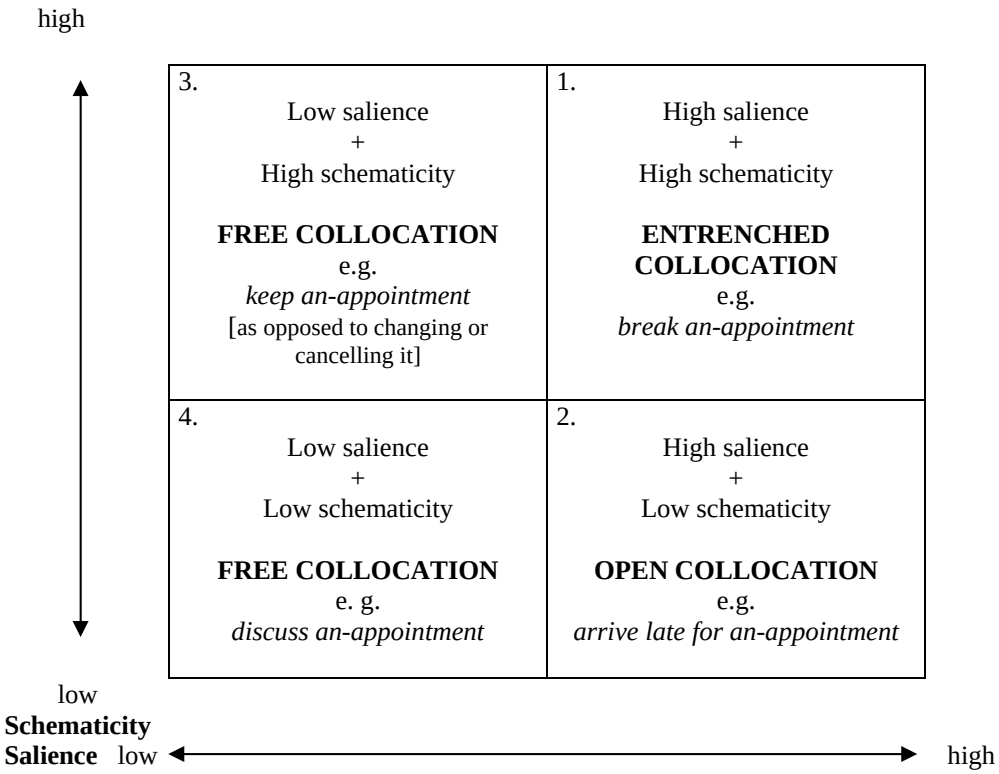


Fig. 46: A typology based on continua of salience and schematicity

Because they use the same event structure and basic level verbs as the most schematic entrenched collocations, there is affinity between the two types, and linguistic evidence will have to be supplemented by an intuitive evaluation of their salience. Finally, expressions that combine low salience with low schematicity (4), which are also characterized as free collocations, are considered to be the kind of combinations that are least likely to be entrenched. The two following examples were found in the data:

- *Diary check - a general discussion of **appointments** and availability for the week followed with Janet, our Co-ordinator.*

- *The King hoped next to see Baldwin, but the Conservative leader could not be found: he was, ironically, lunching with Geoffrey Dawson, the editor of The Times, and discussing Cabinet **appointments** ...*

It appears from the context that the first example refers to the *appointment* reading 'arranging/arrangement for a meeting', whereas the second example refers to the reading '(placing sb in) a position', but in both cases the dominant frame is the DISCUSSION frame evoked by the verb, in which the two kinds of appointment fill the slot of the object being discussed.

How do expressions that are normally classified as **'idioms'** fit into the typology suggested above? For a discussion of this question, I will now return to the expression *break the ice* (cf. section 3.4.2.4 above), which is the only example found in the data:

- *"Fancy a game of darts, lad?" Jos said to Mungo, breaking what was left of **the ice**.*

In the framework of phraseology (e.g. Cowie 1981, Howarth 1996), idioms are construed as one extreme of a continuum with free combinations at the other extreme and restricted collocations as the middle part (cf. section 2.1.4). In that framework, *break the ice* is classified as a figurative idiom, because the expression has a literal equivalent in contrast to a pure idiom like *kick the bucket*, which does not. Both types of idiom are seen as being units of meaning and therefore not compositional.

In contrast to this, cognitive linguists like Gibbs, Nayak and Cutting (1989) have come up with the decomposition hypothesis (cf. section 2.1.5 above). They see idioms as compositional[92] to the extent that the literal meanings of the parts can be related to the figurative meaning of the whole phrase (motivatedness) and to the extent that the figurative meaning can be distributed over its parts (isomorphicity). Corresponding to pure idioms in the phraseological framework, a class of 'non-compositional idioms' is identified, of which *kick the bucket* would again be an example. A conventional expression like *pop the question* is seen as 'normally decomposable', whereas *break the ice* would be classified as an 'abnormally decomposable' idiom, because it contains *ice* as a metaphor for an 'awkward or tense situation'.

[92] As noted in section 2.1.5 above, 'compositionality' is used by psycholinguists to mean 'analysability', which has been defined by Langacker (1987: 457) as an independent parameter that refers to the extent to which speakers are aware of "the contribution that individual component structures make to the composite whole" while compositionality refers to the degree to which the whole was predictable from the parts when the phrase was originally coined.

A similar approach to compositionality is found in Nunberg et al. (1994). Here a distinction is made between 'idiomatically combining expressions' and 'idiomatic phrases', depending on whether "parts of the idiomatic meaning can be put in correspondence with parts of the literal meaning". Whereas idiomatic phrases correspond to noncompositional idioms in the decomposition hypothesis, idiomatically combining expressions like *answer the door* and *spill the beans* correspond to both normally and abnormally decomposable idioms. According to this typology, *break the ice* would be classified as an idiomatically combining expression and also as an 'encoding idiom', because the correspondence between the literal and idiomatic meanings of the parts makes the expression quite transparent. This is not the case in a 'decoding idiom' like *kick the bucket* (cf. the discussion of compositionality and analysability in section 2.2.5 above). Table 28 below sums up the idiom status of the expressions that have been given as examples in the typologies mentioned.

PHRASEOLOGY (Cowie, Howarth)	**pure idioms (noncompositional)** *kick the bucket* *spill the beans*	**figurative idioms (noncompositional)** *break the ice*	**restricted collocation (partly compositional)** *answer the door*
COGNITIVE LINGUISTICS (Gibbs, Nayak, Cutting)	**noncompositional idioms** *kick the bucket*	**abnormally compositional idioms** *break the ice* *spill the beans*	**normally composi-tional idioms** *answer the door*
GENERAL LINGUISTICS (Nunberg, Sag, Wasow)	**idiomatic phrases noncompositional** *kick the bucket*	**idiomatically combining expressions (compositional)** *break the ice, spill the beans, answer the door*	

Table 28: The idiom status of selected expressions in three different typologies

Based on Bauer's (2001: 45) distinction between those established complex lexemes that are lexicalized and those that are institutionalized, and his comparison of lexicalized expressions to idioms (cf. section 2.2.2 above), most of the types of idiom shown in Table 28 would be classified as lexicalized, because they "have diverged from their original form or meaning to the extent that they could not be 'contemporary coinages'" and because of the loss of compositionality. Exceptions are the expressions referred to as restricted collocations, those referred to as normally compositional idioms, and some of those that are referred to as idiomatically combining expressions, which, in Bauer's words, can be said to "still form part of a synchronically productive series, differing only

from potential words in that, by being used, they have come to have a specific reference" (Bauer 2001: 46).

What linguists like Nunberg et al. (1994: 531) are doing, according to Croft and Cruse (2004: 252), is to "dissociate conventionality from noncompositionality". An idiomatically combining expression like *break the ice* is thus seen as both semantically analysable and as semantically compositional. Since it can be analysed according to general syntactic rules, it can further be characterized as 'a grammatical idiom', in contrast to 'extragrammatical idioms' (Fillmore et al. 1988: 505). A further distinction made by Fillmore et al. is that between 'substantive' or 'lexically filled' idioms, in which all elements are fixed, and 'formal' or 'lexically open' idioms', "in which at least part of the idiom can be filled by the usual range of expressions that are syntactically and semantically appropriate for the slot" (Fillmore et al. 1988: 505). Like Croft and Cruse (2004: 233 f.), I will use 'schematic', which is Langacker's term for a more general category, rather than 'formal'. In the example, *ice* is premodified by the quantifying expression *what was left of*, which shows that the idiom *break the ice* is schematic, or lexically open, to some extent. Just as *break the ice* is easily decoded as meaning 'reduce the tension', decoding *break what is left of the ice* as 'removing what is left of the tension' poses no problems; actually the idiom seems to lend a gestalt quality to the situation that makes it quite easy to grasp.

Whereas the continuum posited by cognitive linguists looks quite similar to that posited in phraseology, the underlying principles stand in stark contrast to each other. Since full compositionality is seen as the norm in phraseology, all expressions that are not fully schematic are seen as deviating from this norm, whereas in cognitive linguistics it is taken to be normal for a construction to be substantive, or lexically filled, to some extent. This has inspired 'construction grammar' as an approach that treats the grammar of a language as a collection of constructions ranging from the most substantive to the most schematic ones (cf. e.g. Croft and Cruse 2004: 249).

How is it possible to distinguish between entrenched collocations and idiomatically combining expressions, or abnormally compositional idioms in a cognitive framework that treats both as compositional? In other words, could *break the ice* be classified as an entrenched collocation? This would require that a schematic frame could be identified for *ice* meaning something like 'tension' in

which *break* could be seen as elaborating a salient substructure. It would take an analysis of the complex lexical category ICE to find out if this could be said to be the case.

Another possibility is to see the expression as originating from a conceptual integration of two different frames: one in which an agent physically breaks some ice, possibly with an instrument, and one in which an agent is relieving the tension on some social occasion, e.g. by telling a joke, singing a song or just fooling around. The meaning that emerges in the blend has an agent using his social skills as an instrument in the effort to warm up a party and get it going. As mentioned before (in section 3.4.2.4), the idiom can be seen as drawing on an underlying conceptual metaphor according to which close relations between people are 'warm' and poor relations are 'cold'.

3.4.4.3 Can a support-verb function be posited for *break*?

In this section I will suggest that in entrenched verb + nominal object collocations (as defined in the previous section) the verb can generally be construed as a support verb rather than a full verb, and that this can be seen as a further characteristic of entrenchment. This broad definition implies that a support verb construction is construed as a prototype category with more and less prototypical members. Another question, which will be discussed in section 3.4.4.5 below, is if support verbs have a grammatical rather than a lexical function. Considering that these are far-reaching questions involving the interface between lexical and grammatical meaning, the suggestions that I can make on the basis of my data will have to be rather tentative. Before returning to the notion of support verb as defined in frame semantics (cf. section 3.4.1.1 above), I will look at some other accounts based on data from English as well as Dutch, French and Danish to get an idea of what can be posited as the prototypical features of a support verb construction.

Bache and Davidsen-Nielsen (1997) in their grammar of English do not posit a category of 'support verb', but they note that when the predicator is realized by 'semantically general verbs' like *give, take, do, have* and *make*, **fusion between predicator and direct object** is likely to occur. In an example like *She made a complaint*, which follows the subject-predicator-object pattern, the effect will be to make the fused predicator 'semantically intransitive' (Bache and Davidsen-Nielsen 1997: 204).

Van Durme and van den Eynde (1998) distinguish between 'full verbs' and 'non-full verbs', the latter being in turn subdivided into 'function verbs' (auxiliaries and modals), which have verbal extensions, and 'support verbs', which have nominal extensions. The typology is illustrated by the following examples from Dutch, which is their example language (op. cit. 10):

*Hij **heeft** een fiets* 'He has a bike'	(full verb)
*Hij **heeft** behoefte aan genegenheid* 'He has (a) need for affection'	(non-full verb, support verb)
*Hij **heeft** gewerkt* 'He has worked'	(non-full verb, function verb, auxiliary)
*Hij **heeft** te werken* 'He has to work'	(non-full verb, modal)

Support verb constructions are characterized as **predicators with their own valency scheme**, 'valency' being defined as "the kernel-specific combinatorial potential of lexical elements". Besides, the formal criterion of '**non-proportionality**' is used. While, in the case of full verbs, there is said to be "a constant, proportional relationship between a pronominal paradigm and its corresponding lexicalised constituents" (Van Durme and van den Eynde 1998: 8), constructions with non-full verbs, including support verbs, are not 'proportional to pronominal paradigms'; instead they take 'nominal non-proportional extensions' (Van Durme and van den Eynde 1998: 12 ff.)[93]. The following examples (op. cit.: 14) are used to illustrate the point that support verbs are not proportional to pronominal paradigms:

*De man **heeft behoefte** aan een goed gesprek*
'The man has a need for a good conversation'

De man heeft **dat aan een goed gesprek*
'The man has that for a good conversation'

De man heeft **dat (= behoefte aan een goed gesprek)*
'The man has that (= need for a good conversation)'

Full verbs, on the other hand, are proportional to pronominal paradigms:

*Ik heb **die behoefte** niet uitgevonden*
'I have not invented that need'

[93] For a brief account of the 'pronominal approach' as part of a constructivist approach to syntax, see van den Eynde, 1997.

*Ik heb **ze** niet uitgevonden*
'I have not invented it'

Van Durme & van den Eynde do not characterize support verbs and their nominal objects in terms of their semantics as is often done. However, their example, the Dutch equivalent of *have a need for sth*, is of the kind that is typically offered: delexicalized verb + deverbal noun, which I consider to be a good candidate for the prototype of the category of support verb constructions. Trying to identify support verb constructions by looking for a pattern of combinatorial behaviour that deviates from that of full verbs is a typical approach.

My data do not allow me to test the claim of non-proportionality to pronominal paradigms, but a supplementary search of the BNC for the combination of *break/breaks/broke/broken + it/that* found several examples of what I consider to be entrenched collocations on a par with *break an-appointment*:

- *If it means I break my **contract** here - fine. In fact, why don't I break **it** now? Consider **it** broken.*
- *A ruler is bound by the good old **law**, if he breaks **it** in any serious way, his subjects can rebel, and by formal process compel him to obey the law.*
- *As a **rule**, a striped tie should never be worn with a striped shirt, although Paul Smith frequently breaks **that one** with gusto.*

These examples show that in the collocations *break a contract*, *break the law* and *break a rule*, *break* would not qualify for membership of the category of support verbs construed as a classical category with proportionality as a criterial criterion of membership. However, if support verb constructions are construed as a prototype category, and if non-proportionality to pronominal paradigms is taken to be a prototypical rather than a criterial feature, entrenched collocations with *break* may still be included, although they may not be the most prototypical members. Such a deviation from the prototype may perhaps be explained by *break* being less schematic than most other basic level verbs. As pointed out in section 3.4.4.1 above, *break* can be understood as including an element of negation (‘un-make’ or ‘un-keep’). This was offered as a possible explanation of the use of the past participle as premodifier in examples like *a broken appointment* and may perhaps also explain why support verb constructions with *break* are proportional to pronominal paradigms, as in the examples above.

In her diachronic, corpus-based investigation into valency patterns in French, Schøsler (2003: 401), in addition to the criterion of non-proportionality, uses **constraints on determiners** as a formal criterion to tell support verb constructions apart from 'free constructions':

> *Luc raconte (une/cette/ton) attaque contre la citadelle*
> 'Luc tells about one/this/your attack on the fortress'
>
> *Luc mène une/*cette/*ton) attaque contre la citadelle*
> 'Luc is conducting an/*this/*your attack on the fortress'

The first example indicates that "free constructions have free variation of determiners" while the second example makes the point that support verb constructions "have specific constraints on determiners" (Schøsler 2003: 402 f.).

