Reviews

Claudia Claridge. *Multi-word verbs in Early Modern English: A corpus-based study*. Language and Computers: Studies in Practical Linguistics, 32. Amsterdam and Atlanta, GA, 2000. 317 pp. ISBN 90-420-0459-2. Reviewed by **Laurel J. Brinton**, University of British Columbia.

This book is a long-needed study of certain verbal collocations in English which – on the basis of semantic and syntactic similarities – Claudia Claridge groups together as 'multi-word verbs', including:

> phrasal verbs (*burn* N *down* [transitive]*, fly away* [intransitive]),
>
> prepositional verbs (*agree to* N),
>
> phrasal prepositional verbs (*look down on* N [transitive], *put* N *down to* N [ditransitive]),
>
> verb-adjective combinations (*make* N *known (to* N) [transitive], *hold good* [intransitive]), and
>
> verbo-nominal combinations (*take a walk* [Group I]*, take account of* [Group II]*, take into consideration, bring* N *to light* [Group III]).

Claridge's study is based on the *Lampeter Corpus* (see below), and for the most part presents a 'synchronic snap-shot' of multi-word verbs during the period covered by the corpus.

Chapter 2 contains a brief description of the *Lampeter Corpus of Early Modern English Tracts*, a collection of 120 non-literary prose texts (totaling 1,172,102 words) from the collection of the Founders' Library of the University of Wales, Lampeter. This corpus has been compiled by members of The REAL Centre, Technische Universität Chemnitz.[1] The period covered by the corpus – 1640–1740 – is considered to be an important period historically and linguistically, when standardization was in progress and 'prose style … seems to have undergone an extensive shift' (p 9). Claridge discusses some of the difficulties

involved in defining the categories used – Religion, Politics, Economy, Science, Law, and Miscellaneous – and assigning texts to them. Unlike other historical corpora, such as the *Helsinki Corpus, the Lampeter Corpus* includes entire texts. In most cases, some socio-economic facts concerning the authors are known, though Claridge was not able to make extensive use of this information. More importantly, because the *Lampeter Corpus,* like most other corpora, is not grammatically tagged, and because of the nature of the forms she was studying, Claridge was not able to search the corpus electronically, but had to read the texts and perform manual counts.

In Chapter 3, Claridge discusses difficulties of defining the category of 'multi-word verb'. Her definition must by necessity be quite broad: a 'multi-word verb' consists of two or more words (one of which must be a verb), expresses verbal meaning, allows for an alternative structural analysis, and is stable over time (pp 28–29). Her classification (pp 39–40), as shown above, is based fairly closely on Quirk et al's (1985), and apart from her omission 'for practical reasons' of prepositional verbs of the *turn* N *into* N type, seems intuitively sound. Claridge's justifications for grouping all of the forms into one super-category, 'multi-word verb', are the following: they are all characteristic of the analytic trend of English (eg, making use of prepositions and zero-derivation, allowing for thematic reordering in the face of fixed word order, showing 'semantic spreading'); they constitute a class of composite forms, like idioms, but differ from idioms in forming stable syntactic patterns and in the constituents retaining some vestige of individual meaning; they are lexicalized[2] forms, using their composite nature 'to produce some shifting, changing or enrichment of meaning' (p 43) and perhaps undergoing the same sorts of historical development; and they alternate with (quasi-)synonymous simplex alternatives in most cases. Though I believe that these structures share a number of similarities, the assortment of types does cause some difficulties for Claridge later in the study.

Chapter 4 presents a detailed discussion of the problems of delineating the individual types of 'multi-word verb'. Claridge is adept at reconciling the often conflicting evidence presented by other scholars, though she must frequently resort to the concept of a cline, or gradience, among categories. Phrasal verbs, in part because they have been comparatively well studied in the literature, provide the least difficulty, while prepositional verbs are perhaps most difficult to distinguish (from free combinations of verb plus prepositional phrase). She includes both literal and figurative phrasal and prepositional verbs and uses a combination of semantic/functional and syntactic tests to delineate these collocations. In the case of phrasal-prepositional verbs, Claridge excludes literal combinations since 'almost any phrasal verb can happen to be followed by a prepositional

phrase' (p 64)[3]; she allows only those combinations which are semantically coherent and in which no element is deletable (without a change in meaning). Verb-adjectives, which have received the least attention in the literature, consist of an adjective functioning syntactically as a particle but morpho-semantically an adjective. They are distinguished from complement structures (*paint the room blue)* because the adjective carries the major meaning, can precede a simple noun phrase, and carries 'generally as little modification as possible' (p 69). For the verbo-nominal category, Claridge includes any combination in which the meaning is centred on the noun, including not only deverbal nouns, but also other abstract (non-eventive) nouns such as *eyes* in *set eyes on*, which refers metonymically to the category of seeing[4]. Because of the difficulty in attaining native-speaker agreement, even in Present-day English, she does not rely on syntactic tests to distinguish this category.

Chapter 5 presents a review of the literature on the history of multi-verbs. Issues that arise out of previous work include the relation of phrasal and prepositional verbs to the loss of prefixes in Old and Middle English, the importance of the prepositional passive in determining the unified status of both collocations, the role of reanalysis and the importance of external influence in the development of these forms, evidence for a decline in multi-word verbs in the (prescriptivist) Early Modern English period, and the colloquial or non-colloquial status of multi-word verbs in the Modern English period. Only the last two issues receive treatment here.

