

I Introduction

1.1 Aims and objectives

The aim of this study is to shed light on certain aspects of the noun phrase which over the years have proved problematic and which, as a result, have been the topic of a considerable amount of debate. The aspects dealt with in part I predominantly concern the internal structure of noun phrases containing two nominal elements. At the heart of the discussions in this part is the issue of headedness; other aspects, such as referentiality and predication, definiteness, determination and quantification will, however, also play an important role and will be inextricably woven into the discussion. As such, it is hoped, this part of the study will not only offer plausible and revealing analyses of specific NP constructions, but will also contribute to our understanding of the relations between and functions of the various elements within the NP in general. In part II the focus of attention will shift towards the cognitive and pragmatic factors underlying the production and interpretation of noun phrases. From a pragmatic point of view, information packaging, i.e. the speaker's choice of the most effective linguistic form to achieve his/her communicative objectives, will be explored in detail, while from a cognitive point of view an attempt will be made to explain certain linguistic phenomena in terms of the way knowledge is stored in and retrieved from the mind. The division of labour will, however, not be as strict as these descriptions may suggest. Pragmatic and cognitive factors will be taken into consideration in part I as well; likewise, syntactic and semantic aspects will feature prominently in part II.

1.2 Theoretical framework and overall approach

Over the last three decades or so a large number of linguistic theories have been developed taking what Van Valin and LaPolla (1997: 11) refer to as the 'communication-and-cognition perspective' to language. Examples are Functional Grammar (Dik 1997a, 1997b; more recently Functional Discourse Grammar, e.g. Hengeveld and Mackenzie 2006), Role and Reference Grammar (Van Valin 1993; Van Valin and LaPolla 1997),

Cambridge University Press

978-0-521-18395-6 - The English Noun Phrase: The Nature of Linguistic Categorization

Evelien Keizer

Excerpt

[More information](#)

2

The English noun phrase

Functional-Systemic Grammar (Halliday 1985), Lexical-Functional Grammar (Bresnan 1982, 2001) and Construction Grammar (Fillmore 1988; Croft 2001). Since each of these theories has its own aims and set of underlying assumptions, they differ – sometimes significantly – in approach and emphasis. What these theories have in common, however, is the basic assumption that language is principally a means of communication, and that the form of linguistic utterances is determined first and foremost by their use. In addition, it is recognized that the study of language use must take place within the broader perspective of such general cognitive processes as reasoning, conceptualization and the storage and retrieval of knowledge. For linguists working with these theories, this means that the only viable approach to the study of language is one in which communicative and cognitive factors are not only taken into consideration, but form the basis of any attempt to explain the formal behaviour of linguistic utterances.

The present study, too, has been written in the communication-and-cognitive tradition; the analyses proposed, however, have not been developed within any particular theoretical framework. This has been a very deliberate choice, made for a number of reasons. First of all, although all the functional-cognitive frameworks mentioned certainly have their strong points, they may prove to be of limited use in trying to solve the kind of issues addressed in this study. Naturally, one could choose to solve this problem by adapting or extending the theory in question – typically by applying notions from other, kindred theories. Instead, I have chosen an even more eclectic approach, selecting useful notions used in one or more of the various theories, without favouring any one of these theories. An additional advantage of this approach is that there is no reason to confine oneself to established linguistic notions, and that other disciplines, such as discourse analysis, psycholinguistics and cognitive science, can be resorted to.

A further reason for not working within any particular theoretical framework has to do with the fact that, despite claims of pragmatic or communicative adequacy, the frameworks in question rarely have a truly empirical basis. Instead, there is a tendency to take a predominantly deductive (top-down) approach, with entire systems being developed on the basis of a limited number of isolated examples. At some point, this is likely to create the danger that the internal coherence and consistency of the system (as a system) is considered more important than its ability to account for actual language use. This weakness has typically been associated with the more formal (logical, generative) approaches to language; unfortunately, it also characterizes many communicative-cognitive frameworks.