My data for *break* include only a limited number of finite active transitive constructions that may serve to test whether there are any constraints on determiners where *break* is used in the nonsensori-motor domain as part of an entrenched collocation. Actually, there are only five examples of constructions in which *break* is a full verb combining with the nominal objects *street lighting, glass door,* and [gramophone] *records* and six examples in which *break* may be construed as a support verb in combinations with the nominal objects *agreement, link, contract* and *vows of silence*. It turns out that the same kinds of determiner occur in both groups, viz. definite article and possessive pronoun:

- *... an explosion has broken all **the street lighting**.*
- *Somebody had broken **the glass door** of the shop*
- *He has just broken one of **his records** deliberately*
- *within two days the RPR had broken **the** initial **agreement**.*
- *Had the Government not broken **the link** in 1980; the basic state pension would now be ...*
- *Or suppose that A has broken **his contract** to sell land to B ...*

As argued in the case of the formal criterion of proportionality to pronominal paradigms, if constraints on determiners are taken to be a criterial feature of a classical category of support verb constructions, this would seem to exclude entrenched collocations with *break* from the category, but if it is merely taken to be a prototypical feature, they may still be included as less prototypical members.

Baron and Herslund (1998), in their article about Danish support verb constructions, which they call 'verbo-nominal predicates', distinguish two kinds of object on the basis of different degrees of cohesion with the verb: the 'effected object', which comes into existence as an effect of the verbal action, and the 'affected object' whose referent exists prior to the verbal action.[94] Only **constructions with an effected object** are construed as support verb constructions (Baron and Herslund 1998: 91). Effected objects include deverbal nouns[95] with suffixes, e.g. *undersøgelse* 'examination', which are said to have their own argument structure which merges with the argument structure of the support verb. A second type is deverbal nouns without suffixes, which "no longer are felt as clearly verbal" such as *krav* 'demand' (Baron and Herslund (1998: 97). Like the third type: non-verbal nouns without an argument structure of their own, such as *bro* 'bridge' or *kage* 'cake', these nouns, are said to get their argument structure from the support verb (Baron and Herslund 1998: 97).

The verb is described as typically having a very general or 'bleached' meaning, such as *have* 'have', *være* 'be', *lave* 'do', which is said to be entirely dependent on the meaning of its nominal object. According to Baron and Herslund (1998: 92),

> ... the specific nature of the activity described by the verb follows entirely from the lexical content of the object noun, so that all that is left for the verb to express is a general notion of creation, the 'creation being of course different in each case as a function of the specific object constituent.

With non-verbal objects, they say that the support verb tends to be more specific, "because there is, in this case, little or nothing verbal in the noun itself". Examples are *begå et mord* 'commit a murder, and *bage en kage* 'bake a cake' (Baron and Herslund 1998: 97). Thus, although it seems that they see expressions with a general verb as prototypical, expressions with more specific verbs are also included in their category of support verb constructions: "It is not either a question of special verbs, apart from the necessary and sufficient condition that the verb must, as a consequence of its meaning be able to take an effected object" (Baron and Herslund 1998: 92).

The brief presentations given above of different treatments of the notion 'support verb' have been included to help identify a plausible candidate for a prototypical support verb construction. The

[94] The reference they give for the origin of this distinction is: Jespersen (1924).

[95] Baron and Herslund use the term 'verbal noun' rather than 'deverbal noun'.

prototype they converge on is a verb that is fairly general in meaning followed by a deverbal noun with which it forms a predicative unit. Less prototypical instances of the category are those in which the verb expresses meanings that are not so general and in which the verb is not so clearly deverbal or not deverbal at all. The intuition that verb and nominal object are fused in this construction type is supported by evidence of various types of formal constraint compared to full verb constructions. Assuming that support verbs can be construed as a prototype category, I have argued that these constraints should be seen as symptomatic rather than criterial and thus cannot be expected to apply to all members of the category. Whether specific types of constraint are language specific or not is another matter, which will not be pursued in this study.

Turning now to the notion of support verb constructions in frame semantics, we find the same prototypical features, for example, in the principles laid down by the FrameNet Project for the annotation with nouns as targets (cf. section 3.4.1.1 above). However, in the frame semantic account, the structural elements of the description are embedded in the frame, which is the cognitive model underlying them. Support verb constructions like *take revenge* are characterized as including a frame-bearing noun in which it is the noun that evokes the dominant frame. This is in contrast to full verb constructions like *encourage revenge*, in which the dominant frame is evoked by the verb (Johnson et al. 2003: 2.3.1). Since it is the frame-bearing noun that evokes the frame and typically provides the argument structure, support verb constructions take a frame element as their subject, which is not the case in a full verb construction. Starting from the frame provides a cognitive and functional motivation for the notion of support verb constructions as entrenched collocations that evoke semantic frames and thereby contribute not only to their own 'survival' but also to that of the frame. It also motivates an extension of the range of support verbs beyond the most semantically neutral, resulting in a definition of support verb that "is broader than most others" (Atkins et al. 2003: 280). Ruppenhofer et al. (2002: 368) distinguish between **'verbal meaning'** and additional **'configurational information'**:

> While many such verbs (support verbs and other sorts of lexical functions in the sense of Mel'čuk)[96] add configurational information of one kind or another to the verbal concept (features of aspect,[97] point of view, evaluation, etc.), their main function in many cases is to combine with the nominal object to express verbal

[96] Fillmore et al. (2003b: 250, note 5) describe their notion of support verb as "broader than the traditional notion of 'light verb'" but "narrower than Mel'čuk's 'lexical functions'".

[97] 'Aspect' as used here corresponds to 'Aktionsart' or 'actionality' as well as 'event structure' in other accounts (cf. footnote 102 below).

meaning: all such pairings (verb + object noun) can count as MWEs,[98] those in which the verb is lexically selected by the noun, and should be listed separately in the lexicon.

Recent research by members of the FrameNet team describe these types of 'configurational information' in more detail. Fillmore et al. (2002: 790) state that they "go beyond the familiar list (often called light verbs)", exemplified by combinations with *make, take, have, give*:

1. **Disambiguating an object noun:**
   to have an ***argument*** ('to quarrel') vs. *to make an* ***argument*** (' to reason')
   to keep an ***appointment*** ('to come to a meeting') vs. *to hold an* ***appointment*** ('to have a position')

2. **Profiling different participants:**
   to perform an ***operation*** vs. *to undergo an* ***operation***
   to give an ***appointment*** vs. *to get an* ***appointment***

3. **Profiling different phases of a scenario:**
   to make a ***promise*** vs. *to break a* ***promise*** vs. *to keep a* ***promise***
   to make an ***appointment*** vs. *to break an* ***appointment*** vs. *to keep an* ***appointment***

4. **Selecting a particular register:**
   to make a ***complaint*** vs. *to file, lodge,* or *register a* ***complaint***
   to make an ***appointment*** vs. *to arrange, book,* or *fix an* ***appointment***

5. **Allowing the direct object of a transitive event type to be the subject**
   to require ***examination***

Fig. 47: The frame semantics of support verbs I (based on Fillmore et al. 2002, section 4.2)

"In addition to providing the possibility of converting an event noun into a verbal predication and finding a place for one of its arguments, support verbs serve many important functions" (Fillmore et al. 2002: 790). Five different functions are listed in Fig. 47 above, which all depend on the choice of support verb. The examples in the first line are from the article [my emphasis], and for four out of the five functions I have included examples from my own data in the second line. This broad definition deviates from the prototype in that not only combinations with an effected object are included, but also other combinations in which the dominant frame is that of the noun and which furthermore impose a certain perspective on that frame.

That researchers related to the FrameNet project have produced a broader notion of support verb than others, I believe, is due to the inductive, corpus-based method that is used (Fillmore et al. 2002: 788). Furthermore, the fact that each lexical sense of a keyword is seen as corresponding to its use in a given semantic frame (Atkins et al. 2003: 254) provides the possibility of accounting for

[98] MWE stand for 'multiword expression', a lemma comprised of more than one lexeme (Ruppenhofer et al. 2002: 361).

a complex category of support verb constructions in a principled way, based on the functions that they perform in combinations with frame elements.

An even more complex notion than the one outlined above is found in a power point presentation by Fillmore (2003a) titled *Multiword expressions. An extremist approach*.[99] As shown in Fig. 48 below, support verbs are found to occur not only with subject and object nouns, but also with adjectives and prepositions. The examples that are closest to the prototype are found in **group 2**, support verbs with an object noun, but the two examples *wage war* and *commit a crime* are atypical in that the noun is not deverbal. Also this group includes examples of what Fillmore calls 'ditransitive support verbs', as in *give her a kiss*. However, if *give - a kiss* is construed as a complex, discontinuous predicator with *her* as the object, these constructions could also be classified as monotransitive in contrast to complex intransitive constructions like *make a decision*. The examples in **group 3** deviate from the prototype by including a preposition, as in *take into account*, and by not necessarily including a deverbal noun. The examples listed in **group 1**, in which we find the support verb in intransitive constructions with the frame-bearing noun as the subject are not normally referred to as support verb constructions at all, but since the verb can be said to support the frame evoked by the noun as in the more prototypical cases, I consider these constructions to be a motivated extension of the category. A further extension is found in **group 4**, in which the support verb is found in frames evoked by an adjective, as in *turn red*.

1. **Support verb with subject noun:**
   *the **wind** is blowing, the **fire** is burning, a **riot** occurred, an **accident** happened*

2. **Support verb with object noun:**
   *have an **argument**, make a **decision**, wage **war**, commit a **crime***
   *give sb a **kiss**, pay sb a **bribe,** give sb **advice***

3. **Support verb with preposition + noun:**
   *take sth into **account**, have sth in one's **possession**, take sth under **consideration***

4. **Support verb with adjective:**
   be + any predicative adjective, *go **crazy**, get **rid** of, turn **red***

Fig. 48: The frame semantics of support verbs II (based on Fillmore 2003a: slides 51-66)

The presentation, in addition to broadening the notion of support verb, also extends the notion of support construction itself to include 'support prepositions'. They are found in expressions like *at*

[99] The presentation was given at Pennsylvania University, at the Kick Off Meeting for Robust Semantic Parsing NSF – ITR, November 9-10, 2003.

risk, in danger, on fire in which the noun is described as the 'semantic head' of the lexical unit made up by preposition + noun. This seems to pave the way for including further syntactic types of entrenched collocation (cf. Table 1 in section 1.2 above). The adjective in combinations like *a rapid increase, a deep cut,* and *a wide gap* might thus be classified as 'support adjectives', the adverb in combinations like *to be deeply wounded, widely dispersed* or *commonly known* might be classified as 'support adverbs', and nouns in combinations like *a stroke of luck* and *peal of thunder* might be said to function as 'support nouns'.[100] Such a broad category of support constructions would include most of the syntactic patterns of entrenched collocations (cf. Table 1 in section 1.2 above). However, the idea that support constructions also constitute a prototype category will not be pursued any further in this study; instead I will return to the question of a prototype category of support verb constructions consisting of a verb and its nominal object.

In the part of Fillmore's power point presentation that is called "Beyond 'light' verbs" (Fillmore 2003a: slides 60-65) a **continuum** is presented ranging from 'simple cases' over 'more nuanced cases' to 'more extended cases'. In simple cases the verb is said to have "essentially no meaning except to reveal that its subject is necessarily a participant in the event named by the noun". The subject may play an 'active role', as in *give an exam*, or a 'passive role', as in *sit/take an exam*. In the more nuanced cases the verb is said to contribute "information about register, attitude, Aktionsart, or the like", exemplified by *register a complaint*. In the more extended cases, such as *pass/fail an exam* or *keep/break a promise*, the verb, according to this account, "identifies its subject as a participant in the larger scenario associated with the event named by the verb."

Fillmore's presentation (2003a: slide 57) refers to a test of support verbs according to which "their nominal object can't really be interrogated - meaning that the verb in question isn't functioning as a self-standing verb." The question *What have you made?* followed by an answer like **A decision to go home* would not, it is argued, be a 'natural conversation' just as it would not in most circumstances seem natural to ask **What did you do to/with the appointment?*[101] or **What did you break?* when referring to APPOINTMENT frame 'arranging/arrangement for a meeting'. The point is that

[100] Fillmore et al (2002) use the term 'transparent structures' to refer to refer to support verbs as well as to 'transparent nouns', which are "those nouns that occur initially in a *Noun + Noun* construction for which the surrounding context selects, or is selected by, the second noun rather than the first."

[101] A context in which *appointment* refers to a written document would make the question alright, but then it would not be followed by the answer *I broke it*.

verb and noun are interdependent in support verb constructions. Interdependence between *break* and *an-appointment* has also been found to be a feature of the entrenchment of this combination in terms of A/D-alignment (cf. section 3.4.4.2 above).

My conclusion is that the notion of a prototype category of support verb constructions defined on the basis of frame semantics is compatible with the notion of entrenched verb + nominal object collocations, which is also frame-based (cf. section 3.4.4.2) above. Consequently, I suggest that the two notions are merged, so that an entrenched collocation like *break an-appointment*, in which the noun is frame-bearing, is characterized as a support verb construction. Formal evidence supporting this functional frame-based construal was found in the analysis of the data on *break* (cf. sections 3.4.2.3 and 3.4.3.3 above). Moreover, thinking in terms of Firth's interdependent and mutually congruent levels of meaning (cf. section 2.1.2 above), I find that evidence from the levels of collocation and syntax (including colligation) converge with evidence at the level of context of situation, which I have argued is compatible with the cognitive notion of semantic frames.

3.4.4.4 To what extent is *break an-appointment* grammaticalized?

My findings so far have led me to suggest that the notion of entrenched verb + nominal object collocation and support verb should be merged, which amounts to saying that the verb collocate in an entrenched collocation has a more grammatical function compared to the same verb in a free collocation or in an idiom (cf. section 3.4.4.2 above). I will conclude my analysis of the integrated structure *break an-appointment* by discussing whether *break* as a support verb can be said to have a grammatical meaning also in a more traditional sense, which includes belonging to a (relatively) closed class. As stated in my research question, this involves a discussion of its functional role and the extent to which it is manifested in more general patterns of use (cf. section 3.3.1 above), but it also has to do with what we mean by 'lexical(ization)' and 'grammatical(ization)'.

In Langacker's cognitive grammar both lexical and grammatical morphemes are seen as symbolic units pairing form and meaning, the only difference being that grammatical patterns are more schematic. The false dichotomy between lexicon and grammar, or content words and function words, according to Langacker (1987: 18), has to do with the mistake of using prototypical

characteristics as criterial features: “One way to produce a false dichotomy is to focus solely on representative examples from the two extremes of a continuum: by overlooking intermediate cases, one readily observes discrete classes with sharply contrasting properties." In this view, grammatical structure is seen as forming a continuum with the lexicon, and assuming that meaningful elements cannot be characterized as grammatical is referred to by Langacker as the 'exclusionary fallacy’. However, the idea of a continuum is not to be seen as an escape from rigorous description: “ ... we must still describe the individual structures in explicit detail, even as we articulate their parameters of gradation” (Langacker 1987: 19).

