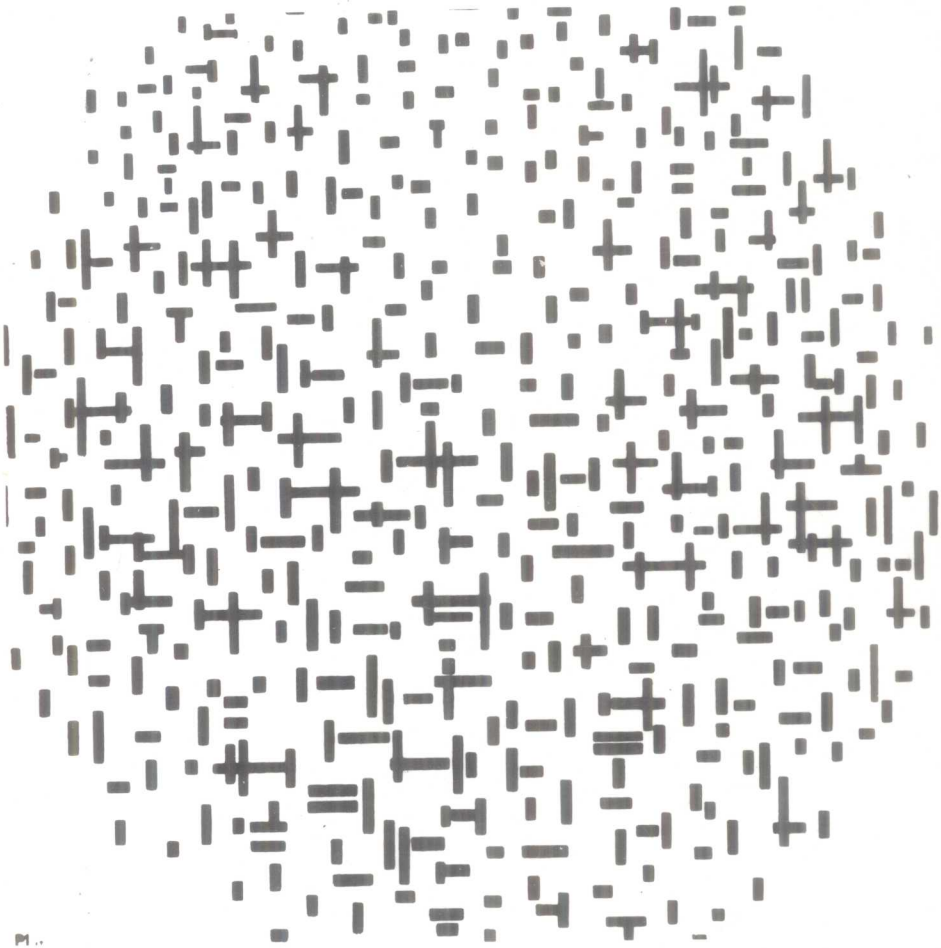
Word Grammar

RICHARD HUDSON



Word Grammar

RICHARD HUDSON

Basil Blackwell

For
Gay, Lucy and Alice

First published 1984
First published in paperback 1986

Basil Blackwell Ltd
108 Cowley Road, Oxford OX4 1JF, UK

Basil Blackwell Inc.
432 Park Avenue South, Suite 1505,
New York, NY 10016, USA

British Library Cataloguing in Publication Data
Hudson, Richard, 19
Word grammar.
1. Semantics
I. Title
422 PE1576

ISBN 0-631-13186-8

Typeset by Getset (BTS) Ltd, Eynsham, Oxford
Printed in Great Britain by
Redwood Burn Ltd, Trowbridge

Preface

This book is about a linguistic theory, word grammar, which is radically different from a theory I advocated in print in 1976, daughter-dependency grammar: notably, it has no place for constituents larger than words or for syntactic features, and it presents language structure as a special case of general cognitive structure. I gather from various friends and colleagues that the world sometimes disapproves of people who change their minds in this way, so I should like to offer the world an apology for littering the intellectual landscape with yet another theory. My only excuse is that it is very difficult to be right first time. I am sure that word grammar is a great improvement on daughter-dependency grammar, so I hope it will survive longer.

