

# *Heads in grammatical theory*

*Edited by*

GREVILLE G. CORBETT

*Professor of Linguistics and of Russian Language  
Department of Linguistic and International Studies  
University of Surrey*

NORMAN M. FRASER

*Logica Cambridge Ltd  
and Department of Sociology  
University of Surrey*

*and*

SCOTT McGLASHAN

*Department of Sociology  
University of Surrey*



CAMBRIDGE  
UNIVERSITY PRESS

---

# 1 Introduction

---

NORMAN M. FRASER, GREVILLE G. CORBETT *and*  
SCOTT McGLASHAN

---

## 1.1 Heads in grammatical theory

The majority of current grammatical theories refer explicitly to the *head* of a phrasal constituent. Yet while the term 'head' has entered the common currency of theoretical linguistics, this does not provide evidence of agreement on what it means. Nor does the term's long and varied career in linguistics guarantee that it identifies a notion which is not already identified by some other, more basic notion or interacting set of notions. The purpose of this volume is twofold: first, it aims to uncover and make explicit the notion (or notions) behind the term 'head'; second, it aims to investigate the status of the notion (or notions) in linguistic theory.

Most linguists would agree with the informal characterization that the head of a phrase is one of its constituents which in some sense dominates and represents the whole phrase. In an important paper published in the *Journal of Linguistics* in 1985, Zwicky drew attention to the fact that use of the term 'head' had been extended from syntax to morphology (for example, by Lieber, 1981; Williams, 1981; and Kiparsky, 1982) in spite of the fact that there was no generally agreed formal definition of the notion in syntax (though an important contribution had already been made by Gazdar and Pullum, 1981 and Gazdar, Pullum and Sag, 1982). Zwicky therefore set out to find a rigorous, generally acceptable definition for 'head'. He proceeded by examining the following eight candidate criteria for the identification of a constituent as a syntactic head.

- 1 Is the constituent the semantic argument, that is, the constituent whose meaning serves as argument to some functor?
- 2 Is it the determinant of concord, that is, the constituent with which co-constituents must agree?
- 3 Is it the morphosyntactic locus, that is, the constituent which bears inflections marking syntactic relations between the whole construct and other syntactic units?

- 4 Is it the **subcategorizand**, that is, the constituent which is subcategorized with respect to its sisters?
- 5 Is it the **governor**, that is, the constituent which selects the morphological form of its sisters?
- 6 Is it the **distributional equivalent**, that is, the constituent whose distribution is identical to that of the whole construct?
- 7 Is it the **obligatory constituent**, that is, the constituent whose removal forces the whole construct to be recategorized?
- 8 Is it the **ruler in dependency theory**, that is, the constituent on which others depend in a dependency analysis?

Application of Zwicky's criteria to a range of common syntactic constructions suggests that headedness is distributed amongst the constituents of a phrase: for example, in the construction NP+VP, VP is the morphosyntactic locus, the governor, and the obligatory constituent, but NP is the semantic argument and the determinant of concord. A harmonious solution cannot be achieved simply by discarding some criteria so as to yield a consistent set of results, since the distribution of successful and failed criteria differs between constructions. Zwicky concluded that 'head' should be identified with 'morphosyntactic locus' since this is all that is needed in order to state the very general principle of percolation which can be found in various forms in a variety of current grammatical theories. Briefly, percolation requires identity of some feature or features (such as syntactic category) between a head and its mother. Zwicky concluded that since headedness could be identified with a single, simple criterion, heads should not be accorded a privileged place in syntactic theory.

Zwicky also argued that a different generalization is required for the purpose of percolation in morphology, so the supposed relevance of heads to morphology is spurious. This conclusion is supported by Bauer (1990). The chapters in this volume deal with heads in syntax rather than morphology.

In a subsequent paper Hudson, whilst agreeing with the logic of Zwicky's argument, challenged his views on the ways in which constructions should be analysed and head-identification criteria applied (Hudson, 1987). On the basis of a set of assumptions derived from his theory of Word Grammar, Hudson succeeded in producing a harmonious analysis in which Zwicky's criteria uniquely identified a single head for each kind of construction surveyed. This led Hudson to conclude that 'head' is a category which unifies a range of different properties.