In cognitive linguistics, grammatical constructions, like other symbolic structures, are construed as gradable prototype categories including central as well as peripheral members. Thus support verb constructions can be included in the category ‘transitive construction’ although they do not share many of the prototypical properties of that construction type (Taylor 1995: 206 f.) and although they function as intransitive constructions (cf. section 3.4.4.3 above). Support verb constructions, in their turn, can also be construed as a prototype category, of which *break an-appointment* is not a prototypical member, nor does it belong to the most peripheral ones. To arrive at a preliminary understanding of the possible grammatical status of *break* when used as a support verb in entrenched collocations, I will look to some of the mechanisms posited in grammaticalization studies, emphasizing those that can be applied to synchronic data on language variability.

The term ‘grammaticalization’ is attributed to the French linguist Antoine Meillet (1912), who defined it as “l’attribution du caractère grammatical à un mot jadis autonome”, ‘the attribution of grammatical character to an erstwhile autonomous word’ (Hopper and Traugott 1993: 18). The idea that grammatical expressions have evolved from lexical expressions can be traced back further, e.g. to the German philosopher and humanist Wilhelm von Humboldt (1761-1835), who suggested that grammar had developed gradually from an early stage at which only concrete objects were denoted, and to the German grammarian Georg von der Gabelentz (1840-1893), who saw grammaticalization as the outcome of competition between the tendency towards ease of articulation on the one hand and the tendency towards distinctness on the other. With time, he said, linguistic forms would ‘fade’ or ‘grow pale’ (*verblassen*) and their colours would ‘bleach’ (*verbleichen*) (Hopper and Traugott 1993: 19 ff.). The same metaphor is found in present day accounts discussing whether grammaticalization involves loss of meaning, or ‘bleaching’.

Because mainstream linguistics came to be dominated by a synchronic, structuralist approach, work on grammaticalization after Meillet was relegated to the special domain of Indo-European studies, according to Hopper and Traugott (1993: 25), who explain its revival as a major theme of general linguistics in the 1970s against the background of a growing interest in pragmatics and typology and a "questioning of autonomous syntactic theory." This revival is often associated with Givón, who coined the much quoted phrase "Today's morphology is yesterday's syntax" (Givón 1971: 413). Twenty years later Heine et al. (1991) offered their own version of this slogan: "Today's syntax is yesterday's pragmatic discourse", thus stressing the continuity between discourse and grammar. In her ph.d. thesis, Lindström (2004: 10-12), gives a detailed account of the origin of the term as well as of the roots of the concept of grammaticalization, which has a much longer and more complex history than implied by the few basic facts that I have included here.

It is generally agreed among linguists of different camps that there is a phenomenon to be studied, but substantial disagreement on the nature of the phenomenon and how it should be approached. While some linguists see grammaticalization as involving the same mechanisms as language change in general, others are more inclined to see it as involving distinct processes that need to be accounted for by a special theory (see for example Newmeyer 2000; Campbell and Janda 2000b). This has turned out to be controversial, also because grammaticalization is sometimes presented by functional linguists as a challenge that cannot be met equally well by a formal approach (cf. Fisher et al 2000: 8-14).

From a nonfunctionalist viewpoint, Newmeyer (2000: 188) refuses to acknowledge grammaticalization as a special theory, claiming that "…(Campbell and Janda 2000) the set of phenomena that fall under its label are a simple consequence of principles that any theory – whether formal or functional – would need to posit anyway." Against this it is argued by Heine (2003, in Lindström 2004: 56) that "... grammaticalization theory is a theory to the extent that it offers an explanatory account of how and why grammatical categories arise and develop." A chronological survey of definitions by Campbell and Janda (2000: 94 ff.) includes one by Haspelmath (1999: 1045), which allows the inclusion of cases in which the outcome of a development is not indisputably 'grammatical':

> The most general definition of grammaticalization would therefore not restrict this notion to changes from a lexical category to a functional category but would say that grammaticalization shifts a linguistic expression further toward the functional pole of the lexical-functional continuum.

Defined so broadly, grammaticalization mechanisms may be said to apply to the development of full verbs into support verbs. Support verbs do not constitute a typical closed class, both because the category includes such a wide range of verbs and because they are more idiosyncratic in their combinatory possibilities than for example auxiliary verbs and articles. On the other hand, they do have many of the characteristics associated with members of closed word classes (cf. Bache and Davidsen-Nielsen 1997: 35 f.). They are function words that combine with open-class words (nouns), and they enter into complementary relationships with each other: within a specific frame, they can be said to "share a functional domain" (Bache and Davidsen-Nielsen 1997: 36). Also, the admission of new members is not 'free', in that it requires the conventionalization of a verb + nominal object combination.

Sweetser (1988) presents a unified view of language change based on 'metaphorically structured meaning transfer' in an attempt "to treat the semantic changes attendant on grammaticalization as describable and explicable in the terms of the same theoretical constructs necessary to describe and explain lexical semantic change in general." Such a unified view seems compatible with the view that lexicon and grammar form a continuum. Furthermore, the fact that linguists sometimes do not fully agree as to whether certain changes should count as grammaticalization or as lexicalization (Fisher et al 2000: 5; Cowie 1995: 185 ff.) indicates that the exact borderline is a matter of construal rather than a matter of fact. Sweetser (1988: 389) claims that "an analysis of meaning change as metaphorically structured will [...] allow us to predict which inferences are preserved across transfers of senses." Metaphorical extension, according to Sweetser, always involves the abstraction from the original, lexical sense of an image-schema, which is then mapped on to a new domain. Image-schematic structure from the source domain will be preserved while more specific aspects of meaning are lost; however, new meaning will be gained from the target domain. In a case of grammaticalization, such as when *be going to* comes to be used as a future marker, the schema is mapped on to a fairly abstract, topological domain of grammatical meaning, and there will be less 'fleshing-out' of meaning than if the target domain is not grammatical (Sweetser 1988: 393). However, grammaticalization, exemplified by the *go*-future, does not merely involve the loss of meaning: "we have rather exchanged the embedding of this image schema in a concrete, spatial

domain of meaning for its embedding in a more abstract and possibly more subjective domain" (Sweetser 1988: 392).

Because of the phenomenon of '**layering'**, i.e. the situation that "the forms reflecting various stages of grammaticalization and non-grammaticalized forms occur side by side" (Fisher et al. 2000: 3), it is possible to identify patterns of metaphorical extension through the analysis of synchronic data, as has been attempted in this study. In the analysis of the lexical category BREAK (cf. sections 3.4.2.1 above), cross-domain metaphorical mappings were seen as an important constitutive factor, as image-schematic structure from the most basic sensorimotor subdomain of 'artefacts and natural things' was used in the less basic sensorimotor subdomains as well as in the nonsensorimotor domain. The analysis found that the lexically specific image-schematic structure could be related to conceptual metaphors underlying event structure in general (cf. sections 3.4.2) and it was argued that *break an-appointment* can be seen as an example of general object event-structure cutting across many different lexical categories (cf. section 3.4.3.2 above).[102]

The source domain for both the object event-structure metaphor and the location event-structure metaphor is the physical, spatial domain, while the target domain is the more abstract domain of causes and events. In the case of *break*, the image-schema that can be abstracted from use in the physical, spatial domain (the basic domain of 'artefacts and natural things') can be said to be the object event-structure, or the action category of the sentences in which it occurs. As pointed out by Bache and Davidsen-Nielsen (1997: 195), action category classifies situations, not verbs, although "Many verbs have a strong potential with respect to the specific values of the action category." As argued above (cf. section 3.4.4.3), I would say that *break* and the other basic level verbs do have such a strong potential, and that their event structure and the action categories with which these verbs are typically associated are closely associated with their functional role in structuring semantic frames.

[102] Lakoff and Johnson relate event structure to the notion of 'aspect', referring to Comrie (1976).

According to Bache (1985), Comrie exemplifies the tendency "to define aspect so broadly that it includes distinctions which belong to Aktionsart." Aktionsart, the type of situation coded by a sentence, is referred to as 'action categories' in Bache and Davidsen-Nielsen (1997: 192), and the "study of situations expressed by sentences" is called the study of 'actionality'. The primary distinction made is that between 'dynamic' and 'stative situations'. 'Aspect', on the other hand, is described as a closely related concept, which is defined as "grammatically expressed assignment of 'situational focus'" (Bache and Davidsen-Nielsen 1997: 285 f.) The main distinction here is between a 'perfective' and an 'imperfective' viewpoint.

In addition to metaphor, which is based on analogy, also metonymy, associated with **reanalysis**, has been found to play an important part as a 'trigger' of grammaticalization. According to Hopper and Traugott (1993: 61 f.), "[r]eanalysis and analogy are the major mechanisms in language change. They do not define grammaticalization, nor are they coextensive with it, but grammaticalization does not occur without them." Langacker (1977: 58) defined reanalysis as "change in the structure of an expression or class of expressions that does not involve any immediate or intrinsic modification of its surface manifestation." Reanalysis, in other words, is 'hidden' and a possible context in which it may first have occurred has to be hypothesized. or 'reconstructed'. Thus the development of *be going to* into a future marker is believed to have started with syntactic reanalysis in a context in which *going* was ambiguous between a full directional verb: *be going* [*to visit Bill*] and a future auxiliary: [*be going to*] *visit Bill*. The reanalysis is an example of local metonymic change, which according to Hopper and Traugott (1993: 87) "involves specifying one meaning in terms of another that is present, even if only covertly, in the context." The reanalysis only becomes 'visible', when, by analogy or metaphorical extension, the expression is used in contexts in which the original analysis would not make sense, such as **be going* [*to like Bill*].

As appears from the examples (from Hopper and Traugott 1993: 88), syntactic reanalysis is shown as '**rebracketing'**. Surface changes that may follow reanalysis, according to Langacker (1977: 58), can be "viewed as the natural and expected result of functionally prior modifications in rules and underlying representations." I believe that the formal constraints on the use of *break* in the nonsensorimotor domain that have been documented in this study (cf. section 3.4.2.3) can plausibly be construed as surface evidence that at some point reanalysis is likely to have taken place. These constraints were found to be compatible with the construal of *break* as a support verb. Also the analysis of *break an-appointment* in terms of A/D alignment, which found interdependence between the constituents (cf. section 3.4.4.2 above), is compatible with an analysis that brackets verb and noun group together. Furthermore, as appeared from the previous section, linguists working within a range of frameworks distinguish between two functional analyses of the same string of word forms (transitive verb + noun group): one in which the noun group is the direct object, and one in which it forms part of a complex predicator. Their findings, I believe, can be taken as independent evidence that construal of such a string as a support verb construction amounts to reanalysis.

Bybee (1985: 11-13) describes a continuum of three major expression types ranging from the most highly fused lexical expressions over inflectional expressions to the most loosely joined syntactic (or 'periphrastic') expressions. Two expression types are described as intermediate between the major types, viz. expressions with derivational morphology and expressions with 'free' grammatical morphemes, which include clitics, particles and auxiliaries. The full range including lexical, derivational, inflectional, free grammatical, and syntactic expressions, is shown in Table 29 below with definitions adapted from Bybee, to which I have added my own examples.

EXPRESSION TYPE	**DEFINITION**	**EXAMPLES**
Lexical	Two or more semantic elements may be expressed in a single monomorphemic lexical item	*break* (combines 'damage' or 'negation' and 'cause')
Derivational	Like lexical expressions, derivational morphemes are often restricted in applicability and idiosyncratic in formation or meaning. As in inflectional expressions, two [or more]distinct morphemes are combined in a single word.	*unkept* *appointment* *sweetly*
Inflectional	Each semantic element is expressed in an individual unit, but these units are bound into a single word. Inflectional expression may be in the form of affixes added to a stem, or a change in the stem itself.	*missed* *broke*
Free grammatical (with clitics, particles, and auxiliaries)	They have properties of grammatical morphemes, i.e. they belong to a closed class and occur in a fixed position, but are not 'bound' to any lexical item, and thus are not inflections. Like inflections, they make up contrast sets that are obligatory in certain environments and have positional restrictions. Like periphrastic expressions, they are not 'bound' to lexical stems.	*I've* *they'll* *break down* *run off* *have gone* *keep going*
Periphrastic/ syntactic	The different semantic elements are expressed by totally separable and independent units, that is in separate words.	*break the vase* *move the piano*

Table 29: A continuum of expression types (based on Bybee 1985: 11-13)

In this continuum, support verb constructions like *break an appointment* may be said to have affinities with periphrastic expressions on the one hand and with free grammatical expressions on the other. Like periphrastic expressions, support verb constructions consist of totally separable units, but they are special in that the units are semantically interdependent, as I have attempted to show in terms of their autonomy/dependence alignment (cf. section 3.4.4.2 above). Also, because

the noun is frame-bearing, the verb tends to be chosen from a lexical set of verbs that elaborate salient substructure of the frame evoked by the noun. In other words, the combination is subject to certain usage-based constraints on paradigmatic substitutability over and above the general selection preferences associated with the constituents (although alternative, more specific conceptualizations are always possible, as argued in section 3.4.2.4 above). Moreover, support verb constructions can be expected to show formal constraints compared to periphrastic expressions that are not entrenched (cf. sections 3.4.2.3 and 3.4.4.3 above). Because of these special features, support verb constructions can also be said to share some of the features of derivational expressions, which Bybee characterizes as “restricted in applicability and idiosyncratic in formation or meaning.”

Finally, support verb constructions can be said to have much in common with free grammatical morphemes, which "make up contrast sets that are obligatory in certain environments and have positional restrictions” (Bybee 1985: 12). It is especially relevant to compare them with auxiliaries, the other kind of ‘non-full verbs’ in the typology of Van Durme and van den Eynde (1998), (cf. section 3.4.4.3 above). According to Bache and Davidsen-Nielsen (1997: 282 ff.), auxiliaries “are also called *grammatical verbs*, because they perform the same kinds of function as verbal inflections, which are indisputably grammatical entities.” Auxiliaries are characterized as part of a scale ranging from a closed class of ‘auxiliaries proper’ (*be, have, do, can, may must, shall* and *will*), over ‘semi-auxiliaries’ (*ought, need, be going to*, etc.) and some ‘verbs with auxiliary-like functions’ that are classified as ‘catenatives’ (e.g. *get* and *keep*) and to ‘clearly lexical verbs’ like *negotiate*. In contrast to lexical verbs, auxiliaries proper “are typically morphologically defective, they share a number of syntactic features, and differ from most lexical verbs in expressing highly general meanings, relating to tense, aspect, modality, or voice.” As regards the class of auxiliaries proper, Bache and Davidsen-Nielsen find that “there are differences with respect to the degree to which they can be considered grammatical words” both in terms of how general they are in meaning and how morphologically defective they are. On the basis of this description, I find that auxiliaries can be construed as a prototype category with ‘auxiliaries proper’ as the central members. This group in turn forms a prototype category, whose central members are *be, have* and particularly *do*, which is described by Bache and Davidsen-Nielsen (1997: 285) as "semantically empty”, compared to the modals, which are said to have more lexical meaning .

In cognitive grammar, generality of meaning is not taken to mean lack of semantic content. Instead of trying to impose a rigid boundary between 'lexical' and 'grammatical morphemes', Langacker (1987: 353 f.) points to the need for a unified account based on variation along two continuous parameters:

(1) the level of specificity at which the dependent predication characterizes a situation

(2) the amount of overlap between standard and target

While lexical morphemes are typically quite specific, grammatical morphemes are highly schematic, but still meaningful. As far as the first parameter is concerned, I have claimed that the verbs in entrenched collocations tend to be highly schematic, which pulls them towards the grammatical end of the continuum (cf. Fig. 46 in section 3.4.4.2 above).