Nevertheless, this study will make use of underlying representations. It needs to be emphasized, however, that these proposed representations are to be regarded merely as notational tools, not as having any psychological status. They are not intended to represent a particular stage in the process of language production or interpretation, nor do they serve to represent the

Introduction**3**

way knowledge is conceptualized in the mind. Instead, they are simply an abstract, idealized means of reflecting syntactic, semantic and pragmatic differences between seemingly similar constructions. In fact, one of the aims of this study is to demonstrate the limitations of underlying representations and the strict classification of construction types they imply. In this sense, it is as much about the nature of linguistic classification as about the specific constructions dealt with in the separate chapters. Thus it is shown that analyses based on a combination of the relevant syntactic, semantic and pragmatic differences do not always lead to neatly distinguishable groups of constructions. And although in most cases it is possible, to some extent at least, to represent differences in linguistic behaviour fairly adequately in underlying representation, through labelling, bracketing, indexing and the use of variables and symbols, these means are often insufficient to reflect more subtle distinctions.

The discussions in part I are designed to show both the advantages and the limitations of strict classification and formal representation. It will be argued that some of the problems in such an approach can be solved by opting for a compromise in which underlying representations are regarded as representing only prototypical cases (best examples of a category). This means that these representations must be regarded as considerable oversimplifications of the complex linguistic reality they are meant to reflect: small differences in degree of category membership (gradience) and the possible convergence of linguistic categories (fuzziness) are, after all, difficult to represent. This in itself need not be a problem; linguistic models, like all models, are by definition oversimplifications. It does, however, leave unanswered the important question of what causes the gradience and/or fuzziness observed. In part II an attempt is made to identify some of these causes and to illustrate how they may affect the behaviour of the (component parts of) noun phrases.

Naturally, the approach chosen has disadvantages as well. In particular, there will be no shared basis to start from, which means that the notions and terminology used cannot be assumed to be familiar. As we all know, however, the only way to avoid confusion and misunderstandings is to clearly specify and define the terms and notions used, no matter what approach is being taken. Therefore, both parts of the book will begin with a brief introduction of the general concepts to be applied at various points in the chapters to follow. Information on the use of more specific terms and notions will be provided whenever necessary.

1.3 The ICE-GB Corpus*1.3.1 ICE-GB: general information*

In view of the fact that in the present study pragmatic factors will play a prominent role in the analyses provided, it will not come as a surprise that

Cambridge University Press

978-0-521-18395-6 - The English Noun Phrase: The Nature of Linguistic Categorization

Evelien Keizer

Excerpt

[More information](#)

4

The English noun phrase

extensive use will be made of authentic examples from written and spoken language. Although a variety of sources have been used, the large majority of examples have been taken from the British component of the International Corpus of English (ICE-GB), a fully tagged, parsed and checked one-million word corpus of written and spoken English, compiled and grammatically analysed at the Survey of English Usage, University College London, between 1990 and 1998 (Nelson, Wallis and Aarts 2002). In exploring this corpus, use was made of the retrieval software ICECUP (the *ICE Corpus Utility Program*), also produced by the Survey of English Usage.¹

With just over one million words (500 texts of approximately 2,000 words each), ICE-GB is small in comparison with, for instance, the *British National Corpus* (BNC; Aston and Burnard 1998), which contains 100 million words. However, since ICE-GB was designed primarily as a resource for syntactic studies, every text unit ('sentence') in ICE-GB has been syntactically parsed, and each unit presented in the form of a syntactic tree.

The texts in ICE-GB date from 1990 to 1993 inclusive. All authors and speakers are British, but differ with regard to gender, age, education and regional background. There are 300 spoken texts (50 of which are scripted) and 200 written. The spoken texts are divided into dialogues (both private, e.g. direct conversations and telephone calls, and public, i.e. with an audience) and monologues (unscripted and scripted). The written component of the corpus consists of 150 printed texts (academic writing, non-academic writing, press reportage, instructional writing, persuasive writing and creative writing), and 50 non-printed texts (non-professional, such as students' essays, and professional, such as correspondence).

