

English in Transition



Topics in English Linguistics

23

Editor

Herman Wekker

Mouton de Gruyter
Berlin · New York

English in Transition

Corpus-based Studies in
Linguistic Variation and Genre Styles

Edited by

Matti Rissanen

Merja Kytö

Kirsi Heikkonen

Mouton de Gruyter

Berlin · New York 1997

Mouton de Gruyter (formerly Mouton, The Hague)
is a Division of Walter de Gruyter & Co., Berlin.

⊗ Printed on acid-free paper which falls within the guidelines of the ANSI to ensure permanence and durability.

Library of Congress Cataloging-in-Publication Data

English in transition : corpus-based studies in linguistic variation and genre styles / edited by Matti Rissanen, Merja Kytö, Kirsi Heikkonen.

p. cm. — (Topics in English linguistics ; 23)

Includes bibliographical references and index.

ISBN 3-11-015632-6 (alk. paper)

1. English language — Early modern, 1500–1700 — Discourse analysis. 2. English language — Early modern, 1500–1700 — Variation. 3. English language — Early modern, 1500–1700 — Style. 4. Literary form. I. Rissanen, Matti. II. Kytö, Merja. III. Heikkonen, Kirsi, 1962–. IV. Series.

PE881.E54 1997

420'.1'41—dc21

97-25407

CIP

Die Deutsche Bibliothek — Cataloging-in-Publication Data

English in transition : corpus based studies in linguistic variation and genre styles / ed. by Matti Rissanen ... — Berlin ; New York : Mouton de Gruyter, 1997

(Topics in English linguistics ; 23)

ISBN 3-11-015632-6

DBN: 95.102774.3⊗

SG: 52

© Copyright 1997 by Walter de Gruyter & Co., D-10785 Berlin

All rights reserved, including those of translation into foreign languages. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher.

Printing: Ratzlow-Druck, Berlin.

Binding: Lüderitz & Bauer, Berlin.

Printed in Germany.

Preface

This book is one of three volumes reporting the results of the project 'English in transition: Change through variation', carried out in the English Department of the University of Helsinki. The first volume, *Early English in the computer age: Explorations through the Helsinki Corpus* (ed. by Matti Rissanen, Merja Kytö and Minna Palander-Collin, Mouton de Gruyter, 1993) is now followed by two volumes, *English in transition: Corpus-based studies in linguistic variation and genre styles* and *Grammaticalization at work: Studies of long-term developments in English*.

Both these volumes approach change in English from the angle of linguistic variation. The articles deal with processes of change in morphology, syntax and lexis, and pay special attention to the role played by textual and discourse factors across the centuries. From the methodological point of view, diachronic variation analysis and the multi-feature approach aiming at the identification of co-occurrence patterns in genres are the main frameworks adopted.

The aim of the present volume is to give new insights into the development of some central verb constructions (with *be* and *have*), expository apposition, and genre-specific features of expressions of affect and attitude in text. The *Grammaticalization at work* volume sheds light on the development of adverbs and indefinite pronouns and on the means of re-flexivization, in relation to various grammaticalization processes.

All the studies in these volumes are based on the Helsinki Corpus of English Texts; supplementary material has been drawn from other corpora and concordances and from primary texts outside the corpora.

The 'English in transition' project was initiated in 1990 as a continuation of an earlier project which produced the Helsinki Corpus. The core team of both projects has been the same, consisting of the authors and editors of the volumes. The editors would like to express their special thanks to all research assistants of the projects and particularly to Arja Nurmi and Päivi Koivisto-Alanko for their excellent work in producing these volumes.

We are most grateful to the Academy of Finland for funding our project for four years. We are indebted to the University of Helsinki for giving us research premises, and to the English Department for up-to-date technical facilities, travel grants and other support. Our thanks are due to Mrs Leena Sadeniemi for giving us expert advice in computer technology and training us to use programs. Finally, we would like to thank the editors of Mouton de Gruyter for accepting the two volumes now published in their Topics in English Linguistics series.

Helsinki, June 1996

M.R.

M.K.

K.H.

Contents

Introduction <i>Matti Rissanen, Matti Kilpiö, Merja Kytö, Anneli Meurman-Solin, Saara Nevanlinna, Päivi Pahta and Irma Taavitsainen</i>	1
<i>Be/have</i> + past participle: The choice of the auxiliary with intransitives from Late Middle to Modern English <i>Merja Kytö</i>	17
On the forms and functions of the verb <i>be</i> from Old to Modern English <i>Matti Kilpiö</i>	87
Re-phrasing in Early English: The use of expository apposition with an explicit marker from 1350 to 1710 <i>Päivi Pahta and Saara Nevanlinna</i>	121
Genre conventions: Personal affect in fiction and non-fiction in Early Modern English <i>Irma Taavitsainen</i>	185
Towards reconstructing a grammar of point of view: Textual roles of adjectives and open-class adverbs in Early Modern English <i>Anneli Meurman-Solin</i>	267
Bibliography	345
Index of subjects	365

List of abbreviations

The following parameter codes included in the Helsinki Corpus appear in the present volume as such or in abbreviated versions (see also Bibliography at the end of the volume).

Prototypical text category:

EX	= 'expository'
IR	= 'instruction religious'
IS	= 'instruction secular'
IS/EX	= 'instruction secular'/'expository'
NI	= 'narration imaginative'
NN	= 'narration non-imaginative'
STA	= 'statutory'

Text type:

BIA	= 'biography, autobiography'
BIBLE	= 'Bible'
BIL	= 'biography, life of a saint'
BIO	= 'biography, other'
COME	= 'drama, comedy'
CORO	= 'correspondence, non-private'
CORP	= 'correspondence, private'
DEPO	= 'proceeding, deposition'
DIARY	= 'diary'
DOC	= 'document'
EDUC	= 'educational treatise'
FICT	= 'fiction'
GEO	= 'geography'
HANDA	= 'handbook, astronomy'
HANDM	= 'handbook, medicine'

x Abbreviations

HANDO	=	'handbook, other'
HIST	=	'history'
HOM	=	'homily'
LAW	=	'law'
LET PRIV	=	'letter, private'
LET NON-PRIV	=	'letter, non-private'
MYST	=	'drama, mystery play'
NEWT	=	'New Testament'
OLDT	=	'Old Testament'
PHILO	=	'philosophy'
PREF	=	'preface' or 'epilogue'
RELT	=	'religious treatise'
ROM	=	'romance'
RULE	=	'rule'
SCIA	=	'science, astronomy'
SCIM	=	'science, medicine'
SCIO	=	'science other'
SERM	=	'sermon'
TRAV	=	'travelogue'
TRI	=	'proceeding, trial'