The second parameter is related to "how close the e-sites come to exhausting the content of the predication." For purposes of analysis, a construction consisting of a verb and a second constituent, verb or noun, is 'exploded' into a **'standard event'**, consisting of: (1) a verb, which is a dependent relational predication with an unelaborated, schematic e-site, corresponding to its landmark, and (2) a **'target event'**, which is the composite structure in which the e-site of the relational predication is elaborated by the second constituent. At the most grammatical end of the continuum, the overlap is substantial and e-sites may come close to 'exhausting the content of the predication'. An extreme example is 'zero predication', which according to Langacker (1987 354 f.) is the proper semantic analysis for *do* as an auxiliary verb. The profile of standard and target are identical, so that in a valence relation with another verb "all the specifications of [DO] (both standard and target) are inherent in the less schematic complement" (Langacker 1987: 355). According to Langacker, while this helps explain why *do* is generally considered to be meaningless, "in cognitive grammar there is no reason whatever to deny it the minimal semantic content characteristic of process predicates."

Also in the case of support verb constructions, the degree of overlap between standard and target, or the degree to which the standard event overlaps with the target event, in which the e-site of the verb is elaborated by a specific noun, can be used as an indicator of the verb's relative position on the lexical-grammatical continuum. By analogy with the above analysis of *do*, I suggest that we can see *make* in combinations with nouns like *appointment* as being at the most grammatical end of support verb constructions. The e-site of *make* comes close to 'exhausting the content of the predication',

and the overlap between standard and target are substantial. However, it cannot be construed as a case of zero predication, as when *do* is used as an auxiliary verb, because, although *appointment* is a noun of action, it is still a noun, and it is the verb that turns the target composite structure into an action. Like *do*, *make* cannot be said to be 'meaningless' although its contribution is limited to (re)imposing a processual profile on the target structure. In the case of *break an-appointment*, the e-site of *break* clearly does not exhaust its specifications, but, as I see it, there is still substantial overlap, since *break* corresponds to a salient substructure of *appointment*. In combinations that are not entrenched and in which the verb is therefore characterized as a full verb rather than a support verb, the overlap would be relatively insignificant (cf. Figs 43 and 44 in section 3.4.4.2 above). If this line of argument is accepted, it supports the case for construing the verb collocates of entrenched collocations as support verbs and for claiming that support verbs are more grammatical than full verbs.

A relatively limited number comprising the most schematic and frequently used basic level verbs, including *make*, *have* and *take* might be isolated from those that are more specific, like *break*, and construed as a 'closed class', but I prefer to construe the whole range of verb collocates that occur in entrenched collocations as a prototype category whose members cover a range of meaning from highly schematic to quite specific. They show characteristics that are similar to auxiliaries: they express verbal categories like finiteness, tense, aspect, actionality (or Aktionsart) and voice, in addition to performing a number of frame-related functions (cf. Fig. 47 in section 3.4.4.2 above), and like auxiliaries they show formal constraints. Finally, support verb constructions can be construed as complex predicators just like combinations of auxiliary and main verb. In both cases, the verb can be said to be functional in the sense that it supports a semantic frame evoked by a clearly lexical element, noun or verb. However, because of the special syntagmatic and paradigmatic constraints on support verbs, they are not in the same league as auxiliaries as far as generality of use is concerned. While some are likely to be used in many different frames, others may occur in only one or two frames.

All things considered, I suggest that support verbs, construed as the verbs in entrenched verb + nominal object collocations like *break an appointment*, can be characterized as a lexico-grammatical prototype category: a category whose members have undergone grammaticalization, and which

may be said to perform a grammatical function, but which are not actually 'grammatical words'. Lindström (2004: 10 f.) refers to grammaticalization as

> ... a concept which exists between two 'worlds', the lexicon and the grammar. It is a story which touches on how the lexicon and the grammar have been and are understood and how this can affect the entities which are in the no man's land between these two, part of both, part of neither.

I believe that support verbs belong in this no man's land, between the lexicon and the grammar, and that this is exactly what makes them such an important resource, which allows language users to combine convention with flexible conceptualization.

3.5 Summary of findings and evaluation of methodology

To conclude my case study, I will now summarize my findings in relation to the four research questions stated in section 3.3.1 above and evaluate the suitability of the framework and the methods used.

3.5.1 Findings related to research questions

The general assumptions underlying all four questions was that entrenched collocations are speech routines allowing speakers to guide hearers by evoking cognitive routines associated with familiar semantic frames, and that their function as linguistic expressions is to further the reproduction, through renewal of connection, of the contexts of situation and their underlying semantic frames.

3.5.1.1 To what extent are entrenched collocations like other composite structures?

In my first research question, the focus was on what entrenched collocations have in common with other composite structures and on a suitable methodology for researching such structures:

1) The contexts of situation to which entrenched collocations contribute and their underlying frames can be identified by analysing the internal structure of component items and their mode of integration. In this entrenched collocations do not differ from other composite structures.

The challenge was to find a general approach that would work both for a highly complex category like *break* and for a much less complex category like APPOINTMENT. A discussion of **polysemy** (in section 3.4.1.1) found that focusing on the basic level of categorization made it possible to link the question of how many readings can be posited to the number of semantic frames evoked by a word, which is feasible in the case of APPOINTMENT. Moreover, combining semantic frames with the network model allows for readings to be linked to higher-level schemas including related frames, which is a suitable approach in the case of a highly polysemous category like BREAK. A further advantage of the network model is that it can accommodate the lack of clear boundaries between different readings and that new readings may emerge over time. Since readings may be distinguished both in terms of what is being categorized and how it is done, I decided to subdivide the analysis for each item into three steps: (1) **domains and referential range**, (2) **image-schemas and event structure,** and (3) **construction type**. The analysis in terms of domain reflects a semasiological approach recording what referential situations are being categorized while the two other analyses reflect an onomasiological approach asking how image-schematic structure and construction type are used for conceptualizing referential content. Rather than summing up the three types of analysis for BREAK and APPOINTMENT separately, following the structure of my thesis, I will summarize the findings for one type of analysis at a time, including results for both BREAK and APPOINTMENT.

The analysis in term of **domains and referential range** resulted in a first hypothesis about the internal structure of the lexical categories BREAK and APPOINTMENT. In the former category, *break* was found to categorize a wide range of experience both in the sensorimotor and in the nonsensorimotor domains (cf. Fig. 9 and Tables 9 and 10 in section 3.4.2.1). The category was construed as a schematic network with ‘artefacts and natural things’ as the most basic sensorimotor domain, which was seen as the source domain for

metaphorical extension to less basic sensorimotor domains ‘body parts’, ‘physical activity’ and ‘the physical environment’. Furthermore, the internal structure of the nonsensorimotor domain was seen as mirroring that of the sensorimotor domain, reflecting that sensorimotor subdomains are systematically exploited to structure experience in the nonsensorimotor domain. Recurring patterns of prototypicality were found as well as overlaps between subdomains, which was seen as normal for a lexical network that has grown over time. As was to be expected, the referential range was much narrower in the case of APPOINT-MENT (cf. Fig. 32 and Table 16 in section 3.4.3.1). 99% of the examples were found in five nonsensorimotor subdomains coding different types of ‘social institutions and constructs’ categorizing action as well as the result of action. The remainder were from three subdomains of the sensorimotor domain ‘artefacts and natural things’. In the APPOINTMENT category, the nonsensorimotor subdomains in which *appointment* is a noun of action seems to be most basic and can be seen as motivating the use of the term to categorize the result of action and maybe even to refer to physical objects somehow associated with the action. As in the case of BREAK, subdomains showed internal complexity (cf. Figs 32 and 33 in section 3.4.3.2), but they were easier to identify, as they did not overlap to the same extent.

While the analysis in terms of domain was meant to trace referential range, it also showed that internal cohesion may be provided by metaphorical mappings between domains, as in the BREAK category. In the second type of analysis the purpose is to show that lexically specific **image-schematic structure** can be related to metaphors underlying the concep-tualization of causation and events in general, as attested by a wide range of entrenched collocations. The object event structure metaphor, which is based on the primary metaphor that ATTRIBUTES ARE POSSESSIONS, is the metaphor underlying much of the use of *break* both in the less basic sensorimotor domains and in the nonsensorimotor domains, while it was argued that, in the case of phrasal verbs, the location event structure metaphor that STATES ARE LOCATIONS provides the underlying image-schematic structure (cf. section 3.4.2.2). In the APPOINTMENT category (cf. section 3.4.3.2) the image-schematic structure of the two subdomains ‘arranging/arrangement for a meeting’ and ‘(placing sb in)

a position' was analysed. The use of basic level verbs like *make* and *break* was construed as reflecting that these domains are conceptualized in terms of the sensorimotor source domain 'artefacts and natural things' and also as exemplifying general event structure. Furthermore, in line with the invariance hypothesis, it was claimed that mappings to the two domains need to respect their inherent image schematic structure, so that for example *make an appointment* will mean 'arrange an appointment' in the former domain (object event structure) and 'place sb in a position' in the latter (location event structure). Along the same line it was argued that the many prepositional phrases reflect general object- and location event structure as well as inherent image-schematic structure in the two domains.

In the third type of analysis carried out for both lexical categories, the focus was on **construction types**. In the case of BREAK, the sensorimotor subdomain 'artefacts & natural things' was compared with the nonsensorimotor domain 'social institutions & constructs' in terms of six construction types (cf. section 3.4.2.3):

(1) transitive verb + noun (obj.), active form
(2) transitive verb + noun (subj.), passive form
(3) noun (subj.) + intransitive verb
(4) past participle of verb + noun
(5) noun + past participle of verb
(6) nominalization of verb

The analysis of construction types in the BREAK domain found a number of differences between combinations coding 'damage' in the sensorimotor domain of 'artefacts and natural things' and combinations coding 'violation' in the nonsensorimotor domain of 'social institutions and constructs'. In the former domain the construction past participle + noun was found in almost three out of every four examples in the data, while in the latter domain, it was only found in about one in every sixteen examples. Conversely, processual predications with a transitive verb accounted for almost nine out of ten examples in the latter domain, but only for a little over one in ten in the former (cf. Table 14 and Fig. 15 in section 3.4.2.3). Furthermore, it was noted that intransitive constructions such as *the cable*

has broken, which can be seen as a figure/ground reversal of transitive constructions, did not occur in the nonsensorimotor domain. The construction in which the noun is followed by the present participle as in *the cello breaking over the piano*, which can be construed as a reduced relative clause in which the verb is intransitive, did not occur in the nonsensorimotor domain either, while there were a few examples in the sensorimotor domain. Finally, nominal predications of the type *cable break*, in which the noun is construed as a count noun, only occurred in the nonsensorimotor domain (cf. Table 15 in section 3.4.2.3 above). It was tentatively suggested that these differences would be compatible with a construal of *break* as a full verb in the sensorimotor domain and as a support verb in the nonsensorimotor domain, in which it is predominantly used in processual predications.

In the APPOINTMENT category, the two largest subdomains 'arranging/arrangement for a meeting' and '(placing sb in) a position' were analysed in the same way (cf. section 3.4.3.3). The results were found to be compatible with the above-mentioned findings for the nonsensorimotor domain 'social institutions & constructs' in the BREAK category (cf. Table 15 in section 3.4.3.3). Again processual predications accounted for over 80% of combinations while the atemporal relation past participle + noun was found in a little under 6% and 1% respectively, of the combinations recorded for the two subdomains. The same analysis applied to ten basic level verbs included in the *appointment* data showed an even stronger domination of processual predications, at over 90%, while only two examples (less than 1%) were found of a past participle premodifying *appointment* (cf. Fig. 38 in section 3.4.3.3). An important point brought out by this analysis was that basic level verbs are not normally used as premodifiers in the past participle form. *Break* is an exception in this respect, but while *broken* as premodifier is very frequent in the sensorimotor domain, it is relatively rare in the nonsensorimotor domain. Assuming that the main function of support verbs is to participate in processual predications rather than in atemporal relations, these findings were found to be compatible with a support verb construal of basic level verbs in the domains analysed.

On the basis of the three analyses carried out for each lexical category, and drawing on the principles laid down in section 3.4.1, the question of how many **readings** should be stipulated for each was addressed in the fourth part of the two analyses (cf. sections 3.4.2.4 and 3.4.3.4). Starting from the initial hypothesis based on domain, 20 possible readings were discussed for the BREAK category: ten for the sensorimotor domain and ten for extensions to the nonsensorimotor domain. Based on the principles for the construal of polysemy discussed in section 3.4.1.1, some of the readings originally proposed were conflated, while others were split up into more readings, but these revisions balanced out so that eventually the number of readings posited was still 20 (cf. fig. 29 in section 3.4.2.4). It was tentatively suggested that a support verb function might be posited for *break* in entrenched collocations in the nonsensorimotor domain, which would make it possible to subsume those readings under a more schematic support verb reading.

As a final hypothesis for the internal structure of BREAK, a network model was drawn up based on the three principles of extension posited to account for the internal structure of the category (cf. Fig. 30 in section 3.4.2.4). One principle was seen as operating within the sensorimotor domain with the most basic domains being exploited as source domains for metaphorical mappings to less basic sensorimotor domains, as when ‘overcoming physical barriers’ is conceptualized as ‘damaging physical objects’. As a second principle, it was found that sensorimotor subdomains are systematically exploited to structure experience in the nonsensorimotor domain, as when ‘overcoming psychological barriers’ is conceptualized as ‘overcoming physical barriers’. Finally, a principle of alternative conceptualization was posited to account for the fact that it is never predictable exactly which subdomain will be recruited as the source domain for a given target domain. Thus ‘overcoming a psychological barrier’ may be conceptualized not only as ‘breaking it’ or ‘breaking through it’ (from ‘overcoming physical barriers’), but also as ‘breaking it down’ (from ‘damaging physical objects’) or ‘breaking out beyond it’ (from ‘opening physical container’ and ‘overcoming physical barriers’). This phenomenon, which was seen as supporting a network model of category structure, was evident from phrasal verbs (combinations with adverbs, prepositions and adjectives). These combinations, which

accounted for over a third of the examples of *break* in the data, were construed as a subordinate level categorizing specific ways of 'breaking'. The only creative collocation found in the data was also seen as supporting the network hypothesis, since it relies on the cognitive ability to integrate two basically incongruous readings of *to break a record*: 'to break a physical object' and 'to break a psychological barrier'.

Whereas in the BREAK category each reading included a range of related semantic frames, readings in the APPOINTMENT category, which is much less complex, could be related to the number of schematic frames identified with specific combinations construed as constituting specific frames imposing each their perspective on the overall frame (e.g. *make* vs. *break an appointment*). The discussion was based on the same general principles as in the case of BREAK as well as on the previous analyses of APPOINTMENT.

Six readings were posited (cf. Table 23 in section 3.4.3.4), including two for the sensorimotor domain, 'accessories for people' and 'accessories for cars, etc.', which were supported by only a few examples. This was also the case for one of the four readings posited for the nonsensorimotor domain: 'action of declaring destination of property'. The remaining readings accounted for the vast majority of examples in the data. One reading was posited for the schematic frame 'arranging/arrangement for a meeting' in which all specific frames use object event structure, while two readings were posited corresponding to the schematic frame '(placing sb in) a position', which was construed as consisting of one subframe using location event structure: 'placing sb in a position', and one subframe using object event structure: the resulting 'position'. Also in the case of APPOINTMENT, a network model was drawn up to illustrate how the different readings might be related in the mind of a contemporary user who is familiar with them all (cf. Fig. 41 in section 3.4.3.4).