I wrote a draft of a completely different book, with the same title, in 1982, and I had extremely helpful comments on either the whole or parts of it from the following colleagues: Noam Chomsky, David Katz, Tokumi Kodama, Jim McCawley, Barbara Prangell, Neil Smith, Mark Steedman, Maureen Taylor, Willy Van Langendonck, Peter Wason, Yorick Wilks, Deirdre Wilson and Mary Wood (plus a number of anonymous publisher's readers). I am most grateful to all of them, and hope they agree that the present book is better than the one they helped to demolish.

I also take this opportunity to acknowledge a grant from the SSRC which allowed me to concentrate for a short three months on the analysis of English grammar underlying the theory.

Contents

1

Introduction

A THEORY OF LANGUAGE STRUCTURE

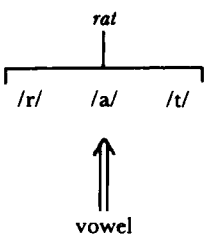
Language as a network

Word grammar is a theory of language structure. At the most general level, it consists of generalization (1):

(1) A language is a network of entities related by propositions.

For example, we might take the entities to include words and phonemes, connected by propositions such as 'The word *rat* consists of /r/ followed by /a/ followed by /t/,' and 'The phoneme /a/ is an instance of the category "vowel"'; in these examples, the two propositions refer to the entities 'the word *rat*', 'the phoneme /r/', 'the phoneme /a/' and 'the phoneme /t/', and 'the category "vowel"'. Further, the two propositions between them connect the word *rat* to the category 'vowel', though indirectly. The two propositions and the five entities to which they refer make up a (very small) subnetwork within the total network of English, and we could diagram this subnetwork as in (2):

(2)



This diagram uses a set of conventions which I shall explain shortly, but the main point of it is to show how the relations among entities in a grammar may be presented as a network. A little imagination leads

from this little subnetwork to the monster network that would be needed for the whole language.

This very general claim does not obviously distinguish word grammar from any other available theory of language structure. Some theories are explicitly formulated in terms of networks, the obvious example being stratificational theory (see Sullivan 1980 for a recent summary). However, it makes no difference whether a theory refers explicitly to the notion of 'networks' or not, because all the available theories of language structure can be interpreted in such a way that they are compatible with (1). For example, a phrase-structure rule can be seen as a proposition about a mother category and its daughters ('a sentence consists of . . .'); a transformation as a proposition about adjacent phrase-markers in a derivation; a lexical entry as a proposition about the syntactic, phonological and semantic structures that are compatible with one another as structures for the same word; and so on. Ever since the notion of 'well-formedness condition' became commonplace in transformational theory (following McCawley 1968a), it has been possible to interpret a grammar in this way, and similar remarks apply to other theories, *mutatis mutandis*. Moreover, any grammar must define a structure in which every entity is related to every other entity directly or indirectly, so it constitutes a network in the sense of (1).

Nevertheless, even at this very general level there are two respects in which (1) does distinguish word grammar from at least some other theories.

(a) In the sense in which I intend the word 'proposition', it does not apply to *procedures*. That is, in terms of the distinction which cognitive scientists commonly make between procedural and declarative knowledge, I see a language as consisting (entirely) of *declarative* knowledge, expressed as propositions. I am aware that the opposite view is widespread among cognitive scientists, including some scholars for whose work I have great admiration (J.R. Anderson 1976, Miller and Johnson-Laird 1976, Winograd 1975, Small 1980, among others); but I have seen no persuasive arguments against the position that we linguists generally take for granted, namely the declarative view of language. (In this respect the cognitive grammar of Lakoff and Thompson 1975 is quite exceptional.) At the same time, though, I recognize the need for procedures that make use of language as part of the solution for practical problems; for example, when you buy a ticket you have to perform various actions, some of which are linguistic. These procedures are what I call 'schemata' in Hudson 1980a, and I

shall assume (at least for the purposes of the present book) that they are a separate part of one's knowledge from the language itself. At any rate, I recognize this interface as an interesting and important area for future research.