Other:

PROF	=	'audience, professional'
NON-PROF	=	'audience, non-professional'
INT	=	'interaction', 'interactive'
INF	=	'informal setting'
FOR	=	'formal setting'
X, XX	=	'unspecified'

Introduction

*Matti Rissanen, Matti Kilpiö, Merja Kytö, Anneli Meurman-Solin,
Saara Nevanlinna, Päivi Pahta and Irma Taavitsainen*

1. Methodological considerations: linguistics and philology in interaction

The last two decades have seen a rapid increase and methodological development in the studies of variation in language. The basic assumption in these studies is “orderly heterogeneity”, i.e. variability which is not random but affected by linguistic and extralinguistic factors or constraints (Weinreich—Labov—Herzog 1968; Samuels 1972; Labov 1994). Language can be seen as meaning potential which is realized in choices between forms and expressions “meaning the same thing” (Halliday 1973: 51; see also Halliday 1978). Within this framework, it has been our aim to establish how linguistic variation is patterned not only socially, regionally and temporally, but according to genres defined by extralinguistic criteria. When the variationist approach is adapted to discourse studies, comparisons of text types defined by their linguistic properties become the key (Schiffrin 1994: 314); the last two chapters of this book extend the application to historical stylistics with the aim of charting genre styles and genre conventions.

In diachronic studies, the variationist approach provides us with a good opportunity to observe the actual process of change. We can trace the birth and death of variant expressions, but perhaps more interestingly, their changing frequencies and distributions within a variant field at subsequent periods of time and in various genres, and we can analyse changes in the intricate mesh of linguistic and extralinguistic factors conditioning the occurrence of these variants. Within this approach our philo-

logical training has been a great asset as it emphasizes the context of expressions and the context of culture in interpreting them.

The importance of extralinguistic factors in the analysis of development and change has necessitated an intensive study and discussion of genre typologies from the point of view of the historical study of language; cf. e.g. Douglas Biber's and Edward Finegan's studies of the 'dimensions' characterizing texts and offering a new basis for their grouping. In the structure of text corpora, the labelling of genres provides a general framework for the discussion of text-related changes at different periods of time. However, the varying importance of conventions or innovative pressures in the evolution of each genre or group of genres may decrease the usefulness of such classifications, as genres are also internally heterogeneous, and the pace and direction of change may be different in individual texts representing a particular genre. A comparison of the occurrences and frequencies of variant expressions in different texts allows the reconstruction of the various levels of past expression: written and speech-based, literary and non-literary, formal and informal, etc. This method is also necessary for all attempts to describe the relationship of the standard(s) to regional or social dialects.

Of the five chapters included in this volume, three discuss morphological, syntactic or lexical questions with special attention paid to variation relating to text type, dialect etc., while two concentrate on the co-occurrence patterns of linguistic features in various types of texts and on questions of genre classification, genre markers and distinguishing features. The main focus is on long-span diachrony, mostly from Early Middle to Modern English, i.e. from the time of the radical reorganization of the structure of the language to the period of its gradual establishment.

The long time span and the wealth of primary data set specific demands for the grammatical models used in this volume. The model should make it possible to compare changing grammatical phenomena across time and, at the same time, be comprehensive enough to provide researchers with analytical tools for the very wide range of morphosyntactic issues involved in classifying linguistic features in genre studies. The model which has proved most useful in this type of research is a structurally oriented one, such as *A comprehensive grammar of the English language* by Quirk et al. (1985) for Present-day English. This type of grammar provides an adequate basis for analysis at a relatively low level of abstraction, enabling the researcher to deal even with the more problematic borderline cases. On the other hand, it is clear that, as a grammar of Present-day English, Quirk et al. cannot be directly applied to different historical phases of

English. In our case, it has provided the basis which the writers of this book have employed in different ways and to different degrees. We have made use of various approaches, from traditional grammar to semantic, pragmatic and textual theories.

2. New openings offered by the Helsinki Corpus and other computerized material

Scholars working on variationist studies benefit from having access to computerized collections of texts. Computerized diachronic corpora, increasingly available in international distribution, make it relatively easy to collect evidence of the occurrence of variant expressions; they also encourage the researcher to approach topics that would earlier have meant an unreasonable amount of time-consuming routine work.

The studies reported in this volume are based on the Helsinki Corpus of English Texts, the first large computerized corpus to cover the time-span of several periods in the history of English. With its 1.5 million words — c. 400 samples of texts dating from the 8th to the 18th century — it offers reliable indicators concerning the structural and lexical developments of English for over a millennium. In many cases, the results are tentative and they must — and fortunately can — be supplemented from other corpora, concordances and primary texts. In the future, the usefulness of the Helsinki Corpus will be further increased by the addition of word-class tagging and syntactic bracketing to the text samples.¹

Each text or group of related texts of the Helsinki Corpus is equipped with a set of parameter values containing information on the text and its author, if known. In Old and Early Middle English, this information is concentrated mainly on the date and dialect of the text and on a fairly loose description of the genre. In Late Middle and Early Modern English, the genre selection is more extensive than in the earlier periods, with samples from drama texts, private letters, law court records, diaries, prose fiction, and so forth.² In these periods, sociolinguistic information is given on the authors of the texts (their rank, sex, and age) and, in the case of letters, on the relationship between the writer and the receiver (intimate family members are distinguished from more distant addressees; addressees superior to authors are distinguished from those inferior to them).

A large and structured corpus such as the Helsinki Corpus, equipped with textual parameter codings, has made it possible for us to combine effectively qualitative and quantitative analysis through variation. This approach has been extensively used in the analysis and interpretation of the data (cf. McEnery—Wilson 1996: 62–63). Our background knowledge of texts, also capitalized on in the compilation of the Helsinki Corpus, has been utilized to place our observations of linguistic phenomena within a larger framework. We have been able to reanalyse and re-evaluate earlier genre and text-type classifications of the older periods of English and to test the validity of the suggestions made concerning the co-occurrence of linguistic features with various types of text. The parameter coding and extensive textual basis have also allowed us to observe the role played by dialectal distribution and the differences between prose and verse texts in the survey of Old and Middle English variant forms. In the discussion of Late Middle and Early Modern evidence, emphasis has been given to the distributions shown by speech-based texts, or texts showing a relatively high degree of orality. In a few cases, it has been possible to draw conclusions from sociolinguistic variables: the genre, degree of interactiveness, level of formality, sex, age, rank or education of the author, or the relationship between the sender and receiver of letters.³ The insights derived from recent trends in sociolinguistics have added a new angle to the discussion of these factors (cf. e.g. Romaine 1982; Milroy—Milroy 1985; Milroy 1992). We believe that, among these variables, genre is the most complex one, as the varying types of genre can be claimed to derive from a combination of other variables such as topic, audience, setting, text category, etc.

