Negation in Non-Standard British English

Gaps, regularizations and asymmetries

Lieselotte Anderwald

ROUTLEDGE
Taylor & Francis Group

London and New York

Negation in Non-Standard British English

Despite the advances of radio and television, despite increasing mobility and urbanization, spoken English is by no means becoming more like the written standard. English dialect grammar, however, is still a new and relatively undeveloped area of research, and most studies to date are either restricted regionally, or based on impressionistic statements. This book provides the first thorough empirical study of the field of non-standard negation across Great Britain.

Based on the British National Corpus, this book investigates a range of morphosyntactic features of negation that can be found in everyday spoken language. Employing the relatively new method of investigating dialect grammar with the help of large corpora, it provides in-depth analyses of contraction types, multiple negation, *ain't*, third-person *don't*, and *wasn't/weren't*.

This book also utilizes a new theoretical frame: the framework of functional typology. Linking variation in one language with variation between languages in this way provides a powerful instrument for the interpretation of dialect data. A comparison of non-standard patterns with cross-linguistic trends and regularities suggests functional explanations that account for the fact that non-standard forms are still spreading, despite the growing influence of the standard variety.

The data discussed in this book offer a new understanding of regional and social variation in British English. The research is relevant for both the construction of a functional theory of language within typology, and the description of individual varieties of languages in dialectology. As such, it will be of essential interest to academics and researchers in the fields of sociolinguistics, syntax and morphology, corpus linguistics, typology, dialectology and English linguistics.

Lieselotte Anderwald is Assistant Professor of English Linguistics at the University of Freiburg, Germany.

Routledge Studies in Germanic Linguistics
Series Editors: Ekkehard König and Johan van der Auwera

Negation in Non-Standard British English

Gaps, regularizations and asymmetries

Lieselotte Anderwald

To Lucian and Eva

Contents

Illustrations

Figures

Maps

Tables

Acknowledgements

Abbreviations

AAVE	African American Vernacular English
AdvP	Adverb Phrase
AmE	American English
AP	Adjective Phrase
aux	auxiliary
aux-neg	auxiliary negation
BNC	British National Corpus
BrE	British English
BROWN	Brown corpus
cop	copula
d	deontic
decl	declarative
e	epistemic
EModE	Early Modern English
FLOB	Freiburg LOB corpus
FRED	Freiburg English Dialect corpus
FROWN	Freiburg Brown corpus
gen	generalization
int	interrogative
IrE	Irish English
LOB	London/Oslo/Bergen corpus
ME	Middle English
Mid	Midlands
ModE	Modern English
MV-neg	Main Verb negation
N	North
neg	negation/negator/negative
negattrac	negative attraction
NITCS	Northern Ireland Transcribed Corpus of Speech
NORM	Non-mobile, Older, Rural Male
NP	Noun Phrase
NPI	Negative Polarity Item
O	Object

OE	Old English
PP	Prepositional Phrase
S	Subject
ScE	Scottish English
SED	Survey of English Dialects
SpS	Spontaneous Speech
SRLM	Somerset Rural Life Museum
StE	Standard English
SW	Southwest
V	Verb
V_{fin}	finite Verb
V_{lex}	lexical Verb
VP	Verb Phrase

Logic symbols

~	not, no
⊃	if . . . then (material implication)
∀	all/every
◇	permission/possibility
□	obligation/necessity

Transcriptions

‖	end of utterance (=speaker change)
		end of clause

Paradigms of verbs are denoted by small capitals (e.g. BE, DO).
Verb forms are denoted by italics (e.g. *am*, *is*, *does*, *be*).

1   Introduction

Dialect grammar

Despite the advances of radio and television, despite increasing mobility and urbanization, spoken English is by no means becoming more like the written standard these days. Although it is true that the old, especially isolated and rural, dialects are dying out, they do not seem to be replaced wholly by standard forms, as many dialectologists and philologists of the past feared. Instead, newer, different, but still non-standard forms can increasingly be heard, and despite all prescriptive uproar and outcries, these non-standard forms are alive and well, and seem quite unstoppable.[1] These often newer developments away from the standard have sometimes been captured under the heading of 'dialect levelling':[2] increasing contact between speakers of different dialects leads to both sides adapting their speech and erasing particularly salient features, and developing new features that are surprisingly pervasive. Although these grammatical features are clearly recognizable as non-standard, and although at least some of them are strongly stigmatized, there is no sign that their use is dwindling under the influence of standard English. The question of what constitutes this remarkable robustness of non-standard forms is the main motivating factor for the following investigation. Several possible paths of motivation have been pursued so far; in particular, in sociolinguistics the notion of 'covert prestige' plays an important role in explaining why heavily stigmatized forms are nevertheless used by their speakers (cf. Labov 1966; Trudgill 1974). Social factors like 'solidarity' or even group pressure should also by no means be discounted (for analyses in terms of network strength cf. Milroy 1980: 19). The main argument of this book will be, however, that there are other – cognitive – factors that have not been considered so far, but that can be shown to play an (additional) important role.

In order to put this book on a sound footing, we shall first of all need a kind of inventory of non-standard forms that would qualify for closer investigation. Fortunately, the interest in dialect grammar (not only of British English) has grown continuously in recent years; a collection of pioneering first studies on individual phenomena appeared in Trudgill and

Chambers (1991), whereas Milroy and Milroy (1993) attempted a more comprehensive regional overview, so that today we are better able to say which phenomena are candidates for widespread regional developments, and which of them may indeed still be on the increase. A first list of probable candidates for non-regional features is provided by Cheshire, Edwards and Whittle's 'Survey of British Dialect Grammar' (1993) where the authors note such nationwide features as *what* used as a relative pronoun, *there's* used with plural subjects, the use of *ain't* and many more. (A full list can be found in the appendix.)

It is clear that even a book-length study like this cannot even attempt to account for all grammatical features of non-standard English, if this investigation is going to be at all detailed and fine-grained. Fortunately, from Cheshire, Edwards and Whittle's list, groups of topics emerge that are interrelated. A number of features have to do with the topic of negation (and a range of features will be added to the list), which already suggests that this field might show interesting variation between the non-standard and standard varieties of English. There are also several other reasons for choosing negation as the area of inquiry.

Negation and typology

Negation is a syntactic and morphosyntactic phenomenon that lends itself extremely well to cross-dialectal (as well as cross-linguistic) investigations because negation is what has been called a 'pragmatic universal':[3] every language (and of course every language variety) must be able to express negative propositions in order to be fully functional.[4] Negation is also probably one of the most basic human concepts – there is no possible way of describing the function of negation without resorting to circularity, as the discussion in Horn (1989: 45ff.) shows, and a semanticist like Wierzbicka acknowledges this by including a negative operator 'not' in her list of 'semantic primes', purported to be irreducible, innate human concepts.[5] The presence of negation is perhaps the one criterion that can distinguish human from animal 'languages' qualitatively (Horn 1989: xiii; Horn and Kato 2000: 1). It is therefore not surprising that the study of negation has featured – for centuries – in the studies of eminent logicians, philosophers, linguists and psychologists.[6] Finally, features of negation have been collected and investigated in many traditional (as well as modern) dialects and sociolects of English, which gives us a reasonable starting point for further comparisons.

A second line of interest of this book lies in the application of a theoretical framework that allows us to ask new questions of non-standard phenomena, and to draw new conclusions from already existing materials. In this way, this book will also present a novel way of looking at the familiar phenomenon of negation. This theoretical approach is the application of functional typology (in the tradition of Greenberg 1966) to the

study of variation in one language. As this kind of application is still in its infancy, the rationale behind it shall be explained in a little more detail.

Language typology is the study of linguistic phenomena across languages in order to determine the range and limits of linguistic variation. To this end, languages are classified according to general criteria that emerge from cross-linguistic comparisons. For example, languages can be classified according to whether they have prepositions or postpositions, and these characteristics correlate in an interesting way with the order of basic sentence constituents, showing that not all logically possible combinations are actually attested. These correlations are typically borne out across representative samples of the languages of the world. It is clear, however, that no linguist will be able to judge the grammaticality of structures in, say, a hundred of (preferably) areally and genetically unrelated languages, so that the typologist typically has to rely on directly questioning native speakers, often in the form of answers to specifically designed questionnaires, as well as investigating all other available material, typically written grammars, wherever possible. In most cases, one or at the most a handful of speakers per language can be questioned, for purely practical reasons. However, this is also one of the shortcomings of this kind of linguistic typology: although an enormous number of languages can be investigated, individual languages can only be discussed in a rather super-ficial way: there simply has to be a trade-off between the breadth and the depth of investigation.

Over the past decades, typological interest has begun to be diverted to more in-depth studies, necessarily of fewer languages, and a whole range were indeed conducted, for example in the international research programme EUROTYP, which concentrated on comparing the languages of Europe. Hawkins (1986) provided another important step in the direc-tion of narrowing the field of inquiry with his detailed comparative study of the two related languages of English and German. The present book is trying to take this approach a logical step further, from geograph-ically and/or historically *related* languages to the investigation of differ-ent varieties of *one* language. Typological studies typically do not take account of dialects – although this is not an issue of principle, but seems to stem from purely practical reasoning: if written grammars do not exist for the largest number of accepted languages of this world, they certainly do not exist for their dialects; dialect speakers are notoriously difficult to come by; and taking account of dialects would increase the number of varieties exponentially. Also, if one concentrated on those dialects where at least some descriptions already exist, i.e. those in the Western world, this would possibly bias samples of languages even more towards Standard Average European ones. On the other hand, taking account of dialectal variation does not seem to pose problems in principle: any dialect system is a naturally evolved variety of language, and predictions made for the languages of this world, as well as general findings from

these large-scale comparisons, should equally hold for individual dialects of one of these languages – quite apart from the fact that the differentiation between what constitutes a language and what constitutes a dialect does not rest on linguistic criteria alone, and it is therefore intrinsically difficult to draw a borderline between these terms.

This new theoretical approach will therefore provide the background for the detailed investigations which follow.[7] Although – or perhaps because? – negation is such a pervasive phenomenon, in linguistic typology there have so far been only a handful of studies that are dedicated to this topic. Dahl (1979), Payne (1985) and Dryer (1988), three studies conducted independently and largely in ignorance of each other, work in the traditional typological framework of comparing a large sample of various languages, and they will provide the typological backbone whenever generalizations are made in this investigation. Dahl and Payne try to classify the different negation strategies in their samples. Reassuringly, they come to similar results, showing that in the languages of the world, negation can be broadly divided into syntactic and morphological processes. Payne's study is slightly more elaborate as it also takes into account negation by negative quantifiers and semantic aspects of negation, whereas Dahl concentrates on sentence negation only. Dryer relates the position of the sentence negator to the basic word order statistically and shows that only a restricted subset of logically possible positions is actually attested. In his functional grammar, Givón (1984) also has a chapter on negation, concentrating, however, on the semantics and pragmatics of negative 'speech acts' (Givón 1984: 321–51). His remarks on the morphology of negation (especially 66–7, 232–3) largely coincide with the studies by Dahl and Payne mentioned above.

On a more restricted topic, Haspelmath (1997) provides a general typological study of *Indefinite Pronouns* where he also looks at what we shall term *negative quantifiers* and their interaction with the sentence negator in a wide range of languages. Most strikingly, he provides new material for the claim that what is sometimes called 'double negation' or 'multiple negation' and what will be termed *negative concord* in this book is the norm rather than the exception if we consider the languages of the world. This will play an important role when we look at the subject of negative concord in non-standard varieties in Chapter 5 – as it seems that it is the (Western European) standard languages that are the odd ones out, rather than the dialects.

A final general typological collection on the subject of negation is Kahrel and van den Berg's *Typological Studies in Negation* (1994), which consists of a number of commissioned papers, each dealing with negation in one language. The wide range of languages investigated here includes European and non-European languages, varying from a few thousand speakers to several million, but this collection cannot be representative. Furthermore, although the individual papers are divided into similar sections, the descrip-

tions are so different that direct comparisons are not really possible and no general conclusions can be drawn from this collection. It will therefore not play a great role in the following analyses.

The most interesting study from our point of view is Bernini and Ramat's *Negative Sentences in the Languages of Europe* (1996). This in-depth study of negation is restricted to the European languages and thus already takes one step towards the narrowing of the field advocated here. The authors develop several criteria for the classification of negative systems in the European languages and arrive at areal typological comparisons and maps. These maps determine a 'core' linguistic area of similar structures in Europe (the 'Charlemagne' area, basically comprising Germany, France and Italy) that has been established as the core area of many linguistic features in Europe by various other EUROTYP projects as well.[8]

The explicit criteria developed and used by Bernini and Ramat (1996) might seem to offer a good starting point for the application of typological findings to dialect data. The tumultuous past which negation (along with many other syntactic phenomena) has undergone in the history of English might also suggest the comparison of dialect features with the present day system of standard English as a particularly interesting field of investigation because striking systematic differences may be expected. However, many basic typological features (as, for example, word order and the position of the negator with respect to it, or the three-fold division in the quantifier system into *some*, *any* and *none*) are shared between standard English and the modern English dialects, as well as the traditional dialects recorded ca. one hundred years ago. The second glance therefore reveals that many criteria employed by Bernini and Ramat on the European languages, for example precisely those that relate negative structures to basic word order, unfortunately have no relevance for an investigation of English dialects, because drastic variations, like differences in the basic word order, are clearly not present here.

Just as some criteria of the typological studies mentioned above cannot be applied to dialect comparison directly, in many cases of the following investigations the reverse is also the case: more specific hypotheses for individual dialect phenomena investigated are not available. Thus, on one level, typological predictions from the studies of negation are too specific, whereas in many others they are far too unspecific to provide testable hypotheses. For this reason, this book will have to have recourse to a more general level of typological principles. In particular, we shall go back to Greenberg's classic *Language Universals* from 1966, which has a short section on negation, where Greenberg finds 'evidence for the marked character of the negative as opposed to the positive' (Greenberg 1966: 50), and to Croft's exposition of Greenberg's ideas in his *Typology and Universals* in particular his chapter on 'Markedness in typology' (Croft 1990: 64–94), where he repeats polarity as one of the basic categories where we can find patterns of markedness

(Croft 1990: 93). These basic patterns will play an important role when we look at the very pervasive non-standard features that constitute the main part of this investigation. First of all, however, we shall have a brief look at the ways negation in standard English can be said to be marked in contrast to affirmative statements.

Markedness

English negation as a feature of morphosyntax should conform to the criteria for morphosyntactic markedness set out in Table 1.1.[9]

Standard English clearly conforms to markedness criterion S1: positive clauses are not marked explicitly by an 'affirmation marker', but by zero, whereas the negative clause is clearly marked by the addition of (at least) the negative morpheme *not/-n't*, as we shall see in detail in Chapter 2, in many cases by the addition of an auxiliary *do* as well.[10]

Another criterion is also clearly fulfilled for standard English, the frequency criterion S8. Negative clauses are much rarer than their positive counterparts in general. Text counts vary considerably, but figures from my investigation suggest a ratio of between 1:7 and 1:10 for contemporary spoken English (one negative clause for ten positive ones), and the figure for written English must be considerably lower, as negation is known to be much rarer in writing than in conversation (for recent corpus-based figures, cf. Biber *et al.* 1999: 159ff., and cf. Chapter 9). The remaining two criteria, however, are not fulfilled: positive and negative paradigms of standard English verbs have an equal number of distinctions, and the positive paradigms are by no means more or less irregular than the negative ones. Strictly speaking, an equal number of distinctions does not constitute a counterexample to a markedness pattern: the exact definition is that the unmarked value has 'at least as many distinct forms in the same paradigm' (Croft 1990: 79) as the marked one. A counterexample is only constituted by a reversal of the expected order (in this case, more distinctions in the negative paradigm than in the positive one). We can say then that for standard English, negation is neutral with respect to the two criteria

Table 1.1 Markedness criteria

	Name	*Explanation*
S1	Zero value	The unmarked value is typically realized by zero
S2	Syncretization	The unmarked value has more distinct forms in the paradigm
S5	Irregularity	The unmarked value has more allomorphs or is more irregular
S8	Frequency	The unmarked value is more frequent in text counts

Source: Adapted from Croft (1990: 70–94).

S2 and S5. It does fulfil criteria S1 and S8 and negation is thus clearly the marked member of the polarity pair affirmative–negative.

This markedness pattern is one we shall have to bear in mind for the remainder of this book, especially when we compare non-standard systems with the standard English one. Markedness patterns are ultimately motivated by functional considerations. Ultimately, then, we want to suggest general typological principles in order to provide functional explanations for the new grammatical phenomena that are on the increase in non-standard English today. Ideally, taking into account general findings from typological studies will also help to specify in which way the English dialects can perhaps be said to behave more 'naturally' than their standard counterpart, for example in the sense of Mayerthaler (1988).

Data

In order to put such a comparative enterprise on a sound basis, the standard of comparison will have to be made explicit. This book therefore begins by looking at the syntax and morphology of negation in *standard* English today. Chapter 2 gives an overview of the relevant features of negation, as well as a brief historical description of the development of negation towards the standard of today, in order to provide a foil for the comparison with non-standard features discussed in the following chapters. The grammatical approach chosen is an 'enlightened' or 'modern' traditional one, mainly following Quirk *et al.* (1985). This has the advantage of providing surface descriptions within a minimally formal apparatus, which most readers will be familiar with from school grammars. This surface description will be sufficient for our purposes, as functional explanations (rather than grammar-internal explanations) will be sought for the investigated phenomena. However, developments in a generative framework will of course also be discussed where relevant. The emphasis throughout will be on 'real' English and wherever possible, actual rather than constructed examples have been used. For standard English, the main source of these examples is the relatively new resource of the British National Corpus (BNC). The BNC is a 100-million-word corpus of British English compiled in the early 1990s. Ninety per cent of the material comes from written sources, which can be expected to represent present day standard English. All examples that were taken from the BNC carry an identifying alphanumerical label (e.g. BM8 183 or CB1 1007) which consists of the text code (the first three digits) and the clause number in which the particular quoted item occurs. In this way, all examples can be retrieved easily and cross-checked.

For some phenomena of standard English, recent diachronic developments have been postulated. These will be investigated with the help of the following four one-million-word corpora that make such a comparison possible.

The Brown corpus (BROWN) is a corpus of written American English; it was compiled from 1961 material of several genres and comprises about one million words. Its British counterpart is the London/Oslo/Bergen corpus (LOB); it was compiled in exact parallel to the BROWN corpus and thus makes comparisons between British and American English possible, which in many cases may indicate diachronic developments. More recent versions of these corpora, again in exact parallel to the originals, have been compiled at Freiburg university in Germany from materials published in 1991. These are called the Freiburg Brown corpus (FROWN) and Freiburg LOB corpus (FLOB).[11] In other words, four-way comparisons are now becoming possible that should at least hint at ongoing processes inside the two standard varieties. Although these four corpora are considerably smaller than the BNC, their perfectly matched structures avoid the pitfalls that comparisons across different corpora (e.g. a comparison of the LOB corpus with the BNC) would entail. Once the background of standard English is established, we shall look at regional developments in non-standard British English in more detail.

From Chapter 3 onwards, our attention will shift to negation in non-standard systems. Chapter 3 is dedicated to an overview of regional variation. In particular, it deals with those aspects of non-standard negation that are specific to individual regions. In the course of this investigation it has emerged that fine-grained regional distinctions are not necessary and in many cases not even possible when one deals with a universal grammatical phenomenon like negation. The regional distinctions in Chapter 3 are therefore rather rough and the main division is between the purported 'Celtic' Englishes – a term that is slowly establishing itself for the English used in Ireland, Scotland and Wales, where substrate influence from the Celtic languages is at least a possibility that should be examined in some detail[12] – and the dialects of England, dealt with in the second part of Chapter 3. The dialects of England are basically divided into the north, the Midlands and the south.[13] We shall concentrate on those features that are distinctive for each particular region; phenomena that are shared with all or most other dialect regions are considered in the remainder of the book. The third chapter is based on dialect studies that are already available, drawing on a wide range of materials, both published and unpublished. When we ask questions about the geographical distribution of traditional dialect features, one very important resource that can provide answers and that is employed is the Survey of English Dialects (SED).

The SED was a typical 'traditional' dialect atlas project. It is based on informant interviews, which were conducted all over England in the 1950s and 1960s on the model of the American dialect atlas projects under way at that time. Following a relatively formal questionnaire of about 1,200 items, the SED fieldworkers typically recorded one answer per question of one informant per location.[14] Like other traditional dialect atlas projects, the SED was designed to elicit the ideal rural, vernacular dialect that, it

was feared, was in danger of dying out, and the ideal informants were therefore typically NORMs, in Chambers and Trudgill's (1998: 29) terms: they were <u>n</u>on-mobile, <u>o</u>lder, <u>r</u>ural <u>m</u>ale speakers (even though in the SED, some females were also interviewed, and some interviews were also conducted in cities). The main interest of the questionnaire is lexical in nature, and will not concern us here further. However, some grammatical and morphological forms were also elicited, and this information will be an important point of comparison for this study, especially in its published form of atlases. In particular, we will examine Orton, Sanderson and Widdowson's *Linguistic Atlas of England* from 1978, where some of the relevant morphological features are displayed in the form of maps.

In general, however, data from questionnaire studies is somewhat unsatisfactory for the investigation of grammatical phenomena. As work in sociolinguistics has shown, for most linguistic phenomena variation is endemic. It is therefore preferable to have longer stretches of speech in order not to miss instances of variable use, which can obviously not be considered in one-word answers to questionnaires. The formal character of a questionnaire-based interview (typically, the fieldworker is a stranger to the informant) might also be an obstacle to obtaining possibly stigmatized forms – again, an obstacle for grammatical items more than for lexical ones. Finally, work in psycholinguistics has shown that while lexical information is readily retrievable (and indeed is often the subject of metalinguistic conversation and comments by non-linguists), grammatical information is much harder to elicit and much more a matter of the unconscious. For these reasons, it would be preferable to have available longer stretches of relatively natural talk by dialect speakers. This method, long employed in sociolinguistic fieldwork, is slowly becoming more established in dialectology as well. Again, this may have to do with a change in interest of the dialectologist that has moved from lexical items (which can easily be elicited with the help of diagrams and questions like 'What do you call . . . ?'), to grammatical ones (which cannot). Fortunately, some – few – corpora of dialect speakers have been made available now also to researchers not originally involved in the projects. The one disadvantage they have, compared with the SED, is that they are restricted to certain regions of the country. On the other hand, the SED itself is restricted to England, stopping at the borders of Wales and Scotland, so that here corpora can at least lead to a complementary coverage of Great Britain. In Chapter 3, then, we shall also investigate evidence from various dialect corpora, where available and where appropriate.

The first corpus of this kind is the Northern Ireland Transcribed Corpus of Speech (NITCS). The NITCS is a corpus of around 300,000 words collected from 1973 to 1980.[15] The conversations recorded are also questionnaire-based interviews and thus are relatively restricted both in their topic (dealing with traditional life, agriculture, children's games, etc.)

and in their format (questions and answers), but the corpus is faithfully transcribed orthographically. For rare grammatical phenomena, 300,000 words possibly does not constitute a corpus of sufficient size, especially as the restricted topics do not allow for the use of some grammatically interesting constructions. (For example, questions dealing mainly with the distant past will only rarely elicit grammatical constructions that indicate current relevance, etc.) Nevertheless, as negation is a relatively frequent phenomenon, the NITCS will serve as an empirical tool in the investigation of negation in Northern Ireland – an area that is not particularly well covered by any other projects.[16]

Another important source, this time for dialect data from the southwest of England, is a range of tape-recordings from the Somerset Rural Life Museum (SRLM), conducted in the 1980s. The elderly informants, basically also NORMs, were interviewed for non-linguistic purposes in the context of oral history projects, but these recordings are a rich source for dialect grammatical features, particularly because the informants' attention was elsewhere.[17] (Parts of this material have been very faithfully (re-)transcribed by Juhani Klemola, to whom I am particularly grateful for providing access to this material.)

From Chapter 4 onwards – the main part of this book – the emphasis will be on those features of non-standard British English that mirror current developments, for a majority of speakers and a majority of regions. Chapter 4 deals with those features of standard English that have emerged in Chapter 2 as gaps or irregularities of the system. Negative concord (or multiple negation) is the subject of Chapter 5. This is a well-known marker of non-standard English worldwide, and it is also the older form historically, so it is perhaps not too surprising that it is still present in non-standard speech. On the other hand, the use of negative concord is so stigmatized that its continued presence calls for a particularly sound functional reason to explain its robustness in the face of pressure from standard, 'correct' English. This robustness can be explained at least in part by referring to typological patterns, in particular the fact that doubling or multiplying of the negator is by far the more usual structure worldwide, which is well motivated functionally, as Chapter 5 will argue.

The final three chapters of this book are dedicated to three negative paradigms that are grouped together because they have much in common, and because their presence calls for a common explanation. These are, in order of increasing frequency of use, the use of *ain't* for forms of present tense BE and present tense HAVE (the subject of Chapter 6), the use of *don't* in the third person singular (the subject of Chapter 7), and the use of non-standard *was/were* as the only non-standard past tense paradigm investigated here (the subject of Chapter 8). These three paradigms (or four, if we want to distinguish *ain't* into forms used for BE and those used for HAVE) are often (wrongly) cited as instances of regularization. This term, however, does not quite apply, as the person distinctions are not

abolished completely, as the term 'regularization' would imply. Rather, these distinctions are (to varying degrees, as we shall see) levelled under negation, leading to a system of asymmetries that diverges greatly from the standard. There are not many direct parallels from typological studies that could explain this pervasive trend, but it can be shown that ultimately very general patterns of markedness, discussed above for standard English, play an important part and provide a functional explanation.

For each of these individual phenomena that are the subject of these detailed investigations in Chapters 4 to 8, a first question will be: where did this feature originate? The answer to this question will come mainly from traditional dialect studies, at the same time providing an overview of what has been done in this field so far. Of course, for regionally comparative phenomena the SED will again be examined. This should provide us with the background against which it will then be possible to measure current developments.

Use of the BNC

The data for these phenomena will come from the spoken sections of the BNC. Apart from the ninety million words of written material already mentioned above, it also contains ten million words of spoken language – although this only accounts for 10 per cent of the whole BNC, it still constitutes a huge subcorpus of spoken language unparalleled by any other available resource. Roughly half of this spoken corpus contains pre-planned speech (for example, television and radio programmes, or lectures and parliamentary speeches), which can be expected to be relatively close to the written norms, and which will for this reason be excluded from this investigation. The remaining five million words, however, constitute a (still enormous) subcorpus of spontaneous speech that is representative of present day spoken British English. The *BNC Handbook* describes this subcorpus as 'a demographic component of informal encounters recorded by a socially-stratified sample of respondents, selected by age group, sex, social class *and geographic region*' (Aston and Burnard 1998: 31; my emphasis). What is particularly fascinating is the fact that the language was not elicited or recorded by outside interviewers. Instead, the participants themselves were asked to record their everyday conversations over a period of time.[18] These respondents were chosen by 'random location sampling procedures . . . from across the United Kingdom' (Aston and Burnard 1998: 32), and as their geographic origin is known, their speech is marked by a 'dialect tag'. In addition, respondents also marked many of their interlocutors for geographic origin, so that in total 1,281 speakers carry one of the twenty dialect tags in the BNC[19]. The BNC software allows the researcher to search for all text contained in the speech of someone headed by one of these dialect tags, and it is these data that constitute the basis for the in-depth investigations in Chapters 4 to 8.

More specifically, most phenomena were investigated in what is here called the BNC-Spontaneous-Speech subsample (SpS) using the software SARA, which makes automatic retrieval of a range of features possible. The SpS subsample includes those texts of the BNC containing spontaneous (=non-context-governed) speech. It is defined *ex negativo* on the basis of the BNC supplementary databases (an automatic collection of the header information of every BNC text, provided by Sebastian Hoffmann from the University of Zurich, Switzerland) as those texts not carrying a text sort label. It mainly comprises the so-called demographic sample texts, but in addition also includes some texts in the context-governed section which contain spontaneous speech such as radio phone-ins, etc. In this subsample, searches were further restricted to those utterances made by a speaker whose geographic origin is known. These (potential) 'dialect' speakers are defined as those speakers for which the dialect tag is filled by an abbreviation for a British dialect (Aston and Burnard 1998: 86–7).[20]

Putting the BNC to this use is new and quite daring. It can, however, be justified for a range of reasons. The BNC is certainly the only resource available today which features natural speech from virtually all regions of Great Britain, obtained by exactly the same methods, and which is perfectly synchronical. With its emphasis on naturally occurring speech, this BNC subcorpus is certainly one of the most valuable corpora for investigations into everyday spoken English that are available today.

Some readers might object that the BNC spoken subcorpus does not contain dialect speakers at all. If the term *dialect speakers* is intended to mean NORMs as discussed above, this is probably true. As Chambers and Trudgill note, 'in common usage . . . a dialect is a substandard, low-status, often rustic form of language, generally associated with the peasantry, the working class, or other groups lacking in prestige' (Chambers and Trudgill 1998: 3). This quotation already makes clear that our everyday use of the term *dialect* is conditioned more by (a lack of) social status of a certain language variety than by a linguistic definition or by linguistic criteria. Even among many linguists, *dialect* is a shorthand form for what could more correctly be termed 'traditional dialect', i.e. exactly the rustic speech of the peasantry that is not considered standard. In this book, however, the term *dialect* is used in its more modern, general sense, defined by Chambers and Trudgill (1998) in this way: ' "Dialect" . . . refers to varieties which are grammatically (and perhaps lexically) as well as phonologically different from other varieties. If two speakers say, respectively, *I done it last night* and *I did it last night*, we can say that they are speaking different dialects' (1998: 5). As the BNC clearly contains instances of this kind of grammatical variation, we would in principle be justified in calling these speakers *dialect speakers* (although of course not necessarily speakers of the traditional dialects). Nevertheless, the terminology we shall adopt here is a more careful one and we will speak of *non-standard*, rather than *dialect* features where these kinds of modern developments are discussed.

We shall keep the terms *dialect tag, dialect area*, etc., however, where these designations refer to the BNC, as they are part of the BNC itself.

Another issue is the question of representativeness. As Chambers and Trudgill note on this topic, 'the greatest proportion of the population is mobile, younger urban and female – in other words, the diametrical opposite of NORMs' (Chambers and Trudgill 1998: 30). Even in the days of the SED, NORMs must have constituted a minute proportion of all speakers – this was probably the reason why they were so enthusiastically sought out in the first place! This must be all the more true today, and the few instances the BNC contains of traditional dialect speech might be representative in the context of the wider population after all.

Finally, we must remember that especially in Great Britain, geographical variation in speech is still intimately correlated with the social standing of a speaker. This is often represented in the form of a triangle (cf. Hughes and Trudgill 1979: 6): 'the higher a person is on the social scale, the less regionally marked will be his [or her] accent, and the less it will differ from RP', and this is even truer where non-standard grammatical features are concerned. Conversely, the strongest local forms can usually be expected in the lower end of the social scale, so that even if a speaker in the BNC does carry a dialect tag, one would expect non-standard regional forms from speakers of the lower classes only, but standard or near-standard forms from speakers of the higher social classes. It is therefore conceivable that a whole group of speakers is correctly marked for their geographical origin (by carrying a dialect tag) but that they nevertheless do not betray their regional affiliation in their speech. In the worst case, this might lead to a whole 'dialect' area (in the BNC material) devoid of any occurrences of non-standard features, i.e. a dialect area with no 'dialect' material at all, because the social distribution for a particular dialect area is skewed. (The converse of course also applies: there might be dialect areas with only speakers of lower socioeconomic status, where regional speech would thus be over-represented.)

When dealing with problems like these it must be borne in mind that the purposes for which the BNC was originally designed were not primarily sociolinguistic in nature, but lexicographic. Despite this, the quotations above from the compilers themselves show that regional as well as social representativeness was certainly one of the aims in the compilation of the spoken part of the BNC, and the BNC will have to allow itself to be measured against these claims. As I have shown elsewhere (Pust 1998), the choice of an orthographic transcription system – although not in principle a problem for investigations into dialect grammar (rather than, say, regional pronunciation) – and in particular the choice of employing 'skilled audio-keyboarders' (e.g. secretaries), rather than linguists, for the transcriptions without subsequent linguistic proof-reading has led to some unfortunate transcription practices that make investigations into at least some morphosyntactic features virtually impossible.[21] As Aston and

Burnard point out, 'a semi-rigorous form of normalization is applied to the spelling of non-conventional forms such as "innit" or "lorra"' (Aston and Burnard 1998: 36). The emphasis (perhaps unintentionally) lies on *semi* – any reader familiar with the BNC demographic sample will know that spelling is unfortunately anything but rigorous. The following detailed studies are therefore also an investigation of the question to what extent material from the BNC can be utilized for these purposes not envisaged by the compilers.

These caveats, however, do not in principle speak against the use of the BNC for these other, unintended, purposes. Another example where linguists have – very profitably – disregarded the editor's injunctions not to use the material for linguistic investigations is the corpus of ex-slave narratives in the United States. Here the editor warns that 'the slave narratives do not generally provide a reliable source for those seeking to study black speech patterns and black English' (Rawick 1977: xxix) – a sensible warning because these narratives are not verbatim records; the interviewers took field notes during the interviews and the narratives were written down afterwards, based on the field notes, and supplemented from memory. Despite this serious handicap, eminent scholars of the field have conducted just those investigations into 'black speech patterns' (cf. the careful studies by Fasold 1976, Pitts 1986, Schneider 1989 or Viereck 1989), and their work has contributed greatly to what is known about the history of African American Vernacular English (AAVE), and therefore also the status of AAVE today. The BNC has none of these serious handicaps: it is a verbatim record that was transcribed faithfully word for word. Charges like differences in spelling or a differing research design seem indeed quite minor in comparison to the corpus of ex-slave narratives. Nevertheless, for all these reasons, the following studies can and should also be regarded as extensive pilot studies into the usefulness of the BNC for purposes other than lexicographical ones.

A note on statistical tests: they were conducted where appropriate. What was mainly tested was a very simple test of significance (by means of Pearson's chi square). All statistical significances were calculated using the SPSS software package. Uniformly, two features were considered to differ significantly if p was below 0.05 ($p < 0.05$),[22] the usual measure employed in sociolinguistic studies. For this reason, the strength of significance is only indicated for the individual results where it differs from this level. In all other cases, 'statistically significant' should be read as a shorthand for 'statistically significant at $p < 0.05$'. This has been left out for purposes of readability. In most cases, however, differences were either clearly not statistically significant (having a probability of $p > 0.3$, or even higher) or very clearly significant, in many cases with $p = 0.001$ or lower. In other words, not too much hangs by the arbitrary cut-off point, which was chosen here in accordance with the majority of other studies.