3.5.1.2 In what respects are entrenched collocations special?

After establishing that entrenched collocations, like other word combinations, can be seen as integrating complex categories, the focus in the second research question was on the specific characteristics of entrenched collocations:

2) In entrenched collocations consisting of a verb and a nominal object, the noun evokes the dominant frame while the verb profiles a specific aspect of the frame. In collocations that are not entrenched, it is the verb that evokes the dominant frame.

It was claimed that entrenchment involves a shift in the autonomy/dependence alignment of verb and noun. On balance, the verb as a relational predication was still seen as the more dependent element, because of its dependence on elaboration by the noun. However, the noun was also construed as having salient substructure elaborated by the verb, which was understood to profile a specific aspect of the frame evoked by the noun. It was argued that the asymmetry between them is reduced compared to a combination in which the verb does not elaborate salient substructure of the noun, and their relationship was consequently characterized as one of interdependence (cf. Figs 43 and 44 in section 3.4.4.2 above). It was pointed out that what is salient substructure is ultimately a matter of construal, and that my claim that *break an appointment* is an entrenched collocation assumes the cognitive validity of my analysis of the data.

In my third research question the focus was on capturing variability as a dimension of entrenchment:

3) Entrenched collocations can be characterized in terms of prototypicality that varies with the schematicity of the verb and its salience in the frame evoked by the noun.

It was argued that combinations with highly schematic basic level verbs, such as *break an appointment*, were most salient and likely to be entrenched. The range of verbs elaborating a specific substructure of an APPOINTMENT frame were construed as a lexical set with the basic level verb as the prototype and less schematic verbs as alternatives that may or may not be entrenched. This notion of the lexical set differs from Halliday's notion of a set

whose members have ‘like privilege of occurrence in collocation’. Rather it was seen as representing a functionally motivated range of variability at the place of the verb and as cutting across the borderline between combinations that are entrenched and those that are not.

A typology of collocations was posited based on continua of salience and schematicity. In this typology, an **entrenched collocation** like *break an appointment* is both highly salient and highly schematic, while an **open collocation** like *arrive late for an appointment* combines high salience with low schematicity. **Free collocations** include the combinations that evoke the frame of the verb rather than that of the noun. They may be highly schematic, but low in salience, like *keep an appointment* meaning ‘not change it’, or they may be low in both schematicity and salience, like *discuss an appointment*. The emphasis in this typology is on the continuity between expressions that are entrenched and alternative expressions that are not entrenched, but still perfectly normal, thereby challenging the Saussurean idea that conventions in language preclude choice.

Finally, an attempt was made to show the continuity between idioms like *break the ice* and the types of collocation mentioned above. Whereas, in the framework of phraseology, idioms are seen as noncompositional in contrast to partly compositional restricted collocations and fully compositional free collocations, cognitive linguists tend to emphasize that most idioms are compositional to some extent, which is linked to their present analysability. If, following Langacker, compositionality is seen as a separate issue from present analysability, even idioms that are no longer analysable can be claimed to be compositional. All expressions can then be placed along a continuum ranging from the most schematic grammatical expressions at one extreme over more and less entrenched collocations to the most substantive or ‘lexically filled’ idioms at the other, without positing full compositionality as the norm from which conventional expressions deviate.

The borderline between entrenched collocations and idioms is not clear-cut. *Break the ice* might be classified as an entrenched collocation if an analysis of the complex category ICE

showed that *ice* could be claimed to have, as one of its meanings, something like a 'tense situation'. Since no such evidence is available in this study, it is classified as an idiom, which is presumed to be a conventionalized blend exploiting an underlying conceptual metaphor associating tension between people with 'coldness'.

3.5.1.3 Verbs in entrenched collocations as grammaticalized support verbs

Finally, I will summarize the discussion, in section 3.4.4.3 above, of whether *break* can be said to be a function word in addition to a content word, as claimed in my fourth research question:

2) The verb in entrenched collocations has a functional, grammaticalized role as support verb.

A prototype category of support verb was posited, drawing on a range of different accounts which testify to the phenomenon that in some cases the relationship between a predicator verb and its grammatical object is felt to be so close that the object is no longer seen as a separate participant. In these accounts, the construction was characterized as a fused or complex predicator, as semantically intransitive, and as having its own valency scheme. The prototype on which these accounts converge is that of a delexical verb with general meaning and a deverbalized noun with a suffix forming a predicative unit, as in *make an appointment*, where the noun is an effected object. In less prototypical cases, the verb is less general in meaning, as in *arrange an appointment*. Also the noun may be less clearly deverbal, because it does not have a suffix, as in *make a promise*, or not deverbal at all, as in *wage war*. It was argued that also cases in which the noun is not an effected object should be included, so that *break an appointment* could also be classified as a support verb construction, though not the most prototypical one.

The claim that certain combinations are support verb constructions is typically accompanied by evidence that they are subject to formal constraints compared to full verb

constructions. Thus Van Durme and van den Eynde (1998) have found that support verb constructions in Dutch are not 'proportional to pronominal paradigms', and Schösler (2003) has found that French support verb constructions show constraints on determiners. Neither of these types of constraint could be shown for *break an appointment*, but it was argued that even if they turn out to apply to prototypical cases of support verbs also in English, it is still possible to categorize *break an appointment* as a less prototypical member of the support verb category, a member which shows constraints of its own (cf. section 3.4.2.3).

The notion of a prototype category of support verbs was found to be fully compatible with a frame semantic account, which moreover provides a functionally and cognitively motivated framework for support verb constructions as entrenched collocations in which the noun is frame-bearing. The prototypical role of the support verb is to express verbal meaning in combination with the noun, but it may have a range of additional functions in profiling specific participants or phases associated with a frame, as in *to break an appointment* as opposed to *making* or *keeping an appointment* (cf. Fig. 47 in section 3.4.4.3). As in the case of entrenched collocations, the relationship between verb and noun in a support verb construction can be characterized as one of interdependence, and I therefore suggested that the two notions are merged, so that an entrenched verb + nominal object collocation, like *break an-appointment*, is characterized as a support verb construction. The possibility of positing a prototype category of support construction was mentioned on the basis of the notion of 'support prepositions' tentatively suggested by Fillmore (2003a). I suggested that such a category might potentially be extended to include also 'support adjectives', as in *a rapid increase*, 'support adverbs', as in *deeply wounded*, and 'support nouns', as in *a stroke of luck*, but this idea has not been pursued in the present study.

Finally, a discussion of the grammatical status of *break* in an entrenched collocation such as *break an appointment* concluded that it could be categorized as belonging to a lexico-grammatical prototype category of support verbs. They are assumed to have undergone a process of grammaticalization compared to full verbs, although they are not prototypical closed class words. The discussion was based on the view that grammatical structure forms

a continuum with the lexicon and that the grammaticalization of a lexical item involves a shift toward the grammatical end of the continuum associated with a more functional role.

Because of the phenomenon of layering, synchronic data provide evidence of variability resulting from metaphorical extension, as in the case of *break*, and, following Sweetser (1988), grammaticalization is seen as involving metaphorical extension to an abstract domain of grammatical meaning. In the case of *break* as a support verb, it was argued that the extension is to the domain of causes and events, where, by analogy with its meaning in the sensorimotor domain of 'artefacts and natural things', *break* expresses object event structure and actionality.

At some point in a grammaticalization process, a local metonymic change associated with syntactic reanalysis is assumed to occur, which only becomes visible when it results in changes in use. The formal constraints demonstrated for support verbs compared to full verbs can be seen as evidence of reanalysis, and I have argued that, in the case of *break*, the differences between construction types found in the sensorimotor and the nonsensorimotor domains provide evidence that, at some point, what was originally analysed as a transitive verb + a nominal object was reanalysed as an intransitive complex predicator. This line of argument was supported by the analysis of *break an appointment* in terms of autonomy/ dependence alignment, which found that a support verb construal of *break* involves a higher degree of interdependence between verb and noun compared to a full verb construal.

The claim that support verbs are more grammatical than full verbs and that some support verbs are more grammatical than others was finally discussed in terms of the two parameters proposed by Langacker (1987): (1) level of specificity at which the dependent predication characterizes a situation and (2) overlap between standard and target. The low level of specificity, or high schematicity, of the basic level verbs found in entrenched collocations was seen as compatible with their use as support verbs and with the claim that a highly schematic verb like *make* is a more prototypical support verb than a relatively less schematic verb like *break*. Also, in a support verb construction like *make an appointment*, it

can be argued that there is substantial overlap between the standard event, in which the e-site of *make* corresponding to its landmark is unelaborated, and the target event, in which it is elaborated by *appointment*; in other words, the e-site can be said to almost exhaust the content of the predication *make*. In the case of *break an appointment*, *appointment* cannot be claimed to exhaust the e-site of *break*, but it was claimed that there is still substantial overlap. By comparison, it was argued that if the relational predication is a full verb as in *break a vase*, the overlap is insignificant.

3.5.2 Evaluation of methodology

The choice of a basically **qualitative approach** was motivated in section 3.2.1 by arguing that while statistical data on relative frequencies of co-occurrence may serve as input for qualitative analysis, they cannot replace it. Corpus data show that familiar collocations are not necessarily very frequent and even if they are, a purely quantitative account does not explain much. The **data** included 1,000 full sentences in the case of *break* (9.5% of all the examples found in the BNC) and 908 sentences in the case of *appointment* (all of the examples in the BNC). The entrenched collocation *break an appointment* occurred only five times, and all five occurrences were found in the data for *appointment*, more specifically they were found as examples of reading 1: 'arranging/arrangement for a meeting', which accounted for 229 of the 908 combinations with *appointment*. Thus the combination *break an appointment* accounted for 2.18% of the examples of *appointment* found for this reading. By comparison *cancel an appointment* with 10 occurrences accounted for 4.37%, and *make an appointment* with 43 occurrences accounted for 18.8%.

While the relatively low frequency of *break an appointment* can be explained in the context of the semantic frame underlying this reading of *appointment*, it is not possible to draw any conclusions as regards the status of this combination as less conventional than *cancel an appointment* or *make an appointment*. The reverse gear in a car may be used much less frequently than the other gears, but a car that does not have one would be considered to be

defective by most people. The point is that, although high frequency of co-occurrence will often be accompanied by high salience, as in the case of *make an appointment*, combinations may be salient although they are relatively infrequent. However, while a quantitative approach cannot be relied on for the identification of entrenched collocations, my analyses of the construction types that *break* and *appointment* were found in, made it clear that it can be very useful in revealing lexicogrammatical patterns that might otherwise go unnoticed. I have argued that such evidence is important in making a case for a support verb construal of verbs in entrenched collocations. My conclusion is therefore that quantitative methods have an important role to play as an input to an overall qualitative approach[103].

The qualitative analyses attempted in this study are defined by a cognitive approach, which is also functional. This implies that words are symbols reflecting conceptualization with communication as its primary function, and that, following Harder (1996a: 91), the function of linguistic expressions is to "further the reproduction of tokens of which they form part". In the case of entrenched collocations, I have argued that their function is to reproduce the contexts of situation and the underlying semantic frames which they evoke. The methodological challenge of explaining linguistic evidence in terms of cognitive models was discussed in section 3.2.2, which concluded that a clear distinction should be made between the linguistic evidence and its interpretation, a principle that I have tried to adhere to throughout this study. On the other hand, it is necessary to make assumptions about cognitive processes of meaning construction to arrive at an understanding of what collocations are about, and I have found that the cognitive framework offers notions and models that make it possible to discuss these assumptions in relatively precise terms such as salience, schematicity, autonomy/dependence alignment, semantic frames, and prototype categories. This makes it possible to build on the experience of other linguists and to express one's own findings in a way that is sufficiently explicit and clear, so that they may be tested and challenged in their turn.

[103] This is in line with work in which corpus linguists investigate the interaction of lexical items and constructions within a cognitive approach (see for example. Stefanowich and Gries 2003).

As a third aspect of methodology, seeing synchronic corpus evidence as resulting from diachronic processes was discussed in section 3.2.3. I have found that this view of the data is compatible with a construal of complex lexical categories as dynamic networks that have grown over time and are still growing, and that it offers a key to understanding polysemy and variability. Conventional collocations like *break an appointment* and idioms like *break the ice* are thus seen as resulting from diachronic processes of entrenchment, while creative collocations are seen as resulting from a process of innovation. Some of them are likely to become entrenched while others will just be ad hoc exploitations of conventional expressions like the one example found in my data (cf. section 3.4.2.4).

4. Collocations as a language resource

I will conclude this study by presenting my conclusion related to the question of what is a suitable framework for the study and categorization of collocations. Also, I will suggest areas for future study of collocations to follow up on the findings presented in Part 3 above.

4.1 Conclusion

In Part 2 of this study I provided a critical review of the general approach to the study of fixed expressions that has been referred to as the traditonal approach to 'phraseology'. The account has drawn extensively on Cowie's anthology with the same title (Cowie 1998b) as well as on comprehensive corpus-based studies (Howarth 1996; Moon 1998). Not only do these sources testify to the pervasiveness of conventional collocations and the role they play in foreign language acquisition; they also provide a valuable insight into the theoretical challenges that researchers have been facing in trying to develop a typology for the categorization of fixed expressions, including collocations. I have argued that the problems of categorization can be related to practical concerns for the problems of the foreign language learner as well as theoretical influences which have a historical explanation, but which are not compatible with the growing understanding that language routines are cognitively based and serve a communicative function.

I therefore propose a unification between the traditional approach and the functionally and cognitively based approach which formed the basis of the empirical part of this study. This would require phraseologists to jettison some of their traditional ideas, such as the

assumption that conventional expressions are arbitrary and deviate from a standard of full compositionality, and that they are subject to strict classification into classical categories on the basis of criterial features (cf. section 2.3). Those who believe that conventional expressions can be identified by measuring frequency of co-occurrence of lexical items would have to realize that relatively infrequent expressions may still be salient so that their statistical evidence cannot stand alone. Would phraseologists be willing to enter into marriage on these terms? I believe that many of them might be, especially as phraseology already includes a cognitive strand, as described in Part 2 (section 2.1.5). Some cognitive studies of idioms and collocations have already been made and cognitive notions like 'continuum' and 'fuzzy boundaries' are making their way into studies that adhere to the Russian model of categorization. The usefulness of Fillmore's notion of semantic frames is becoming generally recognized as a useful tool for lexicographers[104] (see also Cowie 2002 and Fontenelle 2000) so that from the current stage of cohabitation it does not seem to be such a big deal to take the vows and also accept the semantic frame as a cognitive model, which I have argued is compatible with the Firthian notion of 'context of situation' abstracted from use. This would then pave the way for re-embedding phraseology in a functional and cognitive framework, which does not offer strict categorization, but which I find is descriptively more adequate and has much greater potential when it comes to explaining collocation as a language resource. Moreover, a framework that brings the researcher's assumptions about language-related cognitive processes out into the open seems to me to be much preferable to approaches that ignore them, thereby implicitly assuming that we can account for the way we speak without taking into account the way we think.