The corpus-based framework lends itself to various statistical applications by which it is possible to verify the reliability of the results obtained. In addition to traditional frequency surveys, a number of more advanced statistical analyses have been carried out for the purposes of some studies included in the volume. Thus Merja Kytö in her study on the *be/have* variation uses logistic regression analysis to assess the impact of various factors on the use of the variant forms. This analysis tests out how statistical models based on various combinations of factors account for variation in the data and assess the significance of the individual factors and their interaction. Several statistical methods are combined in Irma Taavitsainen's chapter on personal affect and genre styles. Factor analysis is applied first to identify adjoining text types and possible text-type markers. This method serves to indicate the overall patterns of genre styles and how they relate to one another. Then t-tests and f-tests are used

to evaluate the significance of the features in telling fiction apart from the adjoining genres. The combination of these methods yields results which are then assessed in a larger sociohistorical context. In the chapter by Anneli Meurman-Solin on the concept of point of view in texts, the different informative value of mean frequencies and percentages is stressed and the findings are presented by focusing on one feature or factor at a time and, after a detailed analysis of this kind, by clustering them by both syntactic and semantic criteria. This adds to the reliability of data and also allows the mapping of significant correlations in a network-like pattern of dimensions that usefully describe focal features in genre styles and text types.

3. Variation on the level of morphology, syntax and lexis: the verbs *be* and *have*

The studies by Merja Kytö and Matti Kilpiö in this volume focus on the use and development of the verbs *be* and *have*, which have played a vital role in the shaping of the English verb phrase. The paths of *be* and *have* are parallel, both having occurred as lexical and auxiliary verbs providing variant expressions in the auxiliary function (e.g. compound tenses with mutative verbs, and expression of obligation).

Owing to complexities in the development of these verbs, there is no consensus about the role played by such crucial processes as grammaticalization. With the verb *be*, for instance, the status of the verb in progressive forms has clearly changed from the copula in Old English into an auxiliary from Middle English onwards. The uses of *be* and *have* grammaticalize at different stages in the history of English, and the two studies only concentrate on certain key periods in these developments.

The study of the *be/have* variation with mutative intransitives examines the process by which *have* finally supersedes *be* in present and past perfect constructions; the study on *be* focuses on the forms and functions of the verb, with an eye on the developments in its functional load. The former study covers the period from Late Middle to Modern English, the latter from Old to Modern English. Among the extralinguistic factors included in the two studies are chronology, region and dialect, and foreign influence; moreover, with the *be/have* variation, such factors as text type, relationship of the text to spoken language, level of formality, orality and

authorial properties are taken into account. Among the linguistic factors observed in both studies are tense and certain verb constructions (finite/non-finite forms, *-ing* constructions). Furthermore, with the study of the *be/have* variation, attention is paid to the status of the verb (stative and mutative; action and process; frequent and rarer verbs), durative, iterative and conditional contexts, negation, and object-like and other complements. With *be*, additional factors include developments in the morphology of *be*, the function of the verb (its use as an auxiliary or a lexical verb, copular or non-copular), and semantic-syntactic functions of *b*-forms as against non-*b*-forms (for Old English).

These studies have brought up new evidence pinning down trends of development in greater detail than found in previous research, thanks to the combination of quantitative and qualitative corpus linguistics and philological assessments. With the *be/have* variation, the use of *have* is shown to increase gradually from the Late Middle to the Early Modern English period, gain in momentum in the late 1700s and supersede *be* in the early 1800s. The more powerful extralinguistic factors influencing the process of change include chronology and text type, and a number of linguistic factors (relationship to tense, aspect, complementation etc.). In the study dealing with *be*, the most striking morphological developments are the rapid disappearance in the Early Middle English period of the co-existent Old English present tense paradigms (*beon/wesan*), the diversification of forms in Middle English and the subsequent process of simplification and regularization as Early Modern English is reached. In the survey of the main functions of *be*, the remarkable stability seen in the relative share of the three main functions of *be* over the centuries is the most important finding. Within the auxiliary category, the predominance of the use of *be* as a passive auxiliary is the most noticeable feature.

These two studies have shown that the Helsinki Corpus, supplemented by other diachronic corpora and other standard reference works (*LALME*, *MED*, *OED*), offers a solid basis for the empirical approach aimed at accounting for variation and change.

4. Variation in re-phrasing: expository apposition across the centuries

The chapter by Päivi Pahta and Saara Nevanlinna forms a bridge between the studies concentrating on the syntactic and morphological developments described above and those discussing the characteristics of genres and text types. It examines characteristics of expository apposition — the grammatical category connected with re-phrasing. Re-phrasing as a communicative phenomenon occurs in both written and spoken media, and in both planned and unplanned discourse. It can generally be analysed as the writer's or speaker's attempt to reformulate an utterance in order to achieve successful communication. On closer inspection, the decision to re-phrase may be based on various factors, including stylistic and didactic considerations, the author's assessment of the addressee's ability to process given information, or the author's wish to add to the flow of discourse by providing more information about the topic of discussion.

This chapter focuses on the development and use of expository apposition with an explicit marker in Late Middle and Early Modern English. Apposition is seen as a broad notional category containing both nominal and non-nominal phrases, clauses and sentences. There are no previous detailed studies of apposition in this period, and the theoretical framework adopted as the starting-point in this study is the recent discussion of apposition in Present-day English by Meyer (1987 and 1992), where apposition is seen as a syntactic, semantic and pragmatic relation. The main emphasis in Pahta and Nevanlinna's study is on the semantic characteristics of appositional constructions and their distribution across different text types. Attention is also paid to the devices used in linking appositional units, i.e. explicit markers of expository apposition.

The study shows that the use of expository apposition links up with some of the most central lexical phenomena of the Middle English and Early Modern English periods, such as dialectal variation and the adoption and accommodation of loan-words. It also indicates a clear tendency for some text types to favour the use of appositional constructions in general, and certain semantic and syntactic types in particular. This is so throughout the period, although there is internal variation within most text

types. Most of the markers of expository apposition available in Present-day English were found to exist even in the late medieval period, with many others which have since gone out of use. In this respect, the results obtained point to a difference in the use of coordinative apposition, particularly binomial constructions with the markers *and* and *or*.