2 Negation in standard English

Introduction

This chapter is designed to give an overview of the grammar of negation in standard English today, as this will be the relevant point of comparison for our discussion of non-standard features in the following chapters. The discussion of negation in standard English will centre on the 'heart' of negation, i.e. negation effected by the negator *not* and equivalent strategies, for example by *nobody* or *nothing*, and the interaction (or not) of the two. One particularly striking feature of English, the contraction of negator and verb, will be given much room, again as we expect interesting differences in non-standard English. Finally, the situation in English will be compared to other languages of Europe, giving us a measure by which to gauge whether standard English is a typical or untypical European language (with respect to negation). This will again become relevant in the following chapters, when we try to determine whether non-standard English behaves significantly differently from the standard.

Terminology

In this chapter we will be concerned both with the negation of a whole sentence or clause (sentence negation) and the negation of just a constituent of a clause (constituent negation), but not with what has – a little unfortunately – simply been called morphological negation in several treatments of English. The term *morphological negation* has been used in a generalized way to mark word-internal, more precisely, derivational negation, expressed in prefixes like *un-* or *in-* in *unhappy* and *inanimate*. This semantic 'negation' of the meaning of a word, however, has no effect on the syntax of a sentence, and therefore a sentence with a word containing one of these 'negative' morphemes may still be syntactically positive. For example, a sentence like *He is unhappy* requires a negative tag in a reversed polarity question, e.g. *He is unhappy, isn't he?*, as opposed to the *sententially* negative *He is not happy, is he?* (See Klima's criteria for sentence negation in Table 2.1.) For this reason, word-internal negation will not be treated in any depth here, as it is simply not a feature of syntax.

In typological literature, on the other hand, the term *morphological negation* is usually reserved for those languages which express sentence negation as a morpheme on the verb (e.g. Turkish *-me/-ma*). Indeed, it might be helpful to classify this marker of sentence negation more precisely as inflectional negation, distinguishing it from the derivational negation which we find in English *un-*, *dis-*, *in-*, etc. mentioned above. The process of derivation, as in English, creates a new lexeme (*satisfy* > *dissatisfy*; *happy* > *unhappy*, etc.), which can then undergo all the regular inflectional processes (e.g. *dissatisfy* > *dissatisfies, dissatisfied, dissatisfying*; *unhappy* > *unhappier, unhappiest*). Morphological sentence negation as in Turkish on the other hand is an inflectional category on the verb like person, tense, aspect, etc., which Turkish verbs can undergo in a very regular way.[1]

As derivational negation (understood as defined above) shall not concern us here any longer, in the following chapters the term *morphological negation* will be understood as an abbreviation for 'inflectional negation'. Here, negation will be mainly of interest in its syntactic functions – that is, negation is primarily understood as the function of the word *not*, not as derived meanings that are applied to morphemes word-internally ('negative' prefixes like *un-*, *in-*, *dis-*, as discussed above),[2] and certainly not in its psychological function of the description of a certain evaluative statement (*negative* meaning 'bad').

Sentence negation

Criteria

Sentence negation (also called 'sentential negation', 'clause negation', 'nexal negation') has been defined by Klima (1964) for English in purely syntactic terms with the help of several diagnostic tests. His criteria for distinguishing syntactically negative from positive sentences are summarized in Table 2.1.

This gives us the following tests:

Positive sentence:

(1) a. It is quiet down there, *isn't it?*
         (Reverse polarity tag question)

     b. It is quiet down there, *and so it should be.*
         (Same polarity continuation, addition, with *so*)

     c. It is quiet down there . . . , *even in a storm.*
         (Same polarity continuation, focusing, with *even*; CKF 1526)

Table 2.1 Criteria for sentence negation

Criterion		Positive sentence	Negative sentence
A	Reverse polarity tag question ('checking' tag)	Negative tag	Positive tag
B	Same polarity continuation (additive meaning)	With *so*	With *neither*
C	Same polarity continuation (focusing meaning)	With *even*	With *not even*

Source: Adapted from Klima (1964) and Quirk *et al.* (1985: 777f.).

Negative sentence:

(2) a. None of the bones are broken, *are they*?
        (Reverse polarity tag question)

 b. None of the bones are broken, *and neither is the skull*.
     (Same polarity continuation, addition, with *neither*)

 c. None of the bones are broken . . . , *not even the skull*.
     (Same polarity continuation, focusing, with *not even*; B2C 278)

Payne (1985) distinguishes *standard* negation from *sentence* negation. The term *standard negation*, in the case of English negation by *not*, is reserved for the most prototypical way of negating a clause. In the more recent typological literature this negation is also sometimes abbreviated neg_{pred}, because it is typically marked either on the verb phrase (the predicate) or situated in relation to it. Under standard negation Payne understands 'that type of negation that can apply to the most minimal and basic sentences. Such sentences characteristically are main clauses, and consist of a single predicate with as few noun phrases and adverbial modifiers as possible' (Payne 1985: 198). Standard negation in English is effected by the negator *not* or *-n't*,[3] whereas sentence negation is also possible by negative quantifiers like *nothing*, *nobody*, etc., or by an adverb like *hardly*, as we shall see below.

  The English standard negator *not* is usually classified as an adverb. This can in part be justified by the position that *not* can take inside the verb phrase, as this position is usually permitted only for sentence adverbials. Compare the following (a) and (b) examples:

(3) a. A stable environment can *unfortunately* be an unhealthy one.
        (BPB 585)

 b. A stable environment *cannot* be an unhealthy one.

(4) a.   They're *sadly* mistaken. (A4B 165)

    b.   They're *not* mistaken.

Quirk *et al.* argue for a similar analysis for semantic reasons: 'The clausal negative particle *not* could be regarded as a negative restrictive subjunct [i.e. a subcategory of adverbs], excluding the part of the clause that is focused' (Quirk *et al.* 1985: 605). In his comparative work Dryer has shown, however, that sentence negation particles universally do not pattern like adverbs and that 'they at best belong to the class of sentence adverbs and do not really function as dependents of the verb itself' (Dryer 1988: 109). Indeed, negation markers do not occur in Keenan and Comrie's accessibility hierarchy at all (Keenan and Comrie 1977), which otherwise predicts dependents on the verb correctly. Negative markers should therefore perhaps be reclassified as an extra category *negator*. This analysis is also supported by work in a generative framework. Since Pollock (1989) a separate functional category is widely assumed for the sentence negator (NegP). For these reasons, we shall also adopt a more careful terminology and simply speak of the negative marker or the negative particle *not*.

History and syntax of not

In Old English, sentence negation was effected by the standard negator *ne*, which can be traced back to the Indo-European particle of the same form. The syntactic position of *ne* in Old English is preverbal, the structure as follows:

(5) OE   Ic   ne   secge.
            neg_1 V

Jespersen notes that 'this is the prevalent form throughout the OE period' (1917: 9). This standard negation could be strengthened optionally by the 'stronger' negatives *na* 'not a', *nalles* 'not at all', or *noht* (<*nawiht, nowiht*), 'no thing'. These elements are in effect negative polarity items in postverbal position. In Middle English this optional construction became obligatory. The typical form was a combination of *ne* and *not* (<*noht*), resulting in the following discontinuous construction – a construction that is similar to the one found today in French:

(6) ME   I   ne   seye not.
        (S) neg_1 V neg_2

Preverbal *ne* was then further weakened and lost altogether towards Early Modern English. As to the reasons for this development, Jespersen gives a phonetic explanation, speculating that '*ne* was pronounced with so little

stress that it was apt to disappear altogether' (1917: 9). By the fifteenth century we thus find the exclusive use of postverbal *not*:

(7) EModE   I   say   not.
                   V   neg$_2$

This development is clearly parallel to the development in the other Germanic languages. It results in a postverbal negative marker *not* which was originally a strengthening marker in object position. According to Jespersen, there is a strong universal tendency to place the negator before the verb for psychological reasons; Horn (1989: 446) has termed this the 'neg first' tendency and it has been confirmed by the typological samples of Dahl (1979), Payne (1985) and Dryer (1988). Dryer gives a functional explanation for this tendency:

> Negative morphemes carry a large communicative load in the sense that they carry an important part of the message. If the hearer fails to hear the negative morpheme in a sentence, they will have funda- mentally misunderstood the sentence. Given this high communicative importance, it also makes sense that negative morphemes will serve their purpose more effectively if they are not postponed until after the verb. Delaying them increases the risk of misunderstanding, creating a kind of semantic 'garden path', since the apparent meaning of a sentence up to but not including the negative will be the opposite of the intended meaning.
>
> (Dryer 1988: 102)

The development illustrated above for English thus shows the regular development of what has come to be known as *Jespersen's cycle*:

> the original negative adverb is first weakened, then found insufficient and therefore strengthened, generally through some additional word, and this in turn may be felt as the negative proper and may then in course of time be subject to the same development as the original word.
>
> (Jespersen 1917: 4)

Combined with the neg first tendency, we get the following well-known positional cycle for sentence negation:

Stage 1:   neg$_1$ V   (=OE)
Stage 2:   neg$_1$ V   (neg$_2$)   (=OE)
Stage 3:   neg$_1$ V   neg$_2$ (=ME)
Stage 4:   (neg$_1$)   V   neg$_2$ (=ME)
Stage 5:   V   neg$_2$ (=EModE)
Stage 6:   neg$_2$ V

Stages 1 to 5 are attested through the history of standard English, as has been shown above. Stage 6, the movement of the negative marker away from the (functionally) awkward postverbal position, has taken place only in part in English. There are a few instances of structures like *I not say* from Early Modern English, but this phase was short-lived (these constructions were only 'moderately common in the sixteenth and seventeenth centuries', according to Beukema (1999: 22)) and has not resulted in a structure equivalent to stage 1 for Modern English, as Jespersen's cycle might lead one to expect, although it may have been an important stage for the development of the negation we find today (Ukaji (1993) in a generative framework interprets it as a 'bridge' between stage 5 and the constructions of today, and Beukema (1999) shows why this stage was inherently unstable and had to collapse). As we will see, the present situation in English can only be described in a more complex way:

Stage 6a: I   do   not   say.
             V_{fin} neg V_{lex}

Today, the standard negator *not* (or *-n't* respectively) is placed after the first auxiliary (=the operator) of the corresponding positive clause (see Quirk *et al.* 1985: 121ff.) to effect standard negation in standard English. This definition already leads to an interesting and unique feature of English: in clauses where no auxiliary is present in the positive clause, the 'dummy auxiliary' DO is introduced. This semantically empty auxiliary takes over all the inflectional marking of the main verb of the corresponding positive clause (i.e. marking for person, number and tense), leaving the main verb in its infinitival form (base form). This means that different kinds of positive clauses are treated differently in the process of negation, leading to an asymmetrical system.

Positive clause	[+aux]:	[−aux]:
	It *is* finished.	He came downstairs.
	Costs *will* be sunk.	Rob smiles.
	(EX2 573)	
	↓	↓
Negative clause	+*not*:	+inflected DO + *not*:
	It *is not* finished.	He *did not* come downstairs.
	(AO8 1038)	(CD2 202)
	Costs *will not* be sunk.	Rob *does not* smile.
		(CGC 1479)

We can generalize this procedure as in the flowchart in Figure 2.1.

Figure 2.1 shows that for both positive clauses with and without an auxiliary, the end result is the same, namely a negative sentence that does contain an auxiliary.
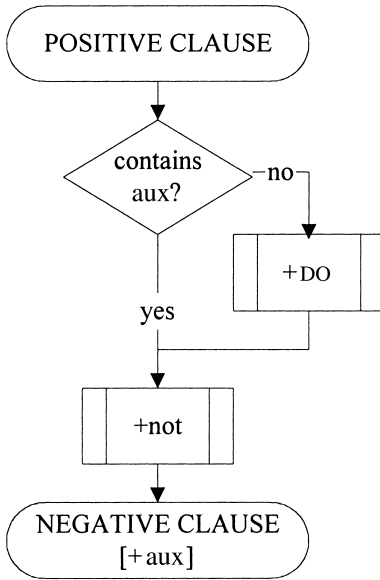
Figure 2.1 Negation (general)

However, there are exceptions to this rule. These are forms of the primary verbs BE and HAVE. Even when BE is used as a full verb (e.g. in copular function), DO-support is not only unnecessary, but ungrammatical. This leads to the exceptional status of a negative clause which does not contain an auxiliary: *This house is haunted* (A4S 129) becomes *This house is not haunted* rather than **This house doesn't be haunted*. Again, the behaviour of BE under negation can be displayed as in Figure 2.2. A comparison with the more general Figure 2.1 makes clear the exceptional status of copular BE under negation.

The only exception to this strict rule is the negative imperative, which is always formed with a form of DO, even for copular BE (e.g. *Don't be afraid to be different*, B21 881) in standard English.[4] For HAVE as a full verb, on the other hand, especially for stative HAVE, both negative strategies (with and without DO) are possible. The two negative strategies result in the following structures: *I had much faith* becomes *I didn't have much faith* or *I hadn't much faith* (ADY 780). This procedure is displayed in a simplified form in Figure 2.3.

Again, the strategy without DO-support leads to a negative sentence which does not contain an auxiliary verb (marked grey in Figure 2.3). In contrast to negation of the verb BE, however, negation with DO-support is not ungrammatical: a more regular alternative to an auxiliary-less negative clause thus exists in this case. The distribution of these two possible
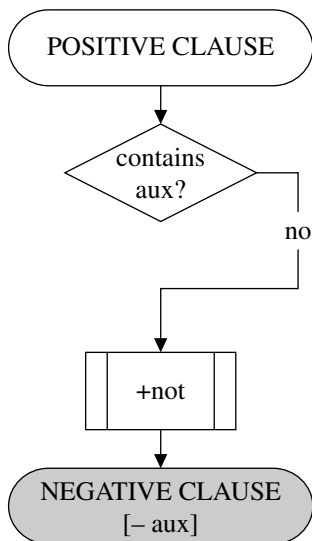
Figure 2.2 Negation of cop BE
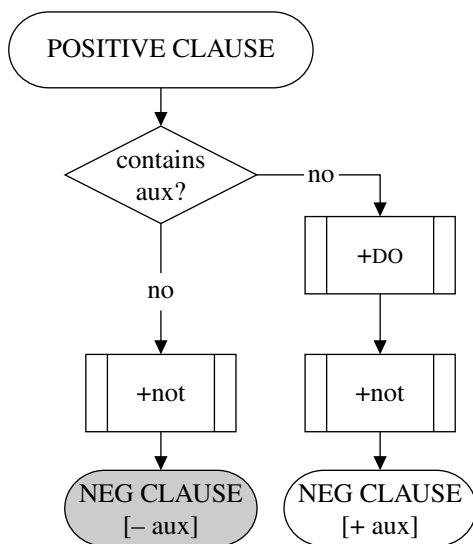
Figure 2.3 Negation of full verb HAVE

Table 2.2 Negation of full verb HAVE

	Total	+DO-support	−DO-support	% −DO-support
BrE 1961 LOB	52	17	35	67.3
AmE 1961 BROWN	79	66	13	16.5

Table 2.3 Negation of full verb HAVE, diachronic comparison

	Total	+DO-support	−DO-support	% −DO-support
BrE 1961 LOB	52	17	35	67.3
BrE 1991 FLOB	47	38	9	19.1
AmE 1961 BROWN	79	66	13	16.5
AmE 1991 FROWN	83	75	8	9.6

constructions merits closer investigation. It has been said that HAVE can only marginally be negated without the use of the 'dummy' DO, and that this is possible only in British English (Quirk *et al.* 1985: 131). To investigate this, negation of the full verb HAVE is compared across the two parallel corpora for British and American standard English, the LOB and the BROWN corpora.[5] The results are presented in Table 2.2.

The figures in Table 2.2 confirm that British English in 1961 still permits postverbal negation (i.e. without a form of DO-support) for forms of the full verb HAVE – British English seems to preserve the older stage (stage 5 in Jespersen's cycle) syntactically and indeed prefers it, whereas the American standard is more innovative and has regularized this anomaly in most cases. Nevertheless, negation without DO-support in over 16 per cent of all cases indicates that this strategy is far from impossible even in the more innovative American English system. One would expect the diachronic shift from HAVE + postverbal neg towards the more regular DO-support to continue. In order to investigate this further, Table 2.3 compares the negation of HAVE in LOB and BROWN with their 1991 counterparts FLOB and FROWN.

As the figures in the last two columns of Table 2.3 ('−DO-support') show, there is indeed a dramatic decrease for postverbal negation with the full verb HAVE in British English between 1961 and 1991, and the preference for postverbal negation (at 67.3 per cent in 1961) has turned into its opposite (at only 19.1 per cent in 1991). American English also shows a decrease, but as it starts from a much lower level, it is necessarily not as dramatic as in British English. Today, then, DO-support is clearly the dominant strategy for negating the full verb HAVE in both British and American standard English. Negation without DO is virtually non-existent in American English today; where it does occur, it bears very strong overtones of antiquated or biblical usage. In British English, the development
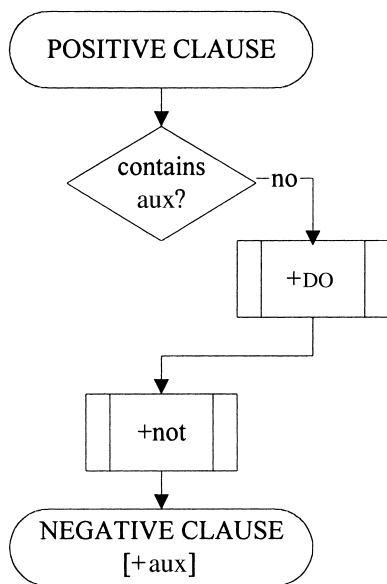
Figure 2.4 Negation of full verb DO

towards DO-support has not quite reached that stage yet, but in 1991 we find roughly the same situation as in the American English of 30 years earlier. This development towards more regularity in the system therefore looks likely to continue towards its natural endpoint (complete loss of postverbal negation for the full verb HAVE) over the next decades.

The third primary verb DO, behaves completely regularly. When used as a full verb (e.g. *They do very well in the marathon*), negation always involves DO-support (e.g. *They don't do very well in the marathon*; AR7 409), as indicated by Figure 2.4.

The three primary verbs can thus be ordered on a scale, according to the degree to which they conform to the general pattern of taking DO-support for full verb uses. At the one end lies the verb BE, which never takes DO-support and is thus very irregular. At the other end lies DO, which obligatorily behaves like all other full verbs. HAVE is situated in the middle, with optional DO-support (but as we have seen, this option is becoming more and more the usual strategy, so today HAVE is closer to DO than to BE). This scale is displayed in Figure 2.5, moving from irregularity to more regularity.

At the same time, this scale mirrors the historical development for the negation of full verbs: they have moved from postverbal negation (stage 5) to the present situation of obligatory DO-support. As we have seen, this change from stage 5 to stage 6a is still going on, as the development for

BE   **>**   **HAVE**   **>**   **DO**
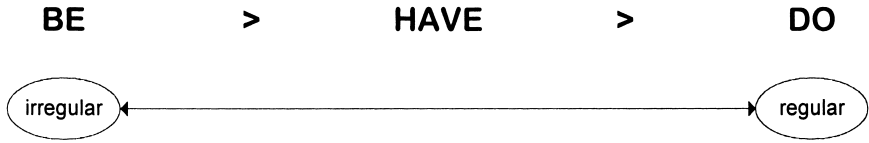
irregular ⟷ regular

Figure 2.5 Irregularity scale for StE primary verbs

the full verb HAVE shows. It is also indicative of a change still in progress that a very high frequency verb like BE has not taken part in it at all. As is well known, highly frequent items are typically able to retain more irregular forms.[6]

For all other full verbs, the strict obligatoriness of DO-support today shows that instead of a full circular movement (as Jespersen's cycle implies), English negation today is similar only in some respects to negation in stage 1.[7] On the one hand, DO-support preserves the Middle English structure: the negative operator still obligatorily follows the finite verb. Because of its obligatory multi-part verbal structure, however, English cannot simply be described as V neg today.[8] As Jespersen points out, the English structure is 'a compromise [with] *not* retaining its place after the verb which indicates tense, number and person, and yet being placed before the really important verb' (1917: 11), the (semantically) 'really important verb' being of course the lexical verb, which contains the lexical content of what is negated. Thus, semantically, we can speak of English as marking negation before the (lexical) verb, and thus as neg first again.

This development will be one of the points to be investigated for spoken English: have all dialects followed the lead of standard English and switched to DO-support for negating full verbs? In particular, how is the full verb HAVE treated in non-standard English today? Is HAVE situated more towards the irregular or the regular pole of the irregularity scale? In other words, is spoken English more conservative or more innovative than the written standard with respect to this feature, or are there perhaps regional differences?

Contraction

Contraction of the negator *not* to *-n't* is a relatively recent phenomenon; its beginnings are usually dated at around the middle of the seventeenth century. It parallels the development in stage 1 of Jespersen's cycle (see above), supporting Jespersen's assumption that the (originally strong) negative marker is weakened phonologically through time. (This might then of course lead to further additional strengthening, putting Jespersen's cycle into swing again.[9]) In standard English today, the original full form *not* can be reduced phonologically to /nt/; it then cliticizes onto the preceding

Table 2.4 Negative contracted verbs

Positive form		Negative form			
		Not affected		Affected	
am	/æm/	*—		*—(aren't)	/ɑːnt/
is	/ɪz/	isn't	/'ɪznt/		
are	/ɑː/	aren't	/ɑːnt/		
was	/wɒz/	wasn't	/'wɒznt/		
were	/wɜː/	weren't	/wɜːnt/		
has	/hæz/	hasn't	/'hæznt/		
have	/hæv/	haven't	/'hævnt/		
had	/hæd/	hadn't	/'hædnt/		
does	/dʌz/	doesn't	/'dʌznt/		
do	/duː/			don't	/dəʊnt/
did	/dɪd/	didn't	/'dɪdnt/		
will	/wɪl/			won't	/wəʊnt/
would	/wʊd/	wouldn't	/'wʊdnt/		
shall	/ʃæl/			shan't	/ʃɑːnt/
should	/ʃʊd/	shouldn't	/'ʃʊdnt/		
can	/kæn/			can't	/kɑːnt/
could	/kʊd/	couldn't	/'kʊdnt/		
may	/meɪ/	(mayn't)	/meɪnt/		
might	/maɪt/	mightn't	/'maɪtnt/		
must	/mʌst/			mustn't	/'mʌsnt/
need	/niːd/	needn't	/'niːdnt/		
ought	/ɔːt/	oughtn't	/'ɔːtnt/		

finite verb form. The irregular morphophonemic behaviour of many lexical stems which in some cases undergo dramatic phonological change after this negative contraction might be an indication that today we are not dealing with straightforward cliticization anymore, as Zwicky and Pullum have argued.[10] On the other hand, the status of the negator (clitic or inflection) does not really alter the perspective of our discussion, so that this point of discussion does not have to be resolved. Under the criterion of contraction, it is possible to distinguish those verbs where the lexical base remains the same phonologically in its positive and negative forms, from those where it is phonologically affected by a cliticized *-n't*.

Table 2.4 shows that the primary verbs DO, HAVE and BE tend to remain unchanged, except for the non-third person form *do – don't*. For the modal auxiliaries, there is a tendency for the original present tense forms to undergo phonological change (*won't*, *shan't*, *can't* – but not *mayn't*), whereas the former past tense forms remain the same.

What is particularly striking in Table 2.4 is the status of *am*. *Am* does not possess a contracted negative form (**amn't*) in standard English. In tag questions, this form is substituted by *are* ('*aren't I?*'), in questions we typically find either the full form *am I not*, or also the contracted *aren't I*.

If contraction occurs in declarative sentences, it is the finite verb which is contracted, not the negator (*I'm not* rather than **I amn't*). Various historical explanations have been advanced for this peculiar gap in the system. Cheshire (1982: 54), for example, suggests that the substitution of what looks like the second person form *aren't* results from the misinterpretation of the form /ɑːnt/, which might have come into existence as a secondary contraction from *amn't > a'n't*, with subsequent lengthening of /æ/ back to /ɑː/ in the eighteenth century.[11] This movement back to /ɑː/ of many /æ/ words (cf. *dance, path*) which also seems to have affected the other irregular forms *can't* and *shan't* in southern standard English might have (wrongly) affected also the form /ænt/ (derived from **amn't*), leading to the cited clash of /ɑːnt/ < **amn't* and /ɑːnt/ < *aren't*. This might ultimately be the reason for the occurrence of *aren't* in the first person singular. *Aren't I* therefore does not necessarily have to be analysed as the plural form *are* extended to *I* due to the systematic gap of a contracted **amn't* form and out of fear of using the non-standard *ain't*, but can be regularly derived from *amn't*. Avoiding the phonologically awkward sequence /mn/ in one syllable may have been a further reason against the establishment of a form **amn't*.[12] One argument against this development is the fact that the usual development for syllable-final or syllable-internal <mn> sequences points to a loss of /n/ rather than the loss of /m/ that Cheshire's development suggests.[13] On the other hand, the different status of *-n't* may have played a role here; in particular the morpheme boundary between /m/ and /n/ in **amn't* may be a reason for the different behaviour. In sum, it must be said that despite different theories, the actual development is still unclear today. This deep-set unclarity may be a reason for the fact that negative contraction of *am not* is avoided at all cost; contraction of personal pronoun and *am* is therefore the more usual strategy.[14] For these reasons it will be very interesting to see how spoken English deals with this 'gap' today: is it filled through regularization, i.e. the creation of a form like *amn't* that has been reported for some traditional dialects, or is this gap dealt with in a different way today?

Contraction of the auxiliary on the other hand (also called non-negative contraction) is possible for all forms as an alternative – as long as the finite verb can in fact be contracted. Table 2.5 gives an overview of the different contractions that are possible for the verb forms mentioned.

From Table 2.5 we can generalize the implicational tendency for standard English that, excluding *am*, wherever auxiliary contraction is possible, negative contraction is always possible, but not vice versa, as summarized in (8).

(8) Implicational tendency for contraction:
     aux contraction ⊃ neg contraction

Table 2.5 Negative vs auxiliary contracted verbs

Positive form	Negative form	
	Aux contraction	Neg contraction
am	'm not	—
is	's not	isn't
are	're not	aren't
was	—	wasn't
were	—	weren't
have	've not	haven't
has	's not	hasn't
had	'd not	hadn't
do	—	don't
does	—	doesn't
did	—	didn't
will	'll not	won't
would	'd not	wouldn't
shall	'll not	shan't
should	'd not	shouldn't
may	—	mayn't
might	—	mightn't
must	—	mustn't
need	—	needn't
ought	—	oughtn't

We can express the same idea by means of a hierarchy for standard English, summarized in (9).

(9) Hierarchy for contraction:
     neg contraction > aux contraction

The only exception to this strong tendency is *am*, which allows auxiliary, but not negative contraction.

Negation by quantifiers

In English, negative quantifiers[15] on their own can make a sentence negative without the support of the standard negator *not*. By Klima's tests for sentence negation, the group of these – lexicalized – negative quantifiers is quite heterogeneous and comprises various adverbial categories, but also nouns, adjectives and other syntactic categories. The inventory of negative quantifiers is displayed in Table 2.6.[16]

   Syntactically, these quantifiers take the same position as their equivalent positive or assertive terms, as a comparison of the (a) and (b) examples shows.[17]

Table 2.6 Inventory of negative quantifiers

Form	Semantic meaning
nobody	animate
no one	animate
nothing	inanimate
never	time
nowhere	place
(in no way)	manner

Source: Adapted from Bernini and Ramat (1996: 164).

(10) Subject position (animate quantifier):
    a. *Nobody* wishes to be exploited. (A06 1554)
    b. *Somebody or something* seemed to be trying to speak to her. (AD9 3950)

(11) Object position (animate quantifier):
    a. They know absolutely *nobody* in common. (HGN 2357)
    b. Let's ask *somebody* else shall we? (F7U 589)

(12) Subject position (inanimate quantifier):
    a. *Nothing* will happen immediately. (AD9 4001)
    b. *Something* will have to be done. (CLP 1257)

(13) Object position (inanimate quantifier):
    a. You'll have achieved *nothing*. (A7N 339)
    b. His mother was bursting to tell him *something*. (A1C 288)

(14) Adverbial position (time):
    a. He *never* told a soul about the Minpins. (CH9 535)
    b. Terriers *sometimes* suffer from this same problem. (BNY 830)

(15) Adverbial position (place):
    a. These titles came into my head from *nowhere*. (CEX 1795)
    b. The money would come from *somewhere*. (FPN 1465)

(16) Adverbial position (manner):
    a. The review is *in no way* comprehensive. (B15 567)
    b. Pilots are *in some way* not 100 per cent on the day in question. (A0H 21)

English is the only language in Europe with a fully developed three-way quantifier system: for every negative form, there are corresponding *some*- and *any*-forms. Following Bernini and Ramat (1996), the meaning of

Table 2.7 Inventory of quantifier system

Semantic meaning	N-quantifier	S-quantifier	A-quantifier
animate	*nobody*	*somebody*	*anybody*
animate	*no one*	*someone*	*anyone*
inanimate	*nothing*	*something*	*anything*
	none/no	*some*	*any*
time	*never*	*sometimes*	*ever*
place	*nowhere*	*somewhere*	*anywhere*
manner	*no way*	*some way*	*any way*

some-forms can be generalized as (specific) existential quantification; *any*-forms indicate generic quantification.[18] Thus, we can complete Table 2.6.

   The difference between the forms in *-one* and in *-body* seems to be negligible in semantic terms. At most, one could adduce a difference in the level of formality. Quirk *et al.* note that 'the pairs of pronouns with personal reference (*e.g. everybody*, *everyone*) are equivalent in function and meaning but the pronouns in *-one* are regarded as more elegant than those in *-body*' (Quirk *et al.* 1985: 377f.). It is therefore all the more interesting that the distribution of these two forms is far from random. Quirk *et al.* (1985: 378) cite the figures in Table 2.8, taken again from the BROWN and LOB corpora.[19]

   The American figures for the forms ending in *-body* are consistently higher than the British ones. Together with Quirk *et al.*'s observation about the difference in stylistic levels (or 'elegance') of the respective forms, this might indicate a change in progress. The forms in *-body* seem to be the newer forms and 'on their way in'. This tendency looks much stronger in American English, and indeed in similar developments American English often is the vanguard, as, for example, in the negation of HAVE as a full verb as we have seen above. A diachronic comparison would therefore lead us to expect a higher percentage for the forms in *-body* for British English today as well. Table 2.9 compares British and American English diachronically across the LOB/FLOB and BROWN/FROWN corpora, but finds a different situation.

   A comparison of Table 2.8 with Table 2.9 suggests that the reverse seems to be the case: whereas the figures for forms in *-body* in British English remain remarkably stable (at an average of around 25 per cent), the American English figures have changed dramatically and are converging with British English. For British English, the only quantifier in *-body* that shows a significant increase in use over the last 30 years is *nobody* (from 36.2 per cent to 47.1 per cent), where both *nobody* and *no one* are now used almost equally. For American English, every single quantifier in *-body* has decreased significantly in use from 1961 to 1991, so that on the basis of the present data it cannot be claimed for British English that the quantifier inventory is currently changing, and that this change is moving

Table 2.8 Comparison of quantifiers in -*body* vs -*one*, 1961

	BrE 1961 LOB			AmE 1961 BROWN		
	Total (-one + -body)	-body	%*	Total	-body	%
any-	173	32	18.5	182	42	23.1
every-	139	33	23.7	166	72	43.4
some-	144	27	18.8	153	57	37.7
no-	174	63	36.2	200	74	37.0
Total	630	155	24.6	699	245	35.1

* Note: Percentages are forms in -*body* (columns 3 and 6) of the total -*one* + -*body* (in columns 2 and 5), e.g. N *anybody*/N (*anybody* + *anyone*).

Table 2.9 Comparison of quantifiers in -*body* vs -*one*, 1991

	FLOB (BrE 1991)			FROWN (AmE 1991)		
	Total (-one + -body)	-body	%	Total	-body	%
any-	191	36	18.8	174	33	19.0
every-	152	37	24.3	248	70	28.2
some-	241	43	17.8	282	57	20.2
no-	136	64	47.1	199	57	28.6
Total	720	180	25.0	903	217	24.0

in the direction of American English. From these figures therefore, no clear working hypotheses or expectations emerge for a diachronic change.

Interaction with not

It has already been mentioned that in standard English, negative quantifiers effect sentence negation on their own, without the presence of the sentence negator *not*. This makes standard English a neg-impermeable language (in the terms of Bernini and Ramat 1996: 218): the sentence negator is not only unnecessary, its presence is ungrammatical. Presence of two negators in one clause is relatively rare and is 'cancelling', i.e. it semantically indicates affirmation instead of negation. This process closely mirrors the situation in logic, where subsequent negations cancel each other out, i.e. one negator operates on the other (compare the law of double negation with the following equivalence: $\sim\sim p = p$), and this kind of double negation (indicating affirmation) will therefore be called *logical double negation* in the following paragraphs. In English, this structure requires

very marked intonation and is only possible in very specific conversational contexts.[20]

(17)   I *didn't* see *nòthíng.* ['I saw something.']

In many cases where this logical double negation occurs in standard English, we find it in the quasi-set formula like *not for nothing*, as in example (18).

(18)   *Not for nothing* are black holes called black. (CET 1405)

Not for nothing can be glossed as the equivalent positive term, 'for something', i.e. there is a good reason why black holes are called black (as not even light escapes from them).

Interaction of quantifiers

A combination of several N-quantifiers in a simple negative clause (i.e. with an overall negative semantic reading rather than a positive one as for logical double negation above) is not permitted in standard English. In this respect, then, English can be characterized as an N1-language.[21] Quantifiers that occur inside the scope of a negator have to take the A-forms (*any*-forms), no matter whether sentence negation is effected by the sentence negator *not* or by an N-quantifier, as examples (19) to (21) show. (Angular brackets are used to indicate the scope of the negator.) *Some*-forms can therefore never occur inside the scope of negation with a simple unmarked negative reading.

(19)   I have *not* [left *anybody* behind]. (CCE 794)

(20)   There's *never* [*anywhere* to move forward to]. (A08 1195)

(21)   The higher direction of MI5 . . . do *not* [have *anything* to do with this], *nor* [have they done so at *any* time]. (CCC 796)

From what has been shown above we can posit the following equivalence between forms with incorporated negation (e.g. *nobody*) and their analytical counterparts (*not . . . anybody*):

(22)   I met *nobody.* <=> I did*n't* meet *anybody.*

    These forms are semantically equivalent; the difference today is mainly a difference in style or formality: incorporated negation is more formal, *not* + A-quantifiers are the more idiomatic expressions.[22] In subject position, however, the neg-incorporated forms (the N-quantifiers) are obligatory.