As mentioned, the ICE-GB corpus is both tagged and parsed. The tagset used was devised by the Survey of English Usage, in collaboration with the TOSCA research group at the University of Nijmegen (Greenbaum and Ni 1996), and was based (with some modifications) on the classifications given in Quirk et al. (1985). In the first instance, tagging was done automatically by means of the TOSCA tagger (Oostdijk 1991); subsequently, the output was manually checked at the Survey of English Usage. Next, the tagged corpus was submitted to the TOSCA parser for syntactic analysis. In many cases, the parser produced several alternative analyses; in these cases, the corpus annotators were given the task of selecting the contextually correct analysis.

¹ ICE-GB can be ordered via the Survey's website (<http://www.ucl.ac.uk/english-usage/>), where a free sample corpus of ten texts, together with ICECUP, is available for download. A detailed and up-to-date manual is now also available (Nelson et al. 2002). For more detailed information about the corpus, its compilation and analysis and the software used and developed by the Survey of English Usage, see also e.g. Nelson (1996), Aarts et al. (1998), Wallis et al. (1999), Wallis and Nelson (2000).

Introduction**5***1.3.2 Use of the corpus*

It is important to emphasize at this point that the present study is not of a corpus-linguistic nature. No extensive use will be made of statistical data to describe and account for the linguistic behaviour of the constructions dealt with. There are various reasons why, for the purposes of this study, a corpus-linguistic approach would have been both inappropriate and inadequate. In the first place, in order to answer the questions addressed in this study, the exact distribution of the constructions in question in terms of frequency of occurrence is but of minor importance. Secondly, no matter how meticulously tagging and syntactic parsing have been performed, the problematic and often ambiguous nature of the kind of constructions examined means that tagging and parsing has not always been done in a consistent manner. Therefore, in consulting the corpus, a large number of different search strategies were used to ensure that every possible instance of a particular construction type was retrieved. Each instance was then examined carefully in its original context and subsequently classified on the basis of the syntactic, semantic, pragmatic and cognitive criteria proposed. Obviously, not all of these examples have found their way into this study. Instead a selection was made of the most relevant examples, ranging from perfectly straightforward cases to the more problematic ones.

1.4 Organization of this study

This study is divided into two parts, both of which start with a chapter on the key notions used in the discussions to follow (chapters 2 and 9). Part I is mainly concerned with the internal structure of the English NP and concentrates on constructions containing two nominal elements. These constructions are notoriously problematic in terms of determining syntactic and semantic headedness, as well as in terms of the function of and relation between the component elements. Other important features are the scope of the determiner, definiteness and quantification, pre- and postmodification, the use of anaphoric pronouns and the referentiality of the two nominal elements. In addition, the discourse function of these constructions will be considered, as well as the relation between this function and the semantic properties and syntactic behaviour of the constructions as a whole and their component parts. Finally, underlying representations will be proposed for each of the construction types distinguished, reflecting (as closely as possible) the semantic, syntactic and pragmatic differences observed. The constructions dealt with in this part are close appositions (*the poet Burns, Burns the poet*; chapter 3); *of*-appositions (*the city of Rome*; chapter 4); binominal noun phrases (*that fool of a doctor*; chapter 5); pseudo-partitive constructions (*a lot of people, a piece of metal, a cup of*

Cambridge University Press

978-0-521-18395-6 - The English Noun Phrase: The Nature of Linguistic Categorization

Evelien Keizer

Excerpt

[More information](#)

6

The English noun phrase

coffee, a group of students; chapter 6); and *sort/kind/type*-constructions (*this sort/kind/type of problem*; chapter 7). Chapter 8 will provide an interim conclusion.