[104] The *International Journal of Lexicography* 16, 2003, Special Issue 3, was devoted to FrameNet and frame semantics.

4.2 Further research

Based on the analyses conducted in relation to my case study of *break an appointment* in the empirical part of this study, I have posited a prototype category of support verb constructions including all verbs in entrenched collocations. This hypothesis seemed plausible and was compatible with my data but will have to be tested in studies of other verb + nominal object collocations. I would especially be interested in applying the same methodology to other collocations with highly schematic basic level verbs expressing location and object event structure. In this context I would like to draw on the methods of 'collostruction' analysis that are currently being developed by corpus linguists in the cognitive framework (see for example Stefanowich and Gries: 2003). I have also made the claim that the potential of creative exploitation should be considered to be part of the functionality of conventional expressions, but as only one example was found in my data, the treatment of creative strategies has been limited, and I would surely like to return to this subject, which is closely linked with a dynamic view of lexical meaning.

Most importantly, the intuitive judgements made when analysing the linguistic data in terms of the cognitive notions should be tested by psycholinguistic studies. How exactly such studies should be designed is a question that I will leave to people with the right background and experience, but since it is a central claim of this study that the functional role of entrenched collocations like *break an appointment* is to call on the frame evoked by the noun, it is vital that this claim can be confirmed by independent psycholinguistic evidence. Moreover, since I have based this study on purely synchronic data, other studies are needed to trace the diachronic development producing the category of grammaticalized collocations posited.

To investigate if there is a possibility of further unification, I would like to extend the discussion of support verbs to include the notion of 'light verbs' in Mel'čuk's meaning-text theory (see for example Mel'čuk 1998) and the notion of participanthood in Systemic Functional Linguistics (see for example Halliday and Matthiessen 1999: 167), according to

which a distinction is made between the ‘Goal’ as a prototypical participant which is impacted by its participation in a process, and ‘Range’, which is not impacted in this way, but which expands or elaborates the process. Finally, it would be interesting to compare the frame-based account of word meaning in the present account with other accounts such as Pustejovsky’s (1995) model of a generative lexicon, which assumes that lexical items have a ‘core set of word senses’ with internal structure, which can be used to generate a larger set of word senses through combination with other items.

Perhaps we are talking about the same elephant after all.

Dansk resumé

Det spørgsmål som afhandlingen stiller sig, er hvad der vil være en velegnet teoretisk ramme for behandlingen af konventionelle kollokationer, idet der fokuseres på forbindelser bestående af et verbum og dets nominelle objekt. Den overordnede tese er at en funktionel og kognitiv tilgang vil være mere velegnet til at beskrive og forklare sådanne forbindelser end den traditionelle tilgang ifølge hvilken fraseologismer er arbitrære netop fordi de er konventionelle og fordi de afviger fra fuldt kompositionelle udtryk, som antages at udgøre normen.

Første del beskriver kort den traditionelle tilgang til fraseologien og præsenterer de forskellige definitioner på 'kollokation' som teknisk term. Der gives en motivation for en alternativ funktionel og kognitiv tilgang baseret på prototypekategorier i stedet for klassiske kategorier. I denne tilgang opfattes kompositionalitet ikke som en addition af betydningselementer, men som en integrering af dynamiske skematiske netværk, hvis kompositionelle betydning er motiveret af de enkelte ords betydning, men ikke kan forudsiges på grundlag af dem. En funktionel og kognitiv tilgang tillader desuden at synkrone data ses i et diakronisk perspektiv, hvilket gør det muligt at diskutere konventionalisering og fornyelse som dynamiske sproglige udviklingsprocesser.

I **anden del** gennemgås den traditionelle tilgang til fraseologien ud fra en funktionel og kognitiv synsvinkel. Denne tilgangs praktiske baggrund inden for fremmedsprogsundervisning og leksikografi anses for at spille en rolle sammen med de teoretiske påvirkninger fra Firth og russisk fraseologi. En klassisk kategoriopfattelse såvel som strukturalistiske og generative principper ses som yderligere påvirkninger, lige som en kognitiv lingvistisk synsvinkel er begyndt at gøre sig gældende. Efter denne beskrivelse følger en kritisk gennemgang af det system der anvendes til kategorisering af fraseologismer. Forsøget på at kombinere klassiske kategorier, som bygger på klare grænser mellem kategorierne, med en kontinuum-model, som bygger på gradvise overgange, anser jeg for at udgøre et problem i den traditionelle tilgang, idet disse to

modeller må siges at være uforenelige. I denne tilgang defineres fraseologismer ved 'restricted compositionality', dvs. at de ikke kun følger sprogets syntaktiske regler, men er underlagt yderligere, konventionelle begrænsninger. Disse begrænsninger defineres syntagmatisk, idet verbet, eller 'kollokatet', beskrives som ensidigt semantisk afhængigt af substantivet eller 'basisordet', og paradigmatisk, idet kollokatet ikke frit kan udskiftes med andre verber med lignende betydning. De konventionelle begrænsninger for hvilke ord der rent faktisk kan indgå i et paradigme anses for at være arbitrære.

Over for disse argumenter anføres det ud fra en kognitiv lingvistisk synsvinkel at det ikke er hensigtsmæssigt at betragte fraseologismer som afvigelser fra et princip om fuld kompositionalitet, som ikke tager hensyn til at ord normalt er polyseme. Hvad angår kollokationers paradigmatiske begrænsninger, er kritikpunktet at de beskrives på grundlag af forhold der er den enkelte kollokation uvedkommende. Andre træk som traditionelt bruges til at karakterisere fraseologismer i modsætning til 'frie forbindelser', nemlig at de ikke er forudsigelige og at deres betydning ikke er konkret, stilles der også spørgsmålstegn ved i denne del af afhandlingen.

Tredje del indeholder afhandlingens empiriske del, som afprøver den antagelse at en funktionel og kognitiv tilgang vil være bedre egnet til studiet af konventionelle kollokationer. Denne del indledes med en diskussion af de funktionelle over for de kognitive sider af sproget. Herefter afklares nogle metodiske spørgsmål om hvordan frekvensbegrebet kan bidrage til en kvalitativ analyse, om hvordan sproglige data kan relateres til kognitive modeller og desuden kan ses som det synkrone resultat af en diakron udvikling. Jeg har valgt *break an appointment* til min undersøgelse som et eksempel på en konventionel kollokation bestående af et 'deleksikalt' verbum, hvis leksikalske betydning er svækket når det bruges I konventionelle kollokationer, og et 'deverbalt' substantiv, som stammer fra et verbum. Mine data, fra British National Corpus, består af 1000 konkordanser med *break* (ca. 9,5 % af alle konkordanser i korpuset) og af 908 konkordanser med *appointment* (samtlige konkordanser med *appointment* i dette korpus). Frasalverber (her forbindelser

med adverbier såvel som præpositioner og adjektiver) tegner sig for ca. en tredjedel af eksemplerne på *break*.

Min problemformulering dækker aspekter som jeg anser for at være centrale for en teori om kollokationer: de to komponenters interne struktur, integrationen mellem dem, samt de træk der er specielle for konventionelle udtryk. Mine overordnede antagelser er at

> … konventionelle kollokationer er rutiner i sproget som giver afsenderen af et sprogligt budskab mulighed for at vejlede modtageren ved at kalde på kognitive rutiner som er forbundet med velkendte kognitive modeller ('semantic frames'). Som sproglige udtryk er det deres funktion, gennem anvendelse i konkrete situationer, at være med til at reproducere velkendte, abstrakte situationstyper ('contexts of situation') og de kognitive modeller ('semantic frames'), som ligger til grund for dem.

Disse antagelser afspejler min funktionelle og kognitive tilgang, men jeg ser dem også som værende i overensstemmelse med Firth's opfattelse af kollokationer som udgørende et betydningsniveau som i praksis er integreret f.eks. med det syntaktiske niveau og det niveau han kalder 'contexts of situation', abstraktioner fra velkendte situationstyper som man har oplevet mange gange. Jeg ser begrebet 'context of situation' som værende kompatibelt med begrebet 'semantic frame' (Fillmore, 1982, 1985) inden for kognitiv lingvistik.

Min problemformulering omfatter følgende fire hypoteser:

1) De situationstyper ('contexts of situation') som konventionelle kollokationer bidrager til og de kognitive modeller ('semantic frames') som de kalder på, kan identificeres ved at analysere komponenternes interne struktur og deres integrationsmåde. Herved afviger konventionelle kollokationer ikke fra andre sammensatte udtryk.

2) I konventionelle kollokationer der består af et verbum og dets substantiviske objekt, er det substantivet der kalder på den dominerende semantiske 'frame' medens verbet profiler et specielt aspekt af denne frame. I kollokationer der ikke er konventionelle, er det verbet der kalder på den dominerende frame.

3) Konventionelle kollokationer kan karakteriseres i forhold til hvor prototypiske de er, hvilket afhænger af hvor skematisk verbet er og dets grad af saliens inden for den frame som substantivet kalder på.

4) Verbet i konventionelle kollokationer har en funktionel, grammatikaliseret rolle som støtteverbum.

Den første hypotese handler om hvad konventionelle kollokationer har til fælles med andre sammensatte udtryk. Den afprøves ved først at analysere BREAK og APPOINTMENT hver for sig som komplekse, polyseme kategorier. En diskussion som sammenkæder spørgsmålet om hvor mange betydninger et ord kan siges at have med det mellemste niveau i et hierarki af betydninger, 'the basic level' (jvf. Taylor 1992) og med antallet af 'semantic frames', danner grundlaget for disse analyser. For en kategori som APPOINTMENT, der udviser moderat polysemi, kan dette lade sig gøre, medens det for en kategori som BREAK, der er polysem i meget højere grad, er hensigtsmæssigt at beskrive betydningerne på et mere skematisk niveau, således at hver betydning omfatter et antal beslægtede 'frames', der anses for at være forbundne i et netværk. Begge kategorier analyseres derefter under følgende overskrifter: (1) domæne og referentiel betydning, (2) billedskemaer ('image schemas') og begivenhedsstruktur ('event structure'), og (3) konstruktionstype.

Derefter analyseres det sammensatte udtryk *break an appointment* på grundlag af Langackers (1987) teori om komponenternes relative autonomi- og dependensforhold. Formålet er at vise at konventionelle kollokationer på den ene side kan analyseres på samme måde som alle andre sammensatte udtryk, men at konventionalisering på den anden side kan siges at indebære et skifte i komponenternes relative autonomi- og dependens-forhold som afspejler at det i en konventionel kollokation er substantivet der kalder på den kognitive model (er 'frame bearing'), som påstået i min anden hypotese.

Den tredje hypotese gående ud på at konventionelle kollokationer kan karakteriseres i forhold til hvor prototypiske de er, afprøves ved hjælp af en analyse der fokuserer på variabilitet som en dimension af konventionalisering. Der opstilles en kollokationstypologi baseret på ordforbindelsernes saliens og på hvor skematiske de anses for at være. Typologien fremhæver kontinuiteten mellem ordforbindelser der er konventionelle og alternative forbindelser der ikke er konventionelle, men fuldstændigt normale. Dette er i modstrid med Saussures tanke om at sproglige konventioner udelukker valgmuligheder.

De verber der kan udfylde en specifik del af en APPOINTMENT frame, opfattes som leksikalske sæt med de skematiske 'basic level' verber som prototyper og de mindre skematiske verber som alternativer, der kan være konventionelle, men ikke nødvendigvis er det. Denne opfattelse af leksikalske sæt afviger fra Hallidays (1966) opfattelse af et leksikalsk sæt som bestående af medlemmer med lige stor mulighed for at optræde i en kollokation ('like privilege of occurrence in collocation'). I stedet opfatter jeg det leksikalske sæt som bestående af et funktionelt motiveret paradigme af verber, som afspejler variabiliteten i sproget og som overskrider grænsen mellem konventionelle og ikke-konventionelle forbindelser.

Diskussion af den sidste hypotese, at verbet i en konventionel kollokation har en funktionel, grammatikaliseret rolle som støtteverbum, falder i to dele. I den første del, som er baseret på en række uafhængige kilder, argumenteres der for en kategori af støtteverber hvis prototype er et 'deleksikalt' verbum med generel betydning som indgår i en prædikativ enhed sammen med et 'deverbaliseret' substantiv med suffiks. Jeg har argumenteret for at konventionelle kollokationer med *break* som kollokat kan siges at indgå i en sådan kategori, skønt ikke som et af de mest prototypiske medlemmer. Dette underbygges med de formelle begrænsninger som analysen af konstruktionstyper påviser for *break* i ikke-konkrete betydninger, som f.eks. i forbindelse med *appointment*.

Den sidste del af denne diskussion omhandler en mulig grammatisk rolle for *break* og andre 'basic level' verber som indgår i konventionelle kollokationer, og jeg konkluderer at de kan siges at tilhøre en leksikogrammatisk kategori af støtteverber. Diskussionen er baseret på den opfattelse at grammatiske strukturer udgør et kontinuum med leksikon og at grammatikaliseringen af et leksikalsk ord indebærer et skift i retning af den grammatiske ende af dette kontinuum og en mere funktionel rolle. Baseret på Sweetser (1988) opfattes grammatikalisering som en metaforisk overførelse af betydning fra det fysiske domæne til et mere abstrakt domæne. I tilfældet *break* betegner det abstrakte domæne kausalitet og begivenhedsstruktur, idet *break* gør det muligt at tale om abstrakte forhold som var de fysiske objekter, analogt med verbets konkrete betydning i det fysiske domæne.

Afslutningsvis diskuteres spørgsmålet om støtteverber kan siges at være mere 'grammatiske' end fuldverber og om nogle støtteverber kan siges at være mere grammatiske end andre ved hjælp af to parametre foreslået af Langacker (1987). Det ene parameter er graden af 'specificitet' ('specificity'), og det andet er graden af overlapning mellem en 'standardkonstruktion' ('standard construction'), som er en verbum + objektkonstruktion hvor objektets plads står tom, og en 'målkonstruktion' ('target construction'), i hvilken denne plads er udfyldt. Selvom *break* er mindre skematisk, eller mere specifik, end nogle af de andre 'basic level' verber, er konklusionen, at verbet kan kategoriseres som et grammatikaliseret støtteverbum, selvom det igen ikke er det mest prototypiske eksempel.

References:

Aisenstadt, E. 1979. Collocability restrictions in dictionaries. In *Dictionaries and their users*, ed. Reinhard R. K. Hartmann, 71-74. Exeter: University of Exeter.

Aitchison, J. 1987. *Words in the mind: An introduction to the mental lexicon*. Oxford: Basil Blackwell.

Alexander, R. 1987. Problems in understanding and teaching idiomaticity in English. *Anglistik und Englischunterricht* 32:105-122.

Alexander, R. J. 1978. Fixed expressions in Englisch: a linguistic, psycholinguistic, sociolinguistic and didactic study (part I). *Anglistik und Englischunterricht* 7. Modernes englisches Drama:181-202.

Amosova, N. N. 1963. *Osnovui anglijskoy frazeologii*. Leningrad: University Press.

Aristotle. 1933. *Metaphysics*. London: Heinemann.

Atkins, S., Fillmore, C. J., Johnson, C. R. 2003. Lexicographic relevance: selecting information from corpus evidence. *International Journal of Lexicography* 16:251-280.

Bache, C. 1985. *Verbal aspect : A general theory and its application to present-day English*: Odense University studies in English; 8. Odense: Odense University Press.