Labov (1972) calls this strict rule of English negattrac, as the negator is attracted to the subject position. It has been claimed that this is due to the scope of the sentence negator *not* which only extends rightwards to the rest of the sentence.[23] The subject is therefore typically outside the scope of the negation.[24] A sentence like **Anybody didn't send any Christmas cards* is ungrammatical, because the subject *anybody* is outside the scope of the sentence negator *not* and can therefore not carry the interpretation 'Nobody sent any Christmas cards'. In other words, the equivalence mentioned above is only valid for quantifiers in syntactic positions other than the subject position. Again, that this is not necessarily the case also for non-standard English is shown by the Irish English example where *any*-forms are also possible in subject position. Whether this is also true of other dialects will be investigated further in Chapter 3.

Constituent negation

The phenomenon

Constituent negation (also called phrasal negation, special negation, local negation, focusing negation, contrastive negation, or correcting negation), the negation of just a part of a sentence, in English is effected by the same negator *not* as sentence negation.[25] For constituent negation, however, *not* can take syntactic positions that are not available to the sentence negator. As the alternative term *focusing negation* indicates, one particular sentence constituent is focused by negation. In standard English, constituent-negating *not* is usually placed immediately in front of the highlighted constituent and thus stands in sharp syntactic contrast to the fixed position of the sentence negator. Semantically, the scope of the constituent negator only extends over the highlighted constituent. Usually, this marked construction is followed by a positive term indicating what should be substituted for the negated constituent. This is usually introduced by *but*, and thus we get the typical *not ... but* structure which surrounds the highlighted constituent that is to be corrected:

(23) Experimental evidence is useful *not* [for corroborating theories], *but* [only insofar as it falsifies theories]. (BM8 183)

(24) Among other metaphors there is a rich cluster based *not* [on sight] *but* [on touch]. (CB1 1007)

(25) The main gains to efficiency will come *not* [from privatization], *but* [from the legal necessity imposed by the EC]. (A8G 15)

Constituent negation can be 'raised' to sentence negation, in other words, the meaning of 'experimental evidence is useful *not* for corroborating

theories, *but* for falsifying theories' (with intended constituent negation) could also be conveyed by (23) a. with sentence negation.

(23)   a.   Experimental evidence is*n't* useful for corroborating theories, *but* for falsifying theories.

Here, the only difference lies in the semantics: although *not* syntactically now has scope over the lexical verb including the subject complement (as all sentence negators have), semantically, it is not *being useful* that is negated (experimental evidence is very useful, after all), but the constituent *for corroborating theories*. This construction with raised negation erases the syntactic differentiation, so that up to the end of the first clause a 'garden path' situation may occur. In spoken language, a contrastive sentence like this therefore has to carry contrastive intonation (with a fall on the contrasted constituent) to avoid this kind of misinterpretation. Despite this possible confusion, constructions with constituent negation are relatively rare. In particular the (constituent) negation of a subject complement after a form of BE may often be indistinguishable from sentence negation, so that constituent negation and raised negation may actually look identical, as in (26) and (27).

(26)   It is *not* a police matter *but* an immigration matter. (A9R 519)

(27)   Her cries are *not* really language at all, *but* instinctive reactions to the environment. (F9W 578)

Neg co-ordination

Sentence constituents, phrases and parts of phrases can be co-ordinated by the use of the negative conjunctions *neither . . . nor* or by a combination of *not . . . nor*, as for example, in (28) to (31) below:

(28)   a.   *Neither* London *nor* Hong Kong can legislate for dreams. (NP in subject position; AA1 16)

   b.   No one wants chaos in East Germany – *not* the West Germans, *nor* the internal opposition, *nor* the Russians. (NP in subject position; A27 148)

(29)   He will have *neither* the money *nor* the thing. (NP in object position; EBW 1369)

(30)   The thrust of these essays is *not* so special, *nor* so self-serving. (AP in subject complement position; A0P 1508)

(31) a. Their work was *neither* exhibited *nor* published. (VP; A7M 1175)

    b. That patient which can*not* control *nor* further influence.
       (VP; A69 886)

In negatively co-ordinated NP constructions, raising to sentence negation is sometimes described as impossible, so that a sentence like *He won't have the money *nor the thing* (from [29]) would be ungrammatical.[26] Instead, in negative co-ordination with *not . . . nor*, *not* occurs in the same positions as for simple constituent negation, i.e. in front of the constituent to be negated, not in its sentence negator position, as the examples (28) to (31) have shown. With VP constructions, however, the negative conjunction *not . . . nor* can be employed in the usual neg$_{pred}$ position, as in examples (31) c. and (32).

(31) c. Their work was*n't* exhibited *nor* published.

(32) We are *not* actively seeking buyers *nor* actively looking to sell the shares. (CH2 3641)

Generally speaking, however, it is co-ordination by *neither . . . nor* (instead of *not . . . nor*) as in examples (33) to (35) that is much more frequent for all kinds of phrases (be they noun phrases NP, adjective phrases AP, verb phrases VP, prepositional phrases PP or adverb phrases AdvP).[27]

(33) He was *neither* apologetic *nor* guilty over it. (APs; A0P 1508)

(34) *Neither* in the Soviet Union *nor* in Yugoslavia have the centralizers blinked yet. (PPs; ABD 1566)

(35) *Neither* sociologically *nor* ideologically did Labour have much to offer. (AdvPs; A66 965)

In a random sample of 357 texts taken from the BNC, totalling over 5.6 million words, co-ordination by *neither . . . nor* occurred 274 times, co-ordination (of phrases) by *not . . . nor* only 38 times – that is, in less than an eighth of all cases (or 12.2 per cent, vs 87.8 per cent for co-ordination by *neither . . . nor*). For the different kinds of phrases, the exact figures are displayed in Table 2.10.

Although we find some variation here – the percentages for *neither . . . nor* vary from 66.7 per cent (for PPs) to 100 per cent (AdvPs) – these figures are consistently far over 50 per cent and co-ordination by *neither . . . nor* is thus clearly the dominant strategy. The relatively low figures for verb phrases, prepositional phrases and adverb phrases may have influenced the percentages. For example, co-ordination of adverb phrases by

Table 2.10 Phrase co-ordination by *neither . . . nor* and *not . . . nor*

	Total	not . . . nor	neither . . . nor	% neither . . . nor *of total*
NP	192	17	175	91.1
AP	59	6	53	89.8
VP	39	8	31	79.5
PP	15	5	10	66.7
AdvP	7	0	7	100.0
Total	312	38	274	Ø87.8

not . . . nor is also possible, but rare: it did not occur in this particular sample.

Negation and modal auxiliaries

Syntax and semantics of modals

Standard English possesses the following nine central modal auxiliaries: *can, could, will, would, shall, should, may, might* and *must*. They can be distinguished from the non-modal auxiliaries by the following four criteria:[28]

- They are followed by the bare infinitive: *she can* dance, *he might be there*, not: **she can to dance, he might to be there.*
- They do not possess non-finite forms, e.g. infinitives or participles: **to can, *canning, *to must, *musting.*
- They do not possess an *-s* form for the third person singular: *he must, she can*, not: **he musts, *she cans.*
- They have abnormal time reference: the (morphological) past tense forms *could, would, should, might* do not have past time reference.

Semantically, two kinds of modal meanings are usually distinguished: deontic (root, intrinsic) meaning and epistemic (extrinsic) meaning. Typical deontic meanings are 'permission', 'obligation' or 'volition'. The secondary, epistemic meaning of a modal expression refers to the judgement or knowledge of the speaker; typical meanings are 'possibility', 'necessity' and 'prediction'.[29] In English, the deontic and epistemic meanings are distributed in the way displayed in Table 2.11.[30]

Negation of modal meaning

These modal meanings (deontic on the one hand, epistemic on the other) interact with negation in a fairly regular way. These interactions can be captured in modal logic by the (Scholastic) square of opposition[31] displayed in Figure 2.6.

Table 2.11 Meaning of StE modals

	Deontic meaning	Epistemic meaning
can	permission	possibility/ability
could	permission	possibility/ability
will	(volition)	prediction (necessity)
would	(volition)	prediction (necessity)
shall	obligation	(prediction)
should	obligation	necessity
may	permission	possibility
might	permission	possibility
must	obligation	necessity
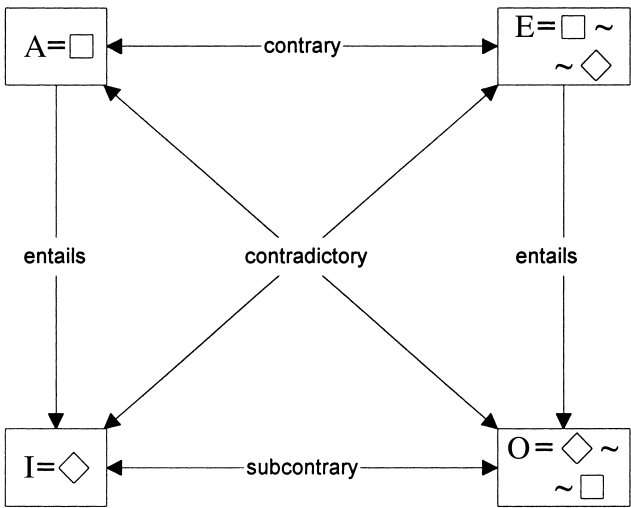
Figure 2.6 Scholastic square of opposition

The four corners can be read as follows:

Deontic meaning:
A: □ = Obligation
I: ◇ = Permission
E: Prohibition   (= □ ~ 'Obligation not to . . .';
                    = ~ ◇ 'No permission to . . .')
O: (= ~ □ 'No obligation to . . .';
                    = ◇~ 'Permission not to . . .')[32]

Epistemic meaning:
A: $\square$ = Necessity
I: $\diamond$ = Possibility
E: Impossibility   (= $\square \sim$ 'Necessity not to . . .';
                          = $\sim\diamond$ 'No possibility to . . .')
O: (= $\sim \square$ 'No necessity to . . .';
                          = $\diamond\sim$ 'Possibility not to . . .')

In other words, depending on the scope of the negator in relation to the modal, interaction of modal meaning and negation can lead to very different results (either the E- or the O-corner). We therefore have to distinguish the following two possible cases: either the modal has scope over the negator, or the negator has scope over the modal. In the first case, where the modal has scope over the negator, and the negator therefore only has scope over the main verb, we speak of main verb negation (MV-neg). In the second case, where the negator has scope over the modal, but only indirectly over the main verb, we speak of auxiliary verb negation (aux-neg). In English, the different modal verbs behave in idiosyncratic ways under negation, as Table 2.12 shows.[33]

   These various possibilities can be summarized as in Table 2.13.[34]

Table 2.12 Examples for the negated modals

Examples	Modal meaning	Deontic/ epistemic	Neg type
He may not enter the country.	$\sim$ permission	$\sim\diamond$ d	aux-neg
You may not be here tomorrow.	possibility $\sim$	$\diamond\sim$ e	MV-neg
People cannot openly speak their minds.	$\sim$ permission	$\sim\diamond$ d	aux-neg
Democracy cannot be about giving politicians a blank cheque.	$\sim$ possibility	$\sim\diamond$ e	aux-neg
The price must not be excessive.	obligation $\sim$	$\square\sim$ d	MV-neg
Thou shalt not steal.	obligation $\sim$	$\square\sim$ d	MV-neg

Table 2.13 System of negation of modals

	Deontic meaning	Symbol	Epistemic meaning	Symbol
can	aux-neg	$\sim\diamond$	aux-neg	$\sim\diamond$
could	aux-neg	$\sim\diamond$	aux-neg	$\sim\diamond$
will	aux or MV-neg		aux or MV-neg	
would	aux or MV-neg		aux or MV-neg	
shall	MV-neg	$\square\sim$	MV-neg	$\square\sim$
should	MV-neg	$\square\sim$	MV-neg	$\square\sim$
may	aux-neg	$\sim\diamond$	MV-neg	$\diamond\sim$
might	aux-neg	$\sim\diamond$	MV-neg	$\diamond\sim$
must	MV-neg	$\square\sim$	– (substituted by *cannot*)	

In other words, those modals that denote the 'weak' I-corner (deontic 'permission', or epistemic 'possibility': *can*, *could*; *may*, *might*) typically effect an auxiliary verb negation reading when they are negated: thus negation promotes the modal meaning from the I-corner to the (strong) E-corner, i.e. to the contradictory meaning – the strongest kind of opposition possible. The only exception to this is the epistemic reading of *may* and *might*, where main verb negation is the unmarked reading. Here, negation semantically leads from the original I-corner to the weak O-corner (rather than the usual E-corner), i.e. to the unlexicalized subcontrary.

Those modals that denote the stronger A-corner (deontic 'obligation', or epistemic 'necessity': *shall*, *should*; *must*) effect a main verb negation reading when negated, and thus move semantically from the A-corner to the equally strong E-corner, i.e. to the contrary meaning. Because of the equivalence mentioned above ($\Box\sim=\sim\Diamond$; $\Diamond\sim=\sim\Box$), which can be translated into 'Obligation not to' = 'No permission to'; 'Permission not to' = 'No obligation to' or, respectively, 'Necessity not to' = 'No possibility to'; 'Possibility not to' = 'No necessity to', there is a rough equivalence of meaning of *mustn't* ($\Box\sim$) and *may not* ($\sim\Diamond$):

(36) a. His older brother . . . must not enter South Africa.
          ['He is obliged not to enter.']

      b. His older brother . . . may not enter South Africa. (AB7 2778)
          ['He is not permitted to enter South Africa.']

In exceptional circumstances, the usual auxiliary negation for *can* and *could*, *may* and *might* can change its scope to main verb negation (in their deontic reading). This exceptional reading requires stress on the negator *not*; therefore, contraction of auxiliary verb and negator are not possible in these cases:

(36) c. He may nòt enter the country, if he doesn't feel like it.
          ['He is permitted not to enter the country.']

      d. He can nòt enter the country, if he doesn't feel like it.
          ['He is permitted not to enter the country.']

Typological characterization of standard English

Sentence negator

Worldwide, the dominant strategies for sentence negation are either (a) morphological negation, as in Turkish (45 per cent)[35] or (b) negation by an invariant particle, as in German (41 per cent). Both types together account for roughly 90 per cent of all languages investigated by Dahl (1979). The third most frequent strategy is negation by a negative verb as

in Finnish (see below), accounting for around 16 per cent of the languages investigated by Dahl. In Europe, negation by an invariant particle is by far the most usual strategy. Only the Finnic languages Lapp, Finnish and Estonian employ a different strategy (negation by a negative auxiliary). Morphological negation does not occur in Europe as a dominant strategy in any language.

The reason for this similar structure is of course the common (Indo-European) history of most of the European languages. Starting from the Indo-European negative marker **ne*, English has taken basically the same historical development as its Indo-European and, in particular, its Germanic sister languages. Due to the weakening/strengthening process generally known as Jespersen's cycle, most Germanic negative markers today can be traced back to an originally optional, extra element first used for extra emphasis, originating in postverbal position. This development has occurred with different speed in the individual languages. Most Germanic languages today are at the fifth stage, where the original particle *ne* has disappeared altogether.[36] This process explains the typologically extremely marked (and therefore highly unlikely) postverbal position of the negative operator in most Germanic languages.[37]

The English development stands out in sharp contrast against this more general trend. The obligatory presence of an extra, semantically empty auxiliary puts it – as a language type – towards the fringes of Europe. Dahl notes that 'dummy auxiliary' constructions are very rare in the languages of the world, although they do occur in a few other languages, but English is probably unique in using this structure for question-formation and emphasis as well (Dahl 1979: 85). Dryer calls this auxiliary construction 'a relatively idiosyncratic quirk of English' (Dryer 1988: 117); for this reason he disregards the obligatory auxiliary completely and describes the English negative construction simply as SNegVO. On the other hand, perhaps the English structure merits more attention typologically precisely because it is almost unique.

A construction similar to the English obligatory DO auxiliary in negative sentences can be found, for example, in Finnish. As already mentioned, the Finnic languages in general use a negative verb for sentence negation. This auxiliary carries at least some of the inflectional marking that would be marked on the finite lexical verb in the positive sentence. The lexical verb of a negative sentence appears in a 'non-finite stem form' (Miestamo 2000), as examples (37) and (38) illustrate.[38]

(37) a. *Minä puhu-n.*
          I   speak-1Sg

      b. *Minä en   puhu.*
          I   neg-1Sg   speak

(38) a.   *He   puhu-vat.*
           They   speak-3Pl

      b.   *He   eivät   puhu.*
           They   neg-3Pl   speak

Payne notes that, diachronically, negative auxiliaries may have originated from higher negative verbs (through a reanalysis of the clause boundary; cf. Payne 1985: 207ff.) and that there seems to be a hierarchy according to which inflectional markings in these constructions are distributed between negative and lexical verb (Payne 1985: 212ff.). In addition, there is a tendency for an inflected negative auxiliary to become invariant, as is indeed the case in Estonian. Bernini and Ramat summarize these tendencies as a development 'from a full Vb_{neg} to an aux_{neg} to a fixed negative marker via numerous intermediate stages' (Bernini and Ramat 1996: 12).

In the Finnish system, the negative auxiliary carries marking for person and number, as examples (35) and (36) have shown. The lexical verb in a negative sentence appears in some kind of invariant form, but still carries tense marking. (This would be imaginable in English as a verb 'to not', and as a structure like **he nots hit my sister*; **I not swam*, etc., or indeed as *he doesn't like my sister*. The translation *I didn't swim*, however, shows that in English, tense is already marked on the auxiliary.)

With respect to the 'dummy auxiliary' structure, we can therefore say that English has started out as a typical Germanic language but now employs a structure that is very different from and untypical of any other Germanic (or indeed Indo-European) language. Bernini and Ramat go so far as to compare the Finnic system explicitly to English: 'Amongst all the other European languages, the one that comes closest to this is English, which does indeed have an explicit neg marker *not/-n't* but which nonetheless adopts a verb (*does*: aux, or other auxiliary)' (1996: 111). It remains to be added, however, that the structure of DO-support itself is not so untypical for a Germanic language in general;[39] compare, for example, the non-standard German structures in (39) and (40) which look parallel:

(39) *Ich   tu'   das   nicht   glauben.*
       I   do   that   neg   believe
       S   aux   O   neg   V

(40) *Ich   tu'   ihn   nicht   anziehen.*
       I   do   him   neg   dress
       S   aux   O   neg   V

The position of the negator is, however, very different due to the different basic word order (TVX for German, SVO for English), which switches

the direct object and the verb between German and English, as the English equivalents in (41) and (42) make clear.

(41) I   don't   believe   that.
     S   aux-neg V   O

(42) I   don't   dress   him.
     S   aux-neg V   O

Neg-permeability

As we have seen above, standard English does not allow more than one negative quantifier per clause, nor a combination of the standard negator *not* with a negative quantifier in the same clause. English is therefore classified as an N1 and a neg-impermeable language by Bernini and Ramat (1996: 183); in a slightly different terminology, Haspelmath assigns English to the typologically very rare type V-NI which disallows co-occurrence (of negative quantifiers) with verbal negation (1997: 203). Again, this is the same for the other standard Germanic languages. Compared with other European languages, these features are rather unusual – the Germanic languages are the only ones that are exclusively neg-impermeable (today). Payne cites the well-known examples of the Slavic languages, where co-occurrence of sentence negation with inherently negative quantifiers is obligatory (Payne 1985: 236f.), as indeed is the co-occurrence of all other quantifiers in the negative form inside the scope of the negator. Bernini and Ramat in their survey of negation in the languages of Europe show an interesting regional distribution for neg-permeability:

> the Germanic group is consistently − [neg-impermeable] and the Slavic group just as consistently + [neg-permeable], while the Romance languages are divided between the three possibilities, although they are mostly +/−: the typology does not coincide with genetic relationship . . . Hungarian and Rumanian, both + languages, do not behave like most of the other languages of the respective families to which they belong, but like the Slavic languages with which they are in close contact.
>
> (1996: 218f.)

In figures, out of forty-five investigated languages, twenty-two are neg-permeable, eight are neg-impermeable (all eight are Germanic languages), and for seven, neg-permeability depends on the syntactic position of the N-quantifier with respect to the sentence negator.[40] Haspelmath makes this point even more strongly:

In my sample, the Latin type (V-NI) is only represented by European languages, suggesting that it is an areal phenomenon. This idea is confirmed by the distribution within Europe, which is almost confined to a contiguous area from Iceland to the Alps and southern France ... type V-NI is so rare that we must ask why it is disfavoured.

(1997: 202)

With respect to neg-impermeability, then, standard English is a typical Germanic language, but a rather untypical European language and certainly a very untypical language of the world.

After this overview of the standard English system, the stage is now set for a closer look at some syntactic and especially morphosyntactic phenomena of negation in non-standard English.

3   Regional variation

Introduction

This chapter concentrates on those features of negation in English dialects that are characteristic of individual regions. For this reason, the British English dialects are mainly divided into the 'Celtic' Englishes, which are very different, and the English English dialects, which show much fewer distinct developments. The 'Celtic' Englishes are those varieties of English spoken in Ireland, Scotland and Wales. Here, Celtic languages were still in wide use only recently and influence from this substrate is, at least possibly, a source for idiosyncratic dialect features. For this reason, short historical introductions will be given that concentrate on when English was introduced and by whom, and what the language situation with respect to the Celtic languages is today. Negation in the respective Celtic language will be compared to the non-standard English variety. The English English dialects are much less differentiated and they are therefore grouped together. Individual developments can be discussed in three broad regional groups – the north of England, the English Midlands, and the south of England. As we shall see, this is a sensible three-fold distinction that will also be useful for the following chapters. Most non-standard features of negation, however, are shared by all, or almost all, dialects today, and these – the main body of this book – will be discussed in detail in the following chapters.

Ireland

History

The island of Ireland, inhabited since ca. 500 BC by northern Celtic tribes speaking a Goidelic (Q-Celtic, Gaelic) language, has had long contact with English speaking settlers (cf. Ó Dochartaigh 2000). The following periods of language contact and thus influence of the English language can be distinguished (cf. also Barry 1996; Görlach 1997). Starting around 1170, the south and east of Ireland was settled by Normans from southwest

England; however, English was only one among a range of languages spoken (besides Norman French spoken by the Norman aristocracy, Latin and Irish Gaelic); English at this time was mainly introduced by English dialect-speaking servants and soldiers. Use of this southwestern dialect declined in the fifteenth century due to increasing 'hibernization' (Hansen, Carls and Lucko 1996: 79), i.e. assimilation to Gaelic. Some survivals of western English are features which are still present today. As Barry notes, however, this period of English 'was not destined to make a permanent impression on Ireland's linguistic landscape' (Barry 1996: ix).[1]

The second phase of contact began in the sixteenth century. The wide-spread Elizabethan and Cromwellian plantations during this time enforced settlement by English immigrants to ensure Ireland's loyalty to the monarchy and to the new Protestant faith (Barry 1996). The majority of speakers came from the west Midlands. Despite these plantations, though, the number of English speakers was generally below 20 per cent.

Finally, the Plantation of Ulster (1610–25) brought Scottish and again English settlers to the north of Ireland; these settlers mainly came from the Scottish Lowlands[2] and from the English northwest Midlands; on the one hand, this brought the distinctive Scottish features to the northeastern coast of Ireland which are still characteristic of Ulster Scots,[3] one of the main northern Irish dialects today, and on the other hand the English dialect that developed into Ulster English, the other main northern Irish dialect today.

The use of English as a first language rose to around 50 per cent ca.1700 and remained relatively stable there until a massive decline in the use of Irish Gaelic in the middle of the nineteenth century. The Great Famine (1845–9) in Ireland, which particularly affected the rural, Gaelic-speaking regions, led to massive emigration, in particular to North America, but also to England, and the shift to English was reinforced by socio-economic pressures, schooling exclusively in English since the introduction of 'national schools' in 1831, the political union since 1801, which created the *United Kingdom of Great Britain and Ireland*, and the like. Thus, by 1891 only 1 per cent of the population was counted as monolingual Irish Gaelic speakers. Today, English is used as their first language by the majority of speakers. English and Irish Gaelic are *de jure* the two official languages of the Republic of Ireland, but the use of English clearly dominates. Even in the Gaeltacht areas, the regions of Irish Gaelic mother tongue speakers, the use of Gaelic is decreasing due to socioeconomic pressures.

In sum, English has been introduced to Ireland mainly by speakers who were dialect speakers themselves, and, at least until the last century, was learned by native Irish Gaelic speakers who became bilingual, so that in this contact situation survival of old dialect features, imperfect language learning and either substral influences and/or other simplifying processes can be expected on a large scale.

Negation in Irish Gaelic

The Celtic languages in general are very different from the Germanic languages typologically, although both belong to the Indo-European group of languages. Most notably, the basic word order in the Celtic languages is VSO (as against SOV for the Germanic languages, or SVO for English). The negative particle is *ní* in Irish Gaelic, and it appears preverbally (as opposed to the predominantly postverbal pattern for the Germanic language group). Quantifier negation in Irish Gaelic is not effected by an inherently negative quantifier (N-forms like *nobody*, *nowhere*) but by a non-assertive quantifier (A-forms like *anybody*, *anywhere*) in combination with the sentence negator. As Acquaviva notes:

> unlike in English, it is not generally possible in Irish to qualify a sentence as within the scope of negation simply by the means of a negative morphological characterization of an adverb or of an argumental NP [i.e. a negative quantifier]. Negation must instead be expressed as a clause-initial particle attached to the main verb, and indefinite NPs or adverbials within its scope may be marked as polarity items.
>
> (1996: 287)

The Goidelic languages are thus very different from other Indo-European languages, which all make use of N-quantifiers, but also from a close sister language like Welsh (belonging to the Brythonic (=British, P-Celtic) branch of the Celtic languages; see below for more details), which also possesses a system of N-quantifiers. The Goidelic languages Irish and Scottish Gaelic only have an A-S-system.[4] In this they resemble a non-Indo-European language like Finnish, with which they form a 'fringe' on the geographic edge of Europe in this respect. Even the A-quantifiers in Irish Gaelic are not fully lexicalized. Complex transparent constructions are formed by means of a generalizing element *ar bith* ('on world', 'on earth') like *duine ar bith* ('person on world': 'anybody') and *ar chor ar bith* ('on time on world': 'ever'), which can be used freely as negative polarity items in appropriate contexts, as examples (1) to (3) show.[5]

(1) *Ní   dhearna   Seán   aon   dearmad   ar   bith.*
    Neg made   Sean   a   mistake   on   world
    'Sean didn't make any mistakes.'

(2) *Ní   raibh   arán   ar   bith   ann.*
    Neg   was   bread on   world there
    'There was no bread at all.'

(3) *D'fhéadfadh   rud   ar   bith   tarlú.*
    could   thing on world happen
    'Anything could happen.'

Celtic languages also have no lexicalized form for 'Yes' or 'No'. Polar questions are answered in the affirmative by repeating subject and verb, usually in an elliptical form; in the negative, this short answer is negated, as in example (4).[6]

(4) *An   bhfaca   tú   Seán? Ní   fhacas.*
    Int see:pret:3Sg   you Sean   neg   saw:1Sg
    'Did you see Sean?   Didn't see.'

These short answers do not exist in Germanic languages in general, but they bear a striking resemblance to the standard English *No, I didn't*, which combines the Germanic particle answer with a more Celtic-style elliptic sentence.

Negation in Irish English

In the literature, negation in Irish English – if mentioned at all – is characterized by the following specific features:

- Especially in Northern Ireland, the Scottish negative clitic *-nae* (for *-n't*) can occur, as in *cannae, isnae.*
- Negative quantifiers also occur with Scottish English *nae-* (for *no-*), as in *naebody, naewhere.*
- A-quantifiers can occur outside the scope of the negator, i.e. to the left of it, as in *Anyone wasn't at home.*[7]

The first two features (occurrence of *-nae/nae-* for *-n't/no-*) are some of the defining grammatical features of Ulster Scots which can clearly be traced back to the seventeenth-century contact situation. Differences in Christian denomination and social status and subsequent social segregation may have helped this immigrant feature to survive; indeed, it has acquired the rank of a (covert) prestige symbol for stressing Ulster Scots (Protestant) identity.[8]

   The third feature, the occurrence of A-quantifiers outside the scope of the negator, is certainly the most striking feature of Irish English negation grammatically, and it is not reported for any other dialect of English worldwide. It violates the obligatoriness of negative attraction (negattrac), proposed by Labov as a rule that holds for the grammar of every English dialect. Labov formulated his observations as follows:

When negattrac is obligatory,[9] it is one of the most mysteriously compelling obligations of all. On the face of it, there seems to be no reason not to say

(17)   *Anybody doesn't sit there any more.[10]

But we can't say it, we don't say it, we won't say it; we reject it without hesitation or reservation. For most listeners, 17 has a curiously ill-formed, fascinatingly perverse character.

We get used to some ungrammatical sentences, and as we repeat them, we can almost see how we might have said them ourselves. We can certainly imagine someone saying

(18)   *This bed was eaten potato chips in.

If this is not already acceptable in some dialect, we would enjoy meeting the man who had the imagination to say it. But we cannot imagine why anyone would say 17, which becomes worse as we repeat it. We cannot even imagine why we cannot imagine it, and if someone did say it, we are not at all anxious to meet him.

(1972: 777f.)

The fact that William Labov does not seem too keen to meet a speaker of Irish English should not disconcert us here; Labov is indeed right to note that the rule of negattrac seems almost inviolable in English. Any deviation from this strict rule in a regular way as in Irish English therefore merits closer investigation and is of course very interesting from a typological point of view. There are two, not necessarily conflicting, explanations for the Irish English pattern:[11]

- Influence from Irish English fronting and clefting by analogy.
- Direct influence from the Irish Gaelic negative construction.

Fronting and clefting are very popular devices in Irish English for emphasizing information, and are much more frequently and more widely used than in other dialects or in the standard.[12] Fronting of objects and prepositional phrases of a negative clause will superficially look like a violation of negattrac, as in the following examples:

(5) I heard about that, but anything further I wouldn't know.
     (NITCS: RF 49)

(6) Anything else I wouldn't lend it eyesight.
     (fronting + shadow pronoun; NITCS: JM 90)

(7) That's only for the school going age, you know, but anybody
     leaving school there's nothing for them at all.
     (fronting + shadow pronoun; NITCS: PM 31)

Fronting itself is a possible candidate for Irish Gaelic substratal influence,
as the fixed intonation in this language does not allow for intonational
emphasis and Irish Gaelic therefore has to rely on syntactic means of reor-
ganization in order to express topicalized constituents.

   Second, we have seen that Irish Gaelic does not possess negative quan-
tifiers – in other words, where standard English has the choice of *nobody*
vs *not . . . anybody* in all positions but the subject, Irish Gaelic can only use
the sentence negator with an A-quantifier. The basic word order VSO for
Irish Gaelic ensures that A-forms in subject position are always formally
inside the scope of the (preverbal) negator. Irish English seems to take this
strategy and to apply it to English with its SVO-structure even when the
A-quantifier is in subject position (the only position in standard English
where this strategy is not permitted). In Irish English, this parallel strategy
results in *any*-quantifiers that are formally outside the scope of the following
negator due to the different basic word order, as, for example:

(8) Anyone doesn't go to mass there.[13]

   This analysis is supported by Harris (1984). He gives the sentence *Anyone
wasn't at home* as an example of a Hiberno-English structure in direct parallel
with the Irish Gaelic construction:

(9) *Ní*   *raibh*   *aon*   *duine*   *sa*   *bhaile.*
     Neg   'be'+past   any   person   in-the   home

Harris stresses that this is one of those features that 'tend to disappear
wherever increased urbanization promotes a severing of links with the use
of Irish . . . [It is] restricted for the most part to conservative rural speech
in predominantly or residually bilingual areas' (Harris 1984: 305), and this
observation is supported by Filppula's recent corpus study of Irish English,
where he compares the rural western dialects of Kerry and Clare with the
eastern dialects of Wicklow and Dublin. Filppula finds that:

> failure of negative attraction appears to be rather infrequent at least
> in the four dialects represented in my database. There were only half
> a dozen instances of structures containing *any/anybody/anyone/anything*
> in subject position within the scope of negation, and they all occurred
> in the Kerry and Clare corpora.
>
> (1999: 179)

That is, today this construction is rather marginal and only found in those areas where 'Irish [Gaelic] can still be said to be "within living memory"' (Filppula 1999: 39).

It is interesting to see that the same phenomenon was apparently part of the negative system in Old English. Mazzon remarks that:

> Labov's neg-attract [*sic*] rule . . . is often contradicted by OE evidence, since sentences . . . are quite common . . . where a non-negated adverb or indefinite precedes *ne* without changing its form, and this was no doubt favoured by the fact that in OE we can still find the declining word order with the verb in final position.
>
> (1994: 164)

As there has never been any prolonged contact in Ireland with speakers of Old English, it seems unlikely that the Irish English system today is the result of direct influence of Old English. The fact that the same structure also occurs – admittedly extremely rarely – in mainland English dialects is sometimes adduced as a counterargument against any Irish Gaelic influence on Irish English. However, this could possibly be due to isolated survivals of the Old English form on the mainland. The historical evidence from Old English certainly shows that even the modern English system, which seems so inviolable today, has not been in place for all that long and that the rule of negattrac is a potential candidate for weakening, or change. In a situation where imperfect language learning may have weakened some of the rigid rules of modern English, the influence from Irish Gaelic and the supporting influence from the extensive use of fronting constructions may have been the decisive factor in the toppling of this otherwise almost general rule.

Data from the NITCS

However this may be, we shall test the actual distribution of some of the typical Irish English features with the help of data from the Northern Ireland Transcribed Corpus of Speech (NITCS).[14] The first feature to be investigated is the use of the Scottish negator *-nae*. The Scottish English clitic negative occurs with a number of verbs in the NITCS, such as *isnae* (3), *wasnae* (4), *werenae* (1), *dinnae* (6), *didnae* (5), *havenae* (2), *hadnae* (4), *cannae* (5), *couldnae* (1), *wouldnae* (9). In comparison with their standard English counterpart *-n't*, however, these Scottish forms only constitute a marginal factor, as Table 3.1 shows.

These figures are plausible if one considers that this Ulster Scots feature does indeed form one end of the continuum. Even Ulster Scots speakers today evidently do not use these forms in 100 per cent of all cases, and especially the relatively formal situation of an interview with a stranger does not seem to promote the use of this language variety.

Table 3.1 Negative contracted verbs in the NITCS

	Total	-n't	-nae	*% -nae of total*
am	2	2	0	0.0
is	140	137	3	2.1
are	43	43	0	0.0
was	296	292	4	1.4
were	106	105	1	1.0
does	104	102	2	1.9
do	1,015	1,009	6	0.6
did	424	418	6	1.4
has	32	32	0	0.0
have	141	139	2	1.4
had	65	59	6	9.2
can	128	121	7	5.5
could	180	178	2	1.1
will	31	31	0	0.0
would	447	435	12	2.7
Total	3,154	3,105	51	∅1.6

The second feature to be investigated is the failure of Labov's negat-trac rule. There are several possible candidates of sentences in the NITCS where an A-quantifier appears to be outside the scope of the sentence negator, as examples (10) to (14) illustrate.

