

Grammaticalization at Work



Topics in English Linguistics

24

Editor

Herman Wekker

Mouton de Gruyter
Berlin · New York

Grammaticalization at Work

Studies of Long-term Developments
in English

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

Grammaticalization at work : studies of long-term develop-
ments in English / edited by Matti Rissanen, Merja Kytö,
Kirsi Heikkonen.

p. cm. — (Topics in English linguistics ; 24)

Includes bibliographical references (p.) and index.

ISBN 3-11-015631-8 (cloth ; alk. paper)

I. English language — Grammar, Historical. I. Rissanen,
Matti. II. Kytö, Merja. III. Heikkonen, Kirsi, 1962—
IV. Series.

PE1075.G65 1997

425'.09—dc20

97-25408

CIP

Die Deutsche Bibliothek — Cataloging-in-Publication Data

Grammaticalization at work : studies of long term developments
in English / ed. by Matti Rissanen ... — Berlin ; New York :
Mouton de Gruyter, 1997

(Topics in English linguistics ; 24)

ISBN 3-11-015631-8

DBN: 95.102826.X⊗

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 mechan-
ical, including photocopy, recording, or any information storage and retrieval system, with-
out 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, *Grammaticalization at work: Studies of long-term developments in English* and *English in transition: Corpus-based studies in linguistic variation and genre styles*.

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 adverbs and indefinite pronouns and in the means of reflexivization, in relation to various grammaticalization processes. The *English in transition* volume sheds light on 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.

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.

vi Preface

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</i>	1
Indefinite pronouns with singular human reference <i>Helena Raumolin-Brunberg and Leena Kahlas-Tarkka</i>	17
The pronominalization of <i>one</i> <i>Matti Rissanen</i>	87
The processes of adverb derivation in Late Middle and Early Modern English <i>Terttu Nevalainen</i>	145
<i>There</i> compounds in the history of English <i>Aune Österman</i>	191
The development of reflexive strategies in English <i>Kirsti Peitsara</i>	277
A medieval case of grammaticalization, <i>methinks</i> <i>Minna Palander-Collin</i>	371
Bibliography	405
Index of subjects	429

Introduction

*Matti Rissanen*¹

1. Introductory

The articles of the present volume deal with the formation of adverbs and indefinite and reflexive pronouns, with particular focus on grammaticalization. Studies of grammaticalization are often based on rather meagre empirical evidence and they tend to concentrate on individual linguistic elements. We hope that the present volume, which reflects the goals and principles of our research project, will add new dimensions to these studies. First of all, our studies take a long-term diachronic perspective, from Old or Middle English to the end of the 17th century or even beyond, always with reference to the Old English background of the developments. Most of them systematically look at long-term changes in a linguistic category or subcategory as a whole: adverbs, indefinite pronouns with personal reference, expression of reflexivity (see section 3 for our variationist approach). Studies dealing with the grammaticalization of individual items (METHINKS, ONE) link their discussion with the more general reorganization of the structure of English.

By adding a more comprehensive diachronic approach to the study of grammaticalization, we hope to be able to place the operations of the various processes of grammaticalization in a wider perspective and in this way open up new insights into the history of English. We do not only trace the development of those grammaticalized items which survive in Present-day English, but we also attempt to analyse the factors which lead to the diminishing popularity and obsolescence of some of the forms and usages in past centuries. Particularly, we aim at illustrating the interde-

pendence of structural systems in a diachronic perspective, not forgetting semantic and pragmatic considerations.

Secondly, our studies are based on systematically collected empirical evidence derived from a large body of computer-readable texts. The improved possibilities for data retrieval and analysis offered by large corpora have encouraged us to apply a method of combining quantitative and qualitative analysis which, we hope, will offer a clearer and more 'life-like' picture of developments which have taken place centuries — or even more than a millennium — ago. From a more heuristic point of view, careful analysis of corpus evidence calls attention to research problems that might not be discovered by less systematic data handling.

The use of electronic corpora facilitates data retrieval but it also makes demands on our analytic tools. Our model of grammar would have to cope with the fuzziest realizations of the categories we are interested in, both changes in progress within these categories and the vagueness inherent in ordinary-language data (see further section 3).

2. On grammaticalization

Of all the Germanic languages, English has undergone the most radical structural and lexical changes during its existence. A detailed survey of these changes is beyond the scope of this introductory note. It is obvious, however, that the levelling and loss of the inflectional endings, traceable even at the earliest stages of Old English and accelerating in Late Old English and Early Middle English, resulted in a gradual rearrangement of the grammatical system of English, which was finally established in the Modern period. The questions of causality and of the interdependence of the changes involved in this rearrangement have been a topic of lively scholarly debate in recent years, and very far-reaching conclusions should be avoided, especially as the diachrony of the various structural systems has not yet been adequately researched. It seems, however, that the co-occurrence of the following developments with the collapse of the system of endings can hardly be a coincidence:

1. Establishment of the S–V–O order, with the reorganization of the arguments of the verb
2. Development of the auxiliary system
3. Development of new pronominal forms, particularly those serving as the head of the noun phrase
4. Development of new adverbial forms and functions
5. Development of the prepositional system
6. Development of new links between the elements of the sentence, both at clause and phrase level

All these developments are in one way or another connected with the concept of grammaticalization. According to the simple definition by Hopper—Traugott (1993: xv, 2), grammaticalization is the process in which lexical items and constructions come to serve grammatical functions, and once grammaticalized, continue to develop new grammatical functions. Grammaticalization is, of course, inseparable from change of meaning; one basic tendency to be noted here is when ‘meanings based in the external or internal described situation become meanings based in the textual and metalinguistic situation’ (Traugott 1989: 35). The direction of the semantic change can in most cases be conveniently described as ‘bleaching’. Hopper—Traugott (1993: 87–93) warn, however, against too straightforward an association of grammaticalization with loss of meaning and point out that traces of the original lexical meanings of the grammaticalized forms often remain. They prefer the expressions ‘pragmatic enrichment’, ‘strengthening’, etc. to ‘bleaching’ or ‘fading’ and point out that grammaticalization results in the development of new uses as often as in the loss of old ones.

In grammaticalization, the speaker’s subjective attitude toward the proposition is of importance (see Traugott 1989: 35 and Traugott—König 1991). Subjectification as such, seems, however, too strong a factor in the analysis of grammaticalization; it is possible that the speaker’s subjective attitude is, to some extent, present in most changes of meaning. For this reason, the more general starting-point of pragmatic inferencing, i.e. the role of the speaker and hearer negotiating meaning in communication

situations, and the dichotomy 'expressiveness/routinization' (Hopper—Traugott 1993: 63ff.) offers a less problematic starting-point.

As mentioned above, all the articles in this volume are concerned with grammaticalization, although this process has been treated as an approach for the analysis and understanding of the changes rather than as a linguistic phenomenon to be studied in its own right. The types of grammaticalization can be divided into 'primary' and 'secondary', on the basis of whether or not they involve radical changes in the word class or structural properties of the items. The changes discussed in the present volume mainly belong to the category of primary grammaticalization: they deal with developments of pronouns and adverbs from other parts of speech. Secondary grammaticalization can be seen in subsequent developments of the pronominal or adverbial forms discussed.

3. Study of change and variation in language

The last two decades have seen a rapid increase and methodological development in the studies of variation in language. The starting-point of these studies is "orderly heterogeneity", i.e. variability which is not random but affected by linguistic and extralinguistic factors (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).

In diachronic studies, the variationist approach gives us a good opportunity to observe the various stages of change: embedding, transition and actuation (Weinreich et al. 1968: 102; cf. also Milroy—Milroy 1985: 341–343; Milroy 1992: 20–21). We can trace not only the birth and death of variant expressions, but perhaps more interestingly, their changing frequencies within a variant field at subsequent periods of time, and we can analyse the changes taking place in the intricate mesh of the factors co-occurring with the variants. The insights derived from recent trends in sociolinguistics have added a new angle to the discussion of these factors, both in view of the language of individual speakers and of larger speech communities (cf. e.g. Romaine 1982; Nevalainen—Raumolin-Brunberg 1996).