Bache, C. and Davidsen-Nielsen, N. 1997. *Mastering English grammar. An advanced grammar for non-native and native speakers.* Berlin, New York: Mouton de Gruyter.

Baranov, A. N.; and Dobrovol'skii, D. O. 1996. Cognitive modeling of actual meaning in the field of phraseology. *Journal of Pragmatics* 25:409-429.

Baron, I. and Herslund, M. 1998. Verbo-nominal predicates and the object relation. In *Studies in valency IV: Valency and verb typology*, ed. Karen Van Durme and Lene Schøsler, 89-127. Odense: Odense University Press.

Bauer, L. 2001. *Morphological productivity*. Cambridge: Cambridge University Press.

Benson, M. 1985. Collocations and idioms. *ELT Documents: Dictionaries, lexicography and language learning*:61-68.

Benson, M., Benson, E. and Ilson, R. 1986. *The BBI combinatory dictionary of English.* Amsterdam and Philadelphia: John Benjamins Publishing Company.

Benson, M., Benson, E. and Ilson, R. 1997. *The BBI dictionary of English word combinations (Revised edition).* Amsterdam, Philadelphia: John Benjamins Publishing Company.

Bierwisch, M. 1981. Basic issues in the development of word meaning. In *The child's construction of language*, ed. Werner Deutsch, 341-387. London: Academic Press.

Bierwisch, M. 1983. Semantische und konzeptuelle Repräsentationen lexikalischer Einheiten. In *Untersuchungen zur Semantik*, ed. R. Ruzicka and W. Motsch, 61-69. Berlin: Akademie-Verlag.

Bobrow, S. and Bell, S. 1973. On catching on to idiomatic expressions. *Memory and cognition* 1:343-346.

Bybee, J. L. 1985. *Morphology. A study of the relation between meaning and form.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

Bybee, J. L., Perkins R. and Pagliuca W. 1994. *The Evolution of grammar. Tense, aspect and modality in the languages of the world*. Chicago and London: The University of Chicago Press.
Bühler, K. 1990. *Theory of language: The representational function of language*.vol. 25: Foundations of semiotics. Amsterdam: Benjamins.
Cacciari, C. and Tabossi, P. 1988. The comprehension of idioms. *Journal of Memory and Language* 27:668-683.
Campbell, L. 2001. The history of linguistics. In *The handbook of linguistics*, ed. Mark Aronoff and Janie Rees-Miller. Oxford, UK and Malden, Massachusetts, USA: Oxford, UK and Malden, Massachusetts, USA.
Campbell, L. and Janda, R. 2000. Introduction: conceptions of grammaticalization and their problems. *Language Sciences* 23:93-112.
Catford, J. C. 1969. J. R. Firth and British Linguistics. In *Linguistics*, ed. Archibald A. Hill. Washington
Chafe, W. 1968. Idiomaticity as an anomaly in the Chomskian paradigm. *Foundations of Language* 4:109-127.
Chomsky , N. 1957. *Syntactic structures*. The Hague: Mouton & Co. [Reprinted in Berlin and New York, 1985].
Chomsky, N. 1980. *Rules and representations*. Oxford: Basil Blackwell.
Chomsky, N. 2000. *New horizons in the study of language and mind*. Cambridge: Cambridge University Press.
Comrie, B. 1976. *Aspect*: Cambridge Textbook in Linguistics. Cambridge: Cambridge University Press.
Cook, V. J. 1988. *Chomsky's universal grammar. An introduction*. Oxford: Basil Blackwell.
Cook, V. J. and Newson, M. 1996. *Chomsky's universal grammar*. Oxford: Blackwell Publishers.
Coulson, S. 2000. *Semantic leaps. Frame-shifting and conceptual blending in meaning construction.* Cambridge: Cambridge University Press.
Cowie, A. P. 1998a. Creativity and formulaic language. *Linguistica e Filologia* 8:159-170.
Cowie, A. P. 1978. The place of illustrative material and collocations in the design of a learner's dictionary. In *In Honour of A.S. Hornby*, ed. Peter Strevens, 127-139. Oxford: Oxford University Press.
Cowie, A. P. 1981. The treatment of collocations and idioms in learners' dictionaries. In *In honour of A.S. Hornby*, ed. Peter Strevens, 223-235. Oxford: Oxford University Press.
Cowie, A. P. 1986. Collocational dictionaries - a comparative view. In *Proceedings of the fourth Anglo-Soviet English studies seminar*, ed. M. Murphy. London: The British Council.
Cowie, A. P. ed. 1989. *Oxford Advanced Learner's Dictionary of Current English*. Oxford: Oxford University Press.
Cowie, A. P. ed. 1998b. *Phraseology. Theory, analysis, and applications*. *Oxford Studies in Lexicography and Lexicology*. Oxford: Oxford University Press.
Cowie, A. P. 1999. *English dictionaries for foreign learners: a history*. Oxford: Oxford University Press.

Cowie, A. P. 2000. The EFL dictionary pioneers and their legacies. In *Kernerman Dictionary News; Editor: Ilan Kernerman*, 1-6.
Cowie, A. P. 2002. Harmonising the vocabulary of Risk. Paper presented at *The Tenth EURALEX International Congress*, Copenhagen, Denmark.
Cowie, A. P. and Howarth, P. 1996. Phraseological competence and written proficiency. In *Language and Education*, ed. G.M. Blue and R. Mitchell, 80-93.
Cowie, A. P., Mackin, R. and McCaig, I. R. 1983. General introduction. In *The Oxford Dictionary of Current Idiomatic English*, ed. A. P. Cowie, Mackin, R. and McCaig, I. R., x-xvii. Oxford: Oxford University Press.
Cowie, C. 1995. Grammaticalisation and the snowball effect. *Language and Communication* 15 (2):181-193.
Croft, W. and Cruse, A. D. 2004. *Cognitive linguistics*. Cambridge: Cambridge University Press.
Cruse, D. A. 1986. *Lexical semantics*. New York: Cambridge University Press.
Cutler, A. 1983. Lexical complexity and sentence processing. In *The process of language understanding*, ed. G. B. Flores d'Arcais and R.J. Jarvella. Chichester, Sussex: Wiley.
Dienhart, J. M. and Kasch, H. 2000. A gentle introduction to regular expressions (for use in searching large corpora). In *Biskit. Bulletins from ISK's IT-Center*.
Dik, S.C. 1989. *The theory of functional grammar*.vol. 1. Dordrecht: Foris.
Everaert, M., van der Linden, E., Schenk, A., Schreuder, R. 1995. Introduction. In *Idioms: structural and psychological perspectives*, ed. Erik-Jan van der Linden Martin Everaert, André Schenk, and Rob Schreuder,, 1-13. Lawrence Erlbaum Associates: Hillsdale, New Jersey; Hove, UK.
Fauconnier, G. 1997. Mappings in thought and language. Cambridge: Cambridge University Press.
Fauconnier, G. and Sweetser, E. 1996. Cognitive links and domains. In *Spaces, worlds and grammar*, ed. Gilles Fauconnier and Eve Sweetser, 1-28. Chicago, IL: university of Chicago Press.
Fauconnier, G. and Turner, M. 1998. Conceptual integration networks. *Cognitive Science* 22:133-187.
Fillmore, C. J. 1968. The Case for Case. In *Universals in Linguistic Theory*, ed. E. Bach, R. Harms (eds.), 1-88. New York: Holt, Reichard and Winston.
Fillmore, C. J. 1977a. The case for case reopened. In *Syntax and Semantics, Volume 8. Grammatical Relations*, ed. Peter Cole & Jerold M. Sadock (eds.), 59-81. New York: Academic Press.
Fillmore, C. J. 1977b. Scenes-and-frames semantics. In *Linguistic Structures Processing*, ed. Antonio Zampolli, 55-81. Amsterdam, New York, Oxford: North-Holland Publishing Company.
Fillmore, C. J. 1982. Frame semantics. In *Linguistics in the morning calm*, ed. Linguistic Society of Korea, 111-137. Seoul: Hanshin.
Fillmore, C. J. 1985. Frames and the semantics of understanding. *Quaderni di Semantica* 6:222-255.

Fillmore, C. J. 1992. "Corpus linguistics" or "computer-aided armchair linguistics". In *Directions in corpus linguistics. Proceedings of Nobel Symposium 82, Stockholm, 4-8 August 1991*, ed. Jan Svartvik, 35-60. Berlin, New york: Mouton de Gruyter.

Fillmore, C. J. 2003a. Multiword expressions: An extremist approach: http://www.cis.upenn.edu/~ace/kick_off_nov2003/fillmore.ppt.

Fillmore, C. J. and Atkins, B.T.S. 1994. Starting where the dictionaries stop: the challenge of corpus lexicography. In *Computational Approaches to the Lexicon*, ed. B.T.S. Atkins and A. Zampolli, 349-393. Oxford: Oxford University Press.

Fillmore, C. J., Baker, C. F., and Sato, H. 2002. Seeing arguments through transparent structures. Paper presented at *The third international conference on language processes and evaluation (LREC)*, Las Palmas.

Fillmore, C. J., Johnson, C. R., and Petruck, M. R. L. 2003b. Background to Framenet. *International Journal of Lexicography* 16:235-250.

Fillmore, C. J., Kay, P., and O'Connor; M. 1988. Regularity and Idiomaticity in grammatical constructions: the case of *let alone*. In *Language Journal of the Linguistic Society of America*, 501-538.

Firth, J. R. 1957. *Papers in linguistics, 1934-1951*. London: Oxford University Press.

Firth, J. R. 1968a. A synopsis of linguistic theory, 1930-1955 (Special volume of the Philological Society, Oxford, 1957, 1-31). In *Selected papers of J. R. Firth 1952-1959*, ed. F. R. Palmer, 168-205. London: Longman.

Firth, J. R. 1968b. Descriptive linguistics and the study of English. In *Selected papers of J. R. Firth 1952-1959*, ed. F. R. Palmer, 96-113. London: Longman.

Firth, J. R. 1968c. Linguistic analysis as a study of meaning. In *Selected papers of J. R. Firth 1952-1959*, ed. F. R. Palmer, 12-26. London: Longman.

Fischer, O., Rosenbach, A, and Stein, D. ed. 2000. *Pathways of change: grammaticalization in English*. *Studies in language companion series*. Amsterdam: Benjamins.

Gabelentz, G. von der. 1891. *Die Sprachwissenschaft. Ihre Aufgaben, Methoden, und bisherigen Ergebnisse.* Leipzig: Weigel.

Gardner, H. 1995. Green ideas sleeping furiously [March 23, 1995]. *The New York Review of Books*: http://cogweb.english.ucsb.edu/Abstracts/Gardner_95.html.

Geeraerts, D. 1990. Soorten van niet-compositionaliteit bij idiomen. Paper presented at *AIO course on Syntax and Semantics of Idiomatic Constructions*, Leiden, the Netherlands.

Geeraerts, D. 1997. *Diachronic prototype semantics : A contribution to historical lexicology*: Oxford studies in lexicography and lexicology. Oxford: Clarendon.

Gibbs, R. W. 1980. Spilling the beans on understanding and memory for idioms in conversation. *Memory and cognition* 8:149-180.

Gibbs, R. W. 1995. Idiomaticity and human cognition. In *Idioms: structural and psychological perspectives*, ed. Erik-Jan van der Linden Martin Everaert, André Schenk, and Rob Schreuder, 97-116. Hillsdale, New Jersey; Hove, UK: Lawrence Erlbaum Associates.

Gibbs, R. W. and Gonzales, G. 1985. Syntactic frozenness in processing and remembering idioms. *Cognition* 20:234-259.

Gibbs, R. W. and Nayak , N.P. 1989a. Psycholinguistic studies on the syntactic behaviour of idioms. *Cognitive Psychology* 21:100 - 138.
Gibbs, R. W., Nayak, N. P. and Cutting, C. 1989b. How to kick the bucket and not decompose: analysability and idiom processing. *Journal of Memory and Language* 28:576-593.
Givón, T. 1971. Historical syntax and synchronic morphology: an archaeologist's field trip. *Chicago Linguistic Society* 7:394-415.
Givón, T. 1989. *Mind, code and context. Essays in pragmatics*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
Gläser, R. 1988. The grading of idiomaticity as a presupposition for a taxonomy of idioms. In *Understanding the lexicon: Meaning, sense and world knowledge in lexical semantics*, ed. W. Hüllen and R. Schultze, 264-279. Tübingen: Max Niemeyer.
Goldberg, A. E. 1995. *Constructions. A Construction Grammar approach to argument structure.* Chicago: University of Chicago Press.
Grady, J. 1997. Foundations of meaning: Primary metaphors and primary scenes, University of California, Berkeley: Ph.D. dissertation.
Gramley, S. and Pätzold, K. M. 1992a. A survey of modern English, 61-70. London: Routledge.
Gramley, S. and Pätzold, K. M. 1992b. Words in combination. In *A Survey of Modern English*. London and New York: Routledge.
Greenbaum, S. 1970. *Verb-intensifier collocations in English*: Janua Linguarum Series Minor 86. The Hague: Mouton.
Halliday, M. A. K. 1994. *An introduction to functional grammar*. London, New York, Sydney, Auckland: Arnold.
Halliday, M. A. K and Matthiessen, C. M. I. M. 1999. *Construing experience through meaning. A language-based approach.* London and New York: Cassell.
Halliday, M. A. K. 1966. Lexis as a linguistic level. In *In Memory of J. R. Firth*, ed. J. C. Catford C. E. Bazell, M. A. K. Halliday, and R. H. Robins, 148-162. London: Longman.
Halliday, M. A. K. 1978. *Language as social semiotic*. London: Edward Arnold.
Hanks, P. 2000. Contributions of lexicography and corpus linguistics to a theory of language performance. Paper presented at *The Ninth EURALEX International Congress, EURALEX 2000*, Stuttgart.
Hansen, M. Mosegaard. 1998. Funktionel og kognitiv lingvistik. In *Seks foredrag om de lingvistiske skoledannelser*. København: Romansk Institut, Københavns Universitet.
Harder, P. 1996a. *Functional semantics: A theory of meaning, structure and tense in English*: Trends in linguistics. Studies and monographs 87. Berlin: Mouton de Gruyter.
Harder, P. 1996b. Linguistic structure in a functional grammar. In *Content, Expression and Structure. Studies in Danish Functional Grammar*, ed. Michael Fortescue Elisabeth Engberg-Pedersen, Peter Harder, Lars Heltoft and Lisbeth Falser Jacobsen, 423-452. Amsterdam and Philadelphia: John Benjamins Publishing Company.
Harder, P. 1999. Partial autonomy. Ontology and methodology in cognitive linguistics. In *Cognitive linguistics: Foundations, scope and methodology*, ed. Theo Jannssen and Gisela Redeker, 195-222. Berlin and New York: Mouton de Gruyter.