(10) There was thrashing with the flails, a lot, like. *Anyone hadn't* a pile of corn. (NITCS: JO 28)

(11) I heard about that, but *anything* further I *wouldn't* know. (NITCS: RF 49)

(12) *Anything* else I *wouldn't* lend it eyesight. (NITCS: JM 90)

(13) *Not anybody* wanting to buy a donkey. (NITCS: JA 11)

(14) That's only for the school going age, you know, but *anybody* leaving school there's *nothing* for them at all. (NITCS: PM 31)

However, a careful analysis shows that (11) is only an example of fronting (StE*: I wouldn't know anything further*), which in unmarked word order would bring the object *anything* into the scope of the negator. (12) equally is an example of a fronted object and additionally contains a shadow pronoun (StE*: I wouldn't lend anything else eyesight*). (13) is a clear example of non-attraction of the negator to the A-quantifier, but because it is a non-finite structure, it could be explained as an elliptical complex construction (StE*:

There was not anybody wanting to buy a donkey). (14) again is an example of a fronted constituent with a shadow pronoun (StE: *There's nothing at all for anybody leaving school*), where again the unmarked word order would bring the A-quantifier into the scope of the negator. Only example (10) remains as a genuine case of the failure of negattrac. The other examples show, however, that on the surface, the two phenomena of failure of negattrac and thematic fronting intersect, and they result in very similar looking constructions. The rarity of clear examples of a failure of negattrac in the NITCS, however, confirms that this feature seems to be rather marginal today in areas where widespread bilingualism has largely disappeared.

Scotland

History

Scotland can be divided into three distinct cultural areas according to Macafee and Ó Baoill (1997): First, moving from north to south, the Northern Isles (Orkney and Shetland Islands, some also include Caithness in this area),[15] which belonged to Denmark (and Norway) until a few centuries ago. They were never settled by Celts; instead, they have a Scandinavian (Norn) rather than a Celtic history and although Norn became extinct as a distinctive vernacular by the eighteenth century, it has 'strongly influenced both literary Scots and especially its northern dialects', according to MacKinnon (2000: 44).[16]

Second, the Highlands and Islands (particularly the western isles), which have a long Celtic background; linguistically the Scottish Gaelic still spoken there belongs to the Goidelic Celtic languages like neighbouring Irish Gaelic, from where it was originally introduced, and indeed some scholars hold that there is still a dialect continuum between Irish and Hebridean Gaelic today.[17]

Third, the Lowlands have been in long contact with Germanic (Anglo-Saxon) language and culture, especially since the northern part of the kingdom of Northumbria was incorporated into the kingdom of Scotland in the twelfth century. The early Anglo-Saxon dialects spoken in this region developed into a sister language of what became English in England. This language is generally known as Scots today,[18] although it used to be simply called *Inglis*. Finally it should be noted in addition that in Glasgow and Galloway there has been a strong influence of immigrant and migrant Irishmen, especially since the middle of the nineteenth century, and, subsequently, of the Irish English they spoke.

The Scots, a Celtic tribe from Ireland, started invading the Pictish territory in the early fifth century, bringing their Gaelic language with them. About the Pictish culture and language they expelled – or, more probably, assimilated – very little is known (apart from the well-known Pictish burial stones and a handful of inscriptions, the linguistic status of which,

however, is debated) and consequently much speculated. It is now widely held that the Picts were of Celtic origin as well and that their language was a southern Celtic (P-Celtic) one;[19] Price sums up various speculations with typical understatement by calling the status of Pictish 'somewhat problematic' (Price 2000a: 1). By the seventh century, the Anglo-Saxon invasion of mainland England had spread to the north of England, and the kingdom of Northumbria was established with the acquisition of Deira. The (Germanic) linguistic influence on the area was strengthened with the arrival of invading Danes in the ninth century in the east of England and the establishment of the Danelaw – at about the same time, the Scots (or Gaels) had conquered all the Pictish territory, and, due to the weakness of the English kingdoms, incorporated Northumbria into Scotland, establishing 'a largely Gaelic-speaking Scottish kingdom, largely coterminous with present-day Scotland, by the eleventh century' (MacKinnon 2000: 44). Similarly, Glauser notes that 'the Scottish/English border has had its present shape since 1482' (2000: 65), when Berwick-upon-Tweed was finally lost to England.

By the sixteenth century, Scots had already replaced Gaelic in the Lowlands, and since then Gaelic has been retreating towards the north-eastern corner of Scotland, such that today mother tongue speakers are only found in the Highlands and Islands district, as noted above. Today, despite revival programmes, it is estimated that only 1.4 per cent of the population are still knowledgeable in Scottish Gaelic.[20]

Negation in Scottish Gaelic

What has been said about Irish Gaelic in linguistic terms also holds for Scottish Gaelic, due to their close connection. However, the social situation in Scotland was very different from that in neighbouring Ireland, and the contact between Gaelic and Scottish speakers must at all times have been very slight. Even traditional Scots does not show Celtic substratum features, as Macafee and Ó Baoill have shown recently (Macafee and Ó Baoill 1997), although Scottish Gaelic on the other hand has a wide range of loanwords from English.[21] It is therefore to be expected that features in the realm of negation like those in Irish English which are possibly due to an Irish Gaelic substratum influence should not appear in Scottish English.

Negation in Scottish English

Scottish English is the most distinct of all varieties when compared to standard English in the realm of negation. This, however, is mainly due to the historical development of Scots, which has in many cases taken slightly different paths from standard English rather than to a strong influence of a Celtic substrate, as has already been noted.[22] The main differences lie in the domain of morphology. For example, the form of the sentence

Table 3.2 ScE negation system

	Standard English	Scottish English
Standard negator	*not*	*no*
Cliticized negator	*-n't*	*-nae*
Negative adjective	*no*	*nae*
Negative quantifier	*no-*(body, where, one)	*nae-*(body, where, one)

negator is *no* (instead of StE *not*), the clitic negator is *-nae* (instead of StE *-n't*). *Nae* is also the form of the negative adjective in isolation (instead of StE *no*), as well as in morphologically complex quantifiers, as in ScE *naebody, naewhere* (instead of StE *nobody, nowhere*). The clitic *-nae* cliticizes mainly onto auxiliary verbs, but the strict division we find in standard English has not been in force in Scottish English for as long. In 1926, Wilson still notes forms like (in his transcription) *keenay* ('know not'), *cairnay* ('care not'), and *looznay* ('loves not'), i.e. full verbs that take the clitic, and Beal equally notes that 'cliticized negatives of other verbs seem to survive longer in Scots than the "main verb + *not*" construction in English' (Beal 1997: 370). Auxiliary verbs in Scottish English can themselves cliticize, just as in standard English, onto the preceding subject (especially pronouns), and double cliticization with the negative clitic cannot occur (Brown and Millar 1980). Table 3.2 summarizes the morphological differences between standard English and Scottish English.

These differences are generally explained as a difference in etymology. Thus Jespersen states that '*na* was very frequent in OE and later as a rival of *not*, and has prevailed in Scotch [*sic*] and the northern dialects, where it is attached to auxiliaries in the same way as *-n't* in the South: *canna, dinna*, etc.' (Jespersen 1917: 17). On the continuum from broad Scots to standard English, it is not surprising to find that the use of the more Scottish forms is variable; the choice of a more standard form is typically governed by the formality of the situation and the higher social class of the speaker.

Although at the beginning of the century, Scots still seemed to have had a wider inventory of verbs that could be negated by the clitic *-nae*, as we have seen above, today the inventory of auxiliary verbs is considerably smaller than the standard English one. The clitic does not only differ from standard English morphologically, it also affects the base of the verb to which it is attached differently. The full inventory is displayed in Table 3.3.

Quite generally one can say that the system of negative contracted verbs is more regular in Scottish English than in standard English. In many cases where the phonological base is affected by the addition of the negative clitic, this is only one option besides one more regular variant. Generally, the base of the negative auxiliary is the same as the strong (emphatic)

Table 3.3 Negative contracted verbs in ScE

StE unchanged	changed	ScE unchanged	changed
*___	*(aren't)*	*amnae*	
isn't		*isnae*	
aren't		*arenae*	
wasn't		*wasnae*	
weren't		*werenae*	
hasn't		*hasnae*	
haven't		*havenae*	*hennae*
hadn't		*hadnae*	
doesn't		*doesnae*	
	don't		*dinnae*
didn't		*didnae*	
	won't	*willnae*	*winnae*
wouldn't		*wouldnae*	
	shan't	—	
shouldn't		*shouldnae*	
	can't	*cannae*	*cannae*
couldn't		*couldnae*	
(mayn't)		—	
mightn't		—	
	mustn't		*musnae*
needn't		—	
oughtn't		—	

form of the non-negated auxiliary.[23] The only exception here is *do* – the strong form is the same as in standard English (/duː/), but the weak form /dɪ/ serves as the basis for the negative clitic form. In the case of *must*, as in standard English, the stem-final /t/ is lost in the combination with *-nae*. This is not surprising, as *must* itself is a relatively recent import from standard English. The historical Scots form is derived from the Norse root *mun* and is still described in Wilson (1926): *Yee mawnay gang* ('you mustn't go'). The presence of *amnae* (cf. Table 3.3) is also noted by Glauser (2000: 69). It will be discussed in more detail in Chapter 4.

Moving to the syntax of the negator, the current literature on negation and auxiliary verbs in present day Scottish English agrees on the point that especially the cliticized negator *-nae* behaves in a strikingly different way from the clitic *-n't* in standard English. Auxiliary verbs which are negated with *-nae* cannot invert with the subject. Forms in *-nae* can therefore never occur in full interrogatives or tag questions. Brown and Millar (1980: 112) restrict their claim to Edinburgh English (E): 'In E the clitic *-nae* never inverts over the subject expression'; Miller (n.d.: 1) extends this quite generally and claims that 'the clitic forms in *-nae* cannot occur in negative interrogatives'. Millar and Brown (1979) also speak of Scottish English in general when they say that 'in negative interrogatives the clitic negative *-nae* cannot undergo subject-operator inversion' (1979: 29). If this

is indeed the case for Scottish English in general today, it must be a relatively new phenomenon: Wilson still records the following forms of interrogatives as an alternative: '*Dinnay yee ken?* or *Div yee noa ken?* (Don't you know?)' (Wilson 1926: 92, his transcription).

   A second phenomenon sometimes noted is the co-occurrence of the clitic *-nae* with the isolate negator *no* in one clause, as, for example, in *He isnae still no working* (in the sense of '(surely) he isn't still out of work' according to Brown and Millar 1980: 106). Brown and Millar in particular claim differences in scope for the two negators in these constructions.[24] However, these negators behave as expected semantically (*not . . . not* is equivalent to a positive statement, perfectly parallel to logical double negation ~~ which is also cancelling), and indeed as a combination of sentence negators would behave in standard English (*he isn't still NOT working*). Of course, this logical double negation has to have its own intonation contour, and is only possible with a strongly stressed *not*. This does not therefore seem to be a particularly Scottish phenomenon – apart from the fact that the specific Scottish negators are employed.

   A more specific Scottish English phenomenon seems to be the use of tag questions. Millar and Brown (1979) note that due to the fact that the clitic *-nae* cannot invert with the subject (as shown above), negative tag questions in Scottish English are either formed as in standard English employing the clitic *-n't*, or using the isolate negator *no*. Negative tags can thus take the two forms as in examples (15) and (16):[25]

(15) She can cook, ca*n't* she?

(16) She can cook, can she *no*?

These two strategies can be combined, resulting in a double negative tag markedly different from standard English, as in example (17).[26]

(17) Your name's no Willie, is*n't* it *no*?

This double negative tag 'will only occur . . . with a negative statement . . . Its meaning is the same as that of the corresponding affirmative tag' (Millar and Brown 1979: 30). In other words, we are again dealing with a case where multiple negation is logically cancelling, and indeed the phenomenon does not seem different in principle from the combination of clitic and isolate negator in statements that were discussed above.

   Millar and Brown also note a specific Scottish tag question, the tag *e*, which functions syntactically and semantically like a full negative tag question.[27] It is only found after positive statements and is therefore called 'a reverse polarity tag particle' by the authors (Millar and Brown 1979: 33). Again, *e* can itself be negated, and it can therefore also give rise to double negative tag questions (*e no?*) like its full syntactic counterparts, as examples (18) and (19) show.[28]

(18)   He likes cheese, *e*?

(19)   He didnae like cheese, *e no*?

The meaning of this double negative tag is characterized as 'negative assumption + negative expectation' (Millar and Brown 1979: 33). Again, the distribution and the meaning is parallel to an affirmative tag, and again the double negation is therefore cancelling (i.e. an instance of logical double negation).

   Another apparent peculiarity of Scottish English is noted by Miller (1993), a difference in scope of the negator when combined with a universal quantifier as in example (20).[29]

(20)   All the hotels don't take British guests. (For: 'Not all the hotels take
       British guests.')

However, Horn's long discussion of the same phenomenon (cf. Horn 1989: 226ff., who devotes a whole subchapter to 'All that glitters is not gold') shows that a difference in the reading of the scope of negator and universal quantifier is itself a rather universal phenomenon and seems to have a general pragmatic motivation. This then does not seem to be a true dialectal feature specific to Scottish English. (This phenomenon is discussed further below, as it also occurs in the north of England.)

   Finally, Scottish English is frequently quoted as allowing double modals. This has been of interest especially in comparison with double modal constructions in the southern United States, which might be derived from Scotch-Irish immigrant constructions (cf. Fennell and Butters 1996).[30] At the same time, evidence is mostly restricted to anecdotes, as corpora of spontaneous speech for Scottish English have not reached a size where a rare conversational phenomenon like double modal constructions could be quantitatively analysed. As there is no objective analysis yet, native speakers and observers differ greatly in their assessment of which combinations of modals are frequent or even possible – one informant even reported the possibility of a triple modal ('You'll might could do that', reported in Brown and Millar 1980). Possible combinations are listed in Miller and Brown (1982); the most frequent combinations involve either *will* or *might* as the first modal and indeed there are indications that *might* 'is acquiring adverbial status' (Miller and Brown 1982: 12). This is supported in particular by the behaviour of these double modal constructions under negation, where *might* can be exchanged for *maybe*, as examples (21) and (22) illustrate.[31]

(21)   a.   He *might no* could do it. =

       b.   He *maybe no* could do it.

(22) a. He *might no* should claim his expenses. =

      b. He *maybe no* should claim his expenses.

The phenomenon of double modals is often linked to a slightly different phenomenon, the extended use of modal verbs, e.g. as infinitives or past participles, as in (23) and (24).[32]

(23) You have to can drive a car to get that job.

(24) He used to would drink black coffee late at night.

Both characteristics indicate a different status of modal verbs in the Scottish English system in general, as they seem to have evolved features more typical of full verbs. These properties are not restricted to negated modals and will therefore not be discussed in more detail here.

Wales

History

The mountains and hills of Wales seem to have been one of the last refuges of those Celts expelled from their own country by the Anglo-Saxon invasion from the sixth century onwards. As has been noted, the original Celtic substratum of Britain (sometimes called British) has left remarkably few traces on the evolution of Old English from the various Germanic dialects spoken at the time. This is sometimes taken as evidence pointing to massive subjugation and expulsion of large parts of the original population after the Anglo-Saxon invasion. Nevertheless, the bordering regions of Wales have probably always been in direct contact with their Germanic speaking neighbours. Apart from a settlement on the southwest coast ('little England beyond Wales'), however, Wales has never been officially settled by settlers from England.[33] Political subjugation began with the Norman invasion after 1066. In 1284, the English law was introduced by the *Statute of Rhuddlan*; in 1301 northern Wales became a principality of the English throne, and the Acts of Union (1535 and 1542) were the beginning of integrating Wales into a British state. It gave the Welsh representation in parliament, 'while Welsh language and customs were abolished in favour of those of England . . . English [became] the official language of Wales' (Thomas 1997: 56), at least for the gentry. Also since 1301 the English heir to the throne traditionally bears the title of Prince of Wales, indicating the close (official) links between the two countries.

   Nevertheless, widespread use of English in Wales only started to advance during the nineteenth century with massive immigration of English workers to the industrialized centres especially in the southeast of Wales, where

linguistically unstable mixed communities evolved. These typically changed to English over the next generation (Awbery 1997). At the same time, however, in-migration from monolingual Welsh-speaking areas to the industrialized centres slowed down the shift to English considerably, as Pryce (1990) points out. Schools also switched to English as the medium of tuition. This change was accelerated in the twentieth century by the spread of mass media. This contact with language from above and with dialect speakers from the neighbouring Midlands has led to the evolution of a variety of English little remarkable for its distinctive dialect features – at least in the area of negation.

Negation in Welsh

Welsh as a member of the Brythonic (southern Celtic) language group is very unusual typologically in that all its sentences are marked for polarity. The preverbal neg marker *nid* contrasts with the preverbal declarative marker *y* and a preverbal interrogative marker *a* (Bernini and Ramat 1996: 12, 110).

(25) *Nid ydyt ti 'n siarad.*
　　 Neg are you in speak
　　 'You are not speaking.'

(26) *Yr ydyt ti 'n siarad.*
　　 Decl are you in speak
　　 'You are speaking.'

(27) *A ydyt ti 'n siarad?*
　　 Int are you in speak
　　 'Are you speaking?'

Typologically it is far more common for a negative sentence to be morphologically more complex than its positive counterpart (see Dahl 1979; Payne 1985); this is in clear accordance with typological principles as negation is quite generally the more marked option (as opposed to positive contexts).[34]

　　 Welsh also makes use of discontinuous negation (for discontinuous negation in the history of English see Chapter 2)[35] in a construction similar to the French *ne ... pas*; here, the preverbal negative particle *nid* can be enforced by a word like *ddim* (literally 'thing', i.e. clearly a former NPI) in postverbal position, as in example (28).[36]

(28) *Nid wyf i ddim yn hoffi coffi.*
　　 Neg_1 am I Neg_2 in like coffee
　　 'I do not like coffee.'

Not surprisingly, considering Jespersen's cycle, the first negator in these constructions can be dropped in informal spoken Welsh, so that today, the negation proper can be expressed postverbally by *ddim* alone.[37]

Welsh does have negative quantifiers, although only a small range of them, in contrast to its Goidelic sister languages, in particular Irish, as example (29) shows.[38]

(29) *A   wyt   ti   wedi   gweld   rhywbeth?   Naddo,   dim byd.*
        Int   are   you   after   see   something   No,   nothing
        'Have you seen something? No, nothing.'

Whereas Irish Gaelic could be characterized as having an A–S system (it has *any-* and *some-* quantifiers, but not negative ones), Welsh has the more usual N–S system. The kind of substratal influence that resulted in very distinctive features of Irish English should therefore not be expected for Welsh English, at least not in the area of negation.

Negation in Welsh English

Welsh English has not been distinguished from other English dialects in the area of negation in the literature so far. In the sparse literature dealing with grammatical features of this dialect, negation is hardly ever dealt with in contrastive terms. Thus, even Thomas (1997), one of the few existing studies, only laconically notes that in general, 'W[elsh] E[nglish] does not differ from standard English in major respects, as far as its syntax goes' (Thomas 1997: 77). This may already hint at the fact that Welsh English (especially in contrast to Irish English and Scottish English) in many respects patterns with neighbouring English dialect areas rather than constituting its own distinctive dialect area. However, in the light of missing comparative studies it cannot be excluded with certainty that this impression may simply be due to a lack of data.

The north of England

History

Beal notes that 'the Northumbrian and Lowland Scots dialects share a common origin in the Anglian dialect of the early kingdom of Northumbria' (Beal 1993: 187). As has already been discussed above, an answer to the question of whether English and Scots should be regarded as separate languages today is ultimately a political matter. Historically, however, it is clear that the present English dialect area of Northumbria in particular is much more closely related to its northern neighbour than to its southern ones. What was left of the early kingdom of Northumbria (extending roughly from the Tees to the Tweed) after the establishment of Scotland was not incorporated into the English kingdom until 1242. Because of the

long connection with Scotland it is therefore not very surprising that also linguistically, these two regions are still relatively closely related and features similar to those found in Scottish English are therefore also encountered in Northumbrian dialects today.

Southern Northumbria, the kingdom of Deira, was invaded by the Danes in the eighth and ninth centuries. This led to a divergent political development from the northern part especially in Yorkshire, Humberside and Cleveland and this divergent political development is paralleled by different linguistic influences (especially from the Scandinavian languages). In addition, in the industrial town of Newcastle there has been a strong influence of Irish working-class immigrants since 1840 (estimates speak of an Irish population in Newcastle in 1851 of as much as 10 per cent). There has also always been significant immigration from Scotland, partly because of the continuing close ties, so that linguistic influences from these two communities might be expected.

Negation in the North

The close neighbourhood of Northumbria with Scotland and the close relations with England are mirrored by some very distinctive grammatical features that these two varieties have in common. These include, for example, the existence of double modals, and this also extends to the grammatical domain of negation. As we have seen above, Scottish English has its own, very distinctive system of negation, and this can, at least partly, also be found in Northumbria. In particular, the notably different form of the standard negator (*-nae*, *no*) seems to have spread south beyond the (mainly political) border, as we find a clitic form without /t/, similar to Scottish English *-nae* in Northumbrian English.

A further difference in morphology is the different negative contracted form of *will* = *winnet* (although this is characterized as 'rare, archaic' by Beal 1993). Again, a comparison with Scottish English shows that this corresponds directly to the Scottish English form *winnae*. Today, the alternative form *'ll not* seems to be gaining ground and is much more widely accepted (McDonald 1981; quoted in Beal 1993: 199).

Moving south a little, Lancashire is notably different from standard English because it has a different standard negator *noan* (for standard English *not*), as noted as early as 1906 by Schilling. These forms also continue further south into the Midlands area (cf. the following section).

For Yorkshire, Petyt notes a striking difference in the realm of secondary contractions of negative contracted forms, where he finds a trend away from standard English. Secondary contractions of verb and negator in the urban centres of West Yorkshire (Bradford, Halifax and Huddersfield) fall into two groups, although Petyt notes that 'the secondary contraction follows the same basic pattern in both groups' (Petyt 1978: 92), but the morphological effects obviously differ:

a. loss of a consonant before *nt*, e.g. [ɪznt] > [ɪnt]. This pattern for secondary contraction can be found in many other non-standard dialects (cf. also Chapter 6), but in West Yorkshire it applies without exception to all verbs with a contracted negative and thus also includes forms like [ʃʊdnt] > [ʃʊnt] or [wɒznt] > [wɒnt].

b. loss of *n* after a vowel, e.g. [aːnt] > [aːt]; [doːnt] > [doːt].[39] This pattern is very distinctive and only found in one other dialect area, the Midlands. Word-final /t/ is realized as a glottal stop in most cases and is likely to be further weakened as well; Britton, for example, claims that 'where [nt] has been reduced to [t] or [ʔ] the low level of sonority of these phones seem to have especially favoured the deletion of the atrophied remnant of the enclitic' (Britton 1992: 39). Complete deletion of the negative clitic is, however, not attested for Yorkshire (for the different situation in the Midlands, see below).

It remains to be stated that the variable rule which governs the secondary contractions in both cases (a. and b.) is remarkable insofar as the morpheme boundary does not constitute a barrier to contraction, so that it results in the creation of an additional allomorph of the negative morpheme, namely (/t/) in this dialect for the forms in b.

As a further feature in addition to the morphological features noted above which characterize Northumbrian English, Beal (1993) also cites a failure of negative attraction. This feature has already been discussed for Irish English and it is usually held to be so odd that it has become a highly distinctive marker of Irish English (cf. the quotation by Labov on page 48). Indeed it is so rare that any occurrence outside of Ireland is usually assigned to Irish English influence. As Newcastle and Tyneside in general have had large-scale immigration from Irish immigrant workers, this explanation would indeed be the first favourite. However, Northumbrian English is not among the best documented, and the examples cited by Beal (here reproduced as examples (30) and (31))[40] as instances of a failure of negattrac are not very clear.

(30) Everyone didn't want to hear them.

(31) Another house wasn't to be seen for miles.

In these examples there are no *any* forms present, and therefore one can obviously not speak of the occurrence of *any* forms outside the scope of the negator – which is after all the defining criterion for the failure of negattrac. Rather, in example (30) we have a universal quantifier (*everyone*) in subject position of a negative clause. This type of construction is not ruled out by the grammar of standard English (in contrast to indefinite –

non-assertive – quantifiers like *anyone*, which cannot occur in subject position of a negative clause). With universal quantifiers, there are two possible semantic readings, because there are two scope-bearing expressions (*everyone* and *not*) present in one clause. Thus, either *everyone* is in the scope of the negator, or the negator is in the scope of *everyone*. In other words, (30) is ambiguous and can be paraphrased as either (30a) or (30b).

(30) a. 'Not everyone wanted to hear them.' = 'Some people didn't want to hear them.'

  b. 'No one wanted to hear them.'

This is a slightly different phenomenon from negative attraction, and we have already touched upon it above (with the Scottish example *all the hotels don't take British guests*). As was already mentioned there, this phenomenon is by no means a feature of specific dialects, neither of Scottish nor of northern English, but a quite general, pragmatically conditioned ambiguity (cf. Horn 1989: 226ff.). What logically is an instance of sentence negation (*everyone ... not*: $\forall \sim$) is transferred semantically to a constituent negation reading (*not everyone*: $\sim \forall$) quite regularly, even in the standard and indeed is quite widespread across other languages.[41] Perhaps we could group this phenomenon with the well-known feature of neg-*raising*, as it also involves a movement from (intended) constituent negation to (syntactic) sentence negation.

   As this discussion shows, example (31) cannot be adduced as an example for a failure of negattrac either. Although it might at first glance look a likely candidate, *another* is not an *any*-form. (At best, it would be an *an*-form.) Quirk *et al.* classify *another* as an indefinite positive assertive pronoun (Quirk *et al.* 1985: 345), which stands in direct contrast to the *non*-assertive *any*-forms, so that, *ceteris paribus*, the same applies to example (31) as to (30). Evidence for a failure of negattrac in Northumbrian English is therefore still wanting. The examples discussed above show that a universal phenomenon like neg-raising is present instead, but this does not mark Northumbrian English as distinct from other dialects or, indeed, languages.

   What seems to be quite distinct, however, is the system of tag questions in Northumbria (Beal 1993: 201ff., quoting from McDonald 1981). A wider range of syntactic combinations than in standard English is paralleled by a more fine-grained semantic differentiation. Same polarity tags in particular are more common and carry distinctive meaning, as Table 3.4 shows.

   In contrast to this, standard English has to resort to intonation rather than syntax to express the semantic difference between information- and confirmation-seeking tags. Quirk *et al.*, for example, state explicitly that in standard English, 'the tag with a rising tone invites verification ... the tag with the falling tone, on the other hand, invites confirmation of the statement' (Quirk *et al.* 1985: 811). Against this, the Northumbrian system of

Table 3.4 System of tag questions in Northumbria

Statement	Tag	Syntax of tag	Semantics
Positive	positive	aux + S	information seeking
	negative	aux + S + *not*	information seeking
	negative	aux*n't* + S	confirmation seeking
Negative	negative	aux + S +*not*	information seeking
	double negative	aux*n't* + S + *not*	confirmation seeking

tag questions seems quite similar in its complexity to the system of tags in Scottish English (cf. Brown and Millar 1980), especially in permitting 'double negative' tags. However, these should not be confused with multiple negation (negative concord) structures as investigated in Chapter 5 – the double negative tags in Northumbrian contain two sentence negators and are semantically instances of logical (or cancelling) double negation.

Shorrocks notes a similar phenomenon for Lancashire, where 'negative tags may be used after negative as well as positive statements' (Shorrocks 1996: 169). In fact, in this system there are two kinds of negative tags: a contracted negative is used after positive statements (e.g. *it's hot, isn't it*) in the usual reverse polarity tags that is no different from the standard, whereas uncontracted forms can occur after negative statements (e.g. *it's not 'ot, is it not*) but do not have to. Shorrocks claims that a reverse polarity tag (again as in standard English) is 'just as likely, and with just the same propositional content' (Shorrocks 1996: 177), thus does not see a possible functional differentiation between these two strategies. He simply comments that 'the existence of negative tags after negative propositions no doubt constitutes a part of the dialect's extensive use of multiple negation, or negative concord' (Shorrocks 1999: 181). The regional distribution, however, seems to be very restricted; Shorrocks states that these forms 'may well be unique to an as-yet-geographically-undefined area around Bolton in the Northwest of England' (Shorrocks 1996: 169).

The Midlands

Negation in the Midlands

Traditional dialect speakers, especially in the northwest Midlands area, have a negative clitic in *-no* as shown in much detail in maps from the SED (Orton, Sanderson and Widdowson 1978). Although there are no maps concentrating on the form of the negative as such, a comparison of the many maps on negated auxiliaries reveals an interesting pattern. The extension of this *-no* area differs from map to map, but a core area can be determined touching Cheshire, Derbyshire, Shropshire and Stafford-shire, as displayed in Map 3.1, in addition to the border area with Scotland

Clitic negator *-no*
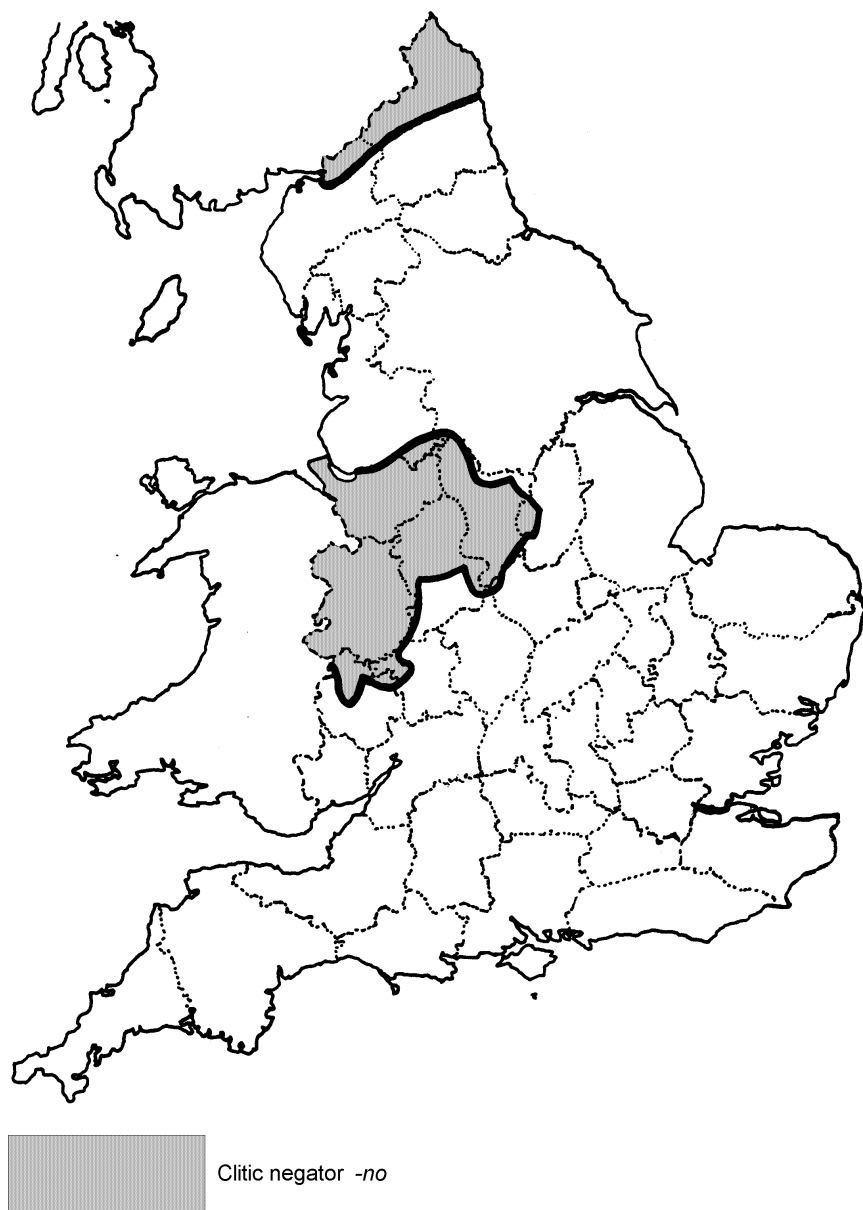
Map 3.1 Clitic *-no* in the SED

Source: Adapted with permission from Orton, H., Sanderson, S. and Widdowson, J. (1978)
The Linguistic Atlas of England, London: Croom Helm.

that has already been identified as closely related to Scottish English above.[42]

The negative clitic *-no* is in most cases, though not always, spelled -NO' in the SED publications. This suggests a derivation of these forms from the standard English negator *not*, with cliticization and loss of /t/, which is certainly one possible way of development. However, there are indications from historical data from Scottish English and Northumbrian English with their similar clitic *-nae* and the isolate negator *no* that these differences in the form of the negator go back to the lexicalization of a different negative element, namely OE *na* rather than the noun *na-wight* (which developed into StE *not*), as, for example, noted by Jespersen and already quoted above: '*na* was very frequent in OE and later as a rival of *not*, and has prevailed in Scotch and the northern dialects, where it is attached to auxiliaries in the same way as *-n't* in the South: *canna*, *dinna*, etc.' (Jespersen 1917: 17).

If we assume the same path of derivation for the Midlands dialect area, this opens up the question of whether we are dealing with an independent parallel development as in Scottish and northern English: namely a simple difference in the etymology of this negator; or whether this should be taken as either direct influence or the remnants ('islands') of a negative strategy that once extended much further than the areas where it has remained until today. This interesting question, however, can only be answered if more historical material becomes available; for the moment it must be left to future research.

Britton (1992) also documents a /nə/ clitic for the northwest Midlands. What is more, this clitic is very well documented in literary representations of dialect speakers of the nineteenth century, for example by George Eliot, and is equally well documented by investigations of traditional dialect speakers that were recorded as late as the 1970s. In the Freiburg dialect corpus FRED, for example, a negative clitic /nə/ is the regular negative strategy for older speakers, and this seems to indicate that it must indeed once have been more widespread than the standard English forms we find today.

Besides this widespread negative clitic which is different from standard English, a different form of the isolate negator (equivalent to StE *not*) is also recorded. Shorrocks, for example, states that 'the negative is formed by the use of the adverb /noːn/ or /nɒt/ "not".' (Shorrocks 1980: 630), i.e. *noan* or *not*.