The emphasis placed on variation has enhanced the role played by the text in the study of change. The distribution of variant forms in texts representing different genres or text types, as defined by extralinguistic factors (topic, the author's background, discourse situation and participant relationship, etc.) offers us a way to analyse register variation of the past. A comparison of the linguistic features in texts representing different distances from spoken language is also practically our only way to form hypotheses of the structures and vocabulary typical of the speech of past centuries. The same method of textual comparison must be used in attempts to describe the relationship of the standard(s) to regional or social dialects.

The long time-span and the wealth of primary data set specific demands for the grammatical models used in all diachronic studies. 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 a very wide range of morphosyntactic issues. Furthermore, the main aim in diachronic studies is less often to develop new theories of grammar but, rather, to provide descriptions of interesting linguistic phenomena in the history of the language in a form that will offer a basis for further studies using different theoretical models.

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 starting-point has been employed by the authors of the present volume in different ways and to different degrees. We could characterize our grammatical orientation as eclectic: we have made use of various approaches, from traditional grammar to semantic, pragmatic and textual theories. This approach has provided a way of categorizing the variant expressions from the diachronic point of view in a way which maintains the main lines of development, while simultaneously making possible an analysis with a relatively high level of detail and even accounting for problematic borderline cases.

4. Computerized corpora

In recent years, computerized corpora have played a decisive role in diachronic studies of variation. They have radically shortened the time needed for collecting evidence of the occurrence of variant expressions and made possible a detailed comparative analysis of the linguistic and text-based factors affecting the choice of the variants. The computer has enabled scholars to study even such topics as have earlier been avoided for the sheer amount of material collecting involved and for the uncertainty as to the relevance of the results.

The Helsinki Corpus of English Texts, which forms the basis of the studies reported in this volume, was 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 gives reliable indicators concerning the structural and lexical developments of English for over a millennium. The results are, however, in many cases, only diagnostic; 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.²

Each text sample of the Helsinki Corpus is equipped with a battery of parameter values containing information on the text and its author, if known. In Old and Early Middle English, this information concentrates 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, etc. In these periods, sociolinguistic information is given on the authors of the texts and, in the case of letters, on the relationship existing between the writer and the receiver.

A large and structured corpus, equipped with textual parameter codings, has made it possible for the authors of the present volume to combine qualitative and quantitative analysis of changes through variation. We have been able to discuss the process of change with reference to different genres or text types within a long time-span, 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 parameter coding and broad textual basis has also allowed us to ob-

serve the role played by dialectal distribution and the differences between prose and verse texts in the survey of our Old and Middle English variant forms. In the discussion of late Middle and Early Modern evidence, special emphasis has been placed on the distributions shown by speech-based texts, or texts showing a relatively high degree of orality.

5. Pronominalization

In pronominalization, the source concepts can either be lexical nouns as is the case with MAN, BODY and THING, or represent other, semantically less definite categories, such as the numeral ONE and the adjectival/pronominal SELF. As Heine et al. (1991: 33–36) and Hopper—Traugott (1993: 97) point out, source concepts are often derived from the basic vocabulary, and even from words belonging to the level of superordinate categories in Roschian terms, as with the lexemes meaning ‘person’ and ‘thing’. Hopper—Traugott’s comment (1993: 97) on the possibility of grammaticalization of formerly fairly specific terms after their generalization shows its validity in the development of BODY, which first meant ‘the material frame of man’, then ‘personal being, individual’ (*OED*, sv. *body*, I, III), the latter meaning becoming the source of the present-day indefinite pronouns in -BODY. On the other hand, the numeral ONE is commonly used in the creation of indefinite pronouns in different languages, which stresses its importance as part of the basic vocabulary along with the plurality quantifier MANY (Lehmann 1982: 52 and Heine et al. 1991: 33). The same is true of the word indicating SELF, which easily becomes part of compound reflexives in various languages.

The lexical and grammatical items which lend themselves to pronominalization typically function as noun phrases or noun phrase heads. There is a resemblance between regular noun phrases and pronouns, which makes it difficult to draw a borderline between them. Quirk et al. (1985: 335–336), who in terms of a prototype framework classify pronouns as central or peripheral, distinguish them from nouns according to different criteria, which as a whole apply to the central pronouns only. Semantically the meaning of pronouns is general and undetermined so that their interpretation depends to a large extent on what information is supplied by the context. Syntactically, most pronouns incorporate their own determiner. Morphologically, some pronouns have characteristics which noun

phrases do not have, such as a contrast between subjective and objective cases, a contrast between first, second and third persons, a distinction between personal and non-personal as well as between masculine and feminine gender, and morphologically unrelated number forms.

The prototypical core formed by elements such as personal pronouns seems relatively stable, while grammaticalization, increasing the number of items or replacing old ones, can be expected in the periphery, for instance among the indefinite and reflexive pronouns (see Raumolin-Brunberg 1994). On the other hand, pronominalization with nominal elements as sources can be characterized as weak grammaticalization or, as Heine et al. (1991: 44) put it, 'structure-preserving abstraction', a process which does not radically affect the categorical status of the linguistic structures concerned.

The character of pronominalization as a weak type of grammaticalization becomes evident if this process is compared with the metaphorically and metonymically conditioned grammaticalization processes appearing for instance in the development of connectives and different verbal systems (cf. e.g. Traugott—Heine 1991; Hopper—Traugott 1993). Cognitive and pragmatic concerns are, nevertheless, important in pronominalization. Subjectification (see e.g. Traugott—König 1991) plays a central role in the development of ONE, and a search for new, more expressive emphatic forms for the universal, negative or reflexive pronouns corresponding to Present-day English EVERYONE/EVERYBODY, NO ONE/NOBODY and MYSELF, YOURSELF, etc., has led to repeated grammaticalizations. Despite the criticism directed at the idea of semantic bleaching (see e.g. Hopper—Traugott 1993: 87–93), it appears to be a conspicuous phenomenon in the process where noun phrases become pronouns, since semantic emptiness is one of the main characteristics of pronouns.

Pronominalization often involves change at different linguistic levels. Among indefinites and reflexives, compounding easily leads to phonological changes, such as the stress placement in the development from ANY BODY to ANYBODY, or HIM SELF(UM) to HIMSELF. Phonological erosion can also take place, e.g. in Old English *ÆFRE ÆLC*, which has developed into *EVERY*. Structure-preserving grammaticalization does not involve major syntactic changes, but, as mentioned above, semantic fading is obvious, and in historical research it seems easier to tackle this topic than the elusive stress placement.

Three articles in the present volume deal with the development of the forms of pronouns. Helena Raumolin-Brunberg and Leena Kahlas-Tarkka discuss the development of the indefinite pronouns with singular human

reference from Old English to the end of the 17th century. Their study focuses on the gradual development and repeated grammaticalization of the complex forms of these pronouns during their long history. This development seems to have been connected with the tendency to mark the head of a noun phrase with a distinctive element, after the loss of the morphological gender and number markers of the adjective (Fischer 1992: 222). The same tendency probably influenced the development of the pronominal uses of ONE, as argued in Matti Rissanen's article. The development of the pronominal ONE also improved the cohesion of the sentence and helped in phrase boundary marking.

Kirsti Peitsara's article on the expression of reflexivity describes another compound pronoun development, the reanalysis of the combination of the personal pronoun and the optional emphasizing element SELF as reflexive pronoun. The factors influencing the rise of the reflexive pronouns are, however, different from those of the compound indefinites: the rise of compound forms may have been supported by the heavy functional load of the oblique forms of the personal pronouns (in addition to 'ordinary' pro-forms, they were used as the so-called possessive or sympathetic dative, as dative of interest, etc.). The study of reflexive pronouns also links pronominalization with verb complementation and the variation of the pronominalized forms with zero forms or with middle voice.

6. Adverbialization

The term *adverbialization* is used here collectively to refer to the various processes of adverb formation in the history of English. Insofar as adverbs can be considered a more grammaticalized category than the other open-class categories, adverbialization is also a process of grammaticalization. This view is morphologically justified by the derived status of adverbs, and the fact that few of them can be inflected (Traugott 1988: 132–134).

