

J. A. VAN EK, *Four complementary structures of predication in contemporary British English: an inventory*. Groningen Studies in English 9. J. B. Wolters, Groningen, 1966. 207 pp. Price: f 27,50.

The most important word in the title of this book is in fact in the sub-heading – ‘inventory’. Over four-fifths of the book is given over to an itemized list of verbs providing information about their frequency, co-occurrence, etc., in relation to the ‘four structures’ of the title. It might seem odd, then, for a reviewer to concentrate on the remaining fifth, but there are very good reasons for doing so. The facts in the inventory are to a large extent beyond dispute. What have to be seriously considered are the theoretical and methodological presuppositions on which this inventory was based. This book – typical of a great deal of research being carried out for theses at present – has to be judged by wider criteria than the facts about English it contains.

The author divides the book into three parts. The first (pp. 1–9) is an introduction, in which he presents definitions of various relevant concepts, and outlines his aims, the nature of his corpus, and his method of description. Part Two is the aforementioned inventory (pp. 19–179). Part Three (pp. 181–204) is entitled ‘summaries and conclusions’. There is a short bibliography. Van Ek claims to have three aims in writing this book. His primary purpose is ‘to provide data’ (primarily frequential and collocational information) about four grammatical patterns in English which he considers (rightly) have been neglected in the standard handbooks; secondly, he wants to ‘provide a large body of well-authenticated citations illustrating the patterns’; and thirdly, he is concerned to ‘ascertain to what extent existing grammars can be said to give adequate ‘coverage’ of the patterns’ (7) – though this is ‘adequacy’ in a very different sense

from its present vogue usage (it refers to pedagogical adequacy, i.e. 'their value as reference works for the student in search of guidance in the field of usage' (198)). There is no reference to other criteria of adequacy. The patterns comprise a *headword* (a transitive verb) and a *structure* functioning as *complement*. There are four types of structure, all with the constituents *noun-stem* (or equivalent, e.g. a pronoun) + *non-finite verb-form*, these existing with a subject-predicate relation between them. Van Ek illustrates as follows (passive transformations are given here in brackets): I let *it happen* (*It was let happen*), I ordered *him to come* (*He was ordered to come*), I saw *him coming* (*He was seen coming*) and I saw *him beaten* (*He was seen beaten*) (3).

The inventory derives from the classification of patterns as found in two sources: a corpus consisting of about 1 million running words, and eight leading grammars. The former comprises both fiction (novels and plays) and non-fiction (newspapers, literary criticism and history, and informative prose) and is a reasonable sample for this kind of work – though of the 255 verbs listed in the summary, as many as 155 are low-frequency items (i.e. they occurred 5 times or less in the corpus) and are consequently ignored by the author in his final discussion. (This raises a number of questions as to the value of corpus work, when so many of the statistics are unusable, but this is not the place for such a digression.) On a number of occasions (e.g. 39, 45, 48, 119, 189) the author goes outside his corpus for examples, and he usually tells us where the information comes from (apart from p. 189, where he talks vaguely about 'elsewhere'). The grammars used are those by Poutsma, Curme, Kruisinga, Jespersen, Kruisinga and Erades, Zandvoort, Hornby, and Scheurweghs, whose titles are presumably sufficiently familiar not to require repetition here. The verbs in the inventory are listed in alphabetical order, followed by details of the patterns used with each verb, information about frequency, sources, and 'limitations' (i.e. 'information on such collocating-possibilities as are thought relevant ... in cases where not all the occurrences have been quoted in full' (15)).

The inventory as such will undoubtedly be of value to the student in search of guidance on matters of usage. It will save him a great deal of time and irritability in that all points of relevance from the available handbooks are collated in one place and additional in-

formation provided in a compact and comprehensible way. I would have liked to see further discussion of limitations, particularly of a stylistic kind, but there is no denying the practical utility of the inventory as it stands. It is a good perspective for studying specific usage problems, and fully justifies the author's criticism of the way in which traditional grammars pay insufficient attention to normal usage, concentrating too much on unusual or exceptional constructions. For those who would like to get more from this data than pedagogic aid, however, the material is of much less use. Van Ek seems to want his book to be taken seriously at an academic, linguistic level as well as in pedagogy. He claims linguistic support at many places (quoting from Hill, Nelson Francis, Fries, and others); he uses incontrovertibly modern linguistic concepts (such as 'collocation'); he bases his verbal classification on Martin Joos; and so on. But as a contribution to linguistics, this book is, regretfully, a failure: it does not provide information which other linguists could readily use. If one wants to incorporate the inventory information into a grammar of some kind, or assess the implications of the author's points for syntactic theory (e.g. how to deal with the selectional restrictions described), or simply obtain a coherent linguistic picture of the whole of this corner of syntax, then one is liable to be disappointed. I found myself unable to make use of this inventory other than in a reference kind of way, and this curtailed its usefulness and interest substantially. The central reason for this is the lack of an adequate theoretical and methodological basis for the description.