Moving to a different morphological phenomenon, based on SED material, Britton (1992) records secondary contracted forms of negative contracted verbs to /t/ rather than the much more usual /n/, similar to the same phenomenon in Yorkshire (noted above). In contrast to this kind of secondary contraction in Yorkshire, however, in the 'Black Country' in southwest Staffordshire, west of Birmingham, the negative clitic /nt/ is not only reduced to /t/, but deleted altogether. There are indications that the loss of the clitic has indeed proceeded by these stages; Britton writes that 'it seems

clear, both from the historical evidence and the evidence of present-day variation between /t/ and ∅ in the enclitic, that /n/-deletion was the antecedent state in the loss of the enclitic in s[outh]w[est] Staffordshire' (Britton 1992: 42). This further shortening results in a situation where the change in vowel quality and quantity (from the positive to the negative form) is the only remaining indication for a negative – positive distinction. In other words, where secondary contraction has gone to this extreme, the distinction 'negative' vs 'positive' has been transferred from the addition of a particle (a free grammatical morpheme) to a change in the root morpheme – thus what in other varieties are allomorphs of the same morpheme have, in this dialect, taken on an additional meaning (namely negation) and have thus acquired morphemic status.

In other words, this contraction, originally determined by purely phonetic criteria, has resulted in a negative system that is very unusual typologically, maybe even unique. Dahl (1979), for example, states that in those languages where negation is expressed morphologically as an inflectional category of the verb, 'in contradistinction to a category such as tense . . . , Neg almost exclusively makes use of . . . affixation. In particular, I have found no examples of infixation . . . , *stem modification* or zero modification' (Dahl 1979: 81, my emphasis). Because of the highly marked status of these verb forms, it is therefore not surprising that in southwest Staffordshire only a subset of negative contracted verbs can be further contracted in this way.

The southwest of England

History

The Cornish language in Cornwall, a member of the southern Celtic language family (closely related to Welsh and Breton), probably died out in the nineteenth century; Payton notes that 'Cornish as a spoken vernacular survived until at least the end of the eighteenth century, with elements (including perhaps individual speakers) lingering into the nineteenth' (Payton 2000: 109), but the comparatively recent presence of this Celtic language means that English was introduced to the extreme southwest relatively late. Any English-based dialect therefore cannot have had a very long tradition in Cornwall; perhaps it is for this reason that 'the Cornish dialect has been subject to almost dismissive treatment by academics who have denied the existence within it of significant Celtic linguistic features' (Payton 1997: 100). Because it is generally held that after the death of (Celtic) Cornish the time was much too short for a regional form of English to evolve, there are no systematic studies of the English dialect of Cornwall to date. Data from traditional dialect speakers in the SED (cf. e.g. Orton, Sanderson and Widdowson 1978) confirm the lack of distinctive regional forms in the extreme southwest for many features investigated.

Somerset on the other hand is a very distinctive dialect area for many phenomena as studies by Ihalainen on periphrastic *do* and the assignment of gender have shown (Ihalainen 1991a, 1991b). The grammatical area of negation, however, does not seem to have evolved many particular regional phenomena in this part of England. The few specific phenomena that have been noted in traditional dialect studies are presented in the following section.

Negation in the southwest

As in most other dialects of England, negative contracted forms (e.g. *weren't, aren't, haven't*) regularly undergo a further contraction or simplification; the usual secondary contraction of negative contracted forms is a contraction to /n/ (with a loss of the word-final alveolar stop), and widespread use of this secondary contraction is already noted by Elworthy as early as 1877 in his study of 'The grammar of the dialect of West Somerset', where all negative verbs carry only *-n* rather than *-n't*, e.g. *bain', wadn'* (Elworthy 1877). However, as has been mentioned before, this is a very general phenomenon that can also be observed regularly in fast speech even in the standard, and is certainly not restricted regionally to Somerset or the southwest.

A second phenomenon which is much more specific to the lower southwest is the form *wadn't* as a past tense form of BE. This is found quite regularly in data from oral history projects from Somerset. This form is usually analysed as a development from the fricative /z/ to a homorganic stop /d/. A similar development is attested in other dialectal forms, in particular for the form /ɪdn/ as a possible realization of /ɪznt/. *Wadn't* would then be an alternative realization of *wasn't*. This analysis is also supported by data from American dialects; Wolfram and Schilling-Estes (1996: 140), for example, cite the realization *wadn't* for *wasn't* (with resultant /d/-flapping – as is usual in American English) as one possible source of confusion for the forms *wasn't* and *weren't* in the dialect of Ocracoke, an island on the Outer Banks in North Carolina.[43] Figures from traditional Somerset dialect speakers (in the transcriptions from Klemola) shall be investigated in more detail to throw some light onto this phenomenon. As will become apparent, the /dn/-realization interacts with the phenomenon of secondary contractions mentioned above. Table 3.5 gives all negative contracted verbs, divided for standard English vs dialectal forms as transcribed in the corpus, listing secondary contracted forms separately.

The secondary contractions of the standard English forms are predictable and of course not specific to the dialect of Somerset, as mentioned above. The dialectal forms, however, are worth investigating in more detail. The two dialectal forms mentioned above are present: *idn't* (for *isn't*) and *wadn't* (for *wasn't*).[44] In these forms the stop /d/ is substituted for the homorganic

Table 3.5 Negative contracted verbs in Somerset

StE forms		Secondary contraction		Dialect forms		Secondary contraction	
				ain't	3		
aren't	0						
isn't	5			*idn't*	1	*idn*	13
wasn't	2			*wadn't*	10	*wadn*	61
				wardn't	1		
weren't	30			*weredn't*	6	*weredn*	1
hasn't	1						
haven't	10						
hadn't	12	*hadn*	1				
don't	88						
didn't	164	*didn*	6				
can't	43						
couldn't	53	*couldn*	1				
won't	30						
wouldn't	35						
shouldn't	1						
mustn't	1						
Total	475		8		21		75

/z/ in what can be described as a process of assimilation (cf. also Wolfram and Schilling-Estes 1998: 47, who also cite parallel realizations, e.g. of *business* with /dn/).[45] The other two forms are not predictable, however: *wardn't* (one occurrence) seems to have added a further /r/, or at least r-colouring to the form *wadn't*, whereas *weredn't* looks like a parallel development to *wadn't* and *idn't*. Indeed, *wardn't* seems to be half-way between the two forms *wadn't* and *weredn't* and perhaps points to a possible continuum between these two forms.

The 'regular' dialectal forms *idn't* and *wadn't* are also characterized by the fact that they occur significantly more frequently with secondary contraction (1:13 for *idn't*; 1:6 for *wadn't*) than in the full form with word-final /t/. For *weredn't*, however, this relation is exactly the opposite (6:1 for *weredn't*). This difference also speaks for a different status of this form – it suggests a later formation, probably after the pattern of *idn't* and *wadn't*, once these forms were established in the dialect. *Weredn't* here has obviously not proceeded to the state of secondary contraction yet. In contrast to data from southern United States American English, however, in the Somerset data, these negative verbs are the only words where a sequence /dn/ is substituted for /zn/; lexical words with the sequence /zn/ like *business* have not changed to /dn/.[46] In other words, in Somerset, this change seems to be specific to the negative verb paradigm. It is striking that for all other verb paradigms except BE, the most frequent negative forms also end in /dnt/ (cf. *hadn't, didn't, couldn't, wouldn't*): of the standard

English forms, these in fact account for 273 of 483 or over 56 per cent of all negative contracted forms in the data from Somerset and this apparent parallel might have acted as a promoting factor in the spread of this 'unphonological' /d/ to a form like *weredn't*. In sum, then, the spread of /d/ through the negative contracted verbs and the fact that this /d/ is restricted to precisely this environment suggests that the negative verbal paradigm is very tight (lexical words with the same phonological sequence /zn/ are not affected, as we have seen), and this /d/ may have arisen through analogy in this case, rather than simple sound change.

The southeast of England

History

The southeast, as the area where the Angles and Saxons originally settled, has the longest history of English, from its Anglo-Saxon and Old English times until today. Since the twelfth century, the southeast has become the centre of influence in all respects for Great Britain. Since the fourteenth century under Edward III, the House of Commons has become a major political instrument. Especially with the accession to the throne by James I (James VI of Scotland), the successor of Elizabeth I, in 1603 and finally with the Acts of Union that united Scotland and England politically in 1707, political influence was concentrated in the southeast almost exclusively of any other region. Canterbury became the centre of religious life very early on. With the foundation of the two universities in Oxford and Cambridge, the southeast also became the centre of intellectual life and of cultural life in general. Numerically and in terms of affluence, the southeast is clearly the dominant region of England today. Not surprisingly, therefore, the southeast has also become the centre of linguistic influence since Early Modern English times. Last but not least, the prestige accent RP also has its basis in the southeast, and the newly remarked upon non-standard pronunciation levelling ('Estuary English') seems to be spreading from the Greater London area to much of England.

Negation in the southeast

Although Cheshire (1982) has devoted a whole chapter to the subject of negation in non-standard speech in the southeast town of Reading, her findings seem to be more characteristic of non-standard speech in general, rather than specific of the dialect of Reading or the wider southeast. This is not surprising, for the southeast does constitute the dominating linguistic influence, not only in the standard, but also in non-standard, non-regional features. Those features investigated by Cheshire like *ain't*, the use of multiple negation, the past tense negator *never* or indeed the use of invariant *don't* are therefore not presented here in any detail. The following chapters

Table 3.6 Summary of distinctive dialect features

	ScE	*IrE*	*N*	*Mid*	*SW*
Negator	*no*	(*no*)	*noan*	*noan*	
Clitic negator	*-nae*	(*-nae*)	*-no*	*-no/-na*	
N-quantifier	*nae-*	(*nae-*)			
No negattrac		√	√?		
Double neg tags	*can't she no*		*isn't it not*		
Secondary contraction	*e no*		>/t/	>/t/>∅	/zn/>/dn/

are devoted to the investigation of these phenomena, and possible regional differentiation will be taken into account there. For the realm of negation, however, the conclusion must be drawn that the southeast does not have any features that are specific to this dialect area, or, perhaps the other way around, that the political and economic dominance of the southeast has extended to linguistic dominance to the degree that its distinctive features have become general features of non-standard speech across the country.

Summary

We can summarize the distinctive features found for the individual regions so far in Table 3.6. It can be clearly shown that Scottish English (and, influenced by Scottish settlers, Irish English) is the most distinct from standard English in the realm of morphology; Irish English with its failure of negattrac is the most different syntactically. Welsh English and the southeast have not been included here as no distinctive features have been noted for the traditional dialects.

4   Filling the gaps?

Introduction

This chapter begins the main part of this investigation which concentrates on those features of modern non-standard English that are not regionally restricted, and that have emerged as potentially interesting areas in our survey of standard English in Chapter 2. The question that will be pursued in this chapter is whether 'gaps' that appeared in the standard English system are filled in spoken English. The most notorious gap is the absence of *amn't from standard English. As we shall see, the presence or absence of a form for amn't in particular is related to the question of auxiliary vs negative contraction, which will therefore be the starting point of this chapter. The two other features investigated are the negation of full verb HAVE and, moving to the realm of modal verbs, the presence or absence of epistemic mustn't.

Negative and auxiliary contraction

Introduction

As Table 2.5 has shown, negative contraction is possible for a much wider range of verbs than auxiliary (or non-negative) contraction in standard English. Practically every verb (except am) has a form with a contracted negative, whereas auxiliary contraction is only possible for a smaller number of verbs. For this reason, speakers have a choice between negative vs auxiliary contraction for the following verb forms only: is, are; have, has, had; will, would; shall, should. Some of the auxiliary contracted forms are ambiguous: he's not is the contracted form of both he is not and he has not (although this use is relatively rare); I'd not can be derived from either I had not, I would not or I should not, and you'll not can – at least in principle – be the contracted form of you will not or you shall not.

In addition, however, one has to consider different syntactic environments. The distinction between auxiliary and negative contraction is only relevant for declarative sentences. Only here and for those verbs listed above

Table 4.1 Options for contractible verbs

	Declarative	Interrogative
Uncontracted	*you are not*	*are you not?*
Neg-contracted	*you aren't*	*aren't you?*
Aux-contracted	*you're not*	—

Table 4.2 Options for non-contractible verbs

	Declarative	Interrogative
Uncontracted	*you can not*	*can you not?*
Neg-contracted	*you can't*	*can't you?*
Aux-contracted	—	—

do speakers have a choice between negative contraction, auxiliary contraction and completely uncontracted forms, as Table 4.1 shows (e.g. between *you aren't*, *you're not* and *you are not*). In full interrogatives as well as in tag questions, however, auxiliary contraction is never possible, as the operator necessarily has to precede the subject, and operators in initial position cannot occur in the contracted form: **'re you not?* In these environments, then, there is no real equivalent of auxiliary contracted forms (of the respective declarative environments) and thus the only alternative to negative contracted forms are completely uncontracted forms, e.g. *aren't you* vs *are you not?* Because auxiliary contracted forms are principally not possible in questions, there is therefore no distinction in interrogative clauses between those verbs that allow auxiliary contraction and those that do not, as Table 4.2 makes clear; in other words, speakers here have the same two options for every verb, namely negative contraction or completely uncontracted forms (e.g. *isn't he* vs *is he not?*, parallel to *can't you* vs *can you not?*).

The choice of one contraction strategy over the other might be a feature that is regionally differentiated – this is commented on quite frequently in the literature.[1] In a typical statement, Trudgill, for example, states that 'speakers of Standard English in the south of England tend to use, in their speech, contracted negatives of the type (3) I haven't done it/I won't do it. In the north of England, the alternative contraction is, in some areas, more common: (4) I've not done it/I'll not do it' (Trudgill 1984: 33). Similarly, Hughes and Trudgill briefly note that:

> speakers of standard English in the south of England tend to use contracted negatives of the type *I haven't got it/She won't go/Doesn't he like it?* The further north one goes, the more likely one is to hear the alternative type: *I've not got it/She'll not go/Does he not like it?* This is particularly true of Derbyshire, Lancashire, . . . Cumbria and Scotland.
>
> (1979: 20)

Cheshire claims after investigating the same phenomenon for her adolescent Reading speakers that her data seems to contradict the very clear tendency postulated by Trudgill: she finds that auxiliary contracted forms like *you're not going anywhere*:

> are preferred by the peer groups for forms of BE: in the recordings they occur 100 per cent of the time for auxiliary BE + not, and 74 per cent of the time for the copular BE + not. This preference is surprising, for the uncontracted forms are usually considered more typical of Northern and Scottish varieties of English (see for example, Hughes and Trudgill, 1979).
>
> (1982: 52)

However, if one reads the sources cited by Cheshire very carefully, it becomes apparent that it is not the preference in Cheshire's data that is surprising. In none of the examples does Trudgill cite a form of present tense BE to support his north–south divide. What is more, Hughes and Trudgill continue the quotation above, explicitly saying that:

> southern speakers . . . use the northern-type contraction in *I'm not*, since *I amn't* does not occur in standard English. They also quite frequently use the *you're not*, *we're not*, *they're not* forms rather than the more typically southern-type forms with *aren't*. Part of the reason for this may lie in the stigmatized non-standard usage of this form with the first person singular, *I aren't*.
>
> (1979: 21)

Present tense BE then seems to behave differently from the other verbs – not only for Reading adolescent peer groups, but for all southern speakers of present day English. Whether these hypotheses are also borne out for present day data will be investigated with the help of material from the spoken sections of the BNC in the following section.

Procedure

The basis for this investigation is the BNC-SpS subsample as defined in the Introduction. All searches were conducted per dialect areas (as defined by the BNC headers); all speech by those speakers who are not carrying a dialect tag was therefore not considered. Specifically, the dialect speakers were searched for all combinations of a pronoun or existential *there* with forms of the auxiliary (and primary) verbs and the negator. The main reason for this procedure was the overwhelming frequencies of the phenomena, which made a principled restriction necessary. After all, the primary verbs are among the most frequent words (not to mention the negator), and the size of some of the BNC dialect areas made it necessary to limit

searches in order to avoid soft- and even hardware breakdowns. It is tech-
nically not possible (yet?) to disambiguate every occurrence of a negative
contracted form for its subject type. Although for the smaller dialect areas
this problem did not arise, the very large dialect areas like London or the
northwest Midlands made this procedure necessary, as only in this way
can comparability across dialect areas be guaranteed. It was therefore
decided to search for those closed-class items that could easily be retrieved
automatically, i.e. the complete range of personal pronouns plus the exis-
tential *there*, but to neglect the open class of full noun phrases, which cannot
be searched specifically.[2]

Negative contracted forms were distinguished from auxiliary contracted
forms, where appropriate, on the one hand, and from (the rare) completely
uncontracted forms on the other hand. The inverted word order was also
searched (of necessity only for negative contracted and for uncontracted
forms) in order to find all forms in interrogative environments; interrog-
atives were further subdivided into full interrogatives and tag questions.
Ambiguous forms (especially for auxiliary contracted forms) were disambig-
uated from the context, wherever possible.[3]

Special status of BE

As has already been noted above, forms of present tense BE are said to
behave differently from all other verbs when it comes to negative contrac-
tion, at least in the south. (Most sources imply that auxiliary contraction
is the dominant contraction type in the north anyway.) Contraction thus
seems to be lexically conditioned, at least in some regions of Great Britain.
We shall now test this hypothesis with data from the BNC. Table 4.3
displays the figures for auxiliary contracted forms, subdivided for (present
tense) BE vs all other forms where these two types of contractions are
possible. (These are: present and past tense HAVE, WILL, WOULD and
CAN with the two forms *can't* and *cannot*.) Table 4.3 only includes data
from declarative environments – as we have seen, auxiliary contraction is
not possible in interrogatives.

Table 4.3 shows the very high average rate of auxiliary contraction for
forms of present tense BE (almost 92 per cent), and the very low rate of
auxiliary contraction for all other verbs that permit auxiliary contraction
(just below 5 per cent). Needless to say, this difference is highly signifi-
cant for every single dialect area (at $p < 0.01$). It is also striking – and
rather unexpected – that the figures for auxiliary contraction for BE are
very consistent for all dialect areas; with Scotland the highest (at 97.4 per
cent) and East Anglia the lowest (at 81.1 per cent), an overall difference
of just over 16 per cent. Auxiliary contraction is thus clearly the domi-
nant strategy for present tense BE in all dialect areas, with a frequency
far beyond 50 per cent in both north and south. Although the extreme
values are still relatively close together, it is interesting to note that Scotland

Table 4.3 Auxiliary contraction for BE vs non-BE in the BNC-SpS

BNC code	Dialect area	Auxiliary contraction/total			
		BE	*%*	*Other verbs*	*%*
XEA	East Anglia	298/367	81.2	18/756	2.4
XHC	Home Counties	840/905	92.8	30/1,306	2.3
XHM	Humberside	106/121	87.6	0/170	0.0
XIR	Ireland	110/118	93.2	11/201	5.5
XLC	Lancashire	347/385	90.1	37/571	6.5
XLO	London	1,438/1,529	94.0	82/1,770	4.6
XMC	Central Midlands	496/515	96.3	88/698	12.6
XMD	Merseyside	108/123	87.8	7/131	5.3
XME	Northeast Midlands	319/358	89.1	12/601	2.0
XMI	Midlands	157/164	95.7	25/290	8.6
XMS	South Midlands	124/132	93.9	61/309	19.7
XMW	Northwest Midlands	644/706	91.2	57/1,108	5.1
XNC	Central northern England	390/429	90.9	37/700	5.3
XNE	Northeast England	297/320	92.8	38/544	7.0
XNO	Northern England	108/116	93.1	6/170	3.5
XSD	Scotland	221/227	97.4	34/312	10.9
XSL	Lower southwest England	207/223	92.8	4/415	1.0
XSS	Central southwest England	526/565	93.1	16/906	1.8
XSU	Upper southwest England	89/98	90.8	0/143	0.0
XWA	Wales	377/433	87.1	14/706	2.0
Total		7,202/7,834	Ø91.9	577/11,807	Ø4.9

is the dialect area the furthest north in Great Britain, and East Anglia is one of the most southern dialect areas (in the southeast). The difference in auxiliary contraction ratio for BE for these two dialect areas is statistically highly significant (at $p < 0.01$), which invites the question whether the other dialect areas in between these extremes form a continuum (in which case we would expect decreasing ratios from north to south, without statistically significant differences between adjacent dialect areas), or whether clear boundaries emerge between these two poles. In order to examine these possibilities, the contraction ratio of every dialect area was compared with its direct neighbours and the ratios were tested for statistical significance. The result is displayed in the schematic map in Figure 4.1 (this figure only aims at accurately depicting which dialect area borders on which other area). Bold lines indicate statistically significant differences.

Figure 4.1 shows that at the one extreme, Scotland does indeed behave significantly differently from the rest of Britain. However, in the north–south direction there are practically no other cross-cutting significance
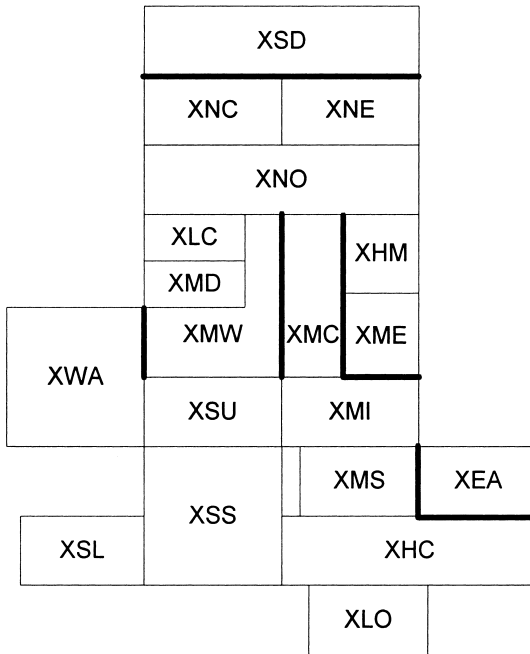
Figure 4.1 Statistically significant differences for aux contraction of BE

boundaries. East Anglia, the other extreme, is similarly cut off from the rest of the country, which explains our initial result of differences between Scotland and East Anglia. All other borders that appear in this map, however, seem to separate areas predominantly in east to west rather than in the north to south direction. Thus, Wales behaves significantly differently from the northwest Midlands, which is significantly different from the central Midlands, which again is significantly different from both the northeast Midlands and Humberside. These differences seem to be mainly due to the exceptional status of the central Midlands, which have a much higher occurrence of auxiliary contracted forms of BE than both neighbours, as a look back at Table 4.3 shows: the central Midlands have a contraction ratio of 96.3 per cent, whereas the northwest Midlands only have 91.2 per cent, and the northeast Midlands only 89.1 per cent. If we compare the northwest and the northeast Midlands directly with each other, we find that there is no significant difference between these two areas.

In a north–south direction, however, we can say that there are in general no significant differences between neighbours. Although this could still be an indication for a dialect continuum (between the two extremes Scotland

and East Anglia), we should in that case expect significant differences between the most northern pole and the most southern one. We find, however, that if we take geographical extremes from this continuous north–south area, there are also no significant differences between, for example, the central north (at 90.9 per cent) and the Home Counties (at 92.8 per cent).[4] In the north–south direction, then, the majority of dialect areas have ratios for auxiliary contraction for present tense BE that are very similar. Despite all indications in the dialect literature so far, data from the BNC do not support any discernible north–south differentiation for the contraction of present tense BE.

Other verbs

When we look at differences between the two kinds of contraction in more detail in this section, the verb BE will be excluded from consideration; as we have seen, BE behaves significantly differently from the rest of the verbs in all dialect areas, and there is no internal regional differentiation of this phenomenon. Table 4.4 therefore details auxiliary and negative

Table 4.4 Negative vs auxiliary contraction in the BNC-SpS (verbs except BE)

BNC code	Dialect area	Total	Neg c.	%	Aux c.	%
XEA	East Anglia	756	733	97.0	18	2.4
XHC	Home Counties	1,306	1,268	97.1	30	2.3
XHM	Humberside	170	170	100.0	0	0.0
XIR	Ireland	201	188	93.5	11	5.5
XLC	Lancashire	571	533	93.3	37	6.5
XLO	London	1,770	1,677	94.7	82	4.6
XMC	Central Midlands	698	609	87.2	88	12.6
XMD	Merseyside	131	124	94.7	7	5.3
XME	Northeast Midlands	601	588	97.8	12	2.0
XMI	Midlands	290	263	90.7	25	8.6
XMS	South Midlands	309	246	79.6	61	19.7
XMW	Northwest Midlands	1,108	1,037	93.6	57	5.1
XNC	Central northern England	700	659	94.1	37	5.3
XNE	Northeast England	544	513	94.3	38	7.0
XNO	Northern England	170	164	96.5	6	3.5
XSD	Scotland	312	275	88.1	34	10.9
XSL	Lower southwest England	415	408	98.3	4	1.0
XSS	Central southwest England	906	888	98.0	16	1.8
XSU	Upper southwest England	143	142	99.3	0	0.0
XWA	Wales	706	687	97.3	14	2.0
Total		11,807	11,172	Ø94.6	577	Ø4.9
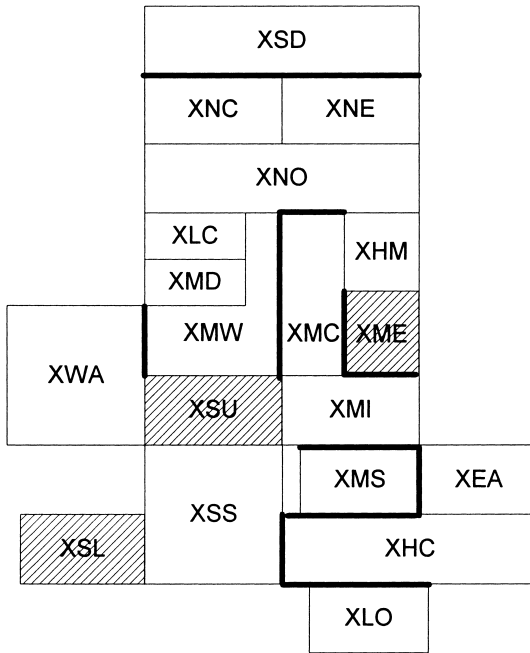
Figure 4.2 Statistically significant differences for neg contraction (excluding BE)

contraction for all verbs except BE. Again, the figures necessarily come from declarative environments only.

One should note that the figures for negative contraction and for auxiliary contraction do not quite add up to the totals (in column one); this difference is caused by fifty-eight total instances of uncontracted forms (ca. 0.5 per cent) in declarative sentences. Uncontracted forms in declaratives will be discussed in more detail below in comparison with uncontracted forms in interrogative environments; for the moment, these figures will be neglected. What emerges from Table 4.4 for negative contraction for verbs other than BE is the striking reverse of the phenomenon of auxiliary contraction for BE dealt with above: negative contraction for all verbs is very much the favoured strategy. Again, the overall average is very high (almost 95 per cent), and the individual dialect areas show values that are again relatively evenly spread, from 100 per cent in Humberside to 79.6 per cent in the south Midlands. Again, a detailed comparison of neighbouring dialect areas in Figure 4.2 investigates the proposed south–north divide (such that northern dialect areas are more likely to use auxiliary contraction than more southern ones).

Figure 4.2 is clearly different from Figure 4.1. Three areas had to be excluded from statistical testing because the figures for auxiliary contracted

forms were below a threshold of five (these excluded areas are Humberside, the upper southwest and the lower southwest, and they are marked grey in Figure 4.2). The analysis of the remaining areas shows that again, Scotland behaves significantly differently from northern England. In the north of England, there is a very homogeneous area of non-significant differences from the border with Scotland extending as far as the north-west Midlands in the west. This might be an indication of a diagonal isogloss for this phenomenon. There is, however, no corresponding homo-geneous area in the south. Here, practically every dialect area shows significant differences from its neighbours. Although this could be an indi-cation of a regional differentiation, if we look at the relative figures it becomes clear that they are not distributed according to increasing frequency from north to south, as one might expect from the literature. Instead, the distribution seems rather random; this can be illustrated by the following exemplary north–south line, extending from the northeast Midlands to London:

XME 97.8% > XMI 90.7% > XMS 79.6 < XHC 97.1 > XLO 94.7

Whereas one would expect increasing ratios from north to south, this comparison indicates that no clear north–south divide – as postulated by Trudgill and others – can be established for this phenomenon either. The figures from the BNC make it very clear that in present day British English, negative contraction is the dominant strategy for all verbs other than BE for all speakers (with the lowest contraction ratio of still 79.6 per cent coming from the south Midlands), just as auxiliary contraction is clearly dominant for present tense BE everywhere (the lowest contraction ratio here is found in East Anglia with 81.1 per cent). Data from the BNC thus show that auxiliary contraction for verbs other than BE is never the domi-nant strategy in absolute terms. Moreover, one can say that even the careful formulation of Hughes and Trudgill (1979) – that one tends to find an increasing likelihood of auxiliary contracted forms in the north – cannot be supported with these data from the BNC.

Instead, one can see that generally, the choice of negative vs auxiliary contraction is lexically conditioned for all dialect areas: negative contrac-tion is the rule for all verbs except BE; only for present tense BE is this relation reversed. These hierarchies are summarized in (1) and (2).

(1) Hierarchy for BE:
     auxiliary contraction > negative contraction

(2) Hierarchy for all verbs (except BE):
     negative contraction > auxiliary contraction

If we recall the systematic possibilities for contractions in Chapter 2, this frequency distribution in (2) corresponds strikingly to the implicational hierarchy posited in Chapter 2 for standard English. The fact that BE behaves very differently from all other contractible verbs is a very important first result that has emerged from our investigation. It is supported by the recent corpus-based grammar of Biber *et al.*, where the authors note that 'apart from the present tense forms of *be*, *not*-contraction is the most common type of reduced form' (Biber *et al.* 1999: 166).

Uncontracted forms

A final point on the topic of auxiliary and negative contraction is a comparison of contracted forms as such (not differentiated for the kind of contraction, i.e. grouping together negative and auxiliary contraction) with completely uncontracted forms. We have seen that because auxiliary contraction is only possible in a declarative environment, the two phenomena of auxiliary contraction and non-contraction intersect: it is generally accepted that the counterpart of auxiliary contraction in declaratives has to be non-contraction in interrogatives (in full as well as tag questions), purportedly especially in the north, as opposed to the more 'southern' negative contraction in all environments. One should therefore expect a

Table 4.5 Uncontracted verbs in the BNC-SpS

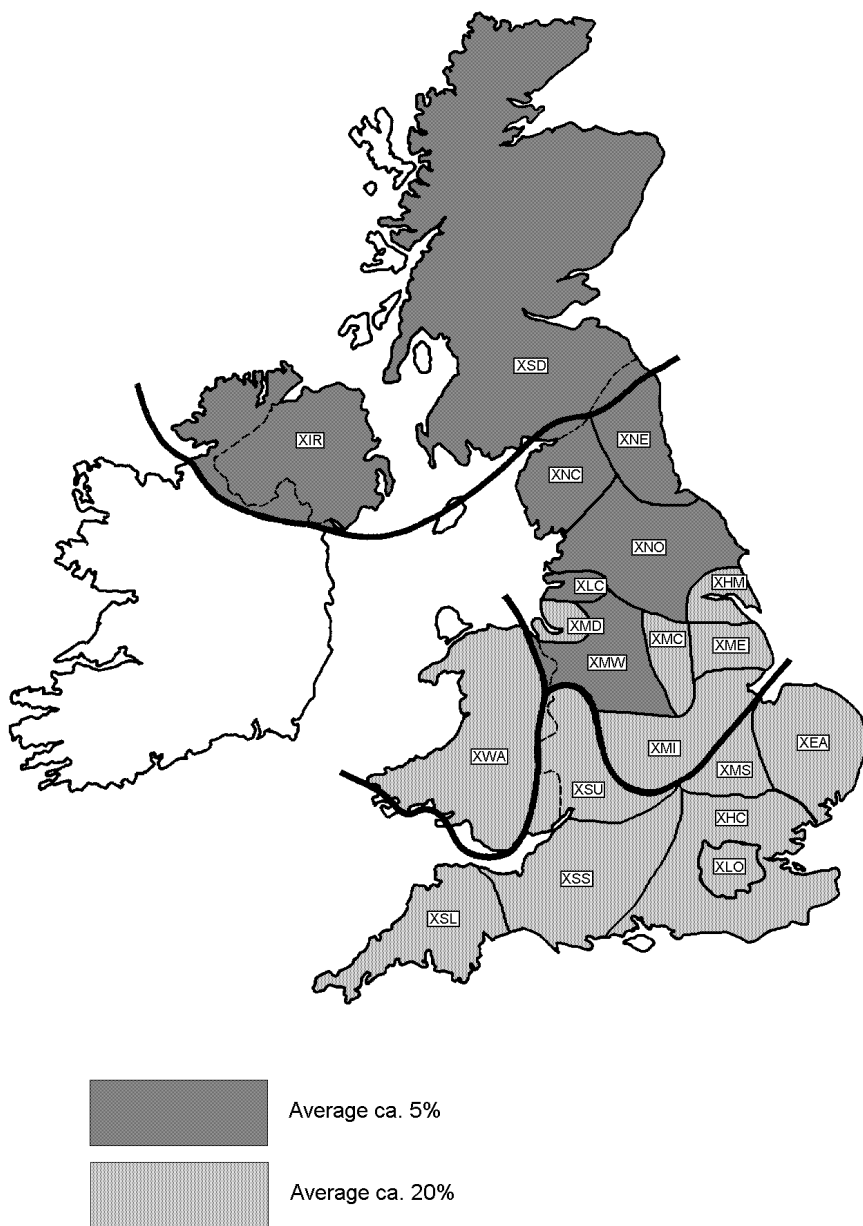| BNC code	Declarative	%	Tag question	%	Interrogative	%
XEA	11/1,123	1.0	5/708	0.7	3/88	3.4
XHC	24/2,211	1.1	2/1,229	0.2	13/224	5.8
XHM	0/291	0.0	0/126	0.0	2/30	6.7
XIR	3/319	0.9	6/150	4.0	9/32	28.1
XLC	3/956	0.3	7/545	1.3	15/89	16.9
XLO	34/3,299	1.0	6/1,807	0.3	26/576	4.5
XMC	4/1,213	0.3	1/702	0.1	5/88	5.7
XMD	0/254	0.0	1/182	0.5	2/27	7.4
XME	2/959	0.2	1/433	0.2	6/108	5.6
XMI	3/454	0.7	0/221	0.0	3/48	6.3
XMS	3/441	0.7	0/290	0.0	1/39	2.6
XMW	22/1,814	1.2	2/1,232	0.2	31/173	17.9
XNC	8/1,129	0.7	2/633	0.3	15/102	14.7
XNE	6/864	0.7	9/334	2.7	11/49	22.4
XNO	0/286	0.0	0/142	0.0	5/27	18.5
XSD	5/539	0.9	13/186	7.0	17/51	33.3
XSL	7/638	1.1	0/473	0.0	5/83	6.0
XSS	5/1,471	0.3	0/1,016	0.0	3/128	2.3
XSU	2/241	0.8	0/126	0.0	0/18	0.0
XWA	9/1,139	0.8	1/754	0.1	4/137	2.9
Total	151/19,641	Ø0.8	56/11,289	Ø0.5	176/2,117	Ø8.3

significantly higher proportion of uncontracted forms in (all) interrogatives than in declaratives in the north. A north–south division should of course also become apparent again.