Most adverbs in the world's languages are indeed derived from other word classes, notably from adjectives, nouns and verbs. Despite their status as a derived category, adverbs must nonetheless be considered open-class lexical items: their number can be freely augmented. As with other open-class categories, more grammaticalized subsystems may arise within the adverbial category as a result of secondary grammaticalization,

that is, further functional specialization and semantic change. This process produces, for instance, deictic compound adverbs (HEREAFTER, THEREUPON, WHEREIN), which themselves consist of closed-class pronominal adverbs (HERE, THERE, WHERE) combined with a preposition (or, from the Old English point of view, prepositional adverb). It also yields focusing adverbs (MERELY, JUST, EVEN), many of which arise from prior intensifier or focusing adverb homomorphs. Neither adverb class can be augmented at will nor can their members be inflected (see Nevalainen 1991: 10–18).

Adverbs can be divided into complex, compound and simple on the basis of their morphological make-up. The vast majority of Present-day English adverbs are morphologically complex. Most of them are formed by adding the suffix -LY to an adjectival base. The process typically produces adverbs of manner ('in an X way', where X corresponds to the adjectival base: RAPID > RAPIDLY). Other adverbial suffixes include -WARD(S) and -WISE (see e.g. Quirk et al. 1985: 438–439, Huddleston 1984: 332–334). Suffixation is by far the most common means by which adverbs were also formed in the past.

Swan (1988) suggests that -LY adverbs provide one of the most important sources for the formation of different kinds of sentence adverb in English. They include various evaluative and modal subclasses and arise from verb-modifiers and intensifiers; compare, for instance, *they denied it frankly* ('in a frank manner') and *frankly, they denied it* ('speaking with candour'). These sentence adverbs result, in our terms, from a process of secondary grammaticalization, which does not alter their category status as adverbs but brings about a change in their function. As they incorporate the speaker's comment on the proposition that they are associated with, sentence adverbs are more subjective and hence more grammaticalized than their verb-modifying homomorphs. Sentence adverbs may also arise through primary grammaticalization, category shift, as in the case of *you know*, *I think* and the by now archaic *methinks*. The process involves routinization and idiomatization of these expressions, which as a result come to be stored and used as simple units (see Hopper—Traugott 1993: 65, 201–203; Palander-Collin 1996). Unlike some other adverb classes, the English sentence adverbial category remains relatively open-ended.

As a means of primary adverbialization, zero-derivation is a much more marginal process in Present-day standard English than suffixation. Most lexicalized zero-adverbs are derived from adjectives. Diachronic study of zero-derivation suggests that in the past functional links between ad-

jectives and adverbs were perhaps closer than in today's standard language (Nevalainen 1994), although functional overlapping between the two classes is still common. This is typically the case of subject-related modifiers, where it is not always easy to determine whether the modifier is to be classified as an adjective or an adverb (*the sun shines **bright**, he fell **flat***). Functionally, zero-derived adverbs range from these weakly codified subgroups to some more specialized and grammaticalized ones, such as process and focusing adverbs and intensifiers (CLOSE, FAST, DIRECT, RIGHT, SLOW, WIDE; Quirk et al. 1985: 405–407).

Compared with the high productivity of the process of suffixation, compound adverbs are not very numerous in Present-day English. Some of them are morphologically opaque historical relics, such as the negative operator NOT, which goes back to the Old English noun phrase NAWIHT. It is questionable whether compounding with WHERE, THERE and HERE is still productive today. Many compound adverbs have become polyfunctional as a result of secondary grammaticalization. Many have also reduced their functional load in the course of time, including the conjuncts HOWEVER and THEREFORE (see Finell 1992 and Österman, below).

In Present-day standard English few adverbs inflect for comparison. Those that do are often formally identical with their adjective homomorphs and share their comparative and superlative inflections (for instance, FAST/-ER/-EST, HARD/-ER/-EST). Some -LY adverbs also have inflectional comparatives and superlatives in standard English (*it's **easier** said than done* 'more easily'; *speak **clearer*** 'more clearly'; Quirk et al. 1985: 406). Most adverbs in -LY were compared periphrastically (MORE/MOST X-LY) in earlier English, too. In terms of morphological fixedness compound adverbs constitute the more grammaticalized end of the adverb category. These adverbs typically do not allow any kind of comparison, inflectional or otherwise.

The three studies by Terttu Nevalainen, Aune Österman and Minna Palander-Collin explore these various processes of adverbialization in the history of English. In terms of productivity and grammaticalization they cover the two extremes of the adverb category from the relatively closed class of compounding at one end to the rather open-ended class of -LY suffixation at the other, with zero-derivation retreating towards the less productive end with time. Nevalainen concentrates on the gradual increase in the derivations with the -LY ending at the cost of zero-derivation in Late Middle and Early Modern English. She also calls attention, among other things, to the questions of the continuum between 'adjectiveness'

and 'adverbness' and the role played by the syntactic function of the adverb in the selection of the type of derivation. Österman's article traces the increasing popularity of the adverbs formed with *THERE*, such as *THEREFORE*, *THEREIN*, *THEREABOUTS*, in Middle English and their rapid decline in the Early Modern English period.

The study by Palander-Collin on *METHINKS* focuses on the primary process of adverbialization which grammaticalizes subject-verb collocations as sentence adverbs. Adverbialization of verb phrases such as *PLEASE*, *PRITHEE* and *YOU KNOW* has been attested from the Middle English period onwards and is not uncommon today (see, e.g., Erman 1987). In the course of time these phrases typically become reanalysed as sentence elements with various adverbial functions in speaker-hearer interaction denoting, for instance, politeness or the speaker's attitude to the proposition. In her article Palander-Collin studies the grammaticalization of *METHINKS* as a sentence adverb expressing the speaker's point of view.

Notes

1. This introduction is based on the ideas and suggestions of all the authors of the present volume. Section 5 (pronominalization) was compiled by Helena Raumolin-Brunberg and section 6 (adverbialization) by Terttu Nevalainen.
2. The authors have done their best to make their articles readable without detailed knowledge of the structure and conventions of the Helsinki Corpus. It is obvious, however, that familiarity with the Corpus will make the appreciation and estimation of the results of the articles easier. For information on the Helsinki Corpus, see Rissanen et al. (1993) and Kytö (1996).

When citing the examples from the Helsinki Corpus, we follow the typographical conventions 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.

The Helsinki Corpus of Older Scots (Meurman-Solin 1995) has decisively improved our possibilities for studying regional variation in late Middle and Early Modern English. At the English Department of the University of Helsinki, there are three other projects in progress which will, among other results, produce new diachronic corpora: the Corpus of Early English Correspondence (see Nevalainen—Raumolin-Brunberg 1996), the Corpus of Early English Medical Writing (see Taavitsainen—Pahta 1997) and the Corpus of Early American English (see Kytö 1993). The Penn-Helsinki Parsed Corpus of Middle English contains a syntactically coded version of the Middle English prose texts included in the Helsinki Corpus. A new ver-

sion of this corpus, with a more many-sided linguistic coding and additional prose texts, is in preparation. The Brooklyn-Geneva-Amsterdam-Helsinki Parsed Corpus of Old English will provide a linguistically coded and glossed version of the Old English section of the Helsinki Corpus.

Kytö et al. (1994), Kytö—Rissanen (1996) and Hickey et al. (1997) give information on other historical corpora of English, either completed or under preparation.

References

Erman, Britt

- 1987 *Pragmatic expressions in English: A study of you know, you see and I mean in face-to-face conversation*. Stockholm: Almqvist and Wiksell International.

Finell, Anne

- 1992 "The repertoire of topic changers in personal, intimate letters: A diachronic study of Osborne and Woolf", in: Matti Rissanen—Ossi Ihalainen—Terttu Nevalainen—Irmu Taavitsainen (eds.), *History of Englishes: New methods and interpretations in historical linguistics*. (Topics in English Linguistics 10.) Berlin: Mouton de Gruyter, 720–735.

Fischer, Olga

- 1992 "Syntax", in: Norman Blake (ed.), *The Cambridge history of the English language*. Vol. 2. Cambridge: Cambridge University Press, 207–408.

Halliday, M. A. K.