The notions of 'head' as this term has been applied in a variety of different theories are not, in fact, as disparate as Zwicky suggests... Different linguists may use the notion 'head' for different purposes – one for percolation, another for government, and so on – but this is to be expected in view of the multiplicity of properties that we have found for heads. (Hudson, 1987: 126)

The number of criteria for the recognition of heads suggested in this exchange (Hudson rejects one of Zwicky's criteria and goes on to offer others which Zwicky does not mention) and the lack of agreement on basic assumptions underlying the analysis of common syntactic constructions should not come as any surprise when it is remembered that the notion 'head' has entered linguistic theory by an unusually diverse set of routes.

The idea that one word may dominate another – that a subordinate word depends on a head word – is the central insight of traditional dependency grammar and its more recent offspring. Early notions of dependency are discernible in the Sanskrit grammar of Pāṇini (between 600 and 300 BC), in the Greek grammar of Apollonius Dyscolus (second century AD) and in the Latin grammar of Priscian (c. AD 450) (Robins, 1979). The medieval Arabic grammarians contrasted the governor or head of a phrase with the governed or dependent elements of the phrase. They stipulated amongst other things that a governor could have many dependents but a dependent could have only one governor; dependency relations were necessarily unidirectional; and dependents were required to be adjacent to their governors (Owens, 1988). Most modern dependency grammarians continue to work with these basic assumptions (for example, Robinson, 1970: 260).

Dependency grammar was first formalized by Tesnière (1959) and further developed by Hays (1964), Gaifman (1965) and Robinson (1970). In recent years there has been an upsurge of interest in dependency grammar, and a significant number of dependency theories have been developed (Anderson, 1971, 1977; Hudson, 1976, 1984, 1990; Miller, 1985; Mel'čuk, 1988; Starosta, 1988). Dependency grammar frameworks are assumed in this volume by Hudson and McGlashan.

The term 'head' is also used by Bloomfield in his seminal discussion of endocentric and exocentric constructions (1933: 195). For him, the head of a constituent is the daughter whose distribution is the same as that of the whole constituent (that is, Zwicky's 'distributional equivalent'). The substitutability of heads for larger constituents

played an important part in the development of structuralist linguistics (Wells, 1947; Harris, 1951; Hockett, 1958). The introduction of X-bar grammar by Chomsky in 1970 (though Harris sketched a recognizable version in 1951) can be interpreted as a response to Lyons' observation that 'the notions of endocentricity, subordination, etc., have no systematic significance in a "rewrite" grammar' (Lyons, 1968: 235). In an X-bar grammar, a phrasal category must be identical except for bar level to a daughter category. The daughter category may be regarded as the head of the phrase. In their recent detailed examination of X-bar grammar, Kornai and Pullum conclude that 'the key concept is headedness' (1990: 46).

Many different versions of X-bar grammar have been proposed since 1970 (see Muysken, 1982: 58). X-bar theory is one of the subtheories of Government-Binding (GB) theory, which is assumed in the chapters by Cann, Payne, Radford and Vincent. X-bar theory is also basic to Generalized Phrase Structure Grammar (GPSG) (Gazdar *et al.*, 1985), whose Head Feature Convention requires that mothers and head daughters, by default, should bear the same features (with the exception of the bar feature). In this way phrase-structure trees become structures for head-to-head feature percolation. Warner (1989) argues within a GPSG framework that certain kinds of phrase are multiply headed, namely those involving co-ordination and certain minor categories. This calls for a reformulation of the Head Feature Convention and of our understanding of the notion 'head'.

Head-driven Phrase Structure Grammar (HPSG) (Pollard and Sag, 1987; Borsley, this volume) places even more importance on heads, since all heads incorporate information about the non-heads with which they combine (in a way inspired by Categorical Grammar). There are very few independent rules other than the basic information-combining rule of unification.