In the course of the analysis, the use of traditional philological tools (knowledge of textual background, cultural status of texts, etc.) proved helpful. With certain limitations, the corpus-based approach offered a fruitful way of collecting data for the study of appositional constructions. Considering the open-class nature of apposition as a linguistic phenomenon, the results obtained in this study show the way for further work on the topic.

5. Expressions of personal affect and stance marking: identifying genre-specific choices

The last two chapters of this volume, by Irma Taavitsainen and Anneli Meurman-Solin, have a somewhat different problem-setting, but the approach combining the quantitative and qualitative methods applies here as well. Their main topic is identifying genre-specific features in the linguistic choices related to participant roles by analysing expressions of affect and attitude in texts. Taavitsainen discusses the use of personal pronouns, exclamations, direct questions and other expressions of personal affect, while Meurman-Solin's study focuses on the frequencies and distributions of adjectives and open-class adverbs as stance markers.

In recent years genres have been looked at from many different perspectives. Besides thorough comparative studies of features of individual genres (such as fiction in Fludernik 1993 and 1996) or a wide range of genres in a particular time period (such as Renaissance genres in Lewalski 1986), we find the interdisciplinary approach (for example in Sell and Verdonk 1994) and the important advances in discourse analysis (Coulthard 1994) particularly relevant. Taavitsainen and Meurman-Solin approach the problems of genre studies from a variationist's and sociolinguist's point of view. Rather than restrict the focus to the dimension of written genres as against genres reflecting usages more typical of spoken language, they aim at pointing out clusters of features which position texts on a number of other dimensions, particularly those which reflect

idiolectal or genre-specific characteristics of participant roles, or focus on genre markers.

Traditional genre labels such as 'history', 'private letter', 'autobiography' or 'sermon' are used in the majority of recently produced text corpora. A typology of this kind has also been adopted as a working tool in the Helsinki Corpus and its supplement, the Helsinki Corpus of Older Scots (see note 1). The labelling in these corpora is based on extralinguistic factors such as the social and communicative function of texts or their subject matter. In a number of studies based on the Helsinki Corpus and/or its supplement of Scots (see the bibliographies in the chapters by Taavitsainen and Meurman-Solin in this volume), medieval and Renaissance prose genres have been shown to be linguistically relatively heterogeneous. This heterogeneity is partly due to the compilers' decision to polarize the samples: they have intentionally selected the representatives of a genre from different stylistic traditions when such stylistic variation has been established in earlier research (Nevalainen—Raumolin-Brunberg 1989: 99). This should always be taken into consideration when generalizing from the results.

Patterns of co-occurring features, illustrated in Taavitsainen's and Meurman-Solin's studies, provide evidence for a classification of texts into text types. Each text type may thus comprise texts which represent different genres; in addition, intertextuality phenomena between genres complicate the issues. Biber's pioneering work (1988) in corpus-based stylistics with its multifeature and multidimensional statistical assessments is strictly linguistic; our innovation is the firm philological anchoring, limiting the comparisons to texts that share common features and that belong to related genres (Taavitsainen) and recategorizing multifunctional linguistic features by means of a thorough analysis of their varying syntactic and semantic properties in different time periods (Meurman-Solin). We have also profited from other studies which tackle related questions. The two studies aim at making it applicable to the analysis of early prose texts by selecting features other than those in Biber's factors, and by introducing a more detailed semantic subcategorization of a more comprehensive set of realizations of some features included in his factor analysis, and by developing the research tasks for different aims. Because of the emphasis on semantic features, the relevant examples are carefully selected by qualitative reading and analysed in the wider context of running text. Both studies thus highlight the importance of combining the quantitative approach of corpus linguistics with a detailed analysis of discourse function and meaning, central in the philological tradition.

To compensate for the low number of texts per genre and subperiod, the Helsinki Corpus offers relatively long samples of texts, so that statistically significant amounts of data can be recorded in them. This is also true of numerous lexical items, which often have an insignificant role in corpora where the sample size is smaller (cf. Biber 1988, Devitt 1989). Further study will show whether the texts in the corpus are prototypical representatives of the different genres, but they can perhaps be claimed to have an established position in their sociocultural context. The findings in the two chapters identify basic similarities and differences between texts and genres, which may be selected as diagnostic features in subsequent studies based on a larger corpus of texts.

Taavitsainen's study tests the methods provided by corpus linguistics for a new aim, using structuralist literary criticism and a variationist approach as her theoretical basis. Her focus is on personal affect which is an optional component of participant relations. Personal affect offers a fruitful object of study since it is realized in various ways in texts, and the co-occurrence patterns of its linguistic features show a great deal of variation. Personal affect has not been defined in an adequate way in the literature, and one of the aims of the study is to define its quality in these texts in more detail.

Several quantifying methods are employed in this study and they are combined with qualitative analysis. Factor analysis is applied first to reveal the co-occurrence patterns of linguistic features marking subjectivity. It arranges these features in a hierarchical order according to their importance in reflecting the underlying dimensions of variation. It also indicates textual affinities and points out genres that are close to one another. Further statistical tests are then applied to a smaller number of genres with close affinities. The aim is to test how significant the features with highest factor loadings are as markers between adjoining text types, with special reference to the corpus parameter of imaginative versus non-imaginative narration. Qualitative analysis is used to complement quantitative assessments by relating the texts to their sociohistorical background and by considering the means available in the language for expressing personal affect at that time. In the conclusions, the quality of personal affect in fiction and the adjoining text types are defined in terms of "surge" features of personal affect (i.e. expressions, such as interjections and expletives, conveying intensified personal charge between the participants of communication), interactive versus egocentric focus, indexical features of proximity, and space-building modality.

In Meurman-Solin's study the hypothesis is that the frequencies and distributions of adjectives and open-class adverbs representing various semantic categories provide evidence of the relative importance of descriptive features as against the importance of stance marking in texts (cf. Biber and Finegan 1989). The study discusses the correlation patterns between these linguistic features and a number of extralinguistic factors such as genre, text category, degree of interactiveness, level of formality, the audience or the author's sex. Moreover, it illustrates how the choices made between adverb and adjective realizations, and, in the case of the latter, between integrated and fragmented structures, reflect varying degrees of informational density, and identify the relative prominence of the author's voice in the individual texts and in the different genres and text categories. Methodologically, the study aims at solving problems related to the traditional ways of grouping texts. It presents an approach in which a detailed analysis of individual features in each idiolect or text leads to the accumulation of evidence as regards co-occurrence patterns of these features; it is the cumulative effect of the carefully analysed patterns that then serves to identify text types. Ideally, these classifications can be meaningfully related to extralinguistic variables, so that ultimately the interrelatedness of the varying social functions of texts and the linguistic expression of participant roles becomes evident.