(b) The view of a language as a network does not commit one to analysing the language as a pair of lists: a list of *rules*, and a list of lexical *entries*. This is of course how a transformational grammar is structured, and likewise most other kinds of grammar, with a few exceptions which include stratificational grammar again, space grammar (Langacker 1982), and lexical generative grammar (Diehl 1981). The list view is a specific interpretation of the network claim, and implies two distinctions which are in fact rather hard to make. One of the distinctions is that between one lexical entry and another – the notorious distinction between polysemy and homonymy being the main culprit. So far as I know, no satisfactory suggestion has yet been made for deciding where the boundaries of a lexical entry lie (Lyons 1977:550 surveys the state of the art), so we have no general principle for deciding, for example, whether *horn* involves just one lexical entry or two in the two meanings 'horn of an animal' and 'musical instrument' (or, worse still, 'part of a motor-car'). My view is that we are trying to make a distinction which corresponds to no kind of reality, so the search for principles on which to base it is bound to be fruitless. The network approach allows us to sidestep the issue completely: we have a set of word-forms, and a set of word-meanings which are related to the word-forms in a complex way, as part of the network. Once we have established these connections, there is no motivation for then going on to chop up the network into discrete chunks called 'lexical entries'.

The other distinction which is generally assumed is that between rules and lexical entries. However, the distinction was absent from the earliest version of transformational grammar (Chomsky 1957), was then denied, or at least blurred, in generative semantics (e.g. McCawley 1968b), and has never been accepted in stratificational grammar. Evidently it is less obvious than one might think if one were to judge by most current discussions in the transformational literature, and elsewhere too. Moreover, there is a generally unrecognized transition area between typical 'rules' and typical 'lexical entries' where their respective properties are mixed up in a confusing way. For example, there are a reasonably large number of constructions which are tied to particular words, such as *What about . . .* ? Should this be treated by means of a rule, on the ground that such sentences contain

patterns which are not permitted by other rules (e.g. *What about the unemployment figures?* contains no verb)? Or should it be included in the lexicon, as a lexical entry, on the ground that it is an idiosyncratic pattern restricted to the word *what* or *how* followed by *about*? Once again, I know of no general principles to which we could refer in deciding questions like this, so I feel that the only reasonable conclusion to be drawn is that we are trying to draw a distinction (rule versus lexical entry) which has no reality. (Other examples of constructions like *What about . . .* ? are mentioned in Lyons 1968:178, Baker 1979 and Carroll 1978; it is easy to extend their lists with examples like *If it weren't for . . ., Not that . . ., If only . . ., No sooner . . . than . . ., The more . . . the more . . ., More fool . . . for . . .-ing . . .*, etc.)

The distinctive characteristic of word grammar, then, is that it is less restrictive than many other theories in that it does not impose on the grammar the requirement that it consist of a list of rules and a (distinct) list of lexical entries. The network structure is compatible with a more diffuse organization than this, and I have argued that such diffuseness is a good thing because it seems to correspond to the reality of language data. Of course, in general one does not boast of the flexibility and power of one's theory of language, because we are all trying to restrict flexibility in order to make claims about language that are as precise as possible. However, it goes without saying that we should put the restrictions of our theory in the same places as language itself is restricted; and what I have tried to show in the above is that the places where other theories restrict language structure, by requiring boundaries, are places where no boundaries actually exist.

Linguistic entities

We can now be more precise about the entities which occur in the network (and in the next subsection we shall do the same for the propositions). What is a linguistic entity, as far as word grammar is concerned? The theory is called this precisely because the notion 'word' is central to the answer. I am aware of the problems that arise in defining the word, especially if we are aiming at an operational definition which will tell us in all cases, for all languages, how to recognize a word boundary. Nevertheless, it seems to be widely agreed among linguists these days that most, or perhaps even all, languages make use of units which could reasonably be called words. Moreover, any reasonably sophisticated theory of language must take account of

cliticization, which I take it involves two words, one of which is contained in the other, so we may expect uncertainty about word-boundaries in precisely these cases. For instance, how many words are there in *We're* (as in *We're ready*)? Similar problems can again be expected from compounding, where two words are combined to form one (e.g. *chocolate box*). We shall have something to say about both these phenomena on pp. 48 – 52, but in the meantime my purpose is to suggest that the problems of identifying word boundaries in these uncertain cases should not prevent us from taking the word as the pillar on which this theory rests.

Assuming, then, that we all know (roughly) what we mean by 'word', the following are the kinds of entity which are referred to in a word grammar:

(a) Single words, which may be taken at any degree of generality (though the different degrees are, of course, kept clearly distinct in the grammar); thus the following would all be treated as entities in a grammar for English:

(i) the word *ran*, with final *b* due to assimilation to a following bilabial (as in *ran back*);
(ii) the word *ran* (irrespective of final consonant);
(iii) the word *run* (irrespective of tense – i.e. the 'lexeme' *run*, in the terminology of Lyons 1977:19);
(iv) the word-type 'past tense verb';
(v) the word-type 'verb';
(vi) the entity 'English word' (it is this entity that brings together all the parts of a grammar for English, as opposed to other languages – a distinction which clearly needs to be made, especially if we want to study bilingual speakers);
(vii) the entity 'word' (which unifies all that we have to say about language in general).