Harder, P. 2001. Kognition i funktionelt perspektiv. Lecture handout and notes. In *Odense Linguistics Circle*. Odense, Denmark: University of Southern Denmark.
Harris, Z. 1991. *A theory of language and information*. Oxford: Oxford University Press.
Haspelmath, M. 1999. Why is grammaticalization irreversible? *Linguistics* 37:1034-1068.
Hausmann, F. J. 1985. Kollokationen im deutschen Wörterbuch. Ein Beitrag zur Theorie des lexikographischen Beispiels. In *Akten des Essener Kolloquiums zur Grammatik im Wörterbuch*, ed. Heinrich Bergenholtz and Joachim Mugdan. Tübingen: Max Niemeyer Verlag.
Heid, U. 1994. On ways words work together - Topics in lexical combinatorics. Paper presented at *Euralex '94*, Amsterdam.
Heine, B. 2003. Grammaticalization. In *The handbook of historical linguistics*, ed. Brian D. Joseph and Richard D. Janda, 575-601. Oxford: Blackwells.
Heine, B., Claudi, U. and Hünnemeyer, F. 1991. *Grammaticalization: A Conceptual Framework*. Chicago: University of Chicago Press.
Hopper, P. J. 1987. Emergent Grammar. *Berkeley Linguistics Society Proceedings* 13:139-157.
Hopper, P. J. and Traugott, E. Closs. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.
Howarth, P. A. 1996. *Phraseology in English academic writing*: Lexicographica Series Maior, 75. Tübingen: Max Niemeyer.
Humboldt, W. von. 1825. Über das Entstehen der grammatikalischen Formen und ihren Einfluss auf die Ideenentwicklung. In *Abhandlungen der Königlichen Akademie der Wissenschaften zu Berlin*, 401-430.
Jackendoff, R. 1995. The boundaries of the lexicon. In *Idioms: Structural and psychological perspectives*, ed. Erik-Jan van der Linden Martin Everaert, André Schenk, and Rob Schreuder,, 133-165. Hillsdale, New Jersey; Hove, UK: Lawrence Erlbaum Associates, Publishers.
Jespersen, O. 1924. *The philosophy of grammar*. London: George Allen & Unwin Ltd.
Johnson, C. R., Fillmore, C. J., Petruck, M. R. L., Baker, C. F., Ellsworth, M., Ruppenhofer, J., and Wood, E. J. 2003. *FrameNet: Theory and Practice, Version 1.1*: http://www.icsi.berkeley.edu/~framenet/book/book.html#whatgetsannotated.
Johnson, C. R., Fillmore, C. J., Wood, E. J., Ruppenhofer, J., Urban, M., Petruck, M. R. L and Baker, C. F. 2001. The frame semantic basis. In *The FrameNet Project: Tools for Lexicon Building*: http://xbean.cs.ccu.edu.tw/~dan/oodbResearch/SemanticWebPapers/FrameNet/FrameNet1R0.75/docs/book.html.
Jones, S. and Sinclair, J. M. 1974. English lexical collocations. *Cahiers de Lexicologie*:15-61.
Katz, J .J. and Postal, P. M. 1964. *An integrated theory of linguistic description*. Cambridge, Mass.: MIT Press.
Katz, J. J. 1973. Compositionality, idiomaticity, and lexical substitution. In *A Festschrift for Morris Halle*, ed. S. Anderson and P. Kiparsky. New York: Holt, Rinehart, and Winston.

Katz, J. J. and Postal, P. M. 1963. Semantic interpretation of idioms and sentences containing them. *MIT Research Laboratory of Electronics, Quarterly Progress Report* 70:275-282.
Kjellmer, G. 1994. *A dictionary of English collocations*. New York: Oxford University Press.
Klappenbach, R. 1968. Probleme der Phraseologie. *Wissenschaftliche Zeitschrift der Karl-Marx-Universität* 17:221-227.
Krenn, B. 2000. Empirical Implications on Lexical Association Measures. Paper presented at *The Ninth EURALEX International Congress, EURALEX 2000*, stuttgart.
Lakoff, G. 1987. *Women, fire and dangerous things. What categories tell us about the mind.* Chicago: University of Chicago Press.
Lakoff, G. 1991. The contemporary theory of metaphor. In *Metaphor and thought, 2nd. ed.*, ed. Andrew Ortony, 202-251. Cambridge: Cambridge University Press.
Lakoff, G. and Johnson, M. 1980. *Metaphors we live by*. Chicago: University of Chicago Press.
Lakoff, G. and Johnson, M. 1999. *Philosophy in the flesh. The embodied mind and its challenge to western thought.* New York: Basic Books.
Langacker, R. W. 1977. Syntactic reanalysis. In *Mechanisms of syntactic change*, ed. C.N. Li, 59-139. Austin: University of Texas Press.
Langacker, R. W. 1987. *Foundations of Cognitive Grammar. Volume I. Theoretical prerequisites*.vol. 1. Stanford, CA: Stanford University Press.
Langacker, R. W. 1988. A view of linguistic semantics. In *Topics in cognitive linguistics*, ed. Brygida Rudzka-Ostyn, 49-90. Amsterdam and Philadelphia: John Benjamins Publishing company.
Langacker, R. W. 1990. *Concept, image, and symbol. The cognitive basis of grammar*: Cognitive linguistics research. Berlin, New York: Mouton de Gruyter.
Langacker, R. W. 1992. The symbolic nature of cognitive grammar: The meaning of of and of of-periphrasis. In *Thirty years of linguistic evolution*, ed. Martin Pütz, 483-502. Philadelphia and Amsterdam: John Benjamins Publishing Company: .
Langacker, R. W. 1999a. Assessing the cognitive linguistic enterprise. In *Foundations and Scope of Cognitive Linguistics*, ed. Theo Janssen and Gisela Redeker, 13-57. Berlin: Mouton de Gruyter.
Langacker, R. W. 1999b. *Grammar and conceptualization*: Cognitive Linguistics Research 14. Berlin and New York: Mouton de Gruyter.
Langendoen, D. T. 1968. *The London School of Linguistics: A study of the linguistic theories of B. Malinowsky and J. R. Firth*. Cambridge, Massachusetts: The M.I.T. Press.
Lea, D. ed. 2002. *Oxford collocations dictionary for students of English*. Oxford: Oxford University Press.
L'Homme, M, and Bertrand, C. 2000. Specialized lexical combinations: Should they be described as collocations or in terms of selectional restrictions. Paper presented at *The Ninth EURALEX International Congress, EURALEX 2000*, Stuttgart.
Lindström, T. 2004. The history of the concept of grammaticalisation, Department of English language and linguistics, Sheffield University: ph.d. thesis.

Lipka, L. 1994. Lexicalization and institutionalization. In *The Encyclopedia of Language and Linguistics*, ed. R.E. Asher, 2164-2167. Oxford: Pergamon.
Makkai, A. 1977. Preface. In *The Early History of the Rumanian Language*: http://www.hungarian-history.hu/lib/dunay/dunay01.htm.
Makkai, A. 1993. Idiomaticity as a reaction to l'arbitraire du signe in the the universal process of semeio-genesis. In *Idioms: processing, structure, and interpretation*, ed. Cristina Cacciari and PatriziaTabossi. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.
Johnson, M. 1987: *The body in the mind*. Chicago: University of Chicago Press
Malinowski, B. 1935. *Coral gardens and their magic*. London: Allen and Unwin.
Meillet, A. 1912. L'évolution des formes grammaticales. *Scientia (Rivista di Scienza)* 12.
Mel'cuk, I. 1988. Semantic description of lexical units in an explanatory, combinatorial dictionary: Basic principles and heuristic criteria. *International Journal of Lexicography* 1:165-188.
Mel'cuk, I. 1998. Collocations and lexical functions. In *Phraseology. Theory, analysis, and applications*, ed. A.P. Cowie, 23-53. Oxford: Clarendon Press.
Mey, Jacob L. 1993. *Pragmatics. An introduction*. Oxford: Blackwell Publishers.
Mitchell, T. F. 1971. Linguistic 'goings-on': Collocations and other lexical matters arising on the syntagmatic record. *Archivum Linguisticum*:35-69.
Moon, R. 1998. *Fixed expressions and idioms in English: A corpus-based approach*: Oxford Studies in Lexicography and Lexicology. Oxford: Oxford University Press.
Moon, R. 2000. Phraseology and early English dictionaries. Paper presented at *The Ninth EURALEX International Congress, EURALEX 2000*, Stuttgart.
Newmeyer, F. J. 1998. *Language form and language function*. Cambridge, Mass.; London, England: The MIT Press.
Newmeyer, F. J. 2000. Deconstructing grammaticalization. *Language Sciences* 23 (2-3):187-229.
Nichols, J. 1984. Functional theories of grammar. *Annual Review of Anthropology* 13:97-117.
Nunberg, G., Sag, I. A. and Wasos, T. 1994. Idioms. *Language* 70:491-538.
Palmer, F. R. ed. 1968. *Selected papers of J. R. Firth 1952-59*. London: Longmans.
Palmer, H. E. 1933a. *Second Interim Report on English Collocations*. Tokyo: Kaitakusha.
Palmer, H. E. 1933b. Some notes on construction-patterns. *IRET Institute Leaflet* 38.
Pauwels, P. 2000. *Put, set, lay and place: A cognitive linguistic approach to verbal meaning*: LINCOM Studies in Theoretical Linguistics 19: Lincom Europa.
Pawley, A. 1985. Lexicalization. In *Language and linguistics: the interdependence of theory, data and application*, ed. D. Tannen and J. Alatis. Washington DC: University of Georgetown.
Pedersen, J. 1995. The identification and selection of collocations in technical dictionaries. *Lexicographica* 11:60-73.
Peters, A. M. 1977. Language-learning strategies: does the whole equal the sum of the parts? *Language* 53:560-573.
Peters, A. M. 1983. *The units of language acquisition*: Cambridge monographs and texts in applied psycholinguistics. Cambridge: Cambridge University Press.

Poulsen, S. 1991. Word combinations as lexical building blocks. A theoretical framework and a possible concept for a specialized dictionary of English collocations for text production. (Not published), Institute for Language and Communication, Odense University: Prize paper.
Poulsen, S. 2002. Creative collocations and conceptual integration. Paper presented at *The Way We Think*, Odense.
Pustejovsky, J. 1995. *The generative lexicon*. Cambridge, Mass: MIT Press.
Rosch, E. 1973. Natural categories. *Cognitive Psychology* 7:532-547.
Rosch, E. 1975. Cognitive reference points. *Cognitive Psychology* 7:532-547.
Rumelhart D., McClelland J., and the PDP Research Group. 1986. *Parallel distributed processing. Explorations in the microstructure of cognition*. Cambridge, MA: MIT Press.
Ruppenhofer, J., Baker, C. F., and Fillmore, C. J. 2002. Collocational information in the FrameNet database. Paper presented at *The Tenth EURALEX International Congress*, Copenhagen.
Saeed, J. I. 1997. *Semantics*. Oxford: Blackwell Publishers Ltd.
Saussure, F. de. 1966. *Course in general linguistics*.vol. McGraw-Hill Book Company. New York.
Schøsler, L. 2003. Grammaticalisation of valency patterns? - An investigation into valency patterns and support verb constructions, based on diachronic corpora. *Forum for Modern Language Studies* 39:400-413.
Searle, J. 1983. *Intentionality: an essay in the philosophy of mind*. Cambridge: Cambridge University Press.
Searle, J. 1999. *Consciousness*. http://philosophy.berkeley.edu/jsearle/Consciousness1.rtf
Sinclair , J. ed. 1995. *Collins Cobuild English dictionary*: HarperCollins Publishers.
Sinclair, J. M. 1991. *Corpus, concordance, collocation*: Describing English Language. Oxford: Oxford University Press.
Sinclair, J. M. 1992. Trust the text. In *Advances in systemic linguistics*, ed. M. Davies and L. Ravelli, 5-19. London: Pinter.
Sinha, C. 2001. Cognitive linguistics, psychology and cognitive science, draft 1. In *Draft Chapter for Handbook of cognitive linguistics*, ed. D. Geeraerts and H. Cuyckens, 1-34. Oxford: Oxford University Press.
Sinha, C. and López, K. Jensen de. 2000. Language, culture and the embodiment of spatial cognition. *Cognitive Linguistics* 11:17-41.
Skinner, B. F. 1957. *Verbal behaviour*. New York: Appleton-Century-Crofts.
Stefanowitsch, A. and Gries, S. Th. 2003. Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics* 8.2:209-243.
Stubbs, M. 1996. *Text and corpus analysis*. Oxford: Blackwell Publishers Ltd.
Swan, M. 1995. *Practical English Usage*. Oxford: Oxford University Press.
Sweetser, E. 1988. Grammaticalization and semantic bleaching. Paper presented at *Berkeley Linguistics Society, Proceedings of the Fourteenth Annual Meeting*, Berkeley.
Swinney, D. 1979a. Lexical access during sentence comprehension: (re)consideration of context effects. *Journal of Verbal Learning and Verbal Behaviour* 18:645-659.

Swinney, D. and Cutler, A. 1979b. The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behaviour* 18:523-534.
Taylor, J. R. 2002. *Cognitive grammar*: Oxford Textbooks in Linguistics. New York: Oxford University Press.
Taylor, J.R. 1992. How many meanings does a word have? *Stellenbosch Papers in Linguistics*:133-168.
Taylor, J.R. 1995. *Linguistic categorization: prototypes in linguistic theory*. New York: Oxford University Press.
Telija, V., Bragina, N., Oparina, E. and Sandomirskaya, I. 1994. Lexical collocations: denominative and cognitive aspects. Paper presented at *Euralex '94*, Amsterdam.
Tolkien, J. R. R. 1974. *The lord of the rings*. London: Allen and Unwin.
Turner, M. 1993. An image-schematic constraint on metaphor. In *Conceptualizations and mental processing in language*, ed. Richard A. Geiger and Rudzka-Ostyn Brygida, 291-306. Berlin, New York: Mouton de Gruyter.
Tversky, B. and Hemenway, K. 1984. Objects, parts and categories. *Journal of Experimental Psychology. General* 113:169-193.
Van de Voort, M. E. C. and Vonk, W. 1995. You don't die immediately when you kick an empty bucket: A processing view on semantic and syntactic characteristics of idioms. In *Idioms: structural and psychological perspectives*, ed. Erik-Jan van der Linden Martin Everaert, André Schenk, and Rob Schreuder, 284-299. Hillsdale, New Jersey; Hove, UK: Lawrence Erlbaum Associates.
van den Eynde, K. 1997. From verbal to nominal valency: Some methodological reflections. In *The valency of nouns, Odense Working Papers in Language and communication, No. 15*, ed. Karen Van Durme, 3-29. Odense: Odense University.
Van Durme, K. and van den Eynde, K. 1998. A verb typology on a distributional basis I: General typology. In *Studies in valency IV: Valency and verb typology*, ed. Karen and Schøsler Van Durme, Lene, 9-19. Odense: Odense University Press.
Van Valin, R. D., Jr. 2001. Functional linguistics. In *The Handbook of Linguistics*, ed. Mark Aronoff and Janie Rees-Miller, 319-336. Malden, Mass.; Oxford, UK: Blackwell Publishers Ltd.
Vinogradov, V. V. 1947. Ob osnovnuikh tipakh frazeologicheskikh edinits v russkom yazuike. In *1864-1920. Sbornik statey i materialov*, ed. A.A. Shakmatov, 339-364. Moscow: Nauka.
Weinreich, U. 1969. Problems in the analysis of idioms. In *Substance and structure of language*, ed. J. Puhvel, 23-81. Berkeley and Los Angeles: University of California Press.
Wittgenstein, Ludwig. 1953. *Philosophical investigations*. Oxford: Basil Blackwell.

Corpus concordance sampler. http://www.collins.co.uk/Corpus/CorpusSearch.aspx: Collins Cobuild.
Oxford English Dictionary Online. http://dictionary.oed.com: Oxford University Press.