Finally, as concerns the theoretical framework adopted, the studies presented in this volume clearly show the power of the two methods, linguistic and philological: accounting for language change in all its complexity becomes a possible — and a rewarding — task when based on the principles of corpus linguistics, variationist approach, and genre studies.

Notes

1. For information on the Helsinki Corpus, see Rissanen et al. (1993) and Kytö (1996).

Among the satellite corpora projects closely related to the Helsinki Corpus is the Helsinki Corpus of Older Scots, compiled by Anneli Meurman-Solin and currently in international distribution (see Meurman-Solin 1995). In addition, there are three projects in progress which will result in new diachronic corpora: the Corpus of Early American English (see Kytö 1993); the Corpus of Early English Correspondence (see Nevalainen—Raumolin-Brunberg 1996) and the Corpus of Early English Medical Writing (Taavitsainen—Pahta 1997).

The Penn-Helsinki Parsed Corpus of Middle English, compiled by Anthony Kroch and Ann Taylor, is based on the prose texts drawn from the Middle English section of the Helsinki Corpus accompanied by a number of supplementary texts. The texts have been annotated for syntactic analysis. A new version of this corpus, with a more versatile linguistic coding and further texts, is in preparation. The Brooklyn-Geneva-Amsterdam-Helsinki Parsed Corpus of Old English, based on the Old English section of the Helsinki Corpus, is another morphologically and syntactically annotated corpus due to appear within the next five years. The team of compilers include Susan Pintzuk, Willem Koopman and others.

Further information on historical corpora of English, either completed or in preparation, is found in Kytö et al. (1994), Kytö — Rissanen (1996), and Hickey et al. (1997).

2. The following text types, defined on the basis on extralinguistic and situational criteria, are represented in the Helsinki Corpus: law, documents, handbooks, science, homilies, sermons, rules, religious treatises, the Bible, philosophy, prefaces, history, geography, travelogue, (auto)biography, fiction, romances, depositions, trial records, drama (mystery plays and comedies), private and official letters, educational treatises, and diaries. There are also texts which have not been given any particular text type definition (e.g. Old and Middle English verse texts).
3. The articles in the present volume do not presuppose detailed knowledge of the contents and structure of the Helsinki Corpus. However, as always, some knowledge of the material will make it easier to appreciate the findings presented. For those interested in learning more about the corpus, Rissanen et al. (1993) and Kytö (1996) will offer a good starting-point.

The typographical conventions in the examples cited from the Helsinki Corpus are explained in Kytö (1996). For convenience, the references to the Helsinki Corpus source texts are listed in the Bibliography at the end of the present volume.

References

- Biber, Douglas
 1988 *Variation across speech and writing*. Cambridge—New York etc: Cambridge University Press.
- Biber, Douglas—Edward Finegan
 1989 "Styles of stance in English: Lexical and grammatical marking of evidentiality and affect", *TEXT* 9/1: 93–124.

- Coulthard, Malcolm (ed.)
 1994 *Advances in written text analysis*. London—New York: Routledge.
- Devitt, Amy J.
 1989 *Standardizing written English. Diffusion in the case of Scotland 1520–1659*. Cambridge—New York etc.: Cambridge University Press.
- Fludernik, Monika
 1993 *The fictions of language and the languages of fiction*. London—New York: Routledge.
 1996 *Towards a 'natural' narratology*. New York: Routledge.
- Halliday, M. A. K.
 1973 *Explorations in the functions of language*. London: Edward Arnold. [Reprinted in 1974.]
 1978 *Language as social semiotic. The social interpretation of language and meaning*. London: Edward Arnold.
- Helsinki Corpus
 1991 *The Helsinki Corpus of English Texts*. Helsinki: Department of English, University of Helsinki.
- Hickey, Raymond—Merja Kytö—Ian Lancashire—Matti Rissanen (eds.)
 1997 *Tracing the trail of time. Proceedings from the Second Diachronic Corpora Workshop*. Amsterdam and Atlanta, GA: Rodopi.
- Kytö, Merja
 1993 "Early American English", in: Matti Rissanen—Merja Kytö—Minna Palander-Collin (eds.), 83–91.
- Kytö, Merja—Matti Rissanen—Susan Wright (eds.)
 1994 *Corpora across the centuries: Proceedings of the First International Colloquium on English Diachronic Corpora, St Catherine's College Cambridge, 25–27 March 1993*. Amsterdam—Atlanta, GA: Rodopi.
- Kytö, Merja (comp.)
 1996 *Manual to the diachronic part of the Helsinki Corpus of English Texts: Coding conventions and lists of source texts*. (Third edition.) Helsinki: Department of English, University of Helsinki.
- Kytö, Merja—Matti Rissanen
 1996 "English historical corpora: Report on developments in 1995", *ICAME Journal* 20: 117–132.
- Labov, William
 1994 *Principles of linguistic change*. Vol. 1. (Language in Society 20.) Oxford: Blackwell.
- Lewalski, Barbara Kiefer (ed.)
 1986 *Renaissance genres: Essays on theory, history and interpretation*. Cambridge, Mass.—London, England: Harvard University Press.