(b) Parts of words, whether these parts correspond to morphemes or to phonological segments (we shall see on pp. 52 – 7 that this distinction is in any case somewhat unclear), and as I have just pointed out, the parts of a word may also include smaller words. Once again, these parts may be taken at any degree of generality, from the particular allophone or allormorph up to the general classes like 'vowel'.

(c) Strings of words involved in co-ordinate structures, whether as conjuncts or as the whole co-ordinate structure. Some of these strings are 'incomplete', in that they could not occur outside a co-ordinate

structure, hence the special connection between word-strings and co-ordinate structures. For example, we recognize all the bracketed strings in the following as linguistic entities, to which the grammar may refer:

({John collects books} and {his wife, records})

The whole of chapter 5 is devoted to the discussion of co-ordinate structures.

(d) Word-meanings – more generally, any element in the semantic structure. It makes little difference whether or not we say that word-meanings are strictly speaking part of the linguistic network as such, because they are one of the two main points at which the (linguistic) network of words makes contact with the (non-linguistic) network of general cognitive structures, and the interface by definition belongs to both networks. In any case, the view for which I shall argue in this book is that the linguistic network is just part of a larger network, without any clear differences between the linguistic and non-linguistic beyond the fact that one is about words and the other is not; so I cannot see any empirical issue at stake in deciding whether word-meanings are part of language or not. The main point is that there are propositions which link words to word-meanings, so we need to take account of these propositions in the theory of language structure.

(e) Elements of the utterance-event, by which I mean the participants of the utterance-event (speaker, addressee), the time, the place and various other factors. For some words, we need to refer to one or more of these elements – for instance, we need to say that *pussy* and *cat* are different with respect to the constraints they place on speaker or addressee (*pussy* being used only by or to young children). Once again, I see no point in arguing about whether or not such elements are really part of the language; since they are part of the interface between words and the rest of the world, they could be taken as part of both. Moreover, if everything is basically contained in a single grand network, it makes little difference what we decide.

Two features of that list of entities are particularly noteworthy: first, there is no mention of phrases, clauses or sentences, as the word constitutes the upper boundary (apart from the units of co-ordinate structures). Chapter 3 will take up this controversial claim, and justify it. Second, the discussion shows that I am not concerned to define the boundaries of language, in contrast with many other linguistic theories. If we assume that words and their parts are indisputably linguistic entities, then there are two ways in which we could use this assumption

as the basis for delimiting language, according to whether we concentrate on the entities or on the propositions that relate them. If we focus on entities, we could say that the language network includes words and their parts (and co-ordinate structures), and nothing else – i.e. excluding word-meanings and elements of the utterance-event. If on the other hand we focus on propositions, we could say that a proposition is linguistic if it refers to a word (or word-part or word-string), in which case word-meanings and so on would be included in the language network, on the ground that they are referred to by some propositions that also refer to words. I see no reason why we should force ourselves to make this choice, because nothing seems to follow from it: I shall try to show that there is little or nothing that we can say about the network centred on words which we could not also say about other parts of the general network.

Propositions connecting linguistic entities

Word grammar allows just five types of proposition to refer to words. (For simplicity, I shall ignore word-parts and word-strings in the following discussion, since they are cumbersome to refer to, and in any case the propositions in which they are involved represent a subset of the types which can refer to words.) Each type of proposition may be identified by means of a particular function, or predicate expression, so we can represent the different propositions as *formulae*. For example, one of the predicate expressions we shall use is 'composition', which refers to the part – whole relation. This is the predicate contained in the proposition that relates the word *rat* to its constituent parts, /r/, /a/ and /t/. A fairly standard formulation of this proposition would put the predicate in front of a pair of brackets, with its arguments inside the brackets;

(1) composition (*rat*,[/r/ + /a/ + /t/])