- 1973 *Explorations in the functions of language*. London: Edward Arnold. [Reprinted 1974].
- 1978 *Language as social semiotic. The social interpretation of language and meaning*. London: Edward Arnold.

Heine, Bernd—Ulrike Claudi—Friederike Hünemeyer

- 1991 *Grammaticalization: A conceptual framework*. Chicago: University of Chicago Press.

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.

Hopper, Paul J.—Elizabeth Closs Traugott

- 1993 *Grammaticalization*. Cambridge: Cambridge University Press.

Huddleston, Rodney

- 1984 *Introduction to the grammar of English*. Cambridge: Cambridge University Press.

Kytö, Merja

- 1993 "Early American English", in: Matti Rissanen—Merja Kytö—Minna Palander-Collin (eds.), 83–91.

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.

Lehmann, Christian

- 1982 *Thoughts on grammaticalization: A programmatic sketch*. (Arbeiten der Kölner Universalien-Projekts 1.) Cologne: Universität zu Köln, Institut für Sprachwissenschaft.

Meurman-Solin, Anneli

- 1995 "A new tool: The Helsinki Corpus of Older Scots (1450–1700)", *ICAME Journal* 19: 49–62.

Milroy, James

- 1992 *Linguistic variation and change: On the historical sociolinguistics of English*. (Language in Society 19.) Oxford: Blackwell.

Milroy, James—Lesley Milroy

- 1985 "Linguistic change, social network and speaker innovation", *Journal of Linguistics* 21: 339–384.

Nevalainen, Terttu

- 1991 *BUT, ONLY, JUST: Focusing adverbial change in Modern English 1500–1900*. (Mémoires de la Société Néophilologique de Helsinki 51.) Helsinki: Société Néophilologique.

- 1994 "Diachronic issues in English adverb derivation", in: Udo Fries—Gunnel Tottie—Peter Schneider (eds.), *Creating and using English language corpora*. Amsterdam: Rodopi, 139–147.

Nevalainen, Terttu—Helena Raumolin-Brunberg (eds.)

- 1996 *Sociolinguistics and language history: Studies based on the Corpus of Early English Correspondence*. Amsterdam: Rodopi.

Palander-Collin, Minna

- 1996 "The rise and fall of METHINKS", in: Terttu Nevalainen—Helena Raumolin-Brunberg (eds.), 131–149.

- Quirk, Randolph—Sidney Greenbaum—Geoffrey Leech—Jan Svartvik
 1985 *A comprehensive grammar of the English language*. London: Longman.
- Raumolin-Brunberg, Helena
 1994 "Prototype categories and variation studies", in: Francisco Fernández—Miguel Fuster—Juan José Calvo (eds.), *English historical linguistics 1992*. (Current Issues in Linguistic Theory 113.) Amsterdam: John Benjamins, 287–303.
- 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: Mouton de Gruyter.
- Romaine, Suzanne
 1982 *Socio-historical linguistics, its status and methodology*. (Cambridge Studies in Linguistics 34.) Cambridge: Cambridge University Press.
- Samuels, M. L.
 1972 *Linguistic evolution, with special reference to English*. Cambridge: Cambridge University Press.
- Swan, Toril
 1988 "The development of sentence adverbs in English", *Studia Linguistica* 42/1: 1–17.
- Taavitsainen, Irma—Päivi Pahta
 1997 "Corpus of Early English Medical Writing 1375–1750", *ICAME Journal* 21: 71–78.
- Traugott, Elizabeth Closs
 1988 "Is internal semantic-pragmatic reconstruction possible?", in: Caroline Duncan-Rose—Theo Vennemann (eds.), *On language, rhetorica, phonologica, syntactica: A festschrift for Robert P. Stockwell*. London: Routledge, 128–144.
- 1989 "On the rise of epistemic meaning in English: An example of subjectification in semantic change", *Language* 65: 31–55.
- Traugott, Elizabeth Closs—Bernd Heine (eds.)
 1991 *Approaches to grammaticalization*. 2 vols. Amsterdam: Benjamins.
- Traugott, Elizabeth Closs—Ekkehard König
 1991 "The semantics-pragmatics of grammaticalization revisited", in: Elizabeth Closs Traugott—Bernd Heine (eds.), 1: 189–218.
- 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: University of Texas Press, 95–195.

Indefinite pronouns with singular human reference

Helena Raumolin-Brunberg and Leena Kahlas-Tarkka

1. Introduction¹

The object of this investigation is a subgroup of pronouns, i.e. items that in traditional terms are called indefinite pronouns with singular human reference, e.g. SOMEONE, ANYBODY, EVERYONE and NOBODY. The traditional characterization is not necessarily accurate, since there are serious doubts about the indefiniteness of some of the items (for further discussion, see 2.2 below). Nevertheless, since it is difficult to invent a more pertinent name for this subcategory of pronouns, we will use the well-established term 'indefinite pronouns'.

The basic data for this study were retrieved from the Helsinki Corpus of English Texts (see Kytö 1996; Rissanen et al. 1993), supplemented by other texts, e.g. the Toronto (Healey—Venezky 1980), Shakespeare (Wells—Taylor [eds.] 1989), and Century corpora (Milić 1990) and a Present-day English Bible translation (*Revised English Bible*, REB, 1989). The full inventory of the pronouns discussed is presented in Table 1. The repertoire and its subdivision follow Quirk et al.'s discussion of indefinite pronouns (1985: 376–392). As time wears on, the four paradigms, assertive 'someone', nonassertive 'anyone', universal 'everyone' and negative 'no one' become more or less symmetrical, all of them having compound variants (ending in MAN, ONE or BODY) and simple items, which in later times became similar to the corresponding determiners, except for NONE (determiner NO). The simple pronouns are called OF-pronouns by Quirk et al. (1985: 379), since in Present-day English they occur in partitive phrases. Semantically they are quantifiers, although this study does not adhere to the logic-based formal tradition from which this term has been derived.

Table 1. Indefinite pronouns with singular human reference.

	Assertive 'someone'	Non-assertive 'anyone'	Negative 'no one'	Universal 'everyone'
Present-day English				
Simple		ANY	NONE	EACH
-ONE	SOMEONE	ANYONE	NO ONE	EVERYONE
-BODY	SOMEBODY	ANYBODY	NOBODY	EVERYBODY
Early Modern English (1500–1700)				
Simple	SOME	ANY	NONE	EACH/EVERY
-MAN	SOME MAN	ANY MAN	NO MAN	EACH/EVERY MAN
-ONE	SOME ONE	ANY ONE		EACH/EVERY ONE
-BODY	SOME BODY	ANY BODY	NO BODY	EVERY BODY
Middle English (1150–1500)				
Simple	SOME	ANY	NONE	EACH/EVERY
-MAN	SOME MAN	ANY MAN	NO MAN	EACH/EVERY MAN
-ONE				EACH/EVERY ONE
-BODY	SOME BODY		NO BODY	
Old English (–1150)				
Simple	SUM	ÆNIG	NÆNIG NAN	GEHWILC/GEHWA ÆGHWILC/ÆGHWA ÆLC ÆGÐER/ÆGHWÆÐER
-MAN	SUM MAN	ÆNIG MAN	NAN MAN	-MAN

The time span of this study is very long, from Old English to the beginning of the eighteenth century. The differences between the data available from different subperiods are considerable, affecting many aspects of the study, from the number of occurrences to the possibilities for textual comparisons. Although our main purpose has been to offer a picture of the longitudinal development of the indefinite pronouns, we have had to be relatively cautious in our analyses of the earliest periods in the history of the English language.

2. The background and theoretical framework

2.1. Method and borderlines

The research was carried out within the variationist framework, the items being chosen to form paradigms in which the members were interchangeable in several, if not all, environments. This procedure made it possible to perform quantitative comparisons along with qualitative ones. Both linguistic and extralinguistic factors constraining the choice of variants could be tackled by this method. Comparisons were carried out not only within the paradigms but between them as well.

As Raumolin-Brunberg (1994a) points out, delimitation of paradigms based on syntactic categories is often problematic because of their prototypical nature. Prototypical and peripheral items may be semantically so different from each other that it is difficult to find underlying constants which would allow items to be regarded as variants of the same set. There were particular problems in this study both at type and token level.