- McEnery, Tony—Andrew Wilson
 1996 *Corpus linguistics*. (Edinburgh Textbooks in Empirical Linguistics.)
 Edinburgh: Edinburgh University Press.
- Meurman-Solin, Anneli
 1995 "A new tool: The Helsinki Corpus of Older Scots (1450–1700)",
ICAME Journal 19: 49–62.
- Meyer, Charles F.
 1987 "Apposition in English", *Journal of English Linguistics* 20/1: 101–
 121.
 1992 *Apposition in contemporary English*. Cambridge: Cambridge Univer-
 sity Press.
- Milroy, James
 1992 *Linguistic variation and change. On the historical sociolinguistics of*
English. (Language in Society 19.) Oxford—Cambridge, Mass.:
 Blackwell.
- Milroy James—Lesley Milroy
 1985 "Linguistic change, social network and speaker innovation", *Journal*
of Linguistics 21: 339–384.
- Nevalainen, Terttu—Helena Raumolin-Brunberg
 1989 "A corpus of Early Modern Standard English in a socio-historical
 perspective", *Neuphilologische Mitteilungen* 90/1: 67–111.
- Nevalainen, Terttu—Helena Raumolin-Brunberg (eds.)
 1996 *Sociolinguistics and language history. Studies based on the Corpus*
of Early English Correspondence. Amsterdam—Atlanta, GA: Rodopi.
- Quirk, Randolph—Sidney Greenbaum—Geoffrey Leech—Jan Svartvik
 1985 *A comprehensive grammar of the English language*. London—New
 York: Longman.
- Rissanen, Matti—Merja Kytö—Minna Palander-Collin (eds.)
 1993 *Early English in the computer age: Explorations through the Helsinki*
Corpus (Topics in English Linguistics 11.) Berlin—New York: Mou-
 ton de Gruyter.
- Romaine, Suzanne
 1982 *Socio-historical linguistics, its status and methodology*. (Cambridge
 Studies in Linguistics 34.) Cambridge—London etc.: Cambridge Uni-
 versity Press.
- Samuels, M. L.
 1972 *Linguistic evolution, with special reference to English*. (Cambridge
 Studies in Linguistics 5.) Cambridge: Cambridge University Press.

Schiffrin, Deborah

1994 *Approaches to discourse*. Oxford—Cambridge, Mass.: Blackwell.

Sell, Roger D.—Peter Verdonk (eds.)

1994 *Literature and the new interdisciplinarity, poetics, linguistics, history*. Amsterdam—Atlanta, GA: Rodopi.

Taavitsainen, Irma—Päivi Pahta

1997 “Corpus of Early English Medical Writing 1375–1750”, *ICAME Journal* 21: 71–78.

Weinreich, Uriel—William Labov—Marvin I. Herzog

1968 “Empirical foundations for a theory of language change”, in: W. P. Lehmann—Yakov Malkiel (eds.), *Directions for historical linguistics: A symposium*. Austin—London: University of Texas Press, 95–195.

***Be/have* + past participle: The choice of the auxiliary with intransitives from Late Middle to Modern English**

Merja Kytö

1. The *be/have* variation in the history of English

In many languages the choice of the auxiliaries *be* and *have* (and their equivalents) has been subject to variation in perfective constructions with intransitives indicating “transition” or “change” (cf. *they are arrived* versus *they have arrived*). In some Germanic and Romance languages the development has resulted in the generalization of one construction (e.g. the almost total dominance of “be” in Present-day Danish and “have” in Spanish and Portuguese); in others both variant forms occur in certain grammatically, stylistically or regionally restricted contexts (as, e.g., in Present-day Swedish, German, French and Italian).¹ This cross-linguistic variation reflects the differences in the systemic realizations of the distinction between state (favouring *be*) and action/process (favouring *have*). Over the successive stages of development, various linguistic and extra-linguistic factors have influenced the choice of one form or the other.

In the Old English period the *be/have* + past participle construction denoted “state” in intransitive and transitive uses (cf. *hie wæron gecumene; hie hæfdon hine gebundenne*). The past participle originally functioned as an adjective and was sometimes inflected; it could also be preceded by an object complement (Mitchell 1985:I, §709). However, the grammatical concord was gradually lost and the past participle was placed immediately after the auxiliary. *Have* originally occurred with transitive verbs only, but early on came to be used with intransitives, too. In Early Middle English *be* prevailed with mutative verbs, but *have* started gaining ground slowly in uses with the emphasis on “action” and the notion of per-

fectivity (Rydén—Brorström 1987: 16–18). Signs of the rise of *have* are apparent from around the 1400s on (for a recent corpus-based study, see Elsness 1989: 100; 1991: 276–283).

The development of *be* and *have* perfects show features attributed to the process of grammaticalization, but scholars disagree about the exact chronology and nature of the process. A recent discussion on the topic can be found in Denison 1993: 340–368 (who uses the term “grammaticalisation”). With *have* perfect the relevant factors have been the loss of inflections in the participle, the word order, and certain VP types (according to the semantics of OE *habban* and the valency of the main verb). The process of grammaticalization would have reached a stage of fulfilment “when the *have* perfect became available for any lexical verb which did not conjugate with *be*”, i.e. when it became an auxiliary verb, the suggested (but debatable) date for this being the late Old English period (Denison 1993: 352). The other possible stages of grammaticalization include the point when the construction became a tense equivalent (probably in late Old English); when it had developed its present-day meaning and superseded *be* (probably in the seventeenth century); when it became used with all non-auxiliary verbs (in late Modern English) (Denison, 1993: 352). There is, similarly, disagreement over the grammaticalization process of the *be* perfect (for a summary, see Denison 1993: 360–361); the factors regarded as having influenced the process include, e.g., the increasing use of *be* as the auxiliary of the passive and the relatively light functional load of *have*, the possibility of neutralizing the present tense third-person singular forms into the clitic *'s*, and the prescriptions of normative grammarians.

In this study the emphasis is on the variational method, which means excluding from analysis the contexts in which there is no choice of one variant form or the other, e.g. the uses with transitive verbs, which have always been associated with the *have* perfect. However, many of the factors pointed out in previous studies have been of great help when specifying the scope of the study and the distributional factors considered.²

2. The scope of the study

This study focuses on the development of the *be/have* paradigm in the history of English from the Late Middle English period up until modern

times. The reasons for setting Late Middle English as a starting point for observing the development are two-fold. Firstly, as pointed out above, a change starts to take place in the use of the auxiliaries in the case of the mutatives in the fifteenth century, a justification for including the data from the Late Middle English period in the analysis. Secondly, the early stages of the increasing use of *have* coincide with the crucial transition from the Middle to the Early Modern English period. As for the time-span covered, pursuing the development up until modern times is of great interest, as it was only in the early 19th century that *have* finally superseded *be* (Rydén—Brorström 1987). *Be* is, moreover, still available with some verbs in present-day received usage (*they are/have gone*, *the sun is/has set*; Rydén—Brorström 1987: 211) and in some dialects (see, e.g. Kallen 1989: 18–23; Melchers 1992; Filppula 1994).

Though the *be/have* variation has been dealt with in a good number of studies, many of the previous works have largely been based on idiolectal corpora of single authors or restricted to cover only certain periods or a few text types. The present study, based on material that covers a period of some 650 years and representative of a variety of authors and text types, should provide new evidence on the role of the linguistic and extra-linguistic factors in the process of change.

The study will be carried out within the framework of socio-historical variation analysis (see the Introduction in this volume). The distribution patterns of the variant forms will be followed across the successive periods of time distinguished for the corpora studied (see below). Conclusions will be based on simple frequency tables and results yielded by the logistic regression analysis used to assess the combined effects of the factors and their interaction (see section 9 below).