Whether one agrees with generative grammar or not, the point made by Chomsky and Halle concerning the use of a corpus is undeniable (see *Journal of Linguistics*, 1 (1965), p. 103): 'Every linguistic description attempts, at least, to extract 'patterns' or 'regularities' from a corpus, or to abstract from it principles that will apply to other linguistic material as well. But statements of 'patterns', 'regularities', and 'underlying principles' go beyond the data. They are based on some assumption about the nature of linguistic patterns or regularities ... which [assumptions], unfortunately, are rarely made explicit'. Putting this another way, a grammatical analysis does not derive from a corpus; a corpus validates (or invalidates) an already partly formulated analysis. No-one approaches a corpus without *some* preconceptions as to the nature of the gram-

mathematical analysis to be carried out, and in any description, these preconceptions should be made explicit, the theoretical and descriptive categories defined and inter-related, and so on. Van Ek, however, seems to want to avoid this: 'I am not concerned here with theories of analysis but with phenomenological description' (18). This is an unrealistic dichotomy, as description presupposes analysis, and the relationship should not simply be ignored. Van Ek considers (18) that all the theory he needs for such a description is contained in his section 1.1; but this must surely be a strange sense of 'theory', as all one gets here is a list of three criteria used to establish his four patterns, and two descriptive labels – no theory (in the sense in which 'linguistic *theory*' is generally intended) at all.

Nor does the author wish to discuss his terminology, because 'the terms are relatively unimportant, they are mere labels and will serve as long as it is clear what they stand for' (18). Quite; but it is *not* always clear. For example, he does not define terms when 'everyone likely to consult this book is fully aware of what they stand for ... There is no point in defining terms used in an accepted sense in every book that happens to employ them in that sense' (4). This is true in principle, but the theoretical implications of one of the terms he cites at this point ('transitive') are by no means self-evident, and there are numerous other terms in the book which are not defined (and which presumably Van Ek thinks have accepted senses?), but which should definitely not be taken for granted, e.g. 'collocate' (one may have items collocating with categories as well as with other items, e.g. 188, 189), 'variant' (e.g. 193), 'categorical' (5). I *think* I know what I mean by 'anomalous finite' (5) and 'statistical bias' (187), though whether my definition of these terms would match Van Ek's is another matter; and I would like to hear a little more about the criteria lying behind the gradient implied between 'well-authenticated' (7), 'authenticated' (198) and 'unauthentic' (198) examples. The phrase 'lexical content' is certainly not being used in its normal linguistic sense: for Van Ek, the 'lexical content' of a sentence is the specific categories or items which co-occur in a surface structure, and has nothing to do with meaning at all.

There are also many cases of loose use of terminology, e.g. 'accusative' and 'nominative' are defined as 'elements of the patterns' (3–4), where surely they should be referred to as categories established in order to account for variation in elements of the patterns

(in this case, pronouns). Again, a case of ambiguity (7) is stated as being able to be 'reduced to a semantic problem' - as if there were ambiguities which could not. Certain fundamental concepts are in no way clear. The *headword* of a pattern, for example, is stated as belonging to 'a group of words which may be defined by listing', but 'This group is not a closed group and therefore an exhaustive list cannot be drawn up' (7). Van Ek tries to get around the contradiction by adding 'it may be expected that if a large enough corpus is examined the resulting list will be exhaustive to such an extent that only semantic variants of items present in it may qualify for further inclusion' - to which I can only ask, How does he know? What are the grounds on which this expectation is based? There are a number of controversial issues being taken for granted here. (To take just one, what is a 'semantic variant'? A synonym? In what sense can one say that of two words displaying the relationship of synonymy, that one is a variant of the other?) Another worrying point concerns the vague discussion in 1.2.2. about 'linguistic feeling', which evidently helps us to make decisions about word-classification: we are not told how this 'feeling' fits into any set of discovery procedures, nor are we told what the 'more satisfactory criteria' (5) are which can be referred to in order to clarify the issue.