As we have seen above, a choice between negative and uncontracted forms is possible in all grammatical contexts, both declarative and interrogative. Table 4.4 has already indicated that for declaratives, the overall figures for uncontracted forms are very low (they account for only 0.5 per cent of all occurrences with personal pronoun subjects). This is of course just as expected, both for the spoken informal contexts that the SpS subsample represents, and for the very frequent pronoun–verb combinations that are considered here. However, so far we have not looked at uncontracted forms in interrogatives, specifically in tag questions. Table 4.5 supplies the figures for each dialect area and each syntactic environment.

A look at the last row ('total') already shows that whereas the ratio for uncontracted forms is roughly as low in tag questions as it is in declaratives, it is much higher in full interrogative environments than in either declaratives or tag questions. Considered from a different perspective, from the absolute figures we can see that 176 out of 383 or almost 46 per cent of all uncontracted forms occur in full interrogatives, although interrogatives only account for 2,217 out of 33,047 or less than 7 per cent of all clauses investigated. If we look at the individual dialect areas, this general trend is apparent everywhere. Although in many dialect areas the relations between tag questions (or declaratives) and interrogatives cannot be tested statistically because the absolute figures very often are too low (below 5), where testing is possible, the difference is highly significant at $p < 0.01$ in every single case.[5] In the other cases, a look at the percentages also points in the same direction. Full interrogatives are thus clearly the preferred environment for uncontracted verb forms, or – putting it the other way around – are clearly the dispreferred environment for negative contracted forms.

If we look at the regional differentiation, a striking pattern emerges. All seven areas where uncontracted forms in interrogatives occur in double figures are situated in the north (Ireland, Scotland, the central north, the northeast, the north, Lancashire, and the northwest Midlands; the total for these dialect areas is 103 out of 523 or an average of almost 20 per cent). In all other dialect areas, interrogatives occur with single figures of under 5 per cent (73/1,594). The difference between these two groups of dialect areas is highly significant statistically. Although the figures for the individual dialect areas are relatively low in many cases and statistical testing is not possible for many dialect area borders, those dialect areas that can be tested against their neighbours show indeed that an isogloss seems to be emerging, running diagonally from south of the central north dialect area to the Welsh border south of the northwest Midlands, as Map 4.1 shows.

If we turn our attention to a comparison of declaratives and tag questions, it becomes apparent that the (very low) figures here look very similar.

XSD
XIR
XNE
XNC
XNO
XLC
XHM
XMD
XMC
XME
XMW
XWA
XMI
XEA
XSU
XMS
XHC
XSS
XLO
XSL

Average ca. 5%

Average ca. 20%

Map 4.1 Averages for uncontracted forms in interrogatives

Source: Adapted with permission from Orton, H., Sanderson, S. and Widdowson, J. (1978) *The Linguistic Atlas of England*, London: Croom Helm.

The bulk of uncontracted forms in the declarative environment is due to the two verbs *may not* and *might not.* Although negative contracted forms are possible (*mayn't* and *mightn't*), these two verbs very strongly prefer uncontracted forms, in contrast to practically all other modal verbs, as Table 4.6 shows. (The peripheral modal *ought* probably has to be disregarded, as it is so very infrequent.) Uncontracted forms of *may* and *might* together already account for 131 occurrences of the total of 151 (cf. Table 4.5) uncontracted forms in declaratives, or almost 87 per cent. This is not true for tag questions: *may* and *might* are not responsible for a single uncontracted form in a tag question, and only for one occurrence in a full interrogative. A possible reason for this striking behaviour of *may* and *might* may lie in their semantics. As shown in Chapter 2, *may* and *might* behave exceptionally in the realm of modal verbs in their epistemic meaning. *May* and *might* are the only verbs that move from the (weak) I-corner ($\Diamond$) to the equally weak O-corner ($\Diamond\sim$) when negated in their epistemic sense (*It may not be the milkman* in the sense of 'it is possible that this is not the milkman'). As the negator therefore has no scope over the modal verb for an epistemic reading, this kind of main verb negation seems to act as a barrier to negative contraction and thus may make an uncontracted form preferable.

The main result then is that (apart from a clearly lexically conditioned 'quirk' caused by *may* and *might*) we find uncontracted forms at any remarkable rate in full interrogatives only. For interrogatives, uncontracted forms are significantly more frequent than in tag questions or declaratives in all dialect areas, and here we find a very interesting regional distribution for uncontracted vs negative contracted forms, as Map 4.1 displays. If we compare the percentages of neighbouring dialects, the clearest regional differences emerge. The only indications of a possible north–south divide in the area of contraction thus come from the figures for uncontracted verb forms. These results run counter to both the relevant dialect literature and perhaps the intuitive impressions of many native speakers. There are several possible explanations that can be adduced for this phenomenon. First of all, Trudgill (1984) and Hughes and Trudgill (1979) might have been mistaken – possibly they have noted the striking regional differences in interrogatives in the rates of uncontracted forms and overgeneralized from this to the supposedly parallel phenomenon of auxiliary vs negative contraction in declarative environments. One argument supporting this analysis is the fact that the dialect literature does not differentiate between auxiliary contracted forms in declaratives and uncontracted forms in interrogatives (as we have seen, these are supposed to be equivalents of each other). Different syntactic environments are usually not distinguished – sometimes with the exception of tag questions, which are occasionally mentioned. On the other hand, it is also possible that since the publication of Hughes and Trudgill (1979), which seems to have been the basis of much of the later dialect literature – e.g. Cheshire (1982), Trudgill (1984), Trudgill (1990), etc. – the contraction patterns in Britain may have

Table 4.6 BNC-SpS: uncontracted modal verbs

Modal	Declarative	%
may	31/32	96.9
might	100/136	73.5
(ought	2/4	50.0)
must	7/110	6.4
shall	2/82	2.4
need	0/42	0.0
dare	0/41	0.0
Total	142/447	Ø31.8

changed radically. This can only be investigated by comparisons with historical data. However, longer narrative passages of dialect speakers from the north (as collected in some of the first texts for the corpus project FRED, for example) indicate that negative contraction has been the preferred strategy for all verbs (except BE) for speakers in the north of England, too, for a long time.[6] No plausible explanation for the divergence of our results from the expectation of previous studies can therefore be given. It remains to be stressed however that data from the BNC strongly suggests that auxiliary and negative contraction are lexically conditioned (such that BE favours auxiliary, all other verbs favour negative contraction, except for MAY and MIGHT, which favour uncontracted forms) and not regionally differentiated. Only in the syntactic environment of full interrogatives do we find a regional distribution such that uncontracted forms tend to occur in the more northern dialect areas. Even here, however, the highest ratio is Scotland with 33.3 per cent: uncontracted forms occur at the most in one-third of all cases. Uncontracted forms are thus never the favoured strategy in any dialect area, not even in interrogatives.

The *amn't* gap

A cursory look at synchronic and diachronic data shows that a filling of the standard English 'gap' for the first person singular BE is not a prominent feature of non-standard dialect systems of English. Most forms can be found in Wright's *English Dialect Dictionary* (1898), where he reports forms of *am* with a contracted negative for a range of counties collected in Table 4.7.

Table 4.7 confirms our first impression that in traditional dialects at the end of the nineteenth century (i.e. recording the use of speakers who will have acquired their speech in the first half of the nineteenth century) the form *amn't* seems to be a Scottish/Irish and north–central phenomenon. Where it does occur, *amn't* is in most cases just one of several alternatives; it is never the only option available to the speaker. We can see that if the 'gap' of the standard system is filled, in most dialect areas

Table 4.7 Forms of *amn't* in Wright (1898)

County	Sentence type		Alternative forms
Antrim	declarative	*a imin't, imnae*	*am no*
South Scotland	interrogative	*ym-n' aa?*	
West Yorkshire	declarative interrogative	*ai, a, i amət amət ai, a, i?*	*aim, am, im not*
Lancashire	interrogative	*am't aw?*	
Shropshire	declarative interrogative	*I amma, amna ammad, amnad I*	*arna, binna, bunna*

this is done not by the introduction of a contracted form of the negator with *am*, but either by substitution of *are* for the first person singular, or by a form deriving from *be* (cf. Shropshire with *binna, bunna*). These alternative strategies are much more frequent than forms of *amn't*.

A similar situation holds for the traditional dialect speakers investigated in the SED, whose speakers can be expected to have acquired their language at the end of the nineteenth century. Maps based on the SED (for example, in Orton, Sanderson and Widdowson 1978) show a distribution of negative forms of the first person singular of present tense BE that is similar to the older data collected in Wright (1898). Map M9, for example, shows for (emphatic) declaratives that a negative contracted form based on the standard English *am* is only present in a very restricted area of Yorkshire/Lancashire (*ammet*) and throughout Derbyshire/Staffordshire (*amno'*). In all cases where the negator is contracted (rather than the auxiliary), the forms of the verb are either substituted forms like *ain't/en't/yun't, aren't* or *isn't*, or they can be traced back to the Old English etymon *be/ben*, as, for example, *bain't, baan't, ben't, byen't, byun't*, or *binno'*.

Forms for interrogative *aren't I?* (Orton, Sanderson and Widdowson 1978: M12) show that *amn't*-forms here cover a slightly wider terrain than for declarative environments. *Arem't* can be found in most of Lancashire, Cheshire and Derbyshire. *Ammet* is recorded in Shropshire, and *anno'* is present in pockets in Staffordshire and along the border with Wales. The distribution of *amn't*-forms in both syntactic environments is displayed in the composite Map 4.2.

As expected, for Northern Ireland, the NITCS records two instances of *I amn't* (vs twenty instances of *I'm not*), as Table 3.1 has displayed. At a ratio of two out of twenty-two or roughly 9 per cent for the first person singular this is not significantly different from the overall average for present tense BE: twenty-one out of 180 instances are forms where the negator is contracted in this corpus; this is an average of around 11 per cent. Irish English today then is clearly a dialect area which has closed the 'gap' of the standard English system in a regular way. Glauser notes *amn't I* as a

Forms of *amn't*
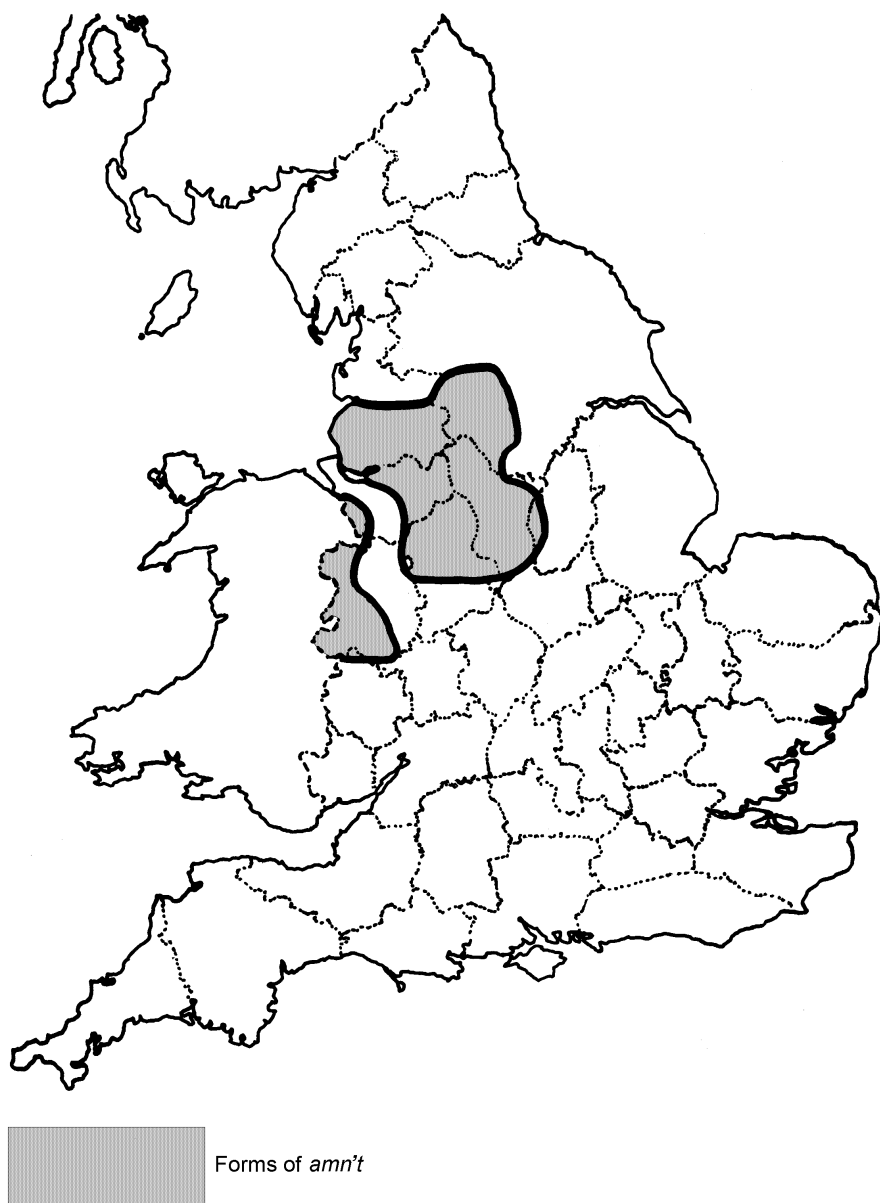
Map 4.2 Amn't in the SED

Source: Adapted with permission from Trudgill, P. (1990) *The Dialects of England*, London: Edward Arnold.

specific feature of Scots (vs Northern English *isn't I*) (Glauser 2000: 69), and indeed we have seen in Chapter 3 that *amnae* is part of the regular Scottish system. At the other end of the country, in the southwest of England, the SRLM records two instances of *I ain't*, as well as one of *I bain't*, but no instances of *amn't*. For present day British English, the BNC does not record a single instance of *amn't* in any dialect area.

Both synchronic data from the BNC as well as diachronic data from Wright and the SED thus show that a negative contracted form based on *am* is not encountered very frequently in non-standard systems (in most cases, it is just one alternative of several); furthermore, its occurrence is very restricted regionally, as Map 4.2 and Table 4.7 have displayed. This gap therefore calls for some explanation, and the absence of *amn't* has recently caused some interesting studies in different theoretical frameworks.

In optimality theory, Bresnan shows that *amn't* is either avoided in standard English or replaced by *aren't* due to a certain order of constraint ranking:[7]

> For declaratives the result is that the syntactic construction *am not* is optimal; for interrogatives, the syntactic construction with *am* inverted and *not* adjoined to VP is optimal. Here syntactic constructions with *am ... not* replace the missing first person singular negative inflected form of *be*.
>
> (2001: 38)

This approach has the advantage that different rankings can explain differences in dialects. It remains, however, largely the typical system-internal explanation of generative grammar and its derivatives, which is not very helpful for outsiders, and unsatisfactory in functional terms. Unfortunately, Bresnan explicitly excludes a discussion of *I'm not* ('the choice between the full verb *am* and the reduced *'m* is an orthogonal issue that is not addressed here', Bresnan 2001: note 28), although this might play a crucial rule, as we shall see below.

Another explanation is presented by Hudson (2000) in the theoretical frame of his (functionalist) word grammar. Hudson argues that the avoidance of *amn't* stems from a conflict of features between those inherited from the more general form *aren't*, which is the default form of present tense BE specified for negation, and those inherited from the form *am*, specified for the first person singular, but unspecified for negation. The presence of these two forms results in a Nixon diamond, a conflict that cannot be resolved (except by stipulation), as neither form is more specific than the other. The problem with this explanation seems to rest with the definitions; in word grammar, negative verbs can only be formed by adding *-n't* to the corresponding positive forms, and only those forms count as positive forms that are arrived at through the subtraction of *-n't* (Hudson 2000: 307). This leads to the analysis of *am* not as a positive form, as it

clearly lacks a corresponding negative, but a WHOLE which by definition prevents the generation of a negative form.[8] This argument thus seems to be rather circular, as it includes in the premiss what should only be contained in the conclusion. In addition, and like Bresnan, Hudson does not take account of auxiliary contracted forms of BE, which are more than just a stylistic alternative. Indeed, I shall argue that the hierarchy reversal as proposed in the previous section holds the key to an explanation of the absence of *amn't*.

As we have seen, BE in general overwhelmingly prefers negative contraction over auxiliary contraction. For the first person singular this means that *I'm not* is much the preferred strategy (over a non-existent *I amn't*), just as *he's not* is preferred over *he isn't* and *we're not* is preferred over *we aren't*. On the basis of the frequency distributions described in Table 4.3 it is reasonable to assume a reversed hierarchy for present tense BE, as we have seen in (1) (p. 80); the hierarchy would then seem to operate lexeme-specifically, as an exception to the more general rule in (2) above. Thus, for present tense BE one would posit auxiliary contraction as the unmarked option, and negative contraction as the marked alternative. Table 4.8 shows the result of this reversal in comparison with the assumed standard English system.

Table 4.8 shows that the assumed 'gap' of standard English, marked grey in column two, simply disappears with the reversed hierarchy for present tense BE (although at the cost of marking the whole present tense paradigm). With the reversal of the hierarchy for present tense BE, expectations will also sink dramatically that non-standard dialect systems should fill the 'gap' of the standard English system (because it is in fact not a gap) and introduce a negative contracted form like *amn't* for the first person singular, in order to bring this paradigm in line with the rest of the present tense system for all other verbs. The fact that we find no negative contracted form of present tense BE for the first person singular (**amn't*) although

Table 4.8 Contractible verbs (primary verbs only)

	Neg c.	*Aux c.*		*Unmarked*	*Marked*
BE	—	'm not		'm not	—
	isn't	's not	*(reversed)*	's not	isn't
	aren't	're not		're not	aren't
	wasn't	—		wasn't	—
	weren't	—		weren't	—
HAVE	haven't	've not		haven't	've not
	hasn't	's not		hasn't	's not
	hadn't	'd not		hadn't	'd not
DO	don't	—		don't	—
	doesn't	—		doesn't	—
	didn't	—		didn't	—

there is a form where the auxiliary is contracted (*I'm not*) will thus fit the reversed hierarchy and does not have to be treated as an exception or a 'gap' in the system. Instead, it would seem that it is the whole paradigm of present tense BE which behaves in an exceptional way. Although this explanation is sufficient reason for the non-existence of **amn't*, it of course raises the further question of why BE should behave in this particular way, different from all other contractible verbs.

There are several possible explanations for this behaviour of BE. Hazen (1996) gives a functional explanation. In a phonotactic analysis of the unusual behaviour of present tense BE he gives a plausible reason why auxiliary contraction is preferred to negative contraction. Based on data from the study of Ocracoke/North Carolina (cf. also Chapter 8) and in accordance with our results from the BNC, Hazen finds that negative contracted forms of present tense BE are clearly the dispreferred option; 'the non-affixed forms [i.e. uncontracted forms], *ain't* and the 'verb plus *not*' forms [i.e. auxiliary contracted forms], have significantly higher frequency than the *-n't* affixed forms (*isn't, aren't*) [i.e. negative contracted forms]' (Hazen 1996: 101). Hazen claims that the reasons for this imbalance are phonological in nature: 'the phonological structure of the variants seems to influence the choice of forms . . . the process of syllabification constrains the choices of phonological forms for negating present tense *be* in English' (Hazen 1996: 110).[9]

Although Hazen argues convincingly for the preference of negative forms that are found in Ocracoke, there are a range of problems when we try to apply his findings to the data from the BNC. First of all, Hazen does not distinguish between auxiliary contracted and completely uncontracted forms, which he treats together. He also does not distinguish syntactic environments, just as the majority of previous studies already quoted. Second, although Hazen manages to supply a plausible explanation for present tense BE, it is difficult to imagine how the same arguments could explain distributions found, for example, for present tense HAVE, which are exactly the reverse. If it is much more natural on phonotactic grounds to say *he's not* rather than *he isn't*, which is also supported by the frequency distributions in all dialect areas of the BNC, as we have seen, then why is it not also much more natural to say *we've not* instead of *we haven't*? For present tense HAVE, the phonotactic arguments applied by Hazen to forms of present tense BE do indeed result in an analysis which shows that *we've not* is much better formed phonotactically than *we haven't*. On phonotactic grounds, then, auxiliary contraction should also be the preferred strategy for present tense HAVE – a prediction that is clearly not borne out for any dialect area, nor, one might assume, for dialect speakers of Ocracoke. Indeed, any phonotactic explanation will presumably run into serious difficulties if the same arguments are used to try to explain these diametrically opposed phenomena.

Hughes and Trudgill hint instead at a socially motivated explanation for a preference of *you're not* over *you aren't* when they say that 'part of the

reason for this may lie in the stigmatized non-standard usage of this form with the first person singular, *I aren't* (Hughes and Trudgill 1979: 21). This is certainly a plausible explanation for the first person singular. *I aren't* is avoided because of its unclear status, perhaps even stigmatization, and *I'm not* is used instead. This explanation presumes, however, that the irregularity of one form (in this case the first person singular) is able to influence the whole paradigm, so that speakers would also prefer *he's not* to *he isn't*, *we're not* to *we aren't*, etc. although these forms are certainly not stigmatized. As we have seen, it is indeed the entire paradigm of present tense BE which behaves 'irregularly', not just the first person singular. While it is not implausible that one form could influence the rest of the paradigm, from the synchronic data it is unfortunately practically impossible to determine cause and effect for this phenomenon.

There just might be another plausible reason that could explain the special status of BE. We have said before that the pattern we find for contraction ratios for BE versus all other verbs are reminiscent of areas of local markedness or a markedness reversal for the complete paradigm of (present tense) BE. This phenomenon of a 'localized hierarchy reversal' or 'markedness reversal' is well attested in linguistic typology for a range of different phenomena. These local 'exceptions' to otherwise general rules are usually cognitively well motivated and therefore do not undermine a claim about the majority of other cases. Particularly well-described markedness reversals come from the category of number. Generally, plural is more marked than singular, which is unmarked (cf. Croft 1990: 66).[10] However, there are objects that typically appear in pairs or as collectives, such as foodstuffs, groups of animals, birds or people, or body parts, and here very often it is the plural that is unmarked whereas a singular form (in this case called 'singulative') is either non-existent or – contrary to the normal singular/plural-morphology – morphologically more complex than the (unmarked) plural. These *pluralia tantum* nouns are known for many languages.[11] The functional reason behind these reversals is clear: as Croft points out, 'this correlation between the category of number and the noun class represents a prototype: . . . some objects are prototypically non-singular (collective)' (Croft 1990: 66). The question then is, in which way can the verb BE be said to be prototypically different from all other verbs, which would justify its exceptional behaviour? In the case of contraction of BE a functional explanation does not seem to be forthcoming very easily, as contractions are obviously much more abstract than plurality in nouns and it is not immediately obvious how negative contraction rather than auxiliary contraction could be less or more 'marked' in any meaningful way. Calling the specific phenomenon we have observed a markedness reversal is of course not in itself explanatory, but merely restates the same in a slightly different framework. However, it can be argued that the verb BE is indeed very different from the other (contractible) primary verb HAVE, or the range of auxiliaries that can be contracted, in

that it carries the least semantic content. The presence or absence of BE does not greatly change the overall meaning of a clause – thus a verbless clause is most naturally interpreted as containing a form of BE. And indeed there are languages that do not possess a form for BE in the present tense, or where it can be variably deleted, as, for example, in many creoles and pidgins (Holm 1988: 174–5).[12] Although Holm calls this a 'very un-European pattern' (1988: 175), copula absence[13] is not that exotic; it is also present in an Indo-European language like present-day Russian. Even in languages where the copula is regularly expressed, dialects or specific contexts may show copula absence. For English, probably the most promi-nent example where we find this phenomenon is AAVE. It is exactly the present tense forms of BE that can principally always be deleted, resulting in sentences like *He tired.*[14] A sentence like *He very nice man* is indeed prob-ably always – unambiguously – interpreted as lacking a copula.

In other words, although the copula BE has an important formal func-tion, its semantic content is extremely low. For this reason it is perhaps not too surprising that, after all, forms of present tense BE tend to be contracted (*He's not*, *we're not*) rather than the negative, simply because their semantic content is so slight, whereas the negator carries a great semantic load. In the extreme case of AAVE this has led to the complete deletion of present tense BE in exactly those positions where BE can be contracted. Seen from this perspective, then, it does indeed make sense to consider the paradigm of present tense BE as a local spot where the prototypical function and thus markedness is reversed. The prototypical function of a verb, the denoting of an 'action' of some kind, is clearly not fulfilled by BE, whose prototypical meaning could perhaps be described by 'existence', 'stativity' or 'habituality'.[15]

In other words, the verb BE is indeed cognitively different from all other verbs, so that a pattern of local markedness can be motivated cognitively. This markedness reversal can then explain the lack of a form like *amn't* in most dialects of English: the markedness reversal shifts the 'gap' towards the end of the system where the presence of a form is optional. In the dialectal systems, then, there is no urgent need to fill this 'gap', because there simply is no gap. **Amn't* is simply one more form that is not present in the system, and as such has the same status as the non-presence of an auxiliary contracted form of *he does not* or *I must not*.

Negation of HAVE (full verb)

Introduction

As we have seen in Chapter 2, the full verb HAVE in standard English can be negated either with the help of DO-support, or without, and the same is true for spoken English, as examples (3) and (4) show. Although this is therefore not necessarily a typical feature of *non-standard* dialects, it

is frequently quoted in the dialect literature and is supposed to have a regional distribution (e.g. Hughes and Trudgill 1979: 25–6).

(3) I personally don't have any children at all. (KC3 304)

(4) We haven't a baby. (KCD 2699)

Both strategies are equally grammatical; the main difference seems to be a stylistic one: full verb HAVE negation without DO seems to be the more conservative form, whereas the more regular strategy with DO-support is no doubt gaining ground and, for example in American English, has practically replaced negation without DO, as has been demonstrated in Chapter 2. A development that is taking place for the written standard can also be expected to take place in the spoken standard and in non-standard speech, and perhaps one might even expect this to have taken place in informal spoken language earlier than in the more formal written styles. On the other hand, for many phenomena, dialects are more conservative than their standardized variety. How British dialects pattern for this phenomenon shall be investigated in the following section.

Procedure

For the following investigation, the BNC-SpS subsample of the BNC spoken texts was searched per dialect area (as defined above) for forms of negated full verb HAVE. In order to find full verb HAVE negated with DO-support, the following combinations were extracted: *don't have*, *doesn't have*, *didn't have*. Uncontracted forms, as shown above, can quite generally be neglected and were therefore not searched in particular. Instances of inversion have also not been searched for. All instances of the semi-modal HAVE *to* were discounted, but since interesting variation appeared for this auxiliary form as well, HAVE *to* will be discussed below in more detail. The results were not restricted to personal pronoun subjects for this part of the study (for HAVE and HAVE *to*). However, of the sixty-eight instances of full verb HAVE negated with DO listed below, only four occurrences had a subject different from a personal pronoun. (Indeed, three of these four instances were elliptical structures without an overt subject, in only one case was the subject a full NP – this also validates the approach taken in the other sections of only searching for pronoun–verb combinations.) For forms of full verb HAVE negated without DO-support, the forms *haven't*, *hasn't* and *hadn't* were searched; the results were then restricted to those forms of negated HAVE followed by a direct object. (This excludes all auxiliary verb uses, especially all instances of *haven't got to*.) The results for negated HAVE used as a full verb are presented in Table 4.9.
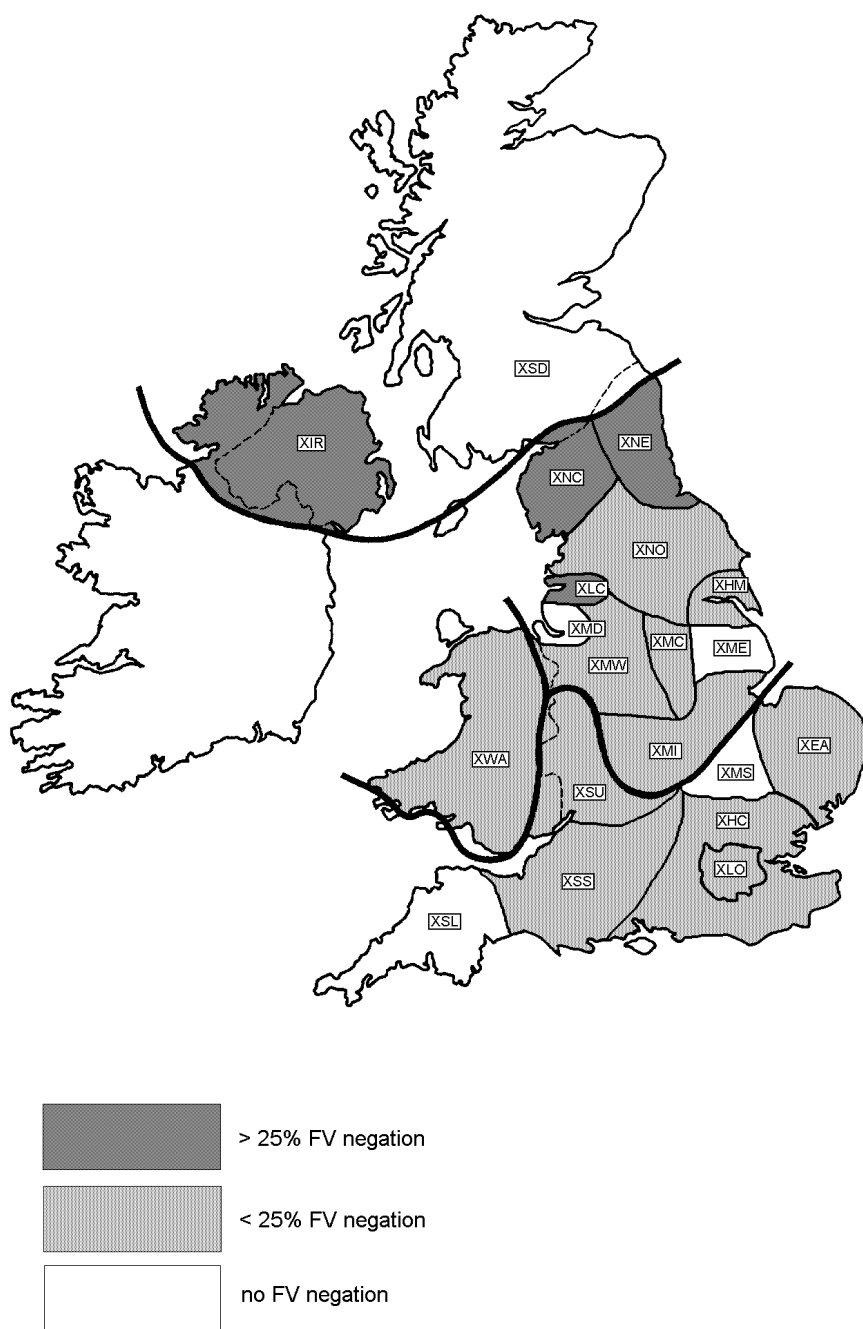
The first point to emerge from Table 4.9 is the fact that although full verb HAVE negation without DO, the more conservative variant, is not

Table 4.9 Negation of full verb HAVE in the BNC-SpS

BNC code	Total	+DO	−DO	% of total
XEA	35	34	1	2.9
XHC	87	82	5	5.7
XHM	17	15	2	11.8
XIR	8	5	3	37.5
XLC	33	24	9	27.3
XLO	103	86	17	16.5
XMC	35	30	5	14.3
XMD	3	3	—	—
XME	33	33	—	—
XMI	12	11	1	8.3
XMS	12	12	—	—
XMW	39	33	6	15.4
XNC	22	16	6	27.3
XNE	25	19	6	24.0
XNO	17	16	1	5.9
XSD	17	17	—	—
XSL	17	17	—	—
XSS	50	46	4	8.0
XSU	6	6	—	—
XWA	38	36	2	5.3
Total	609	541	68	Ø11.2

present in all dialect areas, the overall average of slightly over 11 per cent is relatively high. If we compare this with the figures from the written texts in Chapter 2, where we found an average of 15 per cent for 1991 in British English, the average for the different dialect areas is slightly lower, i.e. spoken language in this respect does indeed seem to be more progressive than the written language.

A second important point is the regional structure. There are six areas where full verb *haven't/hasn't/hadn't* does not occur at all, and these are not only those dialect areas where HAVE is relatively infrequent anyway. These six areas are Scotland, the northeast Midlands, the south Midlands, the lower southwest, the upper southwest and Merseyside. Only the last two have very low occurrences of the negated full verb HAVE (six and three respectively). The rest is concentrated in the south of England – with the exception of Scotland, of course. However, we have seen in the preceding chapters that Scotland often behaves in a markedly different way to its direct neighbours, and perhaps should be discounted in this case as well. The southern area then, not surprisingly, seems to be more progressive than its neighbours with respect to this feature. On the other hand, there are four dialect areas where full verb HAVE is negated without DO-support in around 25 per cent or more of cases. These four areas are Ireland, Lancashire, the northeast and the central north and they form a very homogeneous area in the north. The average for these four dialect

Map labels: XSD, XIR, XNE, XNC, XNO, XLC, XHM, XMD, XMC, XME, XMW, XMI, XEA, XMS, XWA, XSU, XHC, XSS, XLO, XSL

> 25% FV negation

< 25% FV negation

no FV negation

Map 4.3 Progressive and conservative areas for full verb HAVE

Source: Adapted with permission from Trudgill, P. (1990) *The Dialects of England*, London: Edward Arnold.

areas is 27.3 per cent – an average that is significantly higher than the 15 per cent found in written texts. This northern area then seems to be much more conservative with respect to this feature. Both areas are displayed on Map 4.3.

For the negation of full verb HAVE, then, we find an interesting regional differentiation, such that the more southern dialect areas are more progressive than standard English, which seems to be influenced by these varieties, whereas the more northern dialect areas conserve the older use of full verb negation to a greater degree. Again, though, full verb negation is never the dominant strategy; even in Ireland it does not exceed 37.5 per cent and is thus used in just over a third of all cases.

The same searches as for full verb HAVE were conducted for the semi-modal HAVE *to*, because the initial investigation showed that in some – infrequent – cases negation without DO-support did occur, as examples (5) and (6) show.

(5) I hadn't to put my head out, I hadn't to be seen. (KCS 1800)

(6) We hadn't to go. (KSS 1219)

The results are displayed in Table 4.10.