Chapters 6–8 present the core of Claridge's findings concerning multi-word verbs in the *Lampeter Corpus*. Chapter 6 presents data on the individual types of multi-word verbs. Overall, phrasal verbs and verbo-nominal combinations were most common (2,200 and 1,704 tokens, respectively), prepositional verbs and verb-adjective combinations considerably less common (368 and 232 tokens, respectively), and phrasal-prepositional verbs quite rare (53 tokens), preventing much generalization about this last category. Claridge begins by looking at the verbal and non-verbal elements constituting multi-word verbs, and here the diverse nature of the 'multi-word verb' category is apparent: phrasal, prepositional, and phrasal-prepositional verbs exhibit a large variety of verbs and a small number of non-verbal elements, whereas verb-adjective and verbo-nominal combinations, where the verb is a functional element and the non-verbal part is an open class, show the reverse distribution. Generally, phrasal verbs consist of monosyllabic, native verbs and prepositional verbs of borrowed verbs; Romance nouns – primarily without a preceding article – predominate in verbo-nominal combinations. The all-purpose verbs *make, take*, and *give* are the most frequent verbs in verbo-nominal combinations and are common in the other

types; *have* is unique to verbo-nominal combinations. Where possible, Claridge compares results from the *Lampeter Corpus* with results from Present-day English (using the *Lancaster-Oslo/Bergen Corpus* and the *British National Corpus*), finding, for example, that there is some difference in the inventory of particles in phrasal verbs. A section on 'semantic patterns' is generally uninformative, except to show that most multi-word verbs are literal in meaning. In a much larger section on syntactic patterns, Claridge looks at transitivity, the position of the object, modifying elements, the passive, preposition stranding, coordination, and miscellaneous features. Some interesting syntactic findings include the low rate of intransitive constructions (Claridge sees the 'blurring' of the transitive/ intransitive distinction as 'one of the primary purposes' of multi-word verbs, p 148) and the 'surprisingly low' (p 157) occurrence of adjectival modification in verbo-nominal combinations (this being one of the purported reasons for choosing the verbo-nominal combination over the simplex), but overall she finds little difference between Early Modern English and Modern English in respect to syntactic behavior.

Chapter 7 attempts to place the raw data of the previous chapter in context, by examining change over time and the influence of register and socio-economic class. Gross comparisons show phrasal verbs to be less common in Early Modern English than in Present-Day English and verbo-nominal combinations to be about equally common. Divided into decades, the data of the *Lampeter Corpus* do not indicate a unidirectional development, though there appears to be a small decline in the frequency of phrasal verbs. Claridge also looks at the type/ token ratio, the number of single instances, and the number of types per corpus decade as signs of productivity; while the evidence is not overwhelming, multiword verbs seem to be moderately productive. An examination of register seems to show that phrasal verbs abound in the more colloquial register and prepositional verbs (and perhaps verbo-nominal combinations) in the more formal register. In respect to gender Claridge finds that women (there are only two in the corpus) use fewer phrasal and preposition verbs – an aspiration toward the norm? – but in respect to socio-econonic class, she finds no clear pattern. All in all, I found the results of this chapter somewhat disappointing, perhaps because the historical and socio-economic spread of the *Lampeter Corpus* is too confined to permit diachronic and (comparative) sociolinguistic study.

Chapter 8 presents an interesting study of contemporary awareness of and attitudes towards multi-word verbs. Claridge finds a growing awareness of phrasal and prepositional verbs during the period, though not of verb-adjective and verbo-nominal combinations. While attitudes are expressed toward individual forms, there are no clear-cut or systematic attitudes toward multi-word verbs

expressed. She states, moreover, that 'as both awareness and attitudes have been found to be severely under-developed, the majority of choices taken with regard to multi-word verb usage will have to be regarded as not very conscious or elaborate decisions on the basis of the individual instance' (p 220; but cf Chapter 9). The final section of this chapter treats the alternation between the multi-word verb and the Romance verb, but the figures presented seem rather ad hoc and are ultimately unconvincing.

In Chapter 9, Claridge investigates whether motivations that have been suggested in the literature for the choice of multi-word verbs rather than simplex verbs are borne out by the data in the *Lampeter Corpus*. Semantically, she finds that multi-word verbs, because they exhibit 'semantic spreading', or the distribution of meaning of the whole onto individual elements, do offer the possibility of making semantic distinctions not possible with the simplex. They may be more precise than the simplex and may express meaning in addition to that expressed by the simplex, as in the case of phrasal verbs such as *blow down*, where the verb expresses the manner and the particle the result, viz 'down by blowing'. In such cases, the multi-word verb frequently has a transitivizing function Claridge also finds evidence for the expressive potential of multi-word verbs, even in cases in which a particle or preposition seems redundant by modern standards, eg, *assemble together* or *return back*. In respect to the purported *Aktionsart* meaning of the particles, Claridge finds that telic meaning is quite common with *off, up, out, down* (eg, only 11% of phrasal verb tokens with *off* do not express telic meaning), though durative/iterative meaning with *along, away, on* is quite rare, perhaps because of the low frequency of progressives during this period. In verbo-nominal combinations, the indefinite article, which makes the *Aktionsart* distinction in Present-Day English (between, eg, *walk* [an activity] and *take a walk* [an accomplishment]), is typically absent in the *Lampeter Corpus*. For most zero-article items, there is no discernible *Aktionsart* difference between the verbo-nominal combination and the simplex (between, eg, *give chase* and *chase*), except in cases in which the noun is pluralized (eg *give shouts* [activity composed of incremental parts] vs *shout*).