At the type level it was difficult to decide which items to include in the paradigms. The study includes only pronouns containing the indefinite element SOME, ANY, EVERY, EACH and NO. Types like WHOEVER (Old English SWA HWA SWA) and corresponding relative constructions such as SE ðE in Old English and later HE THAT/WHO, as well as reciprocal pronouns have been excluded or touched upon only superficially. The Old English interrogatives HWA or HWILC, also used as indefinites, have been excluded, as they became extinct at quite an early stage, and only those pronouns showing continuity from Old English to Early Modern English have been included. Apart from traditional pronouns, we included such lexical items as later developed into pronouns, e.g. BODY when it was preceded by SOME, ANY, EVERY and NO. These forms may be called prepronouns at the stages preceding prominalization.

The phrases with MAN were particularly problematic. We decided to select the forms with MAN which correspond to Present-day English compound pronouns, e.g. ANY MAN, NO MAN, etc. but not the item A MAN, which Poutsma ([1966]: 1205) includes in the group of quasi-indefinite pronouns. In the discussion of the Bible translations it became evident that A MAN was an alternative to the pronouns in the assertive 'someone' series. We felt, however, that A MAN had too much of its semantic content left to be included among pronouns. It seems to parallel the present-day use of A PERSON in reference to indefinite human beings.

As far as tokens are concerned, analysing which occurrences of SOME, ANY and NONE were singular and which plural was problematic, since only the former were to be included. Also, as regards the tokens of MAN, instances of full semantic content, such as ‘male human being’, ‘human being as opposed to God’, or ‘someone’s servant’ had to be excluded (see also Raumolin-Brunberg 1994b). There is no way of avoiding some degree of subjectivity in deciding where the borderlines lie between full nouns and pronouns. Therefore we would like to point out that the quantification in this study is not to be taken to represent any ‘absolute truth’. Other scholars might make different decisions in the interpretation of the less clear cases. We believe that the figures nevertheless give a good picture of the general developments in the long history of the pronouns in question.

Examples (1) and (2) illustrate the problem of number. The occurrences of SOME in example (1) represent the singular on the basis of the verb form,² while example (2), not included in the corpus, is indeterminate, since neither the verb form nor pronouns or other contextual features reveal the number. Example (3) is a passage of Wyclif’s New Testament. While the form with MAN is included in the corpus as it was later translated by the pronoun NO ONE, it is difficult to say whether *any man* in example (4) or *sumum menn*, *ælces monnes* and *sumes monnes* in (5) refer to a person in general or a male human being only. These examples are excluded from the present corpus.

- (1) What say you to this? that the witches haue their spirits, *some* hath one, *some* hath more, as two, three, foure, or fiue, *some* in one likenesse, and *some* in another, as like cattles, weasils, toads, or mise,... (EModE2 Gifford, *A Dialogue Concerning Witches and Witchcraftes* B4V)³
- (2) And on the Thursday the said Lordes fearing the people, imagined howe to escape out of theyr handes, and caused *some* of their seruauntes to sette fire on certaine houses in the Towne, thinking that... (EModE2 Stow, *The Chronicles of England* 546)
- (3) But we knowun this man, of whennus he is; but whanne Crist schal come, *no man* woot of whennus he is. (ME3 Wycliffe and Purvey, *The New Testament John* 7 27)

- (4) ... wher it is wel known by alle manere of euydences that they konne shewe for hem self or *any man* for hem that they ben fre tenent₃ and fre holderes to the kyng in chief... (ME3, *Petitions in A Book of London English 1384–1425* 201)
- (5) Bisna þe be *sumum menn*, for þam þe *ælces monnes* lif bið *sumes monnes* lar. (OE3/4 *Dicts of Cato* 10)
 ‘set someone as an example for yourself, because everyone’s life is someone else’s source of learning’

2.2. Indefiniteness

The discussion of definiteness has a long tradition, both in philosophy and linguistics (see e.g. Donnellan [1971]; Searle 1969: 26–27; Lyons 1977: 174–192; Hawkins 1978; Givón 1984: 398–412; Reuland—ter Meulen [eds.] 1987; Chesterman 1991). Indefiniteness has raised far less interest; mostly it has represented the binary opposition of definiteness. In later studies definiteness has been considered to be of a composite nature and a continuum, and there are scholars who have claimed that indefiniteness in fact is a more complex phenomenon than its apparent opposite (Givón 1984: 431; de Jong 1987: 271; Chesterman 1991: 40).

The traditional characterization of definiteness and indefiniteness could read as follows: a definite NP has a referent which is assumed by the speaker to be unambiguously identifiable by the hearer; an indefinite NP has a referent which is assumed by the speaker not to be unambiguously identifiable by the hearer (Chesterman 1991: 10). In other words, definite NPs represent familiarity, indefinite ones new information. Although these principles work in a large number of cases, they have their problems, especially in the treatment of nonreferential and uncountable nominals.

In many analyses, the concepts of definiteness, specificity and referentiality have not been kept apart. Definiteness has been defined as “reference to a specific individual or class” (Lyons 1977: 185) or in terms of having “a particular individual in mind” (Ihalainen 1974: 29). The problem here is that both referentiality and specificity form continuums which intersect that of definiteness; in other words, both definite and indefinite NPs can be specific and nonspecific, referential and nonreferential (see e.g. Ihalainen 1974: 81; Huddleston 1984: 254).

- (6) The murderer of Smith must be insane.
- (7) John is the acme of courtesy.
- (8) John is a good man.
- (9) In came a young man.

Examples (6)–(9) illustrate the phenomenon. Example (6) from Donnellan ([1971]: 103–104) and Chesterman (1991: 11) has two readings since either the speaker has a specific individual in mind or not. In the former case the NP is both referential and specific, in the latter nonreferential and nonspecific. Examples (7)–(8) from Chesterman (1991: 12) show that subject complements which are nonreferential, can be either definite or indefinite. Example (9), in turn has a referential and specific indefinite noun phrase as subject.

Hawkins (1978) introduces a pragmatically-oriented theory of definiteness, which he calls location theory. Its main components are shared sets and inclusion versus exclusion. In the case of definiteness the speaker instructs the hearer to locate the referent(s) in some shared set of objects (which are broadly defined) and refers to the totality of the objects or mass within the set (=inclusion). As regards indefiniteness, the speaker is uncommitted about the shared set, but refers to a proper subset, i.e. not-all, of the potential referents (=exclusion; Hawkins 1978: 167–187). In other words, in the case of exclusion there should be at least one other referent that could potentially be referred to by the same expression. This approach bears a resemblance to Givón's (1984: 434–435) claim that new arguments are introduced into the universe of discourse as one member of many within the type. Hawkins (1991) develops these ideas further in relation to recent approaches to pragmatics. The article revises concepts and terminology, but the basic arguments of location theory appear to hold.

Without going too deep into the discussion of the theory of sets, Figure 1 presents a simple model to illustrate the approach used in this study (see Suojanen 1977: 27 and Kahlas-Tarkka 1987: 89). It goes without saying that the size of the set can vary a great deal from two to all people in the world. In text the set may be expressed explicitly, either grammatically, e.g. by an OF-phrase (example (2), *some of their seruauntes*), or lexically (example (19), *The boys*). It can also be inferred from the context, so that in example (28) the set consists of those people who were around, among whom *some body* should call the speaker's wife.

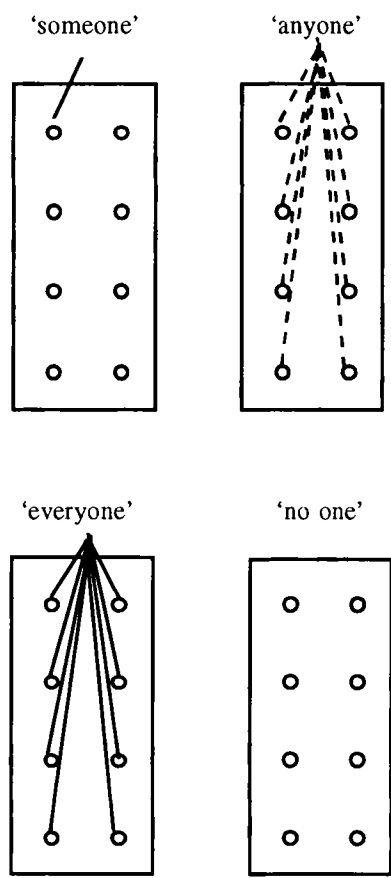


Figure 1. Indefinite pronouns in a framework of sets.