Table 4.10 shows that negation of HAVE *to* without DO is indeed a fairly infrequent phenomenon. Even the central north and Lancashire, which show high relative frequencies for this feature (of over 27 and 15 per cent respectively), have very low absolute figures (eleven and twenty occurrences in total respectively), so that the three occurrences each of *haven't to* do not really constitute an exception. HAVE *to* is clearly used much less frequently in absolute terms than HAVE as a full verb (348 times as opposed to 610 for full verb HAVE, almost a ratio of one in two). In relative terms, however, only 4.3 per cent of all instances of HAVE *to* are negated without the use of DO-support, whereas for the full verb HAVE this figure stands at 11.1 per cent – a highly significant difference (at $p < 0.01$). In other words, the less frequent form (in absolute terms) HAVE *to* also has far fewer occurrences of the negation without DO-support (in relative terms). This supports the hypothesis that we seem to be dealing here with a process of simplification or regularization, which can generally be assumed to proceed through the infrequent paradigms first, whereas very frequent words can retain irregularities for longer.[16] Just as expected, the negation of HAVE as well as of HAVE *to* without DO seems to be in the process of being replaced by the more regular negation with DO-support. And indeed several other indicators support this observation; both the diachronic evidence from written corpora (cf. Chapter 2) and the irregularity scale (cf. Chapter 2) predicted pressure towards change, as negation with DO-support conforms to the vast majority of negated verbs, both main and auxiliary, the only exception being BE, as shown in Chapter 2.

Table 4.10 Negation of semi-modal HAVE *to* in the BNC-SpS

BNC code	Total	+DO	−DO	% of total
XEA	16	16	—	—
XHC	39	39	—	—
XHM	7	7	—	—
XIR	1	1	—	—
XLC	20	17	3	15.0
XLO	79	74	5	6.3
XMC	18	17	1	5.6
XMD	8	8	—	—
XME	17	17	—	—
XMI	9	9	—	—
XMS	6	6	—	—
XMW	29	26	3	10.3
XNC	11	8	3	27.3
XNE	7	7	—	—
XNO	4	4	—	—
XSD	9	9	—	—
XSL	18	18	—	—
XSS	28	28	—	—
XSU	5	5	—	—
XWA	17	17	—	—
Total	348	333	15	Ø4.3

Epistemic *mustn't*

This final section will turn to investigate a further 'gap' of standard English. As Table 2.13 in Chapter 2 has shown, in standard English there is a gap in the system of negated modals: the form *mustn't*, where the main verb is negated, only has deontic meaning (e.g. *You mustn't do that* meaning 'You are obliged not to do that'). The epistemic meaning of positive *must* (necessity) cannot be negated by using *must*; this meaning ('It is necessary that . . . not') is usually substituted by *cannot*, e.g. *His absence must have been noticed* can only be negated in the following way: *His absence can't have been noticed.*[17]

In non-standard dialects, however, just as in American English, epistemic *mustn't* is reported to be possible and present, e.g. *His absence must not have been noticed.* Indeed, Quirk *et al.* note that 'such sentences have been regarded by many commentators as impossible, but are increasingly accepted and used, especially in Am[erican] E[nglish]' (Quirk *et al.* 1985: 225). For the following sections, exceptionally, all instances of *mustn't* were collected (not just those following a personal pronoun, as in the remainder of this book) and disambiguated according to deontic and epistemic meaning. The searches were again restricted to the BNC-SpS texts, and conducted per dialect area. The results are displayed in Table 4.11.

Table 4.11 presents some very interesting results. As the row 'Total' shows, epistemic *mustn't* is far from a marginal phenomenon – it accounts

Table 4.11 Mustn't in the BNC-SpS

BNC code	Total	Deontic	Of which in tag	Epistemic	Of which in tag
XEA	13	9	1	3	3
XHC	12	9	1	2	1
XHM	1	—	1	—	
XIR	1	1		—	
XLC	4	—		4	3
XLO	23	15	2	6	6
XMC	7	3	1	3	3
XMD	5	5		—	
XME	10	9		1	1
XMI	3	2		1	1
XMS	2	—		2	2
XMW	5	3		2	2
XNC	3	—		3	2
XNE	4	3		1	1
XNO	3	—		3	1
XSD	3	3		—	
XSL	1	—	1	—	
XSS	11	5		6	6
XSU	2	2		—	
XWA	5	4		1	1
Total	118	73	7	38	33
		61.9%		32.2%	

for thirty-eight out of 118 cases or almost one-third of all occurrences of *mustn't*. It would be interesting to compare these figures to an overall distribution of deontic vs epistemic uses of *must* – unfortunately, no estimates are available from the literature. A random sample of 190 instances of (positive) *must* from all spoken texts from the BNC was therefore collected. This sample suggests that epistemic *must* is far from a marginal phenomenon. On the contrary, the majority of uses of *must* are in fact epistemic (impressionistically, the semantics of 'obligation' is more idiomatically expressed by HAVE *to*/HAVE *got to* rather than by *must*): out of 190 cases of *must*, ninety-seven or 51.1 per cent were epistemic, even more than half of all occurrences. It is therefore not surprising that in negative contexts, epistemic *mustn't* is, equally, not infrequent. Epistemic *mustn't* does not seem to be regionally restricted; only in those areas where *mustn't* is, overall, rather rare does it not appear. Generally, however, figures are too small to be analysed per dialect area.

A further point to emerge is that epistemic *mustn't* is almost completely restricted to occurrences in tag questions: thirty-three out of thirty-eight instances or almost 87 per cent occur in this grammatical environment. Examples are given in (7) and (8).

(7) He must have been interested in cooking, *mustn't he*?
     (XLC: KBP 0)

(8) Must be ill *mustn't he*, if he can't even have visitors?
     (XNO: KBC 2752)

If we discount this type of construction for the moment, only the following
cases are cases of 'true' epistemic *mustn't* in independent clauses:

(9) *Mustn't* be a very good job then, Gill? (XHC: KP5 1243)

(10) *I mustn't* have been out of that office half an hour.
      (XLC: KCW 7330)

(11) Oh, *I mustn't* have read it [a letter]. (XNC: KB8 8072)

(12) *He mustn't* have liked where it was. (XNO: KBC 1270)

(13) *She mustn't* like fruit. (XNO: KSS 4939)

The distribution of these examples over dialect areas does very vaguely
point in the direction of a regional differentiation. Epistemic *mustn't* in
declaratives seems to be concentrated in the north – only the first example
comes from a southern dialect area (the Home Counties). However, the
absolute frequencies here are so low that this can remain at best a trend
and describe perhaps an impression to be investigated in further studies.
    The majority of instances of epistemic *mustn't* then are found in tag
questions, as we have seen. Although the distribution is reversed for deontic
mustn't, a similar problem arises: there are tag questions that are formed
using this modal (according to the grammatical principles of tag question
formation) – however, employing *mustn't* in this way entails a shift in
meaning (from obligation in the main clause to prohibition in the tag) that
results in a semantically rather odd construction, as in (14):

(14) I must go to Croydon *mustn't I* to change that bra.
      (XHC: KBH 6728)

One possible way out of this dilemma would be the choice of a different
modal or semi-modal, e.g. HAVE *to* (*I have to go to Croydon to change that bra,
don't I?*). Once the speaker has chosen to use *must* in the main clause,
however, the use of a tag necessarily results in an awkward construction.
It is therefore not surprising that all seven tag questions containing *mustn't*
in its deontic use are semantically odd. The use of epistemic *mustn't* –
which, as we have seen, is found almost exclusively in tag questions –
essentially seems to be the consequence of the same conflict between

grammar (tag question formation and use) and semantics (a shift of meaning). Both for deontic and epistemic *must*, grammar seems to be the clear winner.

Epistemic *mustn't* that does not result from this conflict (i.e. epistemic *mustn't* in declaratives, as in examples (9)–(13)) seems to be concentrated in the north. However, it is much more widely distributed than some claims would suggest; epistemic *mustn't* is not a characteristic phenomenon for one dialect area only (e.g. for Irish English). A lack of data, however, must leave a quantitative comparative study to the future.

Summary

In sum, we have seen that non-standard English has various strategies of dealing with standard English gaps and irregularities. In the case of **amn't*, there are convincing arguments that can simply explain the standard English 'gap' away: a hierarchy reversal that takes account of the special (semantically empty) status of the lexeme BE pushes the gap to the end of the hierarchy where filling it becomes optional, and indeed optional forms of *amn't* are what we find in the historical dialects of English. In the case of epistemic *mustn't*, a true gap of standard English is obviously perceived as an impediment to conversation and thus filled, resulting in a regularized system. The irregularity of standard English of permitting full verb HAVE to be negated without a form of DO is preserved in non-standard spoken English to a degree; not surprisingly, we found the north as a more conservative dialect area, whereas the more regular forms (full verb negation with a form of DO-support) are clearly spreading from the south. This also corresponds to results from irregular HAVE *to*, again negated without DO, which only occurs in isolated pockets, mainly in the north, but does not constitute a numerically significant grammatical feature. Clearly, the regular strategy of negation with DO-support is the overall winner for these two phenomena, slowly erasing the special status of the primary verb HAVE and bringing its full verb use in line with all other main verbs.

5 Negative concord

Introduction

This chapter deals with one of the best-known features of non-standard English, the use of double (or even multiple) negation, perhaps better called negative concord, as in examples (1) to (3).

(1) I *couldn't* do *nothing* about it. (KCT 7357)

(2) I *haven't hardly* had *no* fags today. (KC5 2549)

(3) You'd *never* heard *nothing*. (KCP 1775)

After a look at the history of this phenomenon in the English language, this chapter investigates the extent to which neg concord is still present in British English today, and which combinations it favours. After these structural points, we shall then examine whether there is not a regional differentiation after all, despite the claims in the literature that neg concord is so pervasive geographically. Finally, we shall situate English dialects in a typological context and decide whether English dialects are typologically more marked or less than their standard counterpart.

The phenomenon

The term *negative concord* characterizes two slightly different phenomena which will be grouped together here. One is the co-occurrence of a sentence negator, e.g. StE *not* or *-n't*, with a negative quantifier in its scope, but with an overall negative reading to the clause. Bernini and Ramat (1996) give this phenomenon the technical term of *neg-permeability*. As we have seen, standard English is neg-impermeable; only the generic A-quantifiers are allowed in the scope of the sentence negator for a negative reading of the clause (e.g. *I didn't see anything* rather than **I didn't see nothing*, cf. Chapter 2). The second phenomenon has to do with the fact that the sentence negator in English is not obligatory to make a sentence negative;

an N-quantifier on its own can have the same effect, as in: *Nobody saw me.* Again, as we have seen in Chapter 2, in standard English the combination of a (sentence negation effecting) N-quantifier with other N-quantifiers is not permitted for an overall negative reading of the sentence (e.g. *Nobody saw anything* rather than *Nobody saw nothing) – Bernini and Ramat (1996) term these kinds of languages N1 languages. In general, then, any quantifier inside the scope of the negator, whether this is the negative particle or a negative quantifier, has to take the A-form in standard English in order to have an unmarked negative reading, as the examples above have shown. On the other hand, any occurrence of one or more negative quantifiers inside the scope of the negator with an overall negative reading will be taken as an instance of negative concord, or multiple negation (as in the starred examples above).

History

Not surprisingly, the situation we find today for standard English has not always held throughout the history of English. Old English and Middle English were neg-permeable and Nx languages, i.e. in both periods, negative elements were permitted inside the scope of the sentence negator.[1] Indeed, it has been suggested that multiple negation was not only permitted but obligatory even in Middle English times (cf. Jack 1978b, 1978c).[2] This only changed towards Early Modern English – a change which is often attributed to the rise and influence of prescriptive grammars based on classical Latin, where negative concord was generally disqualified as 'illogical'. Indeed, classical Latin is a neg-impermeable and an N1 language where subsequent negators therefore effect logical double (='cancelling') negation.[3]

Recently, however, this established view has been challenged; there are indications that the decline of negative concord was already under way 'naturally' long before the influential grammars of the eighteenth century could enforce their prescriptive dictum. Thus Strang in her *History of the English Language* already notes that 'by Shakespeare's death multiple negation has almost passed out of standard use' (Strang 1970: 152). Mazzon in her investigation of Middle English texts suggests that 'M[ultiple]-Neg[ation] is declining already at Chaucer's time . . . and was kept alive till about 1600 only where it was functional of the expression of emphasis and of other attitudes and subjective connotations' (Mazzon 1994: 164). Nevalainen narrows down the window to the period of 1520–50; on the basis of data from the Corpus of Early English Correspondence she claims that 'the disappearance of multiple negation was well under way . . . in the first half of the 16th century' (Nevalainen 1998: 284). Jespersen provides a procedural explanation for the decline: 'the rarity [of multiple negation structures in Elizabethan English] is probably due to to [*sic*] the fact that the ordinary negation at that time was the comparatively bulky *not*, which

had not yet dwindled down to the less conspicuous *-n't*' (Jespersen 1940: 451).[4]

On a different line of attack, Austin (1984) argues against too large an influence of prescriptivism in general. He shows that even in Lowth's influential prescriptivist grammar (written for the academic, Latin-trained upper classes):

> there is nothing on *not . . . no* or *sentence element* double negatives. Other grammarians, such as Priestly and Baker, who state that two negatives cancel each other and make an affirmative, are also primarily concerned with sentences containing *neither . . . nor*.
>
> (Austin 1984: 140f.)

These additive–correlative contexts are precisely those syntactic contexts where neg concord seems to have survived longest (cf. Nevalainen 1998), even in the upper classes. Austin concludes from this that rather than dying out, 'by the seventeenth century double negatives . . . had already been consigned to the status of sub-standard [*sic*] English' (Austin 1984: 143).

Certainly from the eighteenth century onwards, occurrences of neg concord in the literature are generally used for the portrayal of lower-class speech, as the examples from Jespersen (1917) show. Austin (1984), however, also investigates lesser known school grammars from the 'provinces', writing for those non-academic audiences where he supposes that multiple negation may have survived. In addition, Austin quotes direct evidence from personal letters of relatively uneducated people (the Clift family). From the indirect evidence of the school grammars condemning the use of negative concord and the direct evidence from the Clift family letters one can conclude that indeed by the eighteenth century, negative concord had become a feature of non-standard speech. Tieken-Boon van Ostade (1982) also examines eighteenth-century grammars in some detail and comes to the conclusion that they mostly use (and argue against) examples taken from seventeenth-century writing. As the use of multiple negation in formal prose had already disappeared, the impact of these prescriptivists must have been rather small. The actual spoken usage was naturally relatively unaffected by grammars and indeed Tieken-Boon van Ostade comes to the conclusion that 'double negation continued to be used in the eighteenth century in some informal prose and of course in spoken English' (Tieken-Boon van Ostade 1982: 284).

We can summarize these detailed studies and say that the shift of negative concord from standard to non-standard usage must have occurred at some stage in Early Modern English. It is therefore not surprising to see that features noted as non-standard for the eighteenth century should still be present in dialects today. In fact, negative concord is one of the most common features of non-standard English worldwide. For example, Wolfram and Schilling-Estes state explicitly that:

> there are a number of features of English language variation, such as the use of 'double negatives' (for example, *She ain't been nowhere*) . . . whose distribution among various populations is best explained by starting with considerations of social status difference. In other words, these features tend to be found among lower-status speakers in all dialect regions rather than being confined to speakers in particular areas.
>
> (1998: 31)

More specifically on British English, Edwards and Weltens (1985) in their survey of British non-standard grammar found that:

> the construction is very widespread, as examples of it were given for virtually all dialects that were studied. On the whole, it seems that the use of more than one negative is a matter of concord, not a means of intensification.
>
> (1985: 107)

Negative concord lost its obligatoriness already in Middle English times (in the standard), and with the growing influence of the newer standard neg-impermeable system on non-standard speakers, it is predictable that in this field we should find a great deal of variation, as the trend to use *any*-forms as an alternative to negative concord forms in the scope of the negator is strengthened especially by the growing influence of the standard English system. As Chambers notes, 'when standard speech differs qualitatively from other varieties, it is always the case that those other varieties have variants where the standard allows no variation' (Chambers 1995: 241). This is also the case for non-standard multiple negation: today it is a matter of quantitative variation rather than an 'all or nothing' situation in English dialects, as the standard English system is available to all dialect speakers today as an alternative.

Data from the BNC

Procedure

The following negative elements were searched in the BNC-SpS subsample[5] for all dialect areas: *-n't, not, nobody, no one, nothing, nowt, none, never, nowhere* and *no* (in adjectival function only), as well as the paratactic elements *nor* and *neither*. The occurrence of negative concord was then investigated in this sample. Negative concord was defined as the co-occurrence of two or more negative elements listed above. In order to arrive at a percentage of actual vs possible occurrences, the negative elements listed above were, in addition, searched for in their co-occurrence with generic (i.e. *any*-)

Table 5.1 Neg concord in the BNC-SpS

BNC code	Dialect area	Total (=possible occurrences)	Neg concord	% of total
XEA	East Anglia	362	87	24.0
XHC	Home Counties	619	44	7.1
XHM	Humberside	47	0	0.0
XIR	Ireland	76	4	5.3
XLC	Lancashire	201	10	5.0
XLO	London	842	180	21.4
XMC	Central Midlands	269	15	5.6
XMD	Merseyside	50	2	4.0
XME	Northeast Midlands	248	31	12.5
XMI	Midlands	94	5	5.3
XMS	South Midlands	142	47	33.1
XMW	Northwest Midlands	377	12	3.2
XNC	Central northern England	218	21	9.6
XNE	Northeast England	219	45	20.5
XNO	Northern England	65	5	7.7
XSD	Scotland	101	8	7.9
XSL	Lower southwest England	138	13	9.4
XSS	Central southwest England	424	116	27.4
XSU	Upper southwest England	86	1	1.2
XWA	Wales	263	46	17.5
Total		4,841	692	∅14.3

elements that would substitute the NC negative elements in standard English for the same sample.[6] The general picture that emerges is displayed in Table 5.1.

Table 5.1 shows that negative concord is present in practically all dialect areas (as defined by the BNC). The only exception is Humberside. However, this dialect area should probably be disregarded, as it only contains speech of four speakers, all of whom seem to speak standard English. For practically all non-standard phenomena, Humberside sports no instances at all (cf. also the following chapters). For all other dialect areas, and for the social mix of speakers that are represented in the BNC, the average of negative concord forms (in comparison with the equivalent standard English forms with *any*) of just over 14 per cent is a relatively high figure, and we can thus say that today, negative concord still seems to be well established systematically. On the other hand, the widely divergent ratios with which neg concord occurs (from 33.1 per cent in the south Midlands to only 1.2 per cent in the upper southwest) shows that this feature is not present throughout all dialect areas uniformly. Possible reasons for this uneven spread will have to be considered once we have investigated the structural possibilities of neg concord today in more detail.

The system of neg concord

Let us look at the distribution of negative elements across dialects, as displayed in Table 5.2. Negative elements in columns always indicate the first elements of the respective negative concord structures; rows indicate the respective second elements. Figures in brackets indicate the number of dialect areas where this combination is found. Of the twenty total dialect areas, we have already seen that Humberside has no occurrences of negative concord at all; the maximum figure for dialect areas can therefore only be nineteen. The first cell in Table 5.2 can thus be read as follows: the combination -*n't* . . . *no* occurs 240 times across seventeen dialect areas; *never* . . . *no* occurs nine times in six dialect areas, etc.

In Table 5.2, rows as well as columns have been ordered according to decreasing absolute frequency (cf. the row and column headed 'Total'). It is striking that of the possible 210 combinations investigated, only fifty-one are realized in the sample. This accounts for around 24 per cent, or roughly one in four, logically possible combinations. Of these fifty-one realized combinations, nineteen occur only once. If we take this into account, only about 15 per cent, or less than one in six possible combinations, is realized more than once. Moreover, the realized combinations are not distributed randomly, but structured in a relatively orderly way, as Table 5.2 shows. To indicate this structure more clearly, several rows and columns can be collapsed – such as -*n't* and *not*, *nowt* and *nothing* (as dialectal variants), *nobody* and *no one* (as stylistic variants, cf. Chapter 2). Some other negative elements are of lesser importance when we want to investigate the structure of clausal negative concord, for example the paratactic structures like *nor* or *or nothing*, as well as inherently negative elements like *hardly* (their inherent, but not morphological negativity is an idiosyncrasy of English and therefore difficult to compare cross-linguistically). This leads to Table 5.3 reduced to morphological monoclausal negative concord.

Again, the rows and columns in Table 5.3 have been ordered according to decreasing overall frequency. Particularly striking is the asymmetry between variation in the first and the second negative elements. Only five negators act as the first element in a clause containing a negative concord structure. Of these, -*n't/not* is used in over 88 per cent or almost nine out of ten (!) of all cases as the first element. Of the remaining 12 per cent of negative concord structures, *never* acts as the first element in almost three-quarters of all cases. In other words, -*n't* and *never* together account for almost 97 per cent of all first elements; *no*, *nobody* and *nothing* play a rather marginal role as the first elements of negative concord structures. When we look at realized second elements, *nothing* and *no* are almost equally frequent (they make up 42.2 per cent and 41.4 per cent respectively of all realized second elements). The remaining cases of negative concord are relatively evenly spread over a larger number of negators. Interestingly, -*n't/not* can function as the second element in some (few) cases as well.

Table 5.2 Co-occurring neg elements in the BNC-SpS

2nd / 1st element	-n't	never	not	no	nobody	hardly	nor	no one	nowt	none	Total
no	240 (17)	9 (6)	13 (6)	1 (1)		1 (1)			1 (1)		265
nothing	178 (16)	33 (7)	14 (8)	10 (5)	1 (1)		1 (1)				237
nowt	24 (3)	4 (3)	5 (3)								33
none	28 (9)	3 (3)	1 (1)								32
or nothing	9 (6)	3 (3)	4 (3)	4 (3)							20
no more	8 (2)	4 (3)	7 (4)								19
never	16 (7)					1 (1)					17
hardly	9 (5)	5 (3)								1 (1)	15
nowhere	8 (5)	2 (1)	4 (3)								14
nor	5 (1)	4 (3)									9
nobody	6 (3)		1 (1)	1 (1)							8
neither	5 (2)		1 (1)	1 (1)							7
-n't			1 (1)	1 (1)	3 (3)		1 (1)	1 (1)			7
no one	4 (4)		1 (1)	1 (1)							6
not	2 (2)			1 (1)	1 (1)						3
Total	542	67	52	19	5	2	2	1	1	1	692

Table 5.3 Co-occurring monoclausal morphological neg concord elements in the BNC-SpS

2nd/1st element	-n't/not	never	no	nobody/ no one	nothing/ nowt	*Total*
nothing/nowt	221 (18)	37 (8)	10 (5)	1 (1)		269
no	253 (18)	9 (6)	1 (1)		1 (1)	264
none	29 (10)	3 (3)				32
no more	15 (5)	4 (3)				19
never	16 (8)					16
nobody/no one	12 (5)		3 (3)			15
nowhere	12 (7)	2 (1)				14
-n't/not	3 (3)		1 (1)	4 (4)		8
Total	561	55	15	5	1	637

This construction – Labov (1972) calls it 'Negconcord to pre-verbal position' – has been noted in particular for African American Vernacular English (cf. Mufwene *et al.* 1998) and is often cited as a significant structural difference that sets AAVE apart from other non-standard dialects of English (cf., for example, Henry 1997 for Belfast English). Labov himself states that 'there are many non-standard dialects which do not allow this, and reject such sentences as *Nobody don't like a boss hardly*' (Labov 1972: 786), although he concedes that this is not a marker of ethnicity: apparently there are also white non-standard dialects like northern New Jersey English which permit neg concord to preverbal position (Labov 1972: 786).

   If we consider the data from the BNC against this background, the first thing to note is that in the only cases where the sentence negator appears as the second element, it is either preceded by another instance of the negator itself (but see notes 7 and 8) or, more frequently, by a negative quantifier of person, i.e. *nobody* or *no one*. Indeed, from the published examples for AAVE this tendency seems to be the same as for AAVE. The six dialect areas where a negative element can precede the sentence negator are East Anglia, London, where both *not/-n't* . . . *-n't* and *nobody/no one* . . . *-n't* occur, the northeast, as well as the northeast Midlands, the northwest Midlands and the central southwest, where only the latter construction occurs in the BNC-SpS subsample. All instances from the BNC-SpS where *-n't/not* functions as the second element are supplied in (4) to (11), ordered according to dialect area.

(4)   When I looked she *ain't* got *not* water! It was empty.
      (XEA: KCT 11323)[7]

(5)   Whatever it is, *no one don't* seem to want it. (XEA: KC8 408)

(6)   I hope *nobody ain't* been swearing. (XLO: KCT 5177)

(7)   Mummy *hasn't* got *not* money. (XLO: KPE 3165)[8]

(8)   *Nobody*'s *not* doing cabaret spots with trumpets. (XME: XC2 2999)

(9)   *No one didn't* recognized her. (XMW: KDM 10103)

(10) *Nobody don't* bother with them do they? (XNE: KB7 13659)

(11) Stick it down with sellotape so *no* bugger *can't* open it.
       (XSS: KBE 6782)

These examples make clear that neg concord to preverbal position indeed seems to be restricted to those cases where a personal negator (*no one* or *nobody*) precedes the sentence negator. *No bugger* in example (11) seems to be semantically equivalent to *no one* and therefore does not require classification as a separate category. In example (11), *bugger* functions generically, so that *no* has lost its adjectival function.[9] If we compare the list of dialect areas with the relative frequencies in Table 5.1, it becomes apparent that the choice of -*n't* as a second element in a negative concord construction, although regionally restricted, is not regionally *determined*, but more generally related to the relative frequencies of negative concord constructions as such: of the seven areas where the standard negator appears as the second element, five use negative concord much more frequently than the average of 14.3 per cent. The two areas that go against this trend are the northeast Midlands, which at 12.5 per cent are, however, relatively close to the average, and the northwest Midlands, which at 3.4 per cent rate far below. The reverse also holds: of the six dialect areas that have neg concord ratios above average, five allow -*n't* as the second element in negative concord constructions. Here, the only exception are the south Midlands. The distribution of -*n't*/*not* as the second element thus does not seem to be a feature characteristic of a particular region. Rather, the construction quite generally seems to be relatively infrequent, and therefore only occurs in those samples where the overall frequency of neg concord constructions is high enough to allow a variety of constructions to emerge.

Regional distribution

We have already noted that the figures for neg concord in Table 5.1 are strikingly divergent from one dialect area to another – much in contrast to contraction ratios in Chapter 4, for example, where figures were very high throughout. The investigation of a possible regional differentiation for neg concord shall be the subject of the remainder of this chapter. A first scale ordering the dialect areas in terms of the relative frequency of neg concord does not at first glance suggest a direct correlation, as Table 5.4 illustrates.

Table 5.4 Scale of neg concord in the BNC-SpS

Dialect code	Region	NC-frequency (%)
XMS	South	33.1
XSS	South	27.4
XEA	South	24.0
XLO	South	21.4
XNE	North	20.5
XWA	Mid/Celtic	17.5
XME	Mid	12.5
XNC	North	9.6
XSL	South	9.4
XSD	North/Celtic	7.9
XNO	North	7.7
XHC	South	7.1
XMC	Mid	5.6
XMI	Mid	5.3
XIR	North/Celtic	5.3
XLC	North	5.0
XMD	North	4.0
XMW	Mid	3.2
XSU	South	1.2
XHM	North	0.0

In order to investigate a possible north–south difference, let us look at the very general north/south/Mid assignment. The southern dialect areas seem concentrated at the high end, with the four dialect areas where neg concord is used the most frequently being situated in the south. The remaining southern areas, however, vary considerably in the frequency with which negative concord is employed, from 9.4 per cent in the lower southwest to only 1.2 per cent in the upper southwest. In the Midlands region, if we discount the south Midlands which traditionally belong to the south of England rather than the middle, we find a slightly narrower range from 12.5 per cent in the northeast Midlands to 3.2 per cent in the northwest Midlands. In the north, finally, we have a distribution from 20.5 per cent in the northeast, one area where neg concord constructions are among the most frequent, to zero in Humberside. (The exceptional status of Humberside has already been mentioned.)

This disparate picture poses two theoretical questions: is this distribution determined by actual regional differences, or is it a surface feature which mirrors – for example – an uneven distribution of non-standard speakers in the BNC? Or is it an interaction of the two?[10] As the compilers of the BNC concede, despite the large size of the BNC spoken section, a combination of speaker categories (and thus the creation of subcategories) will not always lead to a representative subsample. It can easily be demonstrated that although the regional distribution of speakers is roughly representative of

the British population as a whole, and although the same holds for the social distribution, a combination of these two criteria indeed leads to non-representativeness. As social and regional variation still correlate strongly in England – the lower the social standing, the more regional features a speaker will have (cf. Trudgill's triangle, e.g. Trudgill 1983: 188) – we can assume that wherever a dialect region is only represented by speakers of a certain social class, the subcorpus will not be representative linguistically for the dialect region as a whole. For the BNC, this holds in particular for dialect areas that are comparatively small, such as Humberside (39,377 words or 0.9 per cent of the spoken sample; four speakers, all of whom speak standard English, as mentioned before), or the upper southwest (it contains 38,911 words or, again, 0.9 per cent of the spoken sample). The lower southwest, although larger (108,126 words, accounting for 2.6 per cent), contains speech of thirty-one speakers; however, twenty-three of these are not assigned to any social class. (The list could be continued.) Another factor possibly skewing dialect results is the uneven length of individual speakers' contributions. (In fact, these range from several ten thousands of words to just two.)

Despite these caveats, the possibility that the figures of negative concord might mirror (at least in part) actual regional differences should not be discarded right away, even though this might seem an odd angle from which to investigate negative concord. As detailed in the previous sections, neg concord is a prime candidate for a non-regional non-standard feature or a general sociolectal phenomenon – neg concord has been taken to be a typical social class marker rather than a candidate for regional variation. Disregarding Humberside for the reasons mentioned above, the data from the BNC certainly supports the fact that multiple negation occurs widely in Great Britain today in every region; the 'Celtic' Englishes (provisionally indicated also by 'Celtic' in Table 5.4) make no difference here. Unfortunately, although neg concord is frequently mentioned for individual, modern or traditional, dialect studies, there are hardly any comparative studies of this phenomenon across dialects to date. Perhaps it is for this reason that a regional differentiation is hardly ever considered as a possibility. The only exception is Cheshire, Edwards and Whittle's questionnaire-based investigation into regional English, where they state tentatively that what seems to be emerging from the data is a regionalization for this feature:

> only 58 of the 80 schools participating in the survey (72.5%) reported hearing sentences such as . . . *I won't do nothing silly* . . . Multiple negation was reported less frequently in the North of Britain than in the Midlands, and most frequently in the South.
>
> (Cheshire, Edwards and Whittle 1993: 76)[11]

How can this possible distribution be investigated with our data from the BNC? Although there are several dialect regions which are not wholly

Table 5.5 South–north grouping of neg concord in the BNC-SpS

	Total	Neg concord	% of total
North (8 areas)	977	95	Ø 9.7
Midlands (5 areas)	1,251	109	Ø 8.7
South (7 areas)	2,613	488	Ø18.8
Total	4,841	692	Ø14.3

representative of regional dialects, as shown above, these unrepresentative regions are themselves relatively evenly distributed over the whole of Great Britain. If we look at average figures for the very general north/south/Midlands regions, we find a distribution as detailed in Table 5.5.[12]

   The percentages in Table 5.5 seem to be striking evidence in support of a regional variation of neg concord, supporting to a degree also the figures collected by Cheshire, Edwards and Whittle (1993). Statistical tests on the figures show that the differences between the south and the Midlands area as well as between the south and the north are highly significant (at $p < 0.01$). In contrast to Cheshire, Edwards and Whittle (1993), however, the difference between Midlands and the north is not statistically significant. These figures would therefore argue not so much for a three-fold south/Mid/north divide but for a Mid–north continuum that is clearly different from the more southern areas.

   In order to investigate this in more detail, the differences in neg concord ratios were compared across neighbouring areas, in a similar procedure as for contraction ratios in Chapter 4. The result is displayed in Figure 5.1.

   Three areas had to be excluded from statistical tests because for them, occurrence of neg concord was below a threshold of five, and statistical tests can therefore not be expected to give reliable results for these dialect areas. These three excluded areas are Humberside, Merseyside and the upper southwest; they are marked grey in Figure 5.1. For all other areas, significant differences are marked by a bold line. The exceptional status of the northeast that has already become apparent from Table 5.1 is confirmed by this comparison of neighbouring dialects: the northeast behaves clearly differently from all its neighbours, while the area around the northeast shows a very homogeneous pattern. Indeed this homogeneous northern area extends as far south as the border towards the south Midlands, and as far west as the border with Wales. The whole of the north (with the exception of the northeast) and the Midlands, then, does indeed constitute one dialect region with respect to neg concord and confirms the general assignment and comparison of Table 5.4.

   The area south of this line, however, cannot exactly be called homogeneous. Significant differences appear at practically every dialect border. A look at the relative frequency figures shows that this is at least in part due to the exceptional status of the Home Counties, which only have an
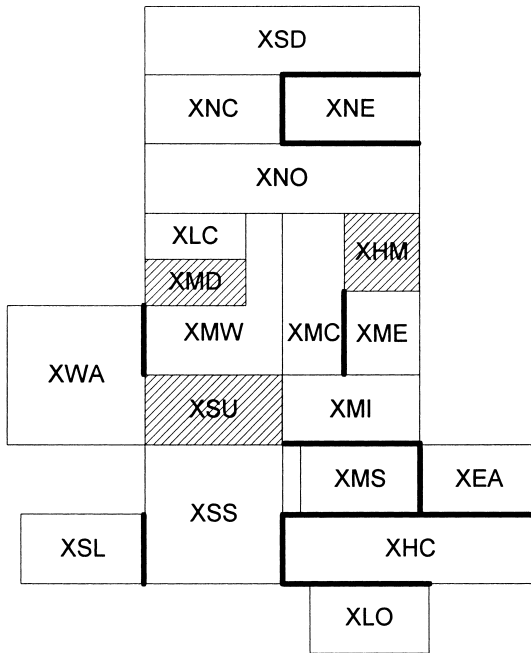
Figure 5.1 Statistically significant differences of neg concord

average ratio of neg concord of 7.1 per cent, as Table 5.1 has shown. The Home Counties behave significantly differently from all their neighbours. If we compare the south Midlands (33.1 per cent) and London (21.4 per cent) directly, disregarding the intervening area of the Home Counties, the difference in neg concord ratios is not statistically significant. If we likewise compare East Anglia (24 per cent) and London (21.4 per cent), the difference between these two dialect areas is also not statistically significant. The Home Counties then seem to play a similarly exceptional role in the southeast as the northeast (with its exceptional 20.5 per cent for neg concord) does in the north, where the average is considerably lower, as we have seen. It is of course well known that the Home Counties, constituting for a large part what is sometimes called the 'stockbroker belt' around London, do indeed often behave linguistically in a much more standard way than the central metropolis, due to the different social composition of the population. This more standard behaviour is clearly mirrored in the distribution of neg concord in the data from the BNC.