In addition to semantic considerations, other stylistic and rhetorical motivations may account for the choice of multi-word verb over simplices. Claridge finds that multi-word verbs incorporating native elements can be used to introduce loanwords or make a text more accessible and 'audience-oriented', that multi-word verbs can be more multifunctional and flexible than simplices, that the relatively predictable effect of multi-word verbs can facilitate the creation of new forms, that multi-word verbs can allow for the spread of stress across the sentence in different ways, that multi-word verbs can provide variation in gen-

eral, and that multi-word verbs, specifically verbo-nominal combinations, can contribute to 'nominal style', though without necessarily increasing the static or stative nature of the discourse.

The final section of the chapter treats the syntactic possibilities afforded by the 'syntactic spreading' of the multi-word verb. One aspect of syntactic spreading is that elements not otherwise permitted in these positions may be end- or front-focused. A second aspect is the modification of elements not otherwise or not easily modified; however, Claridge finds that such modification in verbo-nominal combinations in the *Lampeter Corpus* is less than expected, with the exception of preference for nominal negation (*give no answer* rather than *do not give an answer*). Other syntactic uses of multi-word verbs include an intransitivizing function (*make a discovery* rather than *discover X*), a passive use of the active form (*take warning*), and internal coordination (*set and keep open, preach or pray down, give alarm and outcry*). Finally, Claridge notes that the different types of multi-word verbs provide a range of semantic, stylistic, and syntactic flexibility, with prepositional verbs being the most fixed and verbo-nominal combinations the most free.

Chapter 10 constitutes a brief conclusion to the book. Five appendices present a complete listing of all of the types of multi-word verbs found in the *Lampeter Corpus*.

Overall, the book is well written and well edited, though some minor errors result from the failure to delete words during the process of editing. Some awkward sentences or sentences that are difficult to process result from the author's being a non-native speaker of English, as do, perhaps, some small usage problems, such as the consistently incorrect use of *amount* for *number*. I found only two more serious errors: one table is inadvertently split (pp 185–86) and one paragraph is repeated (p 189).

Claridge's work represents an important contribution to the field. The class of 'multi-word verbs' has, for the most part, been severely understudied, especially in older stages of the language (cf, however, Brinton and Akimoto 1999). This study makes very careful and detailed use of an interesting historical corpus, though the insights gained are more synchronic than diachronic. I cannot say that this work makes a significant advance in computerized corpus linguistics, since the nature of the forms studied required a manual rather than a computerized count. (Claridge does not explain why she chose to study these forms rather than some other forms more amenable to the limitations of a grammatically untagged corpus.) Nonetheless, I recommend this book highly to those interested in verbal collocations in English, and in Early Modern English generally.

Notes

1. Further information concerning the *Lampeter Corpus* can be found at http://www.tu-chemnitz.de/phil/english/real/lampeter/lamphome.htm.
2. Claridge uses the term *lexicalization* at a number of points during the book, but does not, I think clearly define it.
3. The same could be said of any simplex verb, but Claridge does not exclude literal prepositional verbs for this reason.
4. I am uneasy about the inclusion of non-eventive, abstract nouns with no morphological relation to verbs since doing so could, potentially, expand the scope of verbo-nominal category beyond reasonable bounds; for example, would combinations such as *have a flair for, make the most of, put in jeopardy,* or *take to court* be included in the category? Furthermore, Claridge rejects combinations with *be* in both verb-adjective and verbo-nominal combinations except in cases where *be* 'clearly is not present in the semantic structure', as in *be at a loss* (p 81). The distinction is not clear to me.

References

Brinton, Laurel J. and Minoji Akimoto. 1999. *Collocational and idiomatic aspects of composite predicates in the history of English* (Studies in Language Companion Series, 47). Amsterdam and Philadelphia: John Benjamins.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.

Elena Tognini-Bonelli. *Corpus linguistics at work*. Amsterdam – Philadelphia: John Benjamins, 2001. Studies in Corpus Linguistics Series. xii + 223 pp. ISBN 90-272-2276-2. Reviewed by **Inge de Mönnink**, University of Nijmegen.

Corpus linguistics at work is a mixture between a theoretical book on the main issues in corpus linguistics (CL) and a practical introduction to corpus work. The central issue in the theoretical part of the book is a very strictly defined distinction between the corpus-based approach and the corpus-driven approach. In the corpus-based approach, corpus data are used to validate and quantify linguistic theory/description. It is described as 'a methodology that avails itself of

the corpus mainly to expound, test or exemplify theories and descriptions that were formulated before large corpora became available' (p 65). In the corpus-driven approach, a theory is built up step by step in the light of corpus evidence. The observation of certain patterns leads to a hypothesis, which in turn leads to a generalisation in terms of rules of usage and finally finds unification in a theoretical statement.[1] In the practical part of this book, the corpus-driven approach is exemplified in applications in language teaching and translation.

In the first chapter, the issues to be discussed in the rest of the book are introduced. The author very briefly touches upon important issues such as CL as a theory versus CL as a methodology; the definition of a corpus; textual data versus corpus data; the relevance of the Firthian framework for CL; and the role of technology in CL. From the discussion of the corpora and techniques used in the book, it becomes clear that the book focuses on the exploration of raw corpora and thus on lexico-semantical and lexico-grammatical studies.