A further strand of research on heads can be found in work on language typology. The word-order universals of Greenberg (1966) and Vennemann (1975), for example, are expressed in terms of categories such as subject, object and verb. However, the dominant serialization orders of languages have been stated more concisely by reference to the more general categories of head and modifier (Hawkins, 1983). This approach appears first in the writings of the medieval Arabic grammarians (Owens, 1988). Thus, for example, it might be said that the dominant order in Japanese is modifier-head, while the dominant order in Welsh is head-modifier. In chapter 11, Hawkins argues against this approach, an argument which is rendered all the more worthy of attention since it attacks the view which

Hawkins himself previously endorsed. Dryer (1992) bases his rejection of generalizations predicated on consistent ordering of heads and dependents upon detailed study of a sample of more than 600 languages.

Another important typological insight is offered by Nichols (1986), who distinguishes head-marking languages from dependent-marking languages. The fact that different languages place morphosyntactic markers of dependency at different ends of the dependency relation is interesting in itself, but the difference also has important consequences, such as the fact that predominantly head-marking languages appear not to have any exocentric constructions. The chapter by Nichols in this volume extends her typological investigation of centricity into new territory. Nichols' (1986) work focused exclusively on major syntactic categories. Vincent (this volume) carries the investigation forward by considering functional categories as well. (For further work on the role of heads in language typology see Giorgi and Longobardi, 1987, and Clements, 1989.)

So far, the Zwicky–Hudson debate on headedness (including subsequent contributions by other authors) has tended to centre on the English language. There is much scope for re-examination of the arguments and issues in the context of other languages. The chapters by Comrie and Corbett begin this process. Amongst the contributors to this volume, only Nichols (this volume) and Payne (personal communication) maintain what might be termed the 'head of construction constancy principle' which requires the head of a given construction to remain constant across languages. Thus, if *N* is the head of a nominal construction in English, then *N* must also be the head of nominal constructions in all languages which have them. The other authors are prepared to accept that the head of a given construction may change from language to language, and Cann's theory predicts that this will happen.

Much which has been written on the subject of heads has been devoted to identifying the heads of specific constructions and, in particular, of the noun phrase (or determiner phrase, or whatever it turns out to be). The major issue here – and one which runs through almost every chapter of the volume – is whether or not functional categories can be heads. Not very many years ago, the idea that anything other than the noun should be taken as the head of a nominal construction would not have found sympathy anywhere. More recently, however, just such a position has been advanced by people working in a variety of different frameworks. These include Hudson (1984), Hellan (1986), Abney (1987), Szabolcsi (1987),

Löbel (1989), Olsen (1989) and Pollock (1989). Early intimations can be found in Jackendoff (1972) and Hogg (1977). The chapters by Corbett, Payne and Radford focus primarily on the question of what heads the nominal phrase. The chapters by Cann and Vincent provide more general discussions of the role of functional categories as heads.

Whilst each of the chapters in this collection makes a significant contribution to at least one of the areas of endeavour in which the notion 'head' has been adduced, many of them succeed in bringing together insights from more than one theory or field of research. Furthermore, several of the chapters strike out in new directions, themselves defining new areas for future investigation: for example, Hudson takes the debate into the realm of psycholinguistics with his question *Do we have heads in our minds?*; Vincent blends existing research in typology and functional categories and adds an ingredient which has not previously been considered in the context of heads, namely processes of grammaticalization; Zwicky examines the foundations of grammatical theory and comes up with a new category to set alongside head, namely 'base'.

Each chapter in the collection addresses the subject of heads in grammatical theory from a slightly different perspective. Taken together, the chapters define both the points of agreement in the centre and the points of divergence at the margins. In this respect, at least, it is clear that many heads are better than one. Whether the same can ultimately be said to apply to grammatical constructions remains to be seen.

## 1.2 Outline of the book

In their examinations of criteria for the identification of heads, Zwicky and Hudson considered data drawn mainly from English. In chapter 2, Greville Corbett uses Russian numeral phrases to test the competing sets of head-identification criteria. Application of Zwicky's criteria suggests that head-like properties are split between numeral and noun. Application of Hudson's re-interpretation of Zwicky's criteria suggests that head-like properties are unambiguously located on the numeral. Taking the investigation one step further, Corbett applies the two sets of head-identification criteria to adjective-noun combinations. Here it is Zwicky's approach which consistently identifies the noun as head, while Hudson's approach distributes head-like properties between the adjective and the noun. Hudson's approach can only be rescued if the adjective is controversially analysed as the ruler in dependency grammar.