The primary concern of part II is not the internal structure of the noun phrase, but the role of pragmatic and cognitive factors in determining a speaker's choice for a particular construction. Again we will be looking at some problematic constructions, in particular those cases where a speaker seems to have a choice between two or more syntactically and semantically acceptable constructions. The analyses presented in chapters 10 to 12 are intended to show that the choice of the speaker in such cases is influenced by a combination of cognitive and pragmatic factors, which makes strict classification not only more difficult, but also less interesting. Chapter 10 tackles the question of whether it is justified and feasible to distinguish two basic types of noun: those that take complements (often referred to as relational nouns, e.g. the noun *father* in *the father of my friend*) and those that do not (non-relational nouns, e.g. *house* in *the house of my friend*). Chapter 11 deals with the basic principles underlying a speaker's choice between a discontinuous NP (e.g. *no approval has yet been given for the proposal*) and a continuous one (*no approval for the proposal has yet been given*). Chapter 12 is concerned with the question of what determines a speaker's choice between a prenominal possessive (e.g. *the author's opinion*) and a postnominal *of*-construction (e.g. *the opinion of the author*). Every chapter begins with an evaluation of a number of earlier proposals, which is subsequently taken as the starting point for analysis.

Chapter 13, finally, presents a number of general conclusions. Rather than summarizing the proposed analyses in detail, this chapter concentrates on the major themes of this study, pointing out the main tenets of the overall approach taken and the ways in which this approach can contribute to a better understanding of linguistic classification in general and of English nominal constructions in particular.

Cambridge University Press

978-0-521-18395-6 - The English Noun Phrase: The Nature of Linguistic Categorization

Evelien Keizer

Excerpt

[More information](#)

Part I The structural approach: possibilities and limitations

2 Headedness within the NP

2.1 Introduction

This chapter introduces some of the key notions to be used in part I of this study, in particular those relevant to the discussion of headedness within the NP. It reviews those syntactic, semantic and pragmatic notions applied in previous linguistic accounts which play an important role in the analyses presented in subsequent chapters. More detailed descriptions, as well as proposed modifications, can be found in these later chapters; the present chapter is simply meant to provide the necessary background information and serves the additional purpose of dispelling some of the prevailing terminological confusion.

2.2 Internal structure: headedness within the NP

One of the reasons headedness has proved a rather elusive notion is that it can be, and has been, defined at a number of levels. In traditional grammar the term ‘head’ was used to capture linguists’ intuitions about what constituted the most important part of a phrase (its central element or nucleus) and was consequently described in semantic terms. More theoretically inclined linguists, on the other hand, felt the need to couple this notion to the formal behaviour of the elements in question; as a result, certain morphosyntactic tests came to be used to establish headedness. Unfortunately, however, there proved to be numerous cases where the two approaches would select different heads. A third type of test for headedness, based on pronominalization, also failed to provide conclusive evidence. Thus the choice more or less remains between the rather vague, but intuitively appealing, semantic approach and the more systematic, but also more abstract and semantically less revealing, formal approach. Not surprisingly, traditional grammars, as well as functional and cognition-based grammars, largely opted for the former approach; formal theories focusing on syntax, such as generative grammar, opted for the latter. In addition, proposals have been made to regard certain problematic constructions as containing either two heads, or no head at all.

In this section some of the more commonly applied definitions and tests for headedness within the NP are discussed. The list provided does not claim to be exhaustive: it includes those criteria which are relevant to the analyses to be presented in the chapters to follow and whose applicability does not depend on any theory-specific definition.

2.2.1 *Semantic criteria*

The head-modifier distinction was originally used to reflect the intuitive idea that within each phrase one element was somehow more important than the others. Commenting on noun phrases, Jespersen (1924: 96), for instance, describes how '[i]n any composite denomination of a thing or person . . . , we always find that there is *one word of supreme importance* to which the others are adjoined as subordinates'. This idea of the head as the most important or primary element within the noun phrase can still be found in many recent characterizations of the term (e.g. Dik 1997a: 134; Quirk et al. 1985: 60).