Although its approach to the nature of sets is different,⁴ it is interesting to look at Figure 1 in terms of the location theory. While the assertive and nonassertive pronouns clearly fulfil the condition [+indefinite], because there are other referents that could be referred to by the use of the same pronoun, the case is different with the universal paradigm. The universal pronouns behave exactly the way definite expressions do in location theory, referring to all referents within the set.⁵ Suojanen's (1977: 32) apt name for this class of pronouns is inclusive pronouns, or pronouns of totality. Jespersen (1933: 184–187) distinguishes a nonindefinite category of pronouns, those of totality, which includes two subgroups, positive and negative pronouns (ALL, BOTH, EVERY, EACH, and NO, NONE, NEITHER). It is also interesting to observe that Safir (1987: 71) points out how EVERYONE behaves like definite NPs in existential *there*-clauses, while SOMEONE is grouped with indefinites. Pesetsky (1987: 103) also shows some differences between EVERY and the indefinites.

Figure 2 shows that restrictive modification creates subsets. The set in example (10) consists of all the people who overheard certain people talk, and one of them is picked out. The set in (11) comprises those people who were in the church.

- (10) The silly old man willingly paid his penny before hand, and was going ouer; but *some* that ouer-heard their talk, hindered his journey and laughed at the jest, yet pitied his simplicities, and sets him in the right way. (EModE2 Armin, *A Nest of Nimmies* 42)

- (11) Dr. Bolton preached and that you may know that Mr. Hannington is of no ordinary esteeme, I will assure thee hee had such a presence with him as though a stranger to *every one* in the Church, hee was ushered in from his standing amongst severall gentlemen and seated next to the Bishop of London himselfe ... (EModE3 Henry Oxinden, *The Oxinden and Peyton Letters* 277)

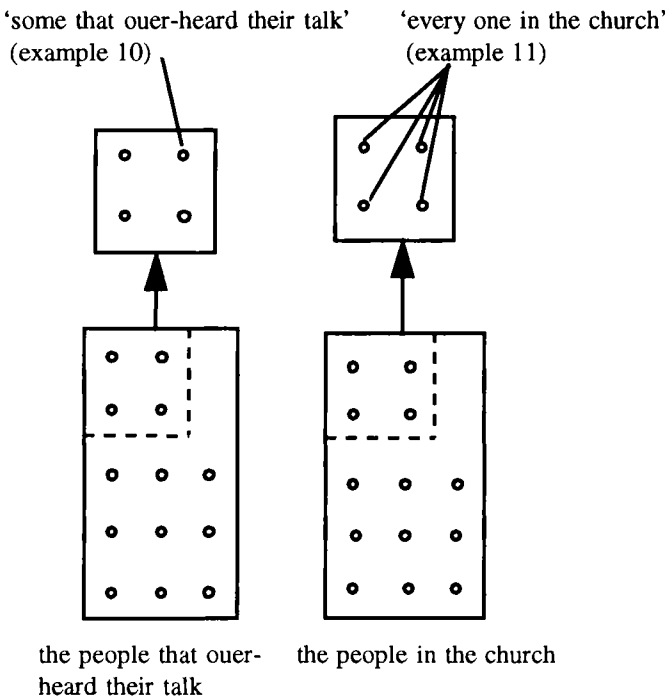


Figure 2. Indefinite pronouns in a framework of subsets.

2.3. Specificity

It seems that the common denominator for these pronouns is the possibility of nonspecific reference. In fact it appears to be specificity, not definiteness that is tested by analysing whether the speaker has a particular or specific individual in mind.

Quirk et al. (1985: 391) argue that the basic difference between the assertive 'someone' and nonassertive 'anyone' series is that the former is specific, though unspecified, and the latter nonspecific. It is, however, not difficult to find examples of the nonspecific use of SOMEONE (examples (12)–(13) to be compared with (10); see also Lyons 1977: 189).

- (12) I found a wallet in the bus. *Someone* must have left it behind.
[-specific]
- (13) And there should be two or three roomes made a little remote from the dwelling house, to which Scholars may be removed and kept apart, in case they be sick, and have *some body* there to look to them. (EModE3 Hoole, *A New Discovery of the Old Art of Teaching Schoole* 226) [-specific]

There are languages where this distinction has been grammaticalized, e.g. Russian has two different pronouns for SOMEONE: *kto-nibud* [-specific] and *kto-to* [+specific]. In Finnish the pronouns *eräs* (standard) and *yksi* (colloquial) tend to be used with specific reference, while *joku* is nonspecific (see also Rissanen 1987).

The distinction between specific and nonspecific reference may, in broad terms, be expressed as follows: if the speaker has a specific individual in mind but the hearer does not know the identity, the reference is specific; if neither the speaker nor the hearer knows the identity of the individual in question, the reference is nonspecific.⁶

2.4. Genericness

Genericness is usually attributed to NPs referring to whole species or subspecies, although their referentiality has been questioned by some scholars who argue that only specific nouns refer (cf. section 2.5, below). At a more general level, genericness concerns propositions which are not bound to specific spatio-temporal situations (Davison 1981: 361).

For many scholars genericness is a property of noun phrases, for others (e.g. Lyons 1977: 194) of whole propositions. In this study our focus is the pronominal noun phrase, although the context inevitably plays an important part in the analysis and semantic interpretation. According to Suojanen (1977: 42), the fact that someone does something often, usually, or always gives rise to different degrees of genericness. What is important is that genericness does not necessarily mean that whole species or sub-species are always involved; the class or type can be relatively limited instead. Restrictive modification creates subsets here too (Chesterman 1991: 80). As Givón (1984: 265) points out, specific individuals may indulge in generic activities in generic times and places (example (14)), and generic groups may together indulge in the same activity at the same specific time and place (example (15)).

(14) John always smokes wherever he goes.

(15) Everybody held their breath as the news came in.

Generic nominals can be either definite singular, indefinite plural or indefinite singular (examples (16)–(18) from Ihalainen 1974: 51). The choice between them is prompted by several syntactic and semantic factors, but there are also a large number of contexts where they are interchangeable (Ihalainen 1974: 51–52; Chesterman 1991: 78).

(16) The beaver builds dams.

(17) Beavers build dams.

(18) A beaver builds dams.

As regards indefinite pronouns, it is the nonassertive, universal and negative pronouns that tend to occur with generic reference, while the assertive paradigm is usually nongeneric. The pronouns referring to all people in the world may be given a generic interpretation (example (3)). The genericness of these pronouns, of course, resembles that of indefinite singular NPs (example (18)), and they cannot be used with verbs denoting the existence of species, etc.

2.5. Reference

Referentiality, like so many other linguistic phenomena, can be seen as a cline (Givón 1984: 390, 430; Chesterman 1991: 188). The following discussion takes up only some aspects of this extensive topic.

While many of the central pronouns are mostly used with textual reference, it is seldom the case with the indefinite pronouns, in particular the compound pronouns (Suojanen 1977: 19–20). If anaphoric relations appear at all, the question is usually not one of coreference, but whether the pronoun has the set as antecedent (example (19)). The use of the pro-forms is an exception, however (example (20)), for a similar use of ONE, see Rissanen's type 'substitute', this volume).

(19) The boys rushed in. Everyone was dirty.