Chapter 2 discusses the way corpora have influenced language teaching. The mismatch that is often felt to exist between theory and linguistic fact is nicely illustrated by an example of the use of a*ny*. 'This type of finding points to the fact that a lot of the mismatch between traditional descriptions and actual usage stems from the fact that the strict interconnection between an item and its environment is more or less ignored.' (pp 17–18) The chapter stresses the correlation between a lexical pattern and a grammatical structure and exemplifies the advantage of using corpora in both vocabulary teaching and grammar teaching, using concordances, collocations and colligations. The chapter is especially interesting for teachers who are new to the field of CL, because of the many concrete examples it contains.

Chapter 3 gives the necessary background to CL, its history, the definition of a corpus, the problem of representativeness, etc. While the discussion does not raise new points, it covers most issues that play an important role in modern CL.

Chapter 4 explains what is meant by 'corpus-based' in this book. The corpus-based approach is described as a position that does not fully accept the direct relevance of actual usage to theoretical statement. The chapter discusses three ways in which a corpus-based linguist can deal with the fact that the data cannot always be described by the theory:

1. Insulation: 'In this approach the data is relegated to a secondary position with respect to the theoretical statement proper.' (p 68) The theory is tested on corpus data, in the form of a grammar-based parser, but at the same time the theory is said to be insulated from the evidence of the corpus. 'The corpus is considered useful because, on occasions, it indicates where minor

corrections and adjustments can be made to the model adopted and, of course, it can also be valuable as a source of quantitative evidence.' (p 66)

2. Simplification/standardisation: This approach is similar to Insulation. The only difference is that the theory tested on the corpus data in this approach is (for some obscure reason) said to be more empirical and data-oriented than the grammars tested in the Insulation approach.

3. Instantiation: In this approach the data is built 'into a system of abstract possibilities, a set of paradigmatic choices available at any point in the text' (p 74). In other words, a pre-existing system is extended with probabilistic data extracted from corpora, but the system is not affected as such.

On pages 72–74 the author clearly rejects the use of annotated corpora for linguistic research. Annotation draws the researcher's attention away from the words proper and thus blurs the tight interconnection between lexical and grammatical patterns. 'Perhaps the most obvious point against annotation, though, is the fact [that] the categories of analysis are provided by the linguist, and these categories, at the outset of a study anyway, have not themselves been derived from corpus data. True, they may be modified by confrontation with corpus evidence, but there seems to be an implication that any modification will be of a minor nature' (p 74).

Chapter 5 discusses the corpus-driven approach. The corpus-driven approach follows a clear methodological path: observation > hypothesis > generalisation > unification in theoretical statement. All four steps are influenced by the intuition of the researcher. His/her intuition is based on experience and knowledge of theory (Firth). Observation should be understood as observation of language data as contained in corpora. The corpus data are accessible through concordances. A concordance displays two types of repeated events: (1) collocation, the recurrent co-occurrence of words, and (2) colligation, the grammatical patterning in which the word is embedded. Colligation is only recognisable through the use of syntactically annotated corpora. However, according to the author, this type of data is not satisfactory, because 'grammatical parsing on its own is not sufficient to account for the crucial evidence in many cases, and unless lexical constraints are built into the picture, the grammatical categories adopted will lack generalisability and replicability. Automatic annotation will, at best, leave some questions unanswered, while manual or semi-manual annotation will often end up stretching the evidence to fit the categories' (p 90). In other words, we have to wait for corpora that have been annotated by means of a (formalised) description of language use as obtained from the corpus-driven approach before we can extract colligational information from corpora. How-

ever, we need colligational information to obtain such a description. This seems a vicious circle. We need information that we cannot get unless we already have that information. What is the way out of this dilemma? This question is not answered in the book. One way out could be to annotate corpora by means of a formalisation of an existing description (which after all is also based on observation and intuition) and to observe where this description (or theory) does not describe the data satisfactorily. However, in this book such an approach would be considered as Insulation and is thus not desirable.

Chapter 6 introduces co-selection. Co-selection is the phenomenon that the formal co-textual features surrounding a word determine its meaning and its function in a specific discourse. In other words, an item and its environment cannot be separated. The chapter provides examples of how the corpus can uncover extended units of meaning, semantic prosody, delexicalisation, and the ideological weight of words.

Chapter 7 discusses the application of the corpus-driven approach to translation and contrastive linguistics. The chapter gives several examples of the use of comparable corpora of English and Italian for the identification of translation equivalence. Starting form the word or pattern to be translated, the first step is to establish its specific meaning/function(s) in context using a corpus of L1. The next step is to identify *prima facie* a translation equivalent for each meaning/ function using a dictionary, a grammar, a translation corpus (if at one's disposal), and/or past experience. Step 3 then establishes the formal realisation of the translation equivalent in L2 using a (comparable) corpus of L2.

Chapter 8 discusses Firth's theory of meaning, which is said to be the basis of corpus-driven linguistics. The chapter exemplifies the view of linguistics that is fundamental to this book: the main concern of descriptive linguistics is to make statements about meaning, and meaning *can* be stated in linguistic terms. The chapter also considers in detail Firth's notions collocation and colligation, which are used throughout the book. Considering the fundamental role of this chapter, I would have preferred it earlier in the book, immediately after the introduction.

Chapter 9 sketches the history of the study of meaning. While it contains an interesting overview of landmarks in the study of meaning, it adds little to the book at hand.

Chapter 10 discusses the basic requirements of corpus-driven linguistics and summarizes the previous chapters.