3. The sources for data

In variational terms, the *be/have* paradigm presents a two-term notional case, involving various aspects of syntactic, semantic and lexical change. For obvious reasons, this paradigm is a particularly rewarding topic within the framework of computer-assisted corpus analysis. Relevant examples with the construction are, if not overwhelming in number, still frequent enough for a thorough-going analysis; they are also relatively easy to retrieve from machine-readable material on the basis of lexical forms. As

there is no single corpus available at the moment that would cover the time-span included in this study, several diachronic corpora will need to be consulted. The differences in the size and structure of the corpora will be taken into consideration when weighing the evidence. The corpora that were used as sources for data are introduced below.

The Helsinki Corpus of English Texts (HC; Late ME, EModE)

The Late Middle English data and a good deal of the Early Modern English data have been drawn from the Helsinki Corpus (see Rissanen—Kytö—Palander-Collin 1993; Kytö 1996); the subperiods included in the Helsinki Corpus material are ME3 (1350–1420) and ME4 (1420–1500), and EModE1 (1500–1570), EModE2 (1570–1640) and EModE3 (1640–1710). All types of texts represented in the corpus will be considered, but priority will be given to those that have counterparts in the other corpora studied.

The Century of Prose Corpus (COPC)

As regards the rise of *have*, the Century of Prose Corpus covers a crucial period of development, extending from 1680 to 1780. Part A of the corpus includes extracts from works of 20 major prose authors of the period (three selections of 5000 words representing each author), while Part B consists of writings (sample length 2000 words) drawn from the pens of one hundred authors writing as journalists, scholars, men of letters and so forth (Milić 1990: 27–29; 1994: 70). Text type definitions are given by the compiler for texts included in Part B only. The first three decades of writing in the Century of Prose Corpus (1680–1710) coincide with the last three decades of the Early Modern English section of the Helsinki Corpus. The whole century covered also coincides with the early decades represented by the ARCHER Corpus (1650–1990, see below). Contrary to the conventions adopted for the other corpora included in this study, the spelling of the Century of Prose Corpus has been modernized.

The ARCHER Corpus

This study owes a great deal to the compilers of the ARCHER Corpus, or A Representative Corpus of Historical English Registers, for access to relevant data drawn from parts of the pilot version of the corpus. When complete, this corpus will include c. 1,000 texts and c. 1.7 million words

(sample size at least 2,000 words). The aim of the corpus is “to enable the analysis of historical change in the range of written and speech-based registers of English from 1650 to the present” (Biber et al. 1994a: 3; see also Biber et al. 1994b). Out of the text types found in the corpus, six have been included in the present study: fiction prose, drama, journals, letters, science and sermons. The six text types are of special interest for this study, as they all bear affinities to those found in the Helsinki Corpus and parts of the Century of Prose Corpus. Texts by British and American authors have been considered; while the British authors are represented more or less evenly over the subperiods distinguished, the texts by American authors are all from three subperiods, 1750–1799, 1850–1899 and 1950–1990 (all scientific writings are by British authors). The corpus is being tagged for grammatical and functional categories; for the purposes of the present study, the data was drawn from the untagged version on the basis of lexical forms.³

Summary of the corpora studied

The main characteristics of the size and the structure of the corpora included in the study are given in Table 1.

Table 1. The structure and size of the (sub)corpora studied.

Helsinki Corpus	Text type	Words	Total
ME3 (1350–1420)	fiction	14,300	
	diary	—	
	private letter	—	
	official letter	5,000	
	drama = mystery play	—	
	science	3,600	
	sermon	18,900	
	document	13,900	
	handbook	19,200	
	philosophy	12,600	
	homily	7,300	
	other	94,900	
Total		189,700	

Table 1. Continued.

Helsinki Corpus	Text type	Words	Total
ME4 (1420–1500)	fiction	8,800	
	diary	–	
	private letter	19,500	
	official letter	3,200	
	drama = mystery play	20,100	
	science	6,400	
	sermon	25,000	
	document	10,500	
	handbook	20,100	
	philosophy	–	
	homily	–	
	other	100,300	
Total		213,900	
Total Late ME			403,600
EModE1 (1500–1570)	fiction	11,600	
	diary	13,100	
	private letter	10,600	
	official letter	6,300	
	comedy	10,600	
	science	12,900	
	sermon	9,500	
	document	–	
	handbook	10,000	
	philosophy	9,900	
	other	95,600	
Total		190,100	

Table 1. Continued.

Helsinki Corpus	Text type	Words	Total
EModE2 (1570–1640)	fiction	12,500	
	diary	12,500	
	private letter	11,600	
	official letter	5,700	
	comedy	11,800	
	science	13,000	
	sermon	10,300	
	document	–	
	handbook	12,300	
	philosophy	6,900	
	other	93,200	
Total		189,800	
EModE3 (1640–1710)	fiction	12,000	
	diary	11,200	
	private letter	13,100	
	official letter	5,900	
	comedy	12,700	
	science	11,300	
	sermon	12,500	
	document	–	
	handbook	11,400	
	philosophy	8,800	
	other	72,100	
Total		171,000	
Total EModE			550,900

Table 1. Continued.

COPC	Text type	Words	Total
Part A (1680–1780)	expository prose by 20 authors (15,000 words each)	300,000	
Total		300,000	
Part B (1680–1780)	fiction	20,000	
	letter	20,000	
	science	20,000	
	biography	20,000	
	periodical	20,000	
	education	20,000	
	essay	20,000	
	history	20,000	
	polemics	20,000	
	travel	20,000	
Total		200,000	
Total COPC			500,000
ARCHER Corpus	Text type	Words	Total
ARCHER1 (1650–1700)	fiction	29,500	
	journal	21,900	
	letter	14,900	
	drama	32,600	
	science	20,900	
	sermon	11,400	
Total		131,200	
Total ARCHER1			131,200

Table 1. Continued.

ARCHER Corpus	Text type	Words	Total
ARCHER2a (1700–1750)	fiction	44,000	
	journal	22,000	
	letter	20,700	
	drama	24,700	
	science	21,400	
	sermon	6,500	
Total		139,300	
ARCHER2b (1750–1800)	fiction	86,200	
	journal	45,300	
	letter	32,200	
	drama	44,400	
	science	20,900	
	sermon	26,900	
Total		255,900	
Total ARCHER2			395,200
ARCHER3a (1800–1850)	fiction	53,900	
	journal	22,800	
	letter	17,200	
	drama	34,400	
	science	21,400	
	sermon	4,600	
Total		154,300	

Table 1. Continued.