(20) So that he bore ill will to no Person, nor hated *any* upon personal accounts. (EModE3 Burnet, *The Life and Death of John Earl of Rochester* 144)

According to Givón (1984: 390) and Ihalainen (1974: 67–71), among others, referentiality is a property of specific NPs only. If this analysis were to be applied here, very few pronouns could be considered referential and Figure 1 would be impossible. In an earlier study, Givón (1978: 293) claims that a referential NP "involves, roughly, the speaker's intent to 'refer to' or 'mean' a nominal expression to have non-empty references, i.e. to 'exist' within a particular universe of discourse." If we look at the matter from the angle we did in Figure 1, it becomes obvious that the existence of the sets and members can be presupposed. The smaller the set is, the more obvious is its existence. The result is that we regard as referential even pronouns which do not select a specific individual (see Suojanen 1977: 27). However, as for negative forms the analysis is problematic. It is of course difficult to claim that the subjects in (21) and (22) are referential, but we can still claim that the sets exist. The existence of a small definite set as in (22) may make it easier to accept this view of referentiality.

(21) No one saw him.

(22) None of them saw him.

There is a clear connection between referentiality and syntactic function. According to Givón (1990: 900), the only clausal participant types that are very likely to be nonreferential are subject complements and expressions of manner or instrument. It is important to notice that the pronouns under examination rarely occur in these inherently nonreferential functions. Givón (1984: 391) also points out that propositional modality has a connection with referentiality. Under the scope of FACT modality (pre-supposition and realis assertion), nominal arguments can only be referential. Under the scope of NON-FACT modality (irrealis assertion, negative assertion), they can also be nonreferential.

2.6. Conclusion

To conclude we present the continuum for indefinite pronouns proposed by Davison (1981). The criterion used has been a scale of indefiniteness, based on the degree to which the word picks out referents from the universe of discourse. "Negative indefinites, which pick out no referent, would be at one extreme, while generics, which pick out indiscriminate referents of a given subclass, would be at the other. In between would be non-specific indefinites which pick out a referent without further identification, and specific indefinites, for which further identification could be supplied" (Davison 1981: 359). Given the similarities between generic and universal pronouns, it may not be inappropriate to classify the universal pronouns with the generics at the same end of the scale (for the similarities and differences, see e.g. Vendler 1967 and Jackendoff 1972).

3. The four paradigms of indefinite pronouns

3.1. Overall characteristics

Two types of the variants under scrutiny, the simple pronouns, including words like EVERY, etymologically a compound, and the compounds with MAN persist throughout the Helsinki Corpus data. The indefinite pronouns in -BODY and -ONE are newcomers, introduced in Middle English.

Both establish themselves in Early Modern English, although the pronoun with ONE finds its way into the negative paradigm relatively late. Table 2 gives an overall picture of the trends in the usage throughout the Helsinki Corpus material. Some figures in the table call for special attention. It is striking, but it is difficult to explain why there seems to be such a great difference between Old and Middle English. For SOME and ANY in particular, there is a considerable decrease from Old to Middle English, whereas NONE represents a contrary development. No obvious explanations are offered by the material, apart from the fact that it is not always comparable, especially in the earlier subperiods.

Relatively high frequencies may be found in some individual texts; out of the 46 instances of SOME in OE3 as many as 23 are found in two poems, *Christ* and *The Fortunes of Men*, and a third of the 120 examples of ANY in OE3 appear in Wulfstan's Homilies. The fact that some texts cause peaks in the frequencies may be somewhat misleading for general conclusions. The same genres are not represented throughout, and the dialectal background is somewhat different. No obvious linear continuum can be pointed out in the available material. On the other hand, there is an obvious rise in the popularity of the -BODY compounds in later subperiods and as obvious a decrease in the -MAN compounds.

As regards the syntax of the indefinite pronouns, it is obvious that they are freely modified by postmodifiers such as relative clauses, prepositional phrases other than OF-phrases, adverbs, BUT-phrases, adjectives and participles. The use of the partitive OF-phrase or the genitive plural form in Old English as a postmodifier (henceforth both usually termed OF-phrases) forms a dividing line between the two. Only the simple pronouns and compounds with ONE (available from ME1 onwards) can appear in the partitive OF/genitive structure. The other two of the pronoun types, compounds in -MAN and -BODY, often occur as unmodified subjects and objects, with reference to human beings in general.

Table 2. Indefinite pronouns with singular human reference. (Occurrences per 100,000 running words).⁷

	OE		ME1-2		ME3-4		EModE1-2		EModE3	
	N	%	N	%	N	%	N	%	N	%
SOME	12.8	80	5.2	73	1.0	28	2.9	53	0.6	11
SOME MAN	3.2	20	1.9	27	2.3	64	1.3	24	0.0	
SOME ONE	0		0		0		0.5	9	1.8	34
SOME BODY	0		0		0.3	8	0.8	14	2.9	55
Totals	16.0		7.1		3.6		5.5		5.3	
ANY	15.5	41	4.3	43	6.9	33	10.5	33	7.0	18
ANY MAN	22.5	59	5.7	57	14.1	54	20.5	63	12.9	32
ANY ONE	0		0		0		0.5	2	7.0	18
ANY BODY	0		0		0		0.5	2	12.9	32
Totals	38.0		10.0		21.0		32.0		39.8	
EACH/EVERY	46.5	65	19.5	33	15.6	39	7.1	24	8.8	34
-MAN	24.9	35	20.0	34	16.8	42	14.2	49	2.9	11
EACH/EVERY ONE	0		19.5	33	7.3	19	7.6	26	11.1	42
EVERY BODY	0		0		0		0.3	1	3.5	13
Totals	71.4		59.0		39.7		29.2		26.3	
NONE	13.6	32	42.3	58	19.3	36	9.5	26	6.4	27
NO MAN	28.8	68	30.1	42	34.9	63	26.3	72	3.5	15
NO BODY	0		0		0.8	2	0.8	2	13.4	58
Totals	42.4		72.4		55.0		36.6		23.4	

Although many grammarians argue that in Present-day English the compound pronouns with ONE and BODY are synonymous (Quirk et al. 1985: 376–377; Jespersen 1914: 444), Bolinger (1976, 1977) claims that there is a subtle semantic contrast between them. According to him (1976: 230), the compounds in -ONE may still reflect the numerical and pronominal values of ONE, closeness to the speaker and individualization, while those in -BODY are unmarked in these two senses. AN/ANRA also had a very special individualizing and emphasizing function with EVERY in Old English, which may still be reflected in this subtle semantic shade.

The early occurrences of the ONE compound, contrasting with ‘many’, and the syntactic priorities in expression or nonexpression of the set corroborate the effect of etymological conditioning on the behaviour of the different variants (see Nevalainen 1991: 256–259). Individualization may characterize both the indefinites with ONE and simple pronouns, while the forms with MAN and BODY are clearly favoured in general discussions of human beings, thus reflecting the original meaning of the headwords ‘human being’, ‘person’.

3.2. The assertive paradigm ‘someone’

The assertive set involves positive statements, in which these pronouns assert the truth of a proposition (Quirk et al. 1985: 83, 383).

3.2.1. The variants

The Helsinki Corpus data in absolute figures and percentages by subperiod are given in Tables 3a and 3b. The corresponding figures per 100,000 running words were given in Table 2 (subperiods conflated).

The simple form is the only one whose use extends over all subperiods, except OE1, and compounds with MAN appear until the late seventeenth century. The first instance of SOME BODY dates from ME4 and SOME ONE⁸ from EModE2. Examples (23)–(28) illustrate the different variants of this paradigm from different subperiods.

- (23) *Sum sloh mid slegge swiðe þa hæpsan, sum heora mid feolan feolode abutan, sum eac underdealf þa duru mid spade, sum heora mid hlæddre wolde unlucan þæt ægðyrl*; (OE3 *Ælfric's Lives of Saints* 32 328)
'One struck the hasps with a hammer, one of them tried to cut through with a file, one also dug under the door with a spade and one of them tried to open the window with a ladder'
- (24) ... and that is but litle to do you any good, for ther is but *some* that will len so long afor the tyme. (EModE1 Isabel Plumpton, *Plumpton Correspondence* 198)
- (25) Nu seið *sum mann*: 'Sceal ic luuiȝe ðane euele mann?' (ME1 *Vices and Virtues* 1 73)
'Somebody says then: "Do I have to love an evil man?"'
- (26) In which mater somtyme they seeme to haue dispensation, for that *som mans* nature is so headstrong & rash, that neede of necessities cause may make him fall into a mischeefe... (EModE2 Elizabeth I, *Queen Elizabeth's Englishings of Boethius, Plutarch, &c* 95)
- (27) But let us grant, that it is possible that *some one* may be able to distinguish betwixt the Good and the Bad... (EModE3 Preston, *Boethius* 195)
- (28) Set downe the basket villaine: *some body* call my wife... (EModE2 Shakespeare, *The Merry Wiues of Windsor* 4 2 102)

The most striking fact is the infrequency of occurrences. The Helsinki Corpus contains only 1 (ME2) to 15 (OE3) instances per 100,000 running words, OE2 and OE3 showing by far the highest frequencies (13 and 15 respectively). The Present-day English data in Tesch (1990: 61–69, 80, 85–86), drawn from the London–Lund and LOB corpora, testify to a far higher joint average frequency of the variants SOMEONE and SOMEBODY: 39 per 100,000 running words.