In this book it is implied that researchers take either the corpus-based approach or the corpus-driven approach, that these approaches are fundamentally different, and that there is no middle course. What is more, it is implied that

the corpus-driven approach is to be preferred over the corpus-based approach, and that researchers following the corpus-based approach are so committed to their theories that they intentionally ignore the facts provided by the language (corpus) data. Although such a coarse representation of corpus linguistic practice may help to clarify the difference between the corpus-based and the corpus-driven approach as sketched by the author, it hardly does justice to reality. It seems to me that for lexical, lexico-semantic and lexico-grammatical studies, the corpus-driven approach described in this book is tenable and even advisable. For researchers and teachers interested in semantic information, this book is an interesting introduction. However, syntactic phenomena are not always accessible through lexical information. At best, colligational information from concordances can be used. But even there, as the author herself observes, we lack corpora that have been appropriately annotated to realize the corpus-driven approach described here. It seems to me that for this type of research the corpus-based approach is justifiable, and indeed the only way forward. I consider the total denial of corpus-based research and research based on annotated corpora to be a major drawback of *Corpus Linguistics at Work*.

Inge de Mönnink. *On the move. The mobility of constituents in the English noun phrase: A multi-method approach.* Amsterdam – Atlanta, GA: Rodopi, 2000. xii + 188 pp. ISBN 90-420-0780-X. Reviewed by **Joybrato Mukherjee**, University of Bonn.

As the title implies, it is not only the corpus-based description of positional mobility in the English noun phrase (NP) that lies at the heart of this book. The author also breaks new ground in pursuing a multi-method approach, ie in combining corpus-linguistic methods with the analysis of intuition-based elicitation data. Thus, both the object of inquiry and the methodology will be discussed in the present review.

The book is organised as follows. Chapter 1 provides an overview of some previous approaches to the English NP. In Chapter 2, de Mönnink gives reasons for the usefulness and viability of the multi-method approach and describes both the corpus analysis and the elicitation experiment. The results of the corpus analysis are then presented and discussed in Chapter 3. In Chapter 4, the corpus data are complemented with the results obtained from the elicitation experiment. In Chapter 5, de Mönnink sets out to explain her findings by drawing on formal

and functional principles which may cause NP constituents to be moved to the left or to the right. Finally, she summarises and evaluates the linguistic evidence and methodology and sketches some prospects for future research (Chapter 6). The appendix comprises detailed information on the corpus design and analysis, the elicitation experiment and abbreviations.

In the opening chapter, the author refers to the description of NPs in classic reference grammars as well as in structuralist, functional and generative grammars. The bottom line is that all those frameworks agree and focus on what de Mönnink calls the 'prototypical NP structure' (p 19): (limiter as adverb phrase) – (determiner as determiner phrase) – (premodifier* as adjective/adverb/noun phrase) – head as noun/pronoun/proform – (postmodifier* as prepositional phrase or clause). Brackets indicate optionality, asterisks mark possible multiple realisations. However, the author is particularly interested in NPs which do not conform to the prototypical NP structure. In such variant NPs, any constituent may be moved to the left (ie 'fronted'), to the right (ie 'deferred') or even be placed outside its mother constituent (ie 'floating', either fronted or deferred). This systematisation makes it possible to identify nine different types of variant NPs, for example 'NPs with a deferred determiner' as in 'Make it quite clear to us all' (p 27). This chapter thus provides a solid and plausible theoretical foundation for the practical analysis of variant NPs.

The second chapter centres around methodological considerations, which are both enlightening and thought-provoking. In particular, de Mönnink points out that a merely corpus-linguistic analysis of variant NPs has two major disadvantages. Firstly, no corpus – however large it may be – will ever cover all possible NP structures. Secondly, it remains unclear whether constructions not occurring in a corpus are in fact ungrammatical/unacceptable, that is whether it is due to the corpus size or to the language as a whole that a certain NP structure is not attested. Therefore, de Mönnink envisages a methodology which makes use of both authentic corpus data and intuition-based judgments of NP structures which are not found in the corpus. To this end, she develops an innovative 'data cycle for descriptive linguistics' (p 34), which is intended to reconcile corpus-based methodology with intuitive data. The cyclical procedure comprises four steps: (1) formulation of initial hypotheses (eg as a result of previous studies and the linguist's own intuition); (2) in-depth analysis of corpus data; (3) confirmation, refinement or revision of the initial hypotheses; (4) elicitation experiments which take into account intuitive data. Having come full circle, those intuitive data may again lead to a reassessment of previous hypotheses. It is obvious that this multi-method approach not only proves to be relevant to the object of inquiry at hand, but is certainly applicable to the analysis of virtually all syntac-

tic phenomena. In a wider setting, de Mönnink argues forcefully and convincingly that corpus linguists should not confine themselves to corpus data alone, but that they should accept the relevance (and inevitability) of intuition. As things stand, corpora may reveal what is probable in language use, but intuition-based judgments can tell us more about what is possible and what is not. The remainder of this chapter is devoted to the detailed description of the two corpora (comprising 170,000 words and 220,000 words respectively) and the design of the elicitation experiments in which 130 native speakers took part.

In the third chapter, the author describes and carefully analyses the data obtained from the corpora. Many authentic corpus examples are discussed in this context. Lack of space forbids a detailed review of de Mönnink's results here. Generally speaking, variant NPs are not a marginal phenomenon since, for example, they represent 5.3 per cent of all complex NPs in the 170,000-word corpus. However, quite a few types of variant NPs are not at all or only sporadically attested in the corpus. For instance, the NP type 'deferred modifier + discontinuous determiner' occurs only twice in the genre of scripted speech. Obviously, such low frequencies pose serious problems for the reliability of statistical analyses such as the chi-square test. Furthermore, NPs with a fronted premodifier exclusively function as subject complements, direct objects and prepositional complements in the corpus. It is here that the multi-method approach turns out to be a promising alternative to a merely corpus-based approach. On the basis of the corpus data, de Mönnink puts forward some hypotheses as to the acceptability of three specific types of variant NPs: the fronted premodifier, the discontinuous adjective phrase and the floating deferred modifier. Among those hypotheses is the following one: 'An NP with a fronted premodifier cannot function as indirect object' (p 83). These corpus-based hypotheses provide the reference points for the elicitation experiment.