However, this is not the notation that I shall use for formulae, though I should find it somewhat hard to produce solid evidence that my notation is better. Instead, I make use of a different notational tradition, that of tagmemic grammar (e.g. Pike 1982), which I also used in earlier theory, daughter-dependency grammar (Hudson 1976a). In this tradition, relations are expressed in terms of 'slots' (corresponding to the predicate expression) and their 'fillers'. This tradition needs to be supplemented, however, because a slot and a single filler cannot express

anything more than a one-place predicate, whereas all the propositions in word grammar express relations and involve two arguments. So I have followed the convention of tying the slot to one argument as its 'possessor', with the filler as the other argument. This notation corresponds to a prose statement: 'the . . . of . . . is . . . .' For example, in the case of *rat* it corresponds to 'the composition of *rat* is /r/ + /a/ + /t/.' To show the 'possessor' of the slot, I shall put it in brackets after the name of the slot (a widespread convention, found in Kac 1978 for example), and I connect the slot and its filler by a colon, which can be verbalized as 'is'. Thus, in place of (1) above, we shall use (2):

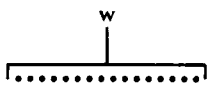
(2) composition(*rat*): /r/ + /a/ + /t/

However, it is helpful to have a notation for *diagrams*, to represent the same propositions. For example, we have already seen how the composition of *rat* can be shown diagrammatrically (in the diagram of (2) on p. 1. One of the advantages of diagrams is that they make the network property of the grammar easier to grasp, whereas formulae have to be presented as a list.

The five types of proposition, then, are as follows (in none of the diagram conventions is the vertical dimension relevant):

(a) *Composition.* This relates the word to its parts.

Diagram:



(Here, as elsewhere, *w* stands for the word concerned.)

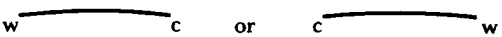
(b) *Model.* This relates the word to the more general entity of which it is an instance; for example, *rat* is an instance of a noun. We shall discuss this relation in more detail in the next section.

Diagram:

(Here *m* stands for the model.)

(c) *Companion.* These propositions relate the word to the other words with which it occurs, and will later be supplemented by more specific propositions which distinguish various grammatical relations, notably 'modifier', 'head' and 'subject' (see chapter 3 for further discussion).

Diagram:



(The *c* stands for 'companion'.)

(d) *Referent.* All the connections between a word and the semantic structure are established via the word's 'referent'. Chapter 4 will explain why I assume that every word has a referent, with a small handful of exceptions (e.g. *and*).

Diagram:

(Here and elsewhere I shall use the asterisk to distinguish a word's referent from the word itself.)

(e) *Utterance-event.* A word is linked to its speakers, addressees, times and places of utterance via the event of uttering the word, which I shall call the 'utterance-event'. These relations are somewhat complex, so I shall devote most of chapter 6 to them and largely ignore them till then.

Diagram:

To give a clearer idea of the relation between the two notations, here is an account of the part of the grammar which is responsible for the verb *pop,* as in *The balloon popped.* First we have it expressed in formulae, with a prose translation on the right which is not part of the grammar, but just a guide for the reader:

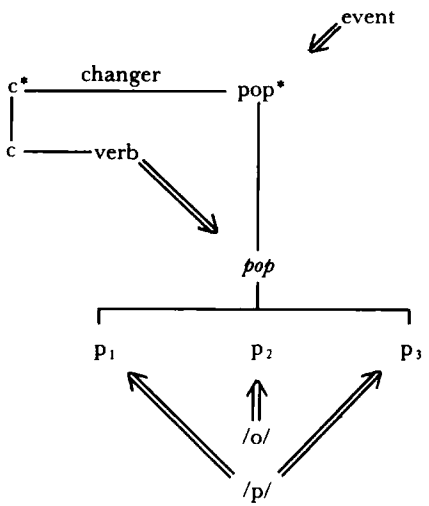
(3)	composition(*pop*):	$p_1 + p_2 + p_3$,	*Pop* consists of
	model(p_1):	/p/,	three phonemes of which the first is an instance of /p/,
	model(p_2):	/o/,	the second is an instance of /o/,
	model(p_3):	/p/,	and the third is another instance of /p/.
	model(*pop*):	verb	*Pop* is an instance of a verb.
	companion(verb):	c,	Verbs (e.g. *pop*)

c < verb	take a companion which precedes them.
referent(*pop*): pop*,	*Pop* refers to pop*,
model(pop*): event,	which is an
changer(pop*): c*	instance of an event, and involves a 'changer', namely the referent of the companion *c*.