In chapter 3, Bernard Comrie offers a criterion for the identification of heads which is quite different from those which have previously featured in the debate. This criterion is discussed in relation to Haruai, a non-Austronesian language of New Guinea, and the results achieved with that language are quite robust. The generalization which drives the criterion is this: in endocentric constructions in Haruai, stress always falls on the dependent rather than the head, in the absence of special pragmatic marking. This generalization is independent of word order. Comrie suggests that searching for language-internal criteria for heads in other languages may be equally successful. Criteria for the identification of heads may belong to two categories: universal or language-particular.

Chapter 4, by Ronnie Cann, reconsiders the criteria discussed by Zwicky and Hudson. Like Hudson, Cann rejects the 'semantic argument' criterion. He also rejects the 'ruler in dependency grammar' criterion. However, unlike Hudson's analysis, Cann's analysis does not find consistent agreement amongst the criteria in their identification of heads. Agreement is obtained where only major categories are involved but not where functional categories are involved. Working within the framework of GB, Cann develops an explanatory account of the variation in distribution of head-like properties across different constructions. This account predicts that distribution will differ between languages as well as between different constructions in the same language.

Chapter 5, by Andrew Radford, and chapter 6, by John Payne, both investigate the number of heads in a nominal phrase. Following distinct lines of argument, the authors reach opposite conclusions. Assuming a GB theoretical framework, Radford argues that a complex nominal like *all these silly theories* should be analysed as a quantifier phrase containing a determiner phrase containing an adjectival phrase containing a noun phrase, and that each of the superordinate phrasal constituents is double-headed (for example, *silly theories* is a constituent whose immediate head is the adjective *silly*, and whose ultimate head is the noun *theories*). Thus, like the Roman god Janus (the mythical guardian of gates and doors, beginnings and ends), complex nominals have two heads (one at the beginning, the other at the end). Payne tests two hypotheses – the single-head hypothesis and the multi-head hypothesis – to see which provides the better account of the facts of incorporation, subcategorization, the position of possessor phrases, appositional noun phrases and agreement and disagreement. *Contra* Radford, he concludes that – unlike the mythical hydra – nominal phrases possess a single head, and this head must be a noun.

Chapter 7, by Nigel Vincent, examines three recent strands of research in syntax and shows some ways in which they shed light on each other. These are: (a) the typological work initiated by Nichols (1986) on head-marking and dependent-marking languages; (b) the increased importance (especially in GB) accorded to functional categories and their projections; and (c) the study of the process of grammaticalization as mechanism of grammatical change. In her 1986 paper, Nichols excluded function words from consideration. Vincent argues that given the recent interest in functional heads, the head-marking/dependent-marking typological research programme should be extended to include them. He further observes that the grammaticalization principle that 'function words are always etymologically derived from lexical words' has implications for the head/dependent status of function words: does grammaticalization lead to category change with or without corresponding change of head/dependent status? These issues are investigated in the context of complement clauses and complementizers.

Chapter 8, by Johanna Nichols, presents a cross-linguistic study of ellipsis in discourse to see whether selection of elements for deletion under ellipsis is governed by universal principles referring to heads. In her sample, Nichols finds evidence for all logical possibilities. That is, she finds languages in which (a) only heads can be removed, (b) only non-heads can be removed, (c) either heads or non-heads can be removed, and (d) neither heads nor non-heads can be removed. In her earlier work on head-marking and dependent-marking languages, Nichols has shown that there is a strong connection between dependent-marking and the presence of the exocentric/endocentric distinction; in languages which are predominantly head-marking, all constituents appear to be endocentric. By contrast, Nichols' chapter in this volume shows that selection of elements to remove under ellipsis is not related to whether the language in question marks heads or dependents. The somewhat surprising conclusion is that deletion of elements under ellipsis is not subject to any universal principles which make use of the notion 'head'. Making this (or any other) cross-linguistic observation requires a constant cross-linguistic definition of what is head in any given construction.