Perhaps the major linguistic criticism concerns the status of the classificatory criteria Van Ek uses - in particular, whether notional or formal considerations are primary. This is not at all clear. His introductory definitions are formal, as are his remarks introducing his final classification (187), which he says is 'primarily' based on 'distributional criteria', though he is not going to lose sight of 'traditional notional classifications'. But it is not obvious what he means by such notional classifications: all he says is that they have a 'high mnemonic value' (187), which is (a) debatable; and (b) in any case the notion of 'value' must be ultimately assessed as to whether what is being memorised is linguistically viable. More serious are the implications of 'primarily' at this point. An unsympathetic reviewer would say that by this word Van Ek means that notional considerations can be allowed to outrank formal ones whenever this produces a nicer description. This is certainly what it looks like. For example, in his discussion of Pattern I (p. 188), he classifies the high frequency verbs which occurred in his corpus into three classes: 'a. verbs of physical perception: to hear, to see,

to watch. b. to help. c. to have, to let, to make'. The first two, he says, 'belong to the same distributional type in that their members (sic) are capable of taking a noun-object and collocate in pattern I with process-verbs. They differ in that only *to help* collocates with *to be* + non-verbal complement'. This is presumably a distributional classification (though the grounds for selecting the differentiating criterion are not explained, and cases where 'see' can take 'to be' - occasionally denoting physical perception, he says on p. 151 - are apparently ignored). Van Ek then goes on to discuss class c, saying that its members 'are incapable of taking a noun-object and collocate in pattern I with process-verbs *and* with status-verbs. In the corpus only *to let* collocates with *to be* + non-verbal complement'. Ah, the reader might ask, Why is *let* not made a separate class on the same grounds as *help* above? The answer, given in a footnote, is that 'notionally ... *to let* can be classed with *to have* and *to make*, whereas *to help* does not belong to verbs of physical perception. For the reason mentioned in the first paragraph of this section [presumably that notional classifications have a high mnemonic value? DC] I prefer to be a little inconsistent here rather than break up the semantic class *to have, to let, to make*'. Such ad-hocism is surely a bad thing. In any case, what is the basis of the so-called semantic class here (which, significantly, Van Ek does not label)?

There are other examples of the same kind in the final summary, and priority is given to notional criteria at other places in the book. Perhaps the most important case is his acceptance of Joos's distinction between status and process verbs, and his subclassification of this, which is based on semantic grounds. This should surely be justified, in view of the demonstrable weaknesses in Joos's general position: see, for example, Palmer's review of Joos in *Lingua* 18.2 (1967), 179-95. There is no absolutely clear distinction between the two categories of verb, as is evident when one considers how best to classify the senses of such polysemic verbs as 'plan', 'forget' and 'see'. But most of the semantic distinctions made by Van Ek are unclear to me (e.g. the distinction between 'mental' and 'physical' perception (151)).

Apart from these notional questions, there are other basic theoretical issues which are not mentioned, though they should have been. There is the whole question of the nature of selectional restrictions

in a grammar. Are Van Ek's collocations grammatical? Or are they lexical, in, say Halliday's sense? Is Van Ek's inventory best viewed as part of a grammar at all? Would it perhaps not be better to incorporate this information into the lexicon? How economical is a grammar involving so many one-member classes? (This area of investigation could have been illuminatingly examined in the light of the data Van Ek has collected.) Again, what justification have substitution frames as a procedure of grammatical analysis (cf. the author's favourable quote from Hill on this point, p. 8, but there is no reference to Quirk's strictures in *Archivum Linguisticum*, 10 (1958), 37-42, or Lees's in *Language*, 36 (1960), 207-21, for instance). Then there is his use of informants, which is highly suspect, but we are given insufficient information to be sure. Van Ek submitted a question (which of two paraphrases was a 'semantic variant' of a given sentence) to four English informants. Apart from the fact that four is hardly a reasonable sample for work of this kind, as recent psycholinguistic research has made clear (see, for example, R. Quirk & J. Svartvik, *Investigating Linguistic Acceptability*, The Hague, 1966), in order to assess the responses we should have been told a great deal more about the experiment, e.g. whether the informants had any linguistic training (if yes, then their judgements might be skewed by awareness of linguistically sophisticated points; if no, could they have understood the question)?, and how the question was presented (whether in spoken or written form). The four informants turn up again on p. 81, this time making tertiary judgements about English (*pace* Bloomfield). To be simplistic in this currently controversial area of linguistics is retrograde.

It seems to me that Van Ek has not established clearly the kind of audience he is writing for. How does he want the book to be judged? As a contribution to linguistics, or to pedagogy? Perhaps the most illuminating indication of this confusion comes towards the end (187): 'It [sc. Van Ek's final summary] is a very rough classification, but has the advantage of great simplicity. A more refined subdivision would need the setting up of a large number of categories and lead to unnecessary and undesirable complexity'. But why is complexity undesirable? Why are a large number of categories a bad thing? On what grounds does simplicity outrank comprehensiveness and accuracy (or whatever the positive antonyms of 'roughness' are)? It seems clear that pedagogic motives are intruding into a de-

scriptive linguistic exercise (cf. the 'mnemonic' justification for notional analysis mentioned above), and this produces an unsatisfactory book from the linguistic point of view. It is to be hoped that the theoretical difficulties will not make things too difficult for the intelligent student primarily concerned in obtaining guidance in usage as well.

DAVID CRYSTAL

*Department of Linguistic Science,
University of Reading, Reading, England*