But what does it mean for an element to be the most important, primary or central part of a noun phrase? The answer to this question is usually given in terms of the designation (or denotation) of a noun phrase. In Givón (2001: 59) we read that 'within the noun phrase, a noun is typically the syntactic and semantic head, defining the type of entity involved'. Similarly, in Langacker's (2002: 12–13) *Cognitive Grammar* it is the profile of the head that prevails at the composite structure level (the head is the 'profile determinant'), where the profile of an element is determined by its designation.¹

On the basis of this general semantic (or notional) characterization, it became possible to define a number of operational tests. The first of these consisted of the requirement that the head be distributionally equivalent to the composite construction; the second defined the head as the obligatory constituent (e.g. Quirk et al. 1985: 60–61; Zwicky 1985: 11; Huddleston and Pullum 2002: 24). In most cases, the operational tests of obligatoriness and distributional equivalence suffice to establish headedness, selecting that part of the construction as the head which also defines the type of entity referred to. In other cases, however, determining headedness may be less straightforward. In applying the test of obligatoriness, for instance, one is faced with the problem that in noun phrases with singular, countable heads the determiner cannot be left out either (e.g. Lyons 1977: 392). This

¹ The terms denotation and designation, as well as such related terms as intension, connotation, meaning and sense, have been used in many different, but often very similar and partially overlapping, ways, both in linguistics and philosophy (e.g. Mill 1856, Frege 1892, Carnap 1956, Lyons 1977), leading to what Geach (1970: 55) describes as 'a sad tale of confusion'. In what follows the notion of denotation plays an important role; it will be given the following, fairly general definition:

For a linguistic element to denote an entity means that this entity belongs to the set of entities to which the linguistic element in question applies.

Headedness within the NP

11

problem applies to other types of phrases as well. In prepositional phrases like *in Amsterdam*, for instance, neither the preposition nor the proper name can be omitted, and neither element can replace the phrase as a whole. This has led to different conclusions concerning headedness, depending on which criterion is taken as decisive. Most traditional linguists have tended to take the proper name or NP following the preposition as the head, this being the element with the highest degree of lexical content. Quirk et al. (1985: 60–1), on the other hand, conclude that since neither element can be omitted, such prepositional phrases are exocentric, while for Langacker (2002: 26) it is the preposition (a relational predication) which lends its profile to the composite structure (i.e. *in Amsterdam* designates a stative relation, not the city itself).

Even more problematic, perhaps, are the kind of constructions to be discussed in the rest of part I, i.e. noun phrases which contain two nominal elements which are either juxtaposed or connected by the functional element *of*. Which element, for instance, functions as the head in constructions like *the lady president*, *the poet Burns* or *the city of Rome*? Or in constructions like *this kind of theory*, *that fool of a doctor* or *a group of people*? According to the characterization given above, and the tests of obligatoriness and distributional equivalence, all these constructions contain two candidates for headedness. Again this has led to a number of analyses, ranging from those which regard either one of the two elements as the head to those which take both or neither of the elements to be the head.

To determine headedness in some of these problematic cases, a third test has sometimes been applied, based on the selection restrictions of the two nominal parts. Akmajian and Lehrer (1976), for instance, point out that in some binominal constructions only one of the two nominal elements complies with the selection restrictions of the verb. In examples (1a) and (1b), for instance, the verbs *spill* and *drink* require a liquid as their subject and direct object, respectively; this would indicate that the second noun, *wine*, functions as the head of the construction. In (2), on the other hand, the verb *break* selects a solid, breakable, subject/direct object, which would indicate that here it is the first noun which functions as the head.²

- (1) a. A bottle of wine spilled.
b. He drank a bottle of wine.
- (2) a. A bottle of wine broke.
b. He broke a bottle of wine.

These facts seem to suggest that in binominal constructions of this kind either of the two elements can function as the (semantic) head, depending

² Similar examples for Dutch can be found in Vos (1999).