I think the analysis should be self-explanatory, but there are some details which I shall discuss and justify in later sections. For the moment we are concerned primarily with notation, rather than content.

Now we can present the same information in terms of a diagram, using the conventions that I have just introduced:

(4)



This diagram should be self-explanatory too, except for one point: I have written the word 'changer' on top of the line which connects c* to pop*. In doing so I am anticipating a convention which I shall introduce

later, when we need to distinguish different semantic relations from one another, to supplement the relation 'companion'.

There are no other types of proposition than these in a word grammar, except for a couple of very elementary ones which always appear as conditions on other propositions, rather than as formulae in their own right. These deal respectively with *identity* and temporal *order*. In formulaic notation, we can by-pass the slot-filler notation in both cases, and use conventional symbols in between the related entities, = in the case of identity, and < for order. Thus we might have two entities, A and B, represented in the grammar as distinct entities, but we might then want to say that they are in fact one and the same entity, meaning that everything which is true of one is also true of the other. We can express this identity by writing 'A = B' in a formula, or by stretching the equals sign as a long double line in a diagram. (The double line with an arrow is in fact just a modification of this convention, to show the asymmetrical 'identity' relation between an instance and its model.)

An example of a part of the grammar in which identity needs to be expressed is the part which describes the equative verb *be*, as in *That's John*. This verb identifies the referent of the subject with that of the complement, so in the structure for *That's John* we should find something like 'That* = John*'. Many parts of the grammar make a somewhat different use of the identity relation, however, in order to express *partial* identity. One such case is the part responsible for gapping (as in *John had porridge for breakfast, and Bill, toast*), which simply requires the second conjunct (the part after *and*) to be interpreted as though it were the same as the first conjunct, except for the two overt elements (*Bill* and *toast*), which are substituted for the corresponding parts of the first conjunct. To formalize this kind of partial identity, we use the 'A = B' notation. Chapter 5 on co-ordinate structures, will explain how this applies to gapped structures, but we shall also use the ' = ' notation at other points for expressing partial identity.

Lastly, we have the question of temporal order, for which we have the symbol < mentioned above. This is to be taken in its mathematical sense 'less than', and can be used in this sense in some parts of the grammar (e.g. the parts dealing with the meanings of gradable adjectives, where we can say that *X is big* means 'X has a size such that the normal size is less than the size of X'). In the case of temporal order, we can imagine points of time as being represented by integers which increase with time, so we can show that one time A is earlier than another time B by writing 'A < B'. We shall exploit this possibility in dealing with the

semantics of tense, but also in relation to temporal order of words in utterances. So if we want to say that the subject precedes its verb, then we can write the equivalent of 'subject < verb'. (We have precisely this rule formulated in (3), where the subject is represented as *c*, a companion of the verb.)

To summarize, we find the following types of proposition in a word grammar: composition, model, class, companion, referent and utterance-event (these are the main ones), plus identity (total or partial) and 'less than'. We also have two sets of notations for representing these types of proposition, according to whether we are using formulae or diagrams to represent the structure of the language. Between them, these propositions serve to link every linguistic entity to every other one, though in some cases the links are very indirect, so we may define the network of a language either by means of a diagram, or by means of a set of formulae. Of course, there are conceptual problems for any linguist who tries to capture anything bigger than a very small part of the total network, because the complexity increases alarmingly (though not exponentially, since there is a great deal of modularity). Each of the notations has its advantages and its snags, but for anyone interested in computerization, the formulaic notation is clearly the better of the two.

Levels and components

Most theories of language structure rest heavily on the notions 'level' and 'component', where 'level' refers to the separate structures that can be found in a sentence (semantic, syntactic, phonological, etc.) and 'component' refers to the separate parts of the grammar which are responsible for these structures. I think one of the reasons why linguists have in the past put such emphasis on these differences is because they offered the hope of modularity in the system – so we can concentrate on the phonology without worrying about the syntax, and so on. However, it is clear from the literature that most of the boundaries concerned are as problematic as those we considered earlier (e.g. polysemy versus homonymy).

In contrast, word grammar puts relatively little emphasis on levels and components, and there is certainly nothing comparable with the organization of a transformational grammar into blocks of rules which can be applied in sequence (first the phrase-structure rules, then the lexical rule, then the transformations, and so on). As a matter of fact, it is possible that this difference is simply a matter of style of presentation,