ARCHER Corpus	Text type	Words	Total
ARCHER3b (1850–1900)	fiction	79,700	
	journal	46,600	
	letter	28,300	
	drama	77,200	
	science	22,400	
	sermon	29,100	
Total		283,300	
Total ARCHER3			437,600
ARCHER4a (1900–1950)	fiction	12,400	
	journal	22,700	
	letter	15,200	
	drama	35,700	
	science	22,500	
	sermon	4,300	
Total		112,800	
ARCHER4b (1950–1990)	fiction	68,100	
	journal	45,100	
	letter	32,100	
	drama	68,400	
	science	22,900	
	sermon	28,000	
Total		264,600	
Total ARCHER4			377,400
Total ARCHER			1 341,400
TOTAL HC+COPC+ARCHER			2 795,900

4. On the criteria of inclusion

Selecting examples for a study of the *be/have* variation with intransitives poses some problems. There are two main questions related to (1) the type and history of the main verb and (2) the rich functional range of the auxiliary *be*. Further selectional criteria to be considered include the role of object-like complements (e.g. nouns of extent or measure), and some relatively infrequent constructions in which the variational role of the auxiliaries is blurred.

The type of the verb

The main verbs are an open-ended class, characterized by a number of semantic features not always too precise to determine. The studies on the history of the *be/have* variation traditionally deal with the use of the more frequent verbs indicating motion (e.g. *arrive, come, go, pass, ride*), process or change (e.g. *alter, change, improve, turn*), happening (e.g. *befall, chance, hap(pen)*), appearing or originating (e.g. *appear, arise, become, begin*) and finishing or disappearing (e.g. *cease, decay, decline, expire, die*).⁴ The inclusion of the less frequent verbs (e.g. *pirouette, penetrate*) has been the researcher's decision.

Most intransitive verbs considered in previous studies have here been included for preliminary analysis (for some individual verbs excluded, see below); along the lines of variational method, some categories dominated by one or other variant form have been excluded from further analysis on the basis of the distribution patterns obtained. Similarly, as seems sensible within the framework of variation analysis, some verbs (e.g. *follow, misgo*), which on the basis of the previous studies have been used with the form *have* only, have been excluded altogether (for a provisional list of verbs and their *be/have* history, see Visser 1952: 661–663).

Some verbs are frequently indeterminate in nature as regards the transitive versus intransitive dimension. For this reason, such verbs as *agree, assemble, end, finish, gather, meet, marry, wed* and *die (be dead)* have been excluded from the analysis.

In some studies attention has also been paid to non-mutative or stative verbs (e.g. *cling, lie, rest, stay*; see Visser 1963–1973: 2044–2084). This category has also been considered at the first stages of the present study.

The functional range of the verb be

The use of *be* as a copula as well as a perfective and a passive marker renders the functional load of the verb heavy (Rydén—Brorström 1987: 24; Elsness 1991: 266–268). With some verbs it is not always possible to distinguish between the copula and the perfective uses (cf. *he is changed* ‘he is different’ versus *he is changed* ‘he has become different’); see example (1).⁵

- (1) The cannon, however, did not stand proof, and the Indians, who made a close attack, were beaten off and the garrison relieved. The fort *is* now totally *decayed*, and Captain Zane, the only inhabitant at or near the place, makes use of it for firewood.
(ARCHER2b/Journal/Peter Muhlenberg)⁶

Nor can one always distinguish between active (perfective) and passive (present/past tense) uses with verbs which can appear in both transitive and intransitive contexts (*he is changed* ‘he has become different’ versus *he is changed* ‘he has been made different’); for a discussion, see Rydén—Brorström (1987: 24), who for this reason exclude an example such as ‘My prospect of getting to London this spring *is* rather *darkened*’ (Wordsworth II: 772, 1836) from their data; for further points, see also Kakietek 1987 [1976]: 310–311); for a corpus example, see (2).

- (2) .. and such is their methode, that rests not so much vppon euidence of truth prooued by arguments, authorities, similitudes, examples; as vpon particular confutations and solutions of euerie scruple, cauillation & obiection: ... so that the Fable and fiction of Scylla seemeth to be a liuely Image of this kinde of Philosophie or knowledge, which *was transformed* into a comely Virgine for the vpper parts; (Bacon, *Advancement of Learning* 20V)

Ambiguous instances such as (1) and (2) have been excluded from this study; the basic criterion of inclusion has been that in its context the construction conveys the notion of perfectivity.⁷

All instances of dynamic intransitives (e.g. *go*, *come*, *rise*, *arrive*) have been included, even though the construction in some examples may come closer to a stative than a perfective meaning (cf. *they are/have gone*, *the sun is/has set*).⁸

Object-like complements

Though this study is primarily concerned with intransitive uses as in (3a), there are verbs that can take complements in the form of nouns of extent or measure as in (3b).⁹

- (3) a. My godfathyr has be syke byt he ys whell mendyd, thankyd be God. [Thys same] day my Loord ys *comyn* to London to aske the Kyng leue to go to the Rodys for he ys sent for.
(Richard Cely, *The Cely Letters* 107)
- b. ... and when mattens vhos done thay whente to a kynnyswhoman off the 3ewnge genttyllwhomane; and I sent to them a pottell of whyte romnay, and thay toke hyt thankefully, for thay *had cwm a myle* a fote that mornyng; (Richard Cely, *The Cely Letters* 152)

With verbs there is variation in whether prepositions or prepositional phrases are used to link the destination or other complement with the verb; cf. (4a) with *enter* + *into* + object pronoun and (4b) with *enter* + object pronoun.¹⁰

- (4) a. They have not perhaps received precisely what they expected when their Christian life began, for the kingdom of heaven cannot be really known until a man *has entered into it*;
(ARCHER3b/Sermon/Robert Dale)
- b. Feb. 26th The works of Ciudad Rodrigo having been completely put in order, and a garrison of Spaniards *having entered it*, the army was ordered to proceed towards Estremadura.
(ARCHER3a/Journal/George Simmons)

In the interest of the emerging individual verb profiles, instances such as (3b) will be included in the analysis. Out of the examples illustrated in (4b), two frequent verbs, *enter* and *pass*, have been sampled systematically for full coverage; as for the other verbs, some of them presenting ambiguous borderline cases between transitive and intransitive uses, the instances with object-like complements have been omitted from the analysis.