In Chapter 4, the author starts off by describing in detail the experimental procedure (including problems encountered while conducting them). It turns out that the elicitation data run counter to corpus-based expectations. For example, NPs with a fronted premodifier are acceptable in indirect object position according to most native-speaker informants: 'You should not tell so jealous a boy that you met his girlfriend in the pub' is generally not considered ungrammatical (p 94). In a similar vein, all other corpus-based hypotheses are put to the test in the elicitation experiment. In general, the data obtained from native speakers' intuitive judgments lead the author to extend the prototypical NP structure step by step. Eventually, she is able to offer a fairly complete description of the structural range of the three variant NPs under scrutiny. De Mönnink also outlines some prospects for a continuation of the data cycle by looking at larger corpora

such as the 100,000,000-word *British National Corpus* (BNC). To pick out one example, the BNC data on variant NPs with a premodifier make the author 'include a prototypical premodifier position which can be realized by a classifying NP or AJP', as for example in 'I had so romantic a long evening yesterday' (p 114). De Mönnink's data cycle thus represents a genuinely heuristic method of linguistic analysis which has no clear-cut point of completion.

In Chapter 5, the author addresses the question as to why variant NPs with fronted, deferred or discontinuous constituents occur in the first place. To this end, she discusses a wide range of theoretical models intended to explain the mobility of constituents in general. De Mönnink's highly perceptive overview includes generative movement rules as well as functionalist and cognitive approaches. All the models are applied to examples of variant NPs obtained from the corpus material and the elicitation experiment, allowing for a critical assessment of the explanatory power of the models. Generally speaking, both pragmatic and syntactic principles turn out to be relevant to a comprehensive theory of mobility.

At the beginning of the last chapter, de Mönnink recaps the two main goals of her study: 'The first goal was to investigate the mobility of the constituents in the English noun phrase and gain insight into the nature and frequency of variant NPs. The second goal, which follows from the first, was to develop a multi-method approach to descriptive studies that combines corpus data and experimental data' (p 147). In my view, both of these goals are achieved. In fact, this book provides a careful and considered analysis of variant NPs and, perhaps more importantly, breaks new ground in corpus-linguistic methodology. It is certainly true that corpus linguistics has become mainstream over the last years. This in itself is no doubt an agreeable development. Unfortunately, too many corpus linguists attempt to ignore intuition altogether, although no corpus covers all that is possible in language use. In this context, de Mönnink's multi-method approach represents a balanced trade-off between the empirically sound observation of large amounts of authentic data and the systematic consideration of native speakers' intuition. That the combination of corpus data and elicitation data is in fact useful for the purpose at hand is vindicated by the quantity and quality of interesting data the author obtains and by the conclusions she is able to draw from their analysis. It should be added that the book is also a pleasure to read because of its easily accessible and truly reader-friendly style. De Mönnink makes use of many discourse organisers so that the reader is never allowed to lose track of the main line of argumentation. Furthermore, she uses many tables and diagrams, which helps the reader to process the wealth of data. The book is

also well written, and apart from very few (and minor) errors the proof-reading has been thorough.

On the Move definitely makes for stimulating reading, not only for corpus linguists, I feel, but for all syntacticians interested in structural variation at the level of NPs. Both corpus linguists and those following intuition-based methods (say, in the generative camp) will certainly profit from closely perusing this book in order to get the best of both worlds. No-one interested in an unbiased approach to linguistic description in general and/or the English NP in particular should miss this excellent study.

Note

1. In this book, 'corpus-driven' is not meant to include some extreme forms of corpus-driven research in which statistical techniques are used to derive/ extract language models from corpora.

Marianne Hundt. *New Zealand English grammar – fact or fiction? A corpus-based study in morphosyntactic variation*. Varieties of English around the World 23. Amsterdam and Philadelphia: John Benjamins, 1998. xiv + 212 pp. ISBN 90–272–4881–8. Reviewed by **Erik Smitterberg**, Uppsala University.

As Marianne Hundt points out, New Zealand English (henceforth NZE) has received less scholarly attention than many other regional varieties of English; Hundt's study is therefore an important contribution to research on varieties of Present-Day English. The main question that Hundt sets out to investigate is whether NZE can be considered a separate national standard. In order to answer this question, she conducts separate analyses of regional variation in the use of linguistic features from the fields of morphology (eg irregular vs regular verb morphology and adjectival comparison), syntax (eg modals and collective nouns), and lexico-grammar (eg uses of the verbs *farewell* and *screen*). Given the large number of features investigated, my account of Hundt's analyses will be selective in the present review, especially as regards the syntactic features. In what follows, I will examine the individual chapters in the order they occur in the study. The review will conclude with an overall assessment of Hundt's study.

The first chapter, 'Introduction', surveys previous research on NZE. It is pointed out that NZE has been recognized as a variety but only codified as regards its lexicon, while American English (AmE) and British English (BrE)

have also been codified with respect to grammar. (In addition to vocabulary, research carried out on NZE has also focused on phonology.) The relation of NZE to other varieties of English is discussed, and Hundt states that one way in which NZE may emerge as a linguistically separate variety is by its containing a unique mix of features found in the other national varieties. She also lists other possible types of differences between NZE and these varieties, and hypothesizes that most genuine NZE features, if any, will 'be found at the interface of grammar and the lexicon' (p 5).