Chapter 9, by Robert Borsley, examines the Head Feature Principle (HFP) of HPSG. This principle requires the value of the HEAD feature in a mother to be identical to the value of the HEAD feature in its head daughter. Unlike GPSG's Head Feature Convention, which supplies default values, HPSG's HFP is absolute and cannot be overridden. Borsley argues that the HFP as currently conceived misses an

important generalization. The solution is to cast the HFP as a default principle which assigns values to features unless other values are explicitly introduced by rule. This revision has serious implications for HPSG since it can no longer be the case that the theory is based purely on unification. Borsley shows how his arguments can be extended to GB, and suggests that they are relevant to any constituency-based framework.

In chapter 10, Scott McGlashan develops a unification-based account of the notion 'head'. Drawing on ideas highlighted in the dependency grammar tradition, categories are partitioned into 'head' and 'modifier' features, roughly corresponding to the distinction between 'take' and 'make' features in Categorical Grammar (CG). He argues that this account is superior to that offered by CG in respect of serialization, category constancy and motivation. However, he also presents evidence which undermines simple accounts of category constancy, and especially those which claim that semantic properties of a head are identical to those of its phrase. For example, in *chocolate soldiers*, the 'material property' of the head is not that of the phrase: *chocolate soldiers* are made of chocolate, not flesh.

Chapter 11, by John Hawkins, is the only contribution in the volume which seriously questions the usefulness of the notion 'head', and its place in Universal Grammar. Hawkins' position is particularly noteworthy since he has in the past made extensive use of the head-modifier distinction in his theories of word-order universals (Hawkins, 1983, 1984). More recently, Hawkins has put forward a parsing theory of word-order universals (Hawkins, 1990), in which serialization patterns are related to such things as the appropriate placement of 'Mother Node Constructing Categories' (MNCCs) so as to minimize cognitive load during processing. MNCCs are constituent-peripheral categories which, by virtue of being unique to a particular kind of phrase, allow a unique mother node to be constructed over the MNCC as early as possible: for example, determiners are MNCCs in English. In this volume, Hawkins argues that many of the properties of heads identified in the Zwicky-Hudson debate are, in fact, properties of MNCCs. The fact that these properties appear in clusters is related to the role of MNCCs as clues for on-line processing, and not to the existence of some putative category 'head'.

The position outlined in chapter 12, by Richard Hudson, is diametrically opposed to the one advanced by Hawkins. Hudson argues that it is necessary to refer explicitly to heads both in grammars and in sentence structures. Drawing on descriptive and

psycholinguistic evidence, Hudson claims that a number of complex phenomena can be explained only if listeners are able to identify heads in utterances during processing. Thus, at the coarsest level, there are three distinct views on heads: (a) heads do not exist in grammars or in sentence structures; (b) heads exist in grammars but not in sentence structures; (c) heads exist in grammars and in sentence structures.

It is fitting that the final chapter, chapter 13, should be written by Arnold Zwicky, whose 1985 paper has helped to focus so many of the issues discussed in this volume. In his latest contribution, Zwicky proposes three binary features to make distinctions which, he argues, are required in any adequate account of dependency functions. These features are: 'F',  $\pm$ semantic functor; 'H',  $\pm$ morphosyntactic locus or head (recall Zwicky's conclusions in his 1985 paper); and a new feature 'B' (Base),  $\pm$ external representative. All constituents bear these features. An Operator Head has the feature structure (+F+H+B); a Modified Head has the feature structure (-F+H+B); and a Specifier has the feature structure (+F+H-B). This account therefore argues against a single definition of heads in which all criteria always identify a single constituent as head. Rather, it predicts that criteria of headedness will not always agree, but that patterns of disagreement are meaningful and signify distinctions between Operator/Argument, Modified/Modifier and Specifier/Base.

## NOTE

Earlier versions of the chapters which make up this collection were presented at the Talking Heads Round Table held at the University of Surrey in March 1991. The editors would like to express their thanks to the Research Committee of the University of Surrey for generous support and to all the contributors for their enthusiastic participation.