Table 3a. The assertive paradigm: 'someone'. (Absolute figures).

	OE1	OE2	OE3	OE4	ME1	ME2	ME3	ME4	E1	E2	E3
SOME	0	7	39	7	6	5	2	2	4	7	1
SOME MAN	0	2	7	4	3	1	4	5	4	1	0
SOME ONE	0	0	0	0	0	0	0	0	0	2	3
SOME BODY	0	0	0	0	0	0	0	1	2	1	5
Totals	0	9	46	11	9	6	6	8	10	11	9

Table 3b. The assertive paradigm: 'someone'. (Percentages).

	OE1	OE2	OE3	OE4	ME1	ME2	ME3	ME4	E1	E2	E3
SOME	0	78	85	64	67	83	33	25	40	64	11
SOME MAN	0	22	15	36	33	17	67	63	40	9	0
SOME ONE	0	0	0	0	0	0	0	0	0	18	33
SOME BODY	0	0	0	0	0	0	0	12	20	9	56

One can hardly expect less need to refer to singular indefinite human beings in assertions in past times than today, even though SOME in the plural is certainly commoner in the whole of the Helsinki Corpus. An explanation for the infrequency must be sought elsewhere, e.g. in the existence of alternative ways of saying the same thing. In this connection we would like to refer to the various uses of ONE (Rissanen, this volume), the noun A MAN and the phrase ANOTHER (MAN), in the sense 'someone else'. It is especially Rissanen's categories 'personal-specific' and 'personal non-specific' that represent the uses of ONE that are semantically equivalent with 'someone'. It is also a fact that the joint instances of SOME OTHER (MAN) and ANOTHER (MAN) for 'someone else' outnumber the occurrences of the paradigm members in all subperiods between OE2 and EModE3.

It is also interesting that the Bible extracts in the Helsinki Corpus do not include a single instance of the four variants of the 'someone' paradigm, although the corresponding passages of the *Revised English Bible* (1989)

have nine occurrences of SOMEONE. The relevant corpus versions have ANY MAN, A MAN, NO MAN and ANOTHER. There seems to have been a change here, but the data point to a post-1700 development.

The infrequency of the 'someone' series is also corroborated by the Shakespeare and Century corpora. The former corpus, including the Complete Works by Shakespeare (c. 900,000 running words when double versions of some texts are excluded), has only 24 occurrences. The Century Corpus, which contains ca 500,000 running words from 1680 to 1780, contains no more than 11 instances of the 'someone' paradigm.

3.2.2. *Syntactic and semantic properties*

The majority of the instances of the assertive paradigm appear as unmodified subjects or objects. Some are modified by relative clauses (SOME and SOME BODY). This is especially the case with existential sentences (example (1)). A genitive modifier may appear with simple SOME in Old English, as in (2). SOME MAN also occurs in the genitive, while SOME ONE is found with the partitive OF-phrase.

Although in the other three paradigms (see 3.3.2, 3.4.2 and 3.5.2) it is usually the simple forms that appear with partitive OF/genitive structures as modifiers, this is the case in only one instance of the clearly singular occurrences of SOME in the subperiods after Old English, where ten instances have been recorded (example (29)). On the other hand, the partitive OF-phrase is one of the most frequent modifiers of those instances of SOME that must be analysed as indeterminate in respect of number (excluded from the corpus). Explicit reference to one person belonging to a set was apparently made with the pronoun ONE (*one of them*) rather than with SOME. The singular interpretation of SOME disappeared later in this context, as elsewhere. Example (30) illustrates a case where explicit difference was made between the singular and the plural at the end of the 17th century.

- (29) And ye Judge said y^t was a great error & a mistake & *some* of ye Justices was in a rage & said whoe has donne this *some body* has donne it of purpose. (EModE3 Fox, *The Journal of George Fox* 80)⁹

- (30) ... give Notice in Writeing to the Master and Wardens of the Company of Frame-Worke-Knitters within the City of London or to theire lawfull Debuty or Debuties for the tyme being *some or one of them* to whom and what Place the same were soe sold disposed or removed... (EModE3 *The Statutes of the Realm* VII 98)

The classification of the ‘someone’ paradigm as assertive is well-founded, since its members mostly occur in affirmative declarative sentences. SOME BODY and SOME ONE are also found with imperative verb forms (example (28)). The simple form SOME occurs once in a conditional clause (for further discussion, see section 3.3.3, below).

The set is expressed either syntactically, or it can be inferred from the context. Cases without an explicit set, in other words where people in general are taken to form the set, appear during all periods, especially in religious writing. SOME MAN favours this usage.

As regards specific versus nonspecific reference, both the simple form and SOME BODY appear with either type, while SOME MAN and the few examples of SOME ONE refer to nonspecific individuals. If SOME MAN appears with specific reference, the reference is normally to a male person, as in (31), but these are excluded from the present study. It is also important to notice that there are cases where both the speaker and the hearer know the identity of the person referred to, as in example (32), but still avoid making direct reference. The reasons for this usage are apparently pragmatic.

- (31) cwom *sum mon* in Norðanhymbra mægðe (‘country’); wæs his noma Eomær (OE2 *Bede’s Ecclesiastical History* 122)
- (32) ... hee doth mee twice as much hurt as good; *some bodie* hath incensed Him very much against mee, you may quessee who hath done it, the partie being not far from you. (EModE3 Henry Oxinden, *The Oxinden and Peyton Letters* 273)

3.3. The nonassertive paradigm ‘anyone’

While the assertive set is connected with expressions asserting the truth of a proposition, the nonassertive forms are not. They mostly occur in spe-

cific structures such as negative statements, questions and conditional clauses (Quirk et al. 1985: 83–84, 389–391; see 3.3.2, below).

3.3.1. *The variants*

As Tables 2, 4a and 4b show, both the simple form ANY and the compound ANY MAN are found all through the data, only ANY MAN missing in OE1. The first instance of ANY ONE dates back to the early days of Early Modern English, and that of ANY BODY to somewhat later times. The growth of the frequency of the BODY variant is particularly conspicuous, from three per cent in EModE2 to 31 per cent in EModE3. The proportion of the MAN compound is high until the latter half of the seventeenth century. Examples of the four variants are found in (33)–(41).

Table 4a. The nonassertive paradigm: 'anyone'. (Absolute figures).

	OE1	OE2	OE3	OE4	ME1	ME2	ME3	ME4	E1	E2	E3
ANY	1	6	52	5	5	4	9	18	17	23	12
ANY MAN	0	14	68	11	6	6	37	19	42	36	22
ANY ONE	0	0	0	0	0	0	0	0	1	1	12
ANY BODY	0	0	0	0	0	0	0	0	0	2	22
Totals	1	20	120	16	11	10	46	37	60	62	68

Table 4b. The nonassertive paradigm: 'anyone'. (Percentages).

	OE1	OE2	OE3	OE4	ME1	ME2	ME3	ME4	E1	E2	E3
ANY	100	30	43	31	45	40	20	49	28	37	18
ANY MAN	0	70	57	69	55	60	80	51	70	58	32
ANY ONE	0	0	0	0	0	0	0	0	2	2	18
ANY BODY	0	0	0	0	0	0	0	0	0	3	32

- (33) ne mæg ænig twæm godum ðeowigan
 (OE3 *Rushworth Gospels Matthew* 6 24)
 'no one can serve two gods'