Chapter 2, 'Theoretical and Methodological Foundations', is devoted to the issues of what constitutes a linguistic standard, and how standards relate to one another; corpus linguistics and variationist frameworks are also discussed. The last topic, though clearly relevant to the study, is somewhat unexpected, as nothing is said about it in the introduction to the chapter (it is, however, mentioned in the 'Aim and Scope' section in Chapter 1). International English is seen as a pluricentric language with a number of interacting national varieties (which need not, however, have equal power to influence the language as a whole). The discussion is well laid out and easy to follow, but some terms, such as 'pluricentric', 'exo-normative', and 'endo-normative' could have been elaborated on for the benefit of readers who do not specialize in the study of national varieties of languages; instead, references are given to works where these terms are discussed.

The material on which the study is based is also introduced in Chapter 2. Most investigations in Hundt's study are corpus-based: data yielded by the Wellington Corpus of Written New Zealand English (WCNZE), the Macquarie Corpus of Australian English (ACE) and the LOB, FLOB, Brown, and Frown corpora are analysed and compared.[1] In addition, data from the spoken Wellington Corpus of New Zealand English (WCSNZE), the spoken component of the British National Corpus (BNC), and the newspapers the *Guardian* (BrE), the *Miami Herald* (AmE), the *Dominion* (NZE), and the *Evening Post* (NZE) are included in some analyses. The study focuses on press language, although some features are investigated from a broader genre perspective. For a number of features, the corpus-based methodology is supplemented with elicitation tests in which speakers of AmE were asked to identify typical BrE features, and NZE speakers typical AmE features, from a list of fifteen sentences. Since Hundt otherwise does not start out from the hypothesis that NZE should be closer to BrE than to AmE, it is not clear why NZE speakers were not also asked to single out typical BrE features. Nevertheless, such additional data are very welcome: in this way, empirical and phenomenological evidence can complement each other, thus providing a fuller picture of regional variation in the use of the features under scru-

tiny. A brief discussion of statistical and linguistic significance is also included in this section. A few statements in this discussion are not clear to me. For instance, it is claimed that statistically significant differences may 'be due to the accumulation of haphazard phenomena, i.e. fluctuation caused by free variation' (p 26); in my opinion, it is the researcher's task to choose a significance level that excludes such variation from the significant results.[2]

In Chapter 3, 'Morphology', Hundt begins to present her results. The Summary aside, the chapter is divided into three main sections: irregular verbs, comparison of adjectives, and the *s*-genitive. In the first section, she shows that NZE, like Australian English (AusE), is more conservative than BrE and, in particular, AmE in the process of change from irregular to regular preterites and past participles of verbs like *burn* (ie from *burnt* to *burned*). As regards the past participle of *prove*, which is changing in the opposite direction (ie from *proved* to *proven*), however, NZE takes an intermediate position with regard to BrE and AmE, the latter being the most advanced variety in this development. The past participle *gotten* (for *got*), which occurs in AmE, does not seem to be part of written NZE. As regards the comparison of adjectives, no variety displays a tendency towards periphrastic (*more*/*most*) comparison of monosyllabic adjectives.[3] Concerning the *s*-genitive, differences in diachrony (as evidenced by a comparison of LOB and Brown with FLOB and Frown) appear to be more pronounced than regional variation, although there are some indications that *s*-genitives are used slightly more in AmE than in BrE, with AusE and NZE exhibiting even lower frequencies. (The discussion is based on the frequency of the *s*-genitive itself rather than its variation with the *of*-construction.) In sum, as regards the morphological parameters where clear regional differentiation was attested, AmE tends to be in the forefront of what seem to be developments in diachrony, while the relative position of NZE, AusE, and BrE varies.

The fourth chapter, 'Syntax', investigates regional differences in the distribution of syntactic variables. The verb *have* is analysed as it occurs universally in English, in the construction *have (got) to*, and as a full verb. Examples of the different patterns in which *have* occurs would have been welcome for the benefit of non-specialist readers. Although not all differences are statistically significant, the universal frequency of *have* appears to be high in NZE, while the verb is comparatively rare in AmE. The construction *have (got) to*, in contrast, is most frequent in BrE, followed by NZE, AmE, and AusE. As regards *have* as a full verb, the other regional varieties appear to be following AmE in preferring *do*-support for full-verb *have*. Concerning the distribution of *shall* and *will*, NZE seems to avoid *shall* to an even greater degree than AmE and AusE. (However, some differences were not significant, and evidence from the elicitation test

showed that many NZE speakers felt sentences like *Will I get you an aspirin?* to be characteristic of AmE.) The increasing use of the progressive is an area where NZE and AusE are more advanced than at least AmE: the construction is significantly more frequent in WCNZE and ACE than in Frown. However, as Hundt points out, significance was calculated without taking, for instance, the proportion of finite to non-finite verb phrases into account. In a study of collective nouns, she shows that, with respect to verbal concord, AmE clearly prefers singular concord (eg *the team was*), while plural concord (eg *the team were*) is still common in BrE, with NZE in between. Similarly, Hundt's investigation of the mandative subjunctive vs the *should*-construction indicates that NZE occupies a middle position relative to AmE, where the subjunctive dominates, and BrE, where the *should*-construction is still quite frequent. Many of the features investigated in Chapter 4, such as the progressive, are undergoing change, something which, Hundt points out, may or may not lead to regional differentiation. In sum, no consistent pattern emerges from the analyses of the syntactic variables[4]; instead, NZE displays a feature-dependent set of similarities with and differences from BrE and AmE.

The fifth chapter, 'Lexico-Grammar', is devoted to analyses of lexical items that may exhibit variation in patterns of use, complementation, etc. The analysis of prepositions following *different* shows (a) that *from* is preferred in all varieties, (b) that *than* appears to be more accepted in AmE, and (c) that *to* may be slightly more frequent in NZE than in BrE. Another investigation indicates that the verbs *protest* and *appeal* tend to be transitive in AmE but intransitive in BrE, with NZE forming a middle ground, although the figures indicate greater similarity to BrE; stylistic differences may also enter into the discussion in this context. The use of *farewell* as a transitive verb is shown to be typical of AusE and NZE. The medio-passive use of *screen* in the sense 'appear on a cinema/television screen' seems to be a feature characteristic of NZE, though it is also found in AusE and AmE; the elicitation test showed that more AmE speakers than NZE speakers thought that this use of *screen* was characteristic of a variety other than their own. As regards the patterns *look to* + infinitive, *look* + object, and *look like* + present participle, NZE is closer to BrE than to AmE.

Chapter 6, 'Statistical Significance and Linguistic Relevance', elaborates on the short discussion of statistical and linguistic significance in Chapter 2. Case studies of three forms – *because*, *would*, and *not* – whose frequency differs significantly among the corpora are presented. The studies show that it may be difficult to find linguistically significant reasons behind the statistically significant differences attested. Conversely, Hundt also claims that variation can be interesting from a linguistic point of view even when it is not statistically significant.

In the seventh and concluding chapter, the results of the study are discussed in the light of previous research and hypotheses formulated in the preceding chapters. The first section addresses the problem of separating regional variation from diachronic change in a situation where the increased dominance of AmE variants is leading to linguistic convergence, and where synchronic, regional differences may constitute different stages of ongoing linguistic changes. The second section is devoted to regional and stylistic variation. The relative positions of the national varieties concerning some of the features investigated are neatly summarized in a figure; it is pointed out that genres may be differently suitable for a description of an emerging national standard. In the third section, empirical/statistical data (significant differences among frequencies) are related to phenomenological data (elicitation tests) in a discussion of different types of linguistic regionalisms. The following section establishes NZE as a separate entity on the basis of the results reached, which, taken together, separate NZE from all other varieties investigated, with the possible exception of AusE: Hundt (p 139) concludes that her study 'has only produced meagre evidence on differences between NZE and AusE'. The penultimate section of the chapter discusses the representativeness of the text samples and consequences for the validity of the results. In the final section, Hundt returns to differences between corpus-based findings and the results of her elicitation tests. She also justifies her study by mentioning four ways in which descriptions of national varieties like NZE are valuable. First, studies such as Hundt's can help teachers choose teaching models. Secondly, NZE may develop into a more influential national variety in the future. Thirdly, even if the differences between NZE and other national varieties of English are relatively minor, they may be psychologically important. Fourthly, the study shows that, in the area of grammar, NZE is not set apart from other national varieties solely by colloquial features.

In sum, Hundt's study is a highly valuable addition to our knowledge of regional varieties of English. By using corpora as well as elicitation tests, Hundt succeeds in comparing perceived and actual usage. In addition, the broad scope of the study, in terms of the large number of features investigated, is truly impressive. There are, nevertheless, a few minor drawbacks. First of all, I felt that the title of the study did not cover the entire field investigated. It is true that the study focuses on NZE and its relation to other regional varieties in terms of the distribution of morphological and syntactic features, as attested in corpora. However, a substantial part of the work is devoted to discussing the distribution of these features in AmE, AusE, and BrE as well, a discussion which can be appreciated independently of the comparisons with NZE; semi-lexical features are also treated; and scholars who are interested in elicitation tests would not

know from the title that such data were used to supplement some of the corpus-based analyses. Secondly, in some cases a fuller discussion of why some linguistic features but not others were selected for investigation would have been appreciated (although the broad scope of the study justifies some brevity in the description of individual features). Lastly, the borderline between syntax and lexico-grammar could have been discussed further. For instance, it is not clear to me why the issue of variation between the preterite and the perfect when co-occurring with the three adverbials *yet*, *since*, and *just* belongs to the field of syntax, while variation between transitive and intransitive uses of the two verbs *protest* and *appeal* belongs to the field of lexico-grammar.

These question marks notwithstanding, Hundt's study is both insightful and innovative in its use of different types of linguistic data, and clearly admirable in its inclusion of many different linguistic features. The author has succeeded in her aim of showing whether there is such a thing as New Zealand English.

Notes
1. The press sections of WCNZE, ACE, FLOB, and Frown were compiled on the same basis as those of the well-known LOB and Brown corpora, but comprise material published in 1986 (WCNZE and ACE), in 1991 (FLOB), or in 1992 (Frown).
2. Another matter is that variation among linguistic features may be constrained by so many factors that tests for statistical significance, which can only tell us whether the variation is random or not, are not of much help in accounting for the causes underlying the differences attested in the corpora studied.
3. Disyllabic adjectives, where variation between periphrastic and inflectional comparison would be more expected, are not investigated; instead, reference is made to another corpus-based study.
4. The other syntactic variables investigated include marginal modals, the preterite and the perfect, the *get*-passive, and *for-to*-infinitive constructions.