The northeast on the other hand consists mainly of the Tyneside conurbation – a region that has played a particularly innovative role (in linguistic terms) in recent times which has set it apart from its direct neighbours

and made it much more similar to more southern dialect areas in many respects. Again, then, significant differences that have become apparent from the BNC material can be plausibly interpreted to mirror (at least in part) actual differences in linguistic behaviour rather than to be freak occurrences due to faulty or unreliable underlying data.

The linguistic status of Wales remains to be discussed. Although Table 5.4 assigns Wales to the neighbouring Midlands dialect areas, as is traditionally done (if Wales is considered at all), the very high figures for neg concord in this dialect area (of 17.5 per cent) and the clear statistical differences between Wales and the neighbouring northwest Midlands (with neg concord ratio of only 3.2 per cent) seem to indicate that Wales for this feature patterns with the south rather than with the homogeneous north–Mid area. Confirmation of this hypothesis, however, crucially hinges on the status of the upper southwest, which would link Wales directly with the southern region. As the data situation for the upper southwest is so unclear, however, an isogloss cannot be drawn: it cannot be determined from data from the BNC with any certainty whether this isogloss would run south or north of the upper southwest. Despite this drawback, a regional analysis for the feature of neg concord has brought to light an interesting patterning that was in part suspected (by Cheshire, Edwards and Whittle 1993), if mistrusted, but that otherwise certainly goes against the received opinion that negative concord has no regional differentiation. These results also confirm the exceptional status of several dialect areas (the northeast, the Home Counties) that has been noted before. On the other hand, they also throw light on some rather unexpected phenomena (the linguistic behaviour of Wales, for example) which may open areas of research for the future.

Summary

Why should a feature like multiple negation still be present so very widely, despite the considerable stigma attached to this construction? Jespersen (1917: 68ff.) gives a good overview of the various theories that have been adduced as explanations for the phenomenon of multiple negation. His own explanation is as follows:

> logically one negative suffices, but two or three in the same sentence cannot be termed illogical; they are simply a redundancy, that may be superfluous from a stylistic point of view, just as any repetition in a positive sentence . . . , but is otherwise unobjectionable . . . it requires greater mental energy to content oneself with one negative, which has to be remembered during the whole length of the utterance both by the speaker and by the hearer, than to repeat the negative idea (and have it repeated) whenever an occasion offers itself.
>
> (Jespersen 1917: 71)

Bernini and Ramat give a similar functional explanation when they state that the rarity of non-neg concord languages in Europe 'seems due to the lack of explicitness of the negative sense in similar structures and hence the need to reinforce it in order to clarify the message' (Bernini and Ramat 1996: 188). It is therefore not surprising that most European languages make use of this construction. More precisely, in Bernini and Ramat's study, of the thirty-one languages under investigation, a total of eighteen allow the combination of sentence negator and a negative quantifier, for seven this is not permitted. (The remaining six languages do not possess N-quantifiers and can therefore not be classified. These are the Gaelic languages, Basque and the Finnic languages.) The situation for the co-occurrence of N-quantifiers is even more dramatic: 'only the Germanic languages, of the type <N and −neg> but with only postverbal negative elements, are N1 languages' (Bernini and Ramat 1996: 187). In other words, neg-impermeable N1 languages are very rare in Europe and mainly restricted to the Germanic languages. Although standard English is there-fore indeed a typical Germanic language in this respect, it is not a typical European language (and the same applies to its Germanic sister languages, at least the standard varieties). Compared to a geographically wider sample, this is even more striking: as Haspelmath points out, it is the strict N1 languages of Europe that are in need of explanation, rather than the reverse, as they are typologically extremely unusual (Haspelmath 1997: 202). As we have seen, for practically all non-standard dialects of Great Britain today, negative concord is at least possible, though not obligatory any more. Even in the very modern spoken component of the BNC, the occurrence of negative concord for some dialect regions ranges as high as 33 per cent. Although the distribution of neg concord seems to be region-ally determined in terms of relative frequency, the systematic possibility is still present for all dialect areas; even the Celtic Englishes are not excepted from this general feature.

From a typological point of view, then, it is not surprising that all non-standard dialects in Great Britain can optionally make use of neg concord, in particular as the functionally important negator is practically always reduced to *-n't*, as we have seen in Chapter 4. A system which does not permit negative concord, such as standard English, seems to be the 'odd one out', because, as Edwards points out, 'standard English is the only British dialect which does not express negation in this way' (Edwards 1993: 226). However, for the Germanic languages in general it seems to be the case that non-permittance of neg concord is a common feature of the stan-dardized varieties. The option we find for practically all dialect areas for Great Britain, as the data from the BNC has shown, is, however, both more widely spread geographically and typologically and makes much more sense in cognitive terms, as it helps to reduce ambiguities and possible misunderstanding.

6 AIN'T

Introduction

Together with negative concord, *ain't* is perhaps the best-known shibboleth of non-standard English, and this already implies that it is highly stigmatized.[1] *Ain't* is a negative form of unclear historical origin and of very wide usage – both grammatically and geographically. Probably due to a historical coincidence, *ain't* functions as the negative form of both present tense BE and present tense HAVE in non-standard English today.

This chapter starts from the history of *ain't* and investigates the distribution in the historical (rural) dialects, before moving to the present situation in British non-standard speech. The structural discussion will centre on the comparison of AIN'T for BE and for HAVE, and especially on the relation of *ain't* to *in't*. We will also investigate whether AIN'T is still regionally distributed today. Finally, we will see that AIN'T embodies the principle of asymmetry, a pervasive principle that we encounter in language typology, which will be further illustrated in the remaining chapters.

The phenomenon

AIN'T is used as the negative form for present tense BE as well as present tense HAVE. AIN'T does not make person distinctions; thus there is only one form across all persons and numbers. Typical realizations are given in examples (1) to (3).[2]

(1) It *ain't* there.

(2) We *ain't* going.

(3) I *ain't* done it.

It is generally observed that where AIN'T functions as the negative of HAVE, it can only be used for the auxiliary, it cannot negate the full verb

HAVE. Trudgill, for example, explicitly excludes sentences such as (4) and (5) as unacceptable.

(4) *I ain't a clue.[3]

(5) *I ain't my breakfast at 8.[4]

This restriction does not hold for BE, where AIN'T can function as the negative form both of the auxiliary and the main verb use (i.e. the copula), as in (6) and (7).

(6)   I ain't coming.[5]

(7)   It ain't there.[6]

   In her detailed study of adolescent non-standard English in Reading, Cheshire also stresses this distribution of AIN'T: 'it functions . . . as the present tense negative form of auxiliary BE, the present tense negative form of the BE copula, and the present tense negative form of auxiliary HAVE (but not of full verb HAVE)' (Cheshire 1982: 51). Cheshire also notes an interesting pattern in the distribution of AIN'T across these three functions: in her data, AIN'T occurs 'most often as auxiliary HAVE . . . and least often as auxiliary BE' (Cheshire 1982: 51). This can be diagrammatically displayed as a scale in (8).

(8)   Hierarchy for *ain't*:
      aux HAVE > cop BE > aux BE

This scale can serve as a useful working hypothesis for testing the BNC material further below.

   Cheshire also speculates that on the basis of their differing etymology, 'it is reasonable to expect . . . that nonstandard *in't* would correspond to standard English *isn't*, and nonstandard *ain't* to the other standard English forms' (Cheshire 1982: 52), but finds herself that this expectation is not borne out by her results. In order to investigate the use of AIN'T in spoken present day British English, these various claims shall be tested in more detail with the help of the BNC. First, however, we shall examine the history of this curious all-purpose verb form.

History

As Jespersen has shown, *ain't* (for present tense HAVE) can be derived from variation in the Early Modern English pronunciation of *have* in negative clauses: 'for *have* we had two pronunciations, with short and long vowel . . . the former prevailed in positive sentences [hæv], the latter with *-n't*

became [heɪnt] . . . instead of *han't* the spelling *ain't* also occurs as a vulgarism (*h* dropped)' (Jespersen 1940: 431).[7] *Ain't* for present tense BE has also been derived from a contraction of *am + -n't* with a simplification of the nasal cluster /mn/ to /n/ with subsequent diphthongization.[8] There are phonotactic reasons, however, that speak against this path of development. The only other words where a /mn/ cluster is reduced to /n/ derive without exception from Greek *mnemo*; here the cluster is word-initial (cf. *mnemonic* [niːˈmɒnɪk], *Mnemosyne*); moreover, simplification is not obligatory (Wells 1990, for example, notes alternative pronunciation variants in /mn-/). For all other words, a reduction of /mn/ to /m/ rather than to /n/ is much more frequent and in non-initial positions probably obligatory (cf. words as diverse as *damn, hymn, column; damned*) and could therefore have been expected for **amn't* as well. A simplification of **amn't* to /æmt/ would also have avoided the otherwise homophonic clash with *ant*. On the other hand, this contraction would have deleted an important part of the negative morpheme, and this morphological status might have been a barrier to this contraction. The more usual way of development is illustrated by dialects that possess a contracted form of *am + not*; these have usually neutralized the nasal cluster by inserting an epenthetic vowel, resulting in forms like *amment*, which would present a third possible way out of the /mn/ cluster. For these reasons it is not clear how plausible a development of *amn't* to *ain't* really is.[9]

Ain't is also, perhaps more naturally, derived from *aren't*. In fact, Jespersen speculates that '*ain't* in the first person sg. probably has arisen through morphological analogy' (1940: 433). After the loss of post-vocalic *r* in British English, the vowel in /ɑːnt/ would only have had to undergo diphthongization to result in /eɪnt/. Finally, *ain't*, as well as the related *in't*, is also derived regularly from *isn't*:

> The *s* [z] was frequently dropped in *isn't* . . . for *isn't* we find *'ent* . . . and in the 18th c[entury] the form *i'n't* . . . but the vowel is unstable; Swift writes *e'n't*, and if we imagine a lowering and lengthening of the vowel . . . , this would result in a pronunciation [eint]; now this must be written *an't* or *ain't*, and would fall together with the form mentioned above as possibly developed from *aren't*.
>
> (Jespersen 1940: 433f.)

In sum, the possible derivations can be reproduced diagrammatically as in Figure 6.1.

The *OED* interestingly gives quotations of *ain't* used for BE from much earlier sources (around 1778) than *ain't* used for HAVE (the first quote under the headword *ain't* is from 1845 from an American source, but is predated by quotes from Haliburton from 1835). In fact, if we distinguish the individual pronouns, we find the following first quotations:
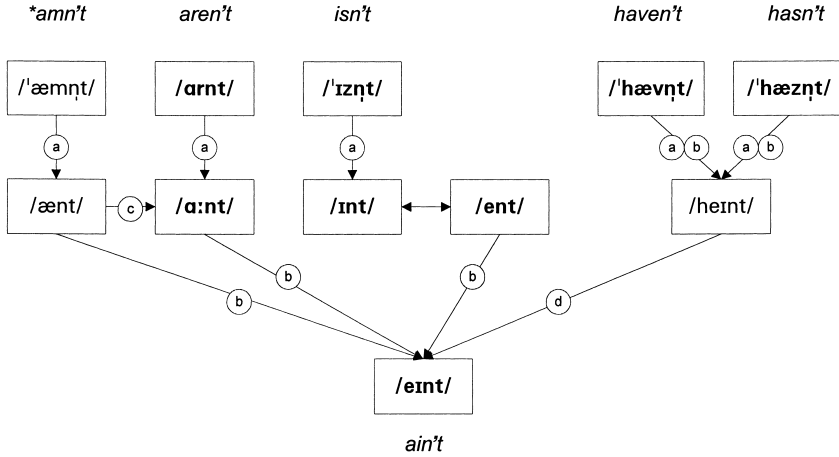
Figure 6.1 Paths of development for *ain't/in't*

For *aren't*:

(9) 1695 *a'n't* = I *aren't*
     I an't Calf enough to lick your chalk'd Face, you Cheese-Curd you.
     (*OED* headword *cheese*)

(10) 1706 *a'n't* = NP *aren't*
     But if *your Eyes a'n't* quick of Motion, They'll play the Rogue, that
     gave the Caution. (*OED* headword *an't*)

(11) 1725 *a'n't* = *aren't you*
     An't you weary of wifeing? (*OED* headword *wife* v.)

(12) 1734 *a'n't* = *aren't we*
     Ha, ha, ha! *an't we*? no! How ignorant it is! (*OED* headword *an't*)

For *en't*:

(13) 1710 *en't* = *there isn't*
     The politest atheist can't be sure that *their* [*sic*] *e'nt* a God. (*OED*
     headword *be* v.)

For *in't*:

(14) 1742 *in't = it isn't*
      No indeed; *it i'n't* worth while. (*OED* headword *i'n't*)

For *ain't*:

(15) 1778 *ain't =* NP *aren't*
      Those you are engaged to ain't half so near related to you as we are.
      (*OED* headword *ain't*)

(16) 1830 *ain't = I am not*
      But *I ain't* a pumpkin, the Squire he knows that. (*OED* headword
      pumpkin)

(17) 1833 *ain't = they aren't*
      A dicker's a dicker I allays concate, where people's upon honor, but
      not where *they aint.* (*OED* headword *dicker* n.)

Contracted forms for *haven't/hasn't* on the whole are attested earlier, but
these early forms still retain the initial /h/:

For *han't*:

(18) 1662 *ha'n't = I haven't*
      I *ha'n't* seen her since my last mischance. (*OED* headword *rigging*
      (vbl.) n.)

(19) 1697 *han't =* NP *hasn't*
      Pray Heaven that old rogue Coupler *han't* sent us to fetch milk out
      of the gunroom. (*OED* headword *gun-room*)

(20) 1835 *ain't = they haven't*
      They *ain't* got two ideas to bless themselves with. (*OED* headword
      pumpkin-head)

Although this considerable difference in first attestations of almost 60 years
might, at first glance, point to a gradual diffusion of this form through these
two verb paradigms, the presence of *han't* from a very early point of time
onwards argues for a historical coincidence whereby possibly as many as
five verb forms have gradually merged in the one form *ain't.* Although the
OED material is seriously skewed in several respects, and can certainly not
be taken as a representative historical corpus, nevertheless, the history of
ain't is still largely unclear and remains to be written, where the material
from the *OED* quoted here might play a minor role in the exploration of

```
          *amn't      aren't       isn't              haven't    hasn't

 1650 |                                                   |
      |                                                 han't
      |
      |      an't
 1700 |                  an't                                       han't
      |                             en't
      |
      |                                          in't
 1750 |
      |
      |                  ain't
 1800 |
      |      ain't                  ain't                ain't     ain't
 1850 |
```
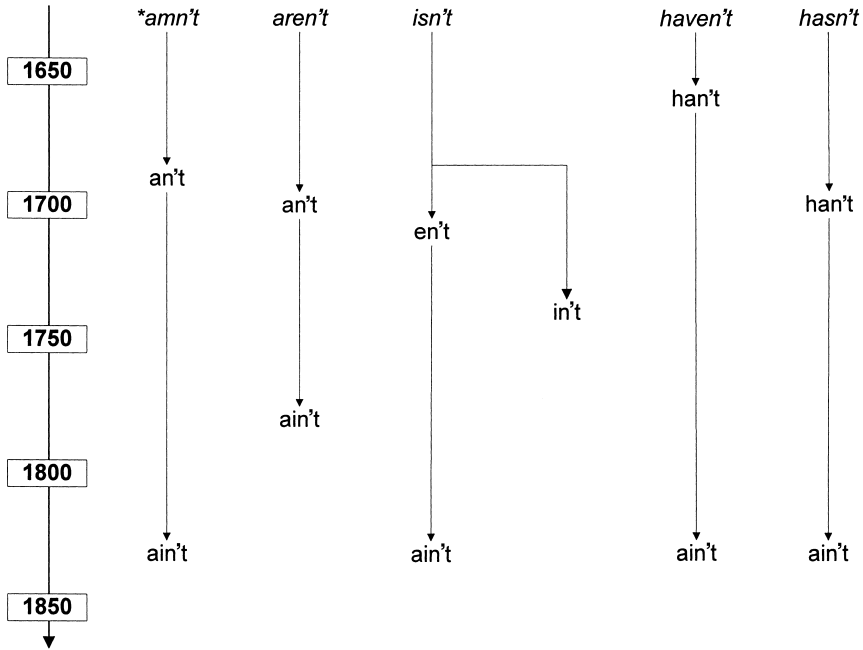
Figure 6.2 Temporal development of *ain't* (based on the *OED*)

this topic. Assuming either that the dates are basically correct, or on the other hand that they are all equally incorrect, we can graphically display a timeline tracing the development of *ain't* from the various forms, as they are attested.

By 1837, *ain't* must have been so widely used that Charles Dickens (the source for almost all these early attestations) could employ this form to distinguish his (London) characters. The one earlier attestation of *ain't* used for *aren't* points at the long development this form must have taken. *In't* on the other hand is much older and has evolved independently, and again this is as expected, as far fewer steps are required to derive *in't* from *isn't*. This process of /z/-deletion as soon as the negator is contracted to /nt/ is quite naturally explained through English phonotactics, as it avoids the relatively highly marked syllabic /n/, making the word monosyllabic. The same is true for all other verb forms, so that we can assume that this second contraction must have taken place relatively quickly after the negator began to be contracted. And indeed very early attestations of secondary contractions are recorded for practically all forms for times at which it is generally held that the negator itself was only starting to be contracted.
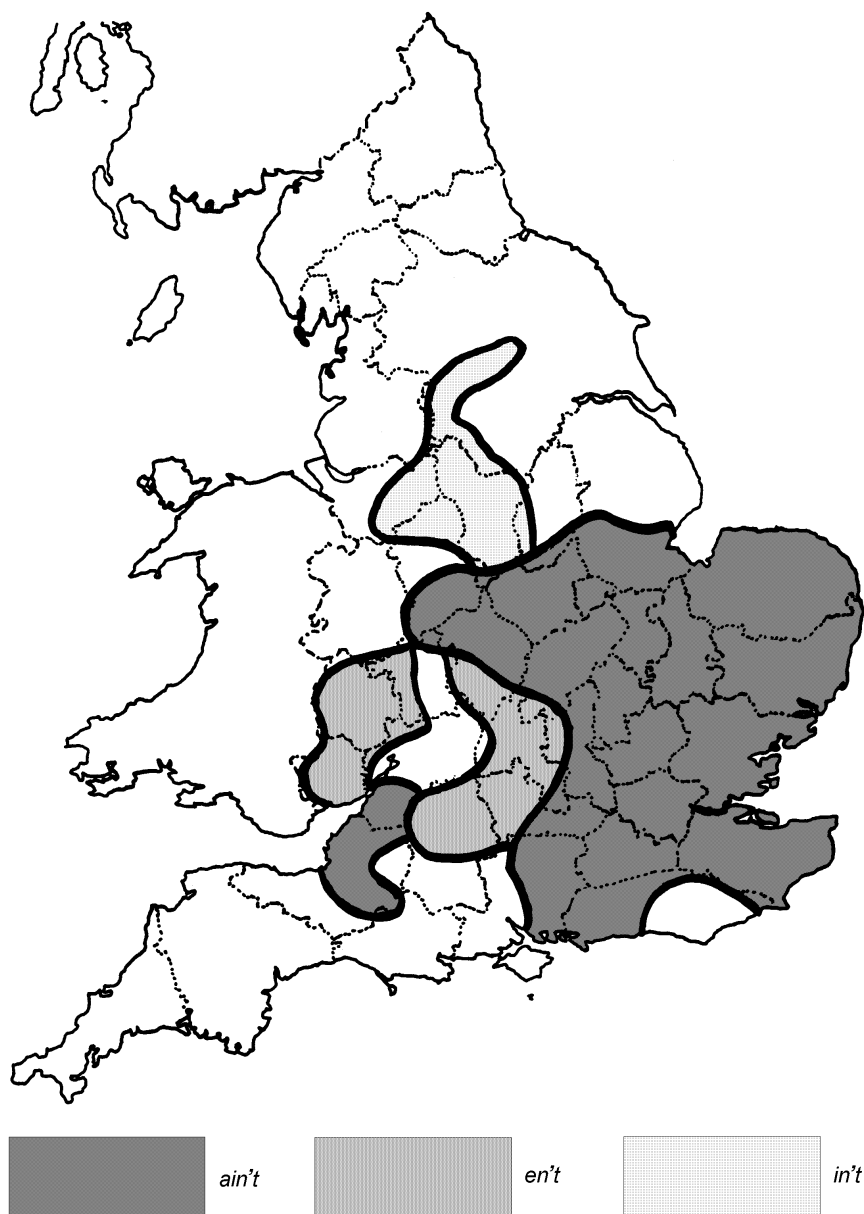
Map 6.1 AIN'T in the SED

Source: Adapted with permission from Orton, H., Sanderson, S. and Widdowson, J. (1978) *The Linguistic Atlas of England*, London: Croom Helm.

It is possible to investigate the distribution of AIN'T for traditional English dialect speakers by drawing on data from the *SED* (in particular the maps in Orton, Sanderson and Widdowson 1978). *The Linguistic Atlas of England* contains several maps dealing with present tense BE for all the personal pronouns. Among the recorded answers are – besides the standard English forms and some variation in the root morphemes – the two forms *ain't* and *en't*, and, restricted to the third person only, *in't*. Due to the traditional questionnaire method of collecting the data, the occurrences are based only on one sentence each (per personal pronoun per informant), and *ain't* and *en't* are therefore of necessity in complementary distribution. (As is obvious from the research design, questionnaire elicitations cannot record intra-speaker variation.) Nevertheless, the composite Map 6.1 should be able to show the central characteristics of any occurrence of *ain't* vs *en't* for the traditional English dialects.

Map 6.1 indicates that there is a large area in the southeast where *ain't* is part of the traditional dialect system. This area includes East Anglia and the south Midlands, and extends a little north of the Wash. *En't* on the other hand is concentrated in the upper southwest and central southwest. No significant occurrence of *ain't* can be found in the north of England; occurrences of *in't* are only recorded for the third person singular and form a pocket in the central Midlands and the central north.

Data from the BNC

Procedure

The same variation in pronunciation noted by Jespersen from the eighteenth century onwards can still be observed today. This is even reflected in the spelling. Although *ain't* has become a relatively standardized orthographic representation, the alternatives *in't* and, more rarely, *en't*, show that differences in vowel quality are still perceived as significant. In the BNC transcriptions, *ain't* is by far the most frequent form. *In't* (and *int*) occurs more regionally, and a form with a lowered vowel *en't* occurs in some literary representations of non-standard speakers. For the purposes of this chapter, *ain't* is distinguished from *in't* (including spelling variants such as *int*; the spelling *aint* does not occur) only where relevant. When we look at the general rate of non-standard vs standard English forms, it is not necessary to distinguish separate subclasses of non-standard terms, and *ain't* and *in't* have therefore been grouped together. To avoid terminological confusion, this undifferentiated *ain't/in't* group will be referred to by small capitalized AIN'T, indicating the non-standard paradigm as such. Simple italic *ain't* on the other hand refers to the lexical form *ain't* itself, as opposed to *in't*.

In order to arrive at some regionally comparable results from the BNC, searches were once again restricted to combinations of a personal pronoun

(or *there*) with forms of present tense BE and HAVE. As in the previous chapters, all searches were conducted per dialect areas, and examples not belonging to the Spontaneous Speech subsample (SpS) were excluded manually. All examples of AIN'T were manually disambiguated according to the following criteria: person, number (where appropriate), kind of verb (copular BE, auxiliary BE, or HAVE) and type of clause (declarative, full question or tag question). The search results are displayed in the following sections.

BE vs HAVE

This section compares the overall results for AIN'T (i.e. *ain't* and *in't* taken together) for BE to AIN'T for HAVE. Table 6.1 is a summary of the overall figures for these two primary verbs, and also the use in tag questions. The column headed 'Total' is the sum of all standard and non-standard occurrences (i.e. the number of possible occurrences), and the frequency index indicates the simple percentage of AIN'T forms of these totals. The distribution across dialect areas will be discussed further below.

As expected, AIN'T is a feature of non-standard British English that is well established in British non-standard English today. With averages of only 8.9 per cent for BE and a slightly more impressive 14.1 per cent for HAVE, AIN'T is roughly as frequent as negative concord (cf. Chapter 5), probably the best known other indicator of non-standard speech in the English-speaking world, and indeed this status of high stigma might explain the relatively low occurrence of AIN'T as well as of neg concord, compared to other non-standard features.

If one compares the 'Total' figures for the two verbs, it emerges that the average frequency of AIN'T used for BE is much lower than the frequency of AIN'T used for HAVE, although in absolute terms, BE occurs much more frequently than HAVE (BE has 13,195 occurrences vs 'only' 5,021 for HAVE, i.e. more than 2.5 times as many). This relation seems to point to a very general tendency: the rarer a phenomenon is (in absolute terms), the more frequently do we find the form AIN'T. In other words, this first very general overview indicates that AIN'T seems to be spreading as a strategy of simplification.[10] This correlation also seems connected with cross-linguistic regularities. The same correlation between low frequency and simpler paradigm is the basis for describing relations of markedness in the typological literature: the unmarked member of a pair (or group)

Table 6.1 AIN'T for BE vs HAVE in the BNC-SpS

	Total	AIN'T	%	Tags	AIN'T	%
BE	13,195	1,172	8.9	3,847	668	14.8
HAVE	5,021	707	18.2	1,111	202	18.2

Table 6.2 AIN'T in tags vs non-tags in the BNC-SpS

	Non-tags	AIN'T	%	Tags	AIN'T	%
BE	8,680	504	5.8	3,847	668	14.8
HAVE	3,910	505	12.9	1,111	202	18.2

of elements is typically more frequent and also more complex, or more irregular.[11] The marked member on the other hand therefore combines lower frequency with higher regularity. Although one cannot very plausibly describe the difference between the two primary verbs BE and HAVE as one of markedness, as they do not stand in a paradigmatic grammatical relationship, the underlying trend certainly seems to point to the same phenomenon.

If we compare the overall figures for BE and HAVE with their counterparts in tag questions, a second interesting fact should be noted. The average frequency of AIN'T in tag questions is significantly higher than the overall average for both verbs. If we compare tag questions with the rest,[12] the figures are even clearer, as Table 6.2 shows.

For BE, AIN'T occurs as the negative form in almost 6 per cent of all cases in environments other than tags, vs almost 15 per cent of all cases in tags, where it is thus 2.6 times as frequent. For HAVE, the negative form AIN'T is used in almost 13 per cent of all non-tags, but in over 18 per cent of all tag questions; this is an increase by a factor of 1.4 which is still statistically significant. Again, by the pattern of simplification this increase is as expected, because by the criteria of word order, intonation and absolute frequency of use, tag questions are decidedly marked in comparison with declarative sentences. The occurrence of AIN'T then seems to follow a very general typological trend where grammatical distinctions that hold in the unmarked case have a strong tendency to be simplified for the marked option. Cheshire's hypothesis with regard to the use of *ain't/in't* that 'tag questions . . . strongly favour the use of a nonstandard form' (Cheshire 1982: 55) can also be supported by data from the BNC-SpS subsample on a very general scale.

The following sections will discuss the data in more detail, in particular in relation to a possible regional distribution. Table 6.3 details the figures for the use of AIN'T for BE per dialect region. To facilitate a regional reading, the rows are in order of decreasing (relative) frequency of AIN'T.

A close investigation of the percentages in the final column reveals a striking regional distribution: there is one group of dialect areas at the top of the table – the majority – where the frequency of AIN'T is persistently high (at least in double figures, in fact clustering around 17 per cent). There are no significant differences between dialect areas adjacent in this table, and there are hardly any significant differences between geographically adjacent dialect areas. For the second group of dialect areas, AIN'T

Table 6.3 AIN'T for present tense BE in the BNC-SpS

BNC code	Dialect area	Total	AIN'T for BE	% AIN'T of total
XNC	Central northern England	767	134	17.5
XMS	South Midlands	253	44	17.4
XMC	Central Midlands	840	141	16.8
XMI	Midlands	285	46	16.1
XLC	Lancashire	656	99	15.1
XNO	Northern England	186	22	11.8
XEA	East Anglia	715	75	10.5
XSS	Central southwest England	953	96	10.1
XSL	Lower southwest England	409	39	9.5
XME	Northeast Midlands	616	58	9.4
XLO	London	2,447	169	6.9
XHC	Home Counties	1,564	100	6.4
XMW	Northwest Midlands	1,250	79	6.3
XMD	Merseyside	202	12	5.9
XWA	Wales	756	31	4.1
XNE	Northeast England	468	14	3.0
XIR	Ireland	200	6	3.0
XSU	Upper southwest England	150	3	2.0
XSD	Scotland	306	4	1.3
XHM	Humberside	172	0	0.0
Total		13,195	1,172	∅8.9

only occurs with a frequency of around 5 per cent, significantly less frequently than for the first group. Again, there are no significant differences within this group. At the very bottom of the table are the 'Celtic' Englishes: Ireland with only 3.0 per cent and Scotland at 1.3 per cent seem to have only negligible occurrences of AIN'T, and this is of course as expected (AIN'T forms are not recorded to occur in the traditional dialects of either Scotland or Ireland). Wales on the other hand clearly patterns with its neighbours, the northwest Midlands and the upper southwest. Whereas the first (high frequency) group extends throughout England in a homogeneous area, extending from the central north over the east and central Midlands to East Anglia and to the lower southwest, the second (low frequency) group constitutes isolated areas or 'pockets'. The main region is the mid-west (Wales, west Midlands, Merseyside and the upper southwest); other areas are, most notably, the northeast, and finally the southeast (London and the Home Counties), as Map 6.2 details.

Table 6.4 similarly details the figures for AIN'T for HAVE per dialect region. The rows have again been ordered in decreasing (relative) frequency of AIN'T.

Table 6.4 shows that just as for BE, AIN'T used for HAVE is present throughout the British Isles, but in significantly higher numbers. The range is also much wider, ranging from almost one in three occurrences (31.3

XSD

XIR

XNE

XNC

XNO

XLC

XHM

XMD

XMC

XME

XMW

XMI

XEA

XWA

XMS

XSU

XHC

XLO

XSS

XSL

AIN'T > 9%

AIN'T < 9%
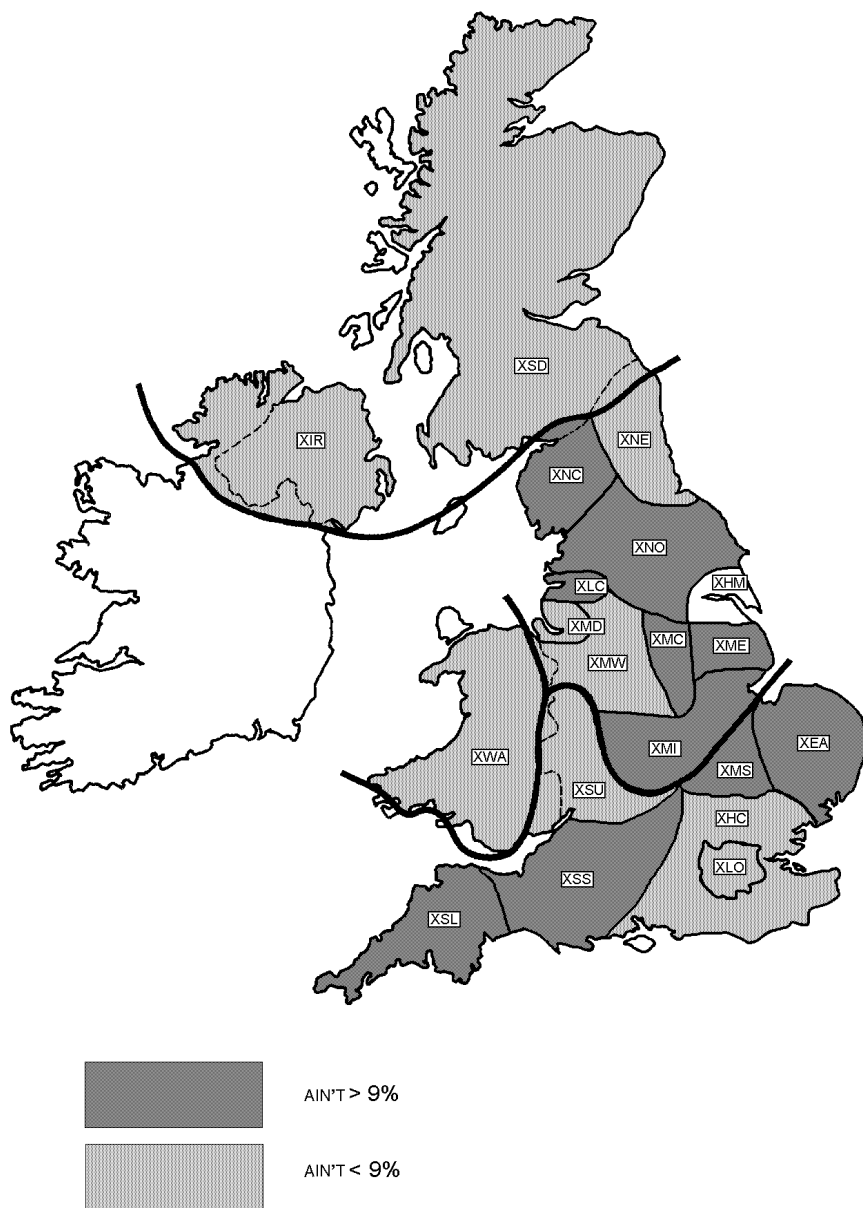
Map 6.2 AIN'T for BE in the BNC

Source: Adapted with permission from Trudgill, P. (1990) *The Dialects of England*, London: Edward Arnold.

Table 6.4 AIN'T for HAVE in the BNC-SpS

BNC code	Total	AIN'T for HAVE	% AIN'T of total
XEA	351	110	31.3
XSL	277	76	27.4
XMS	173	44	25.4
XNO	51	11	21.6
XSS	425	80	18.8
XMI	127	23	18.1
XME	230	39	17.0
XMC	320	53	16.6
XLO	805	121	15.0
XNE	187	28	15.0
XSU	51	7	13.7
XNC	267	25	9.4
XLC	197	16	8.1
XHC	530	40	7.5
XWA	275	15	5.5
XMD	64	3	4.7
XMW	433	14	3.2
XIR	70	1	1.4
XSD	118	1	0.8
XHM	70	0	0.0
Total	5,021	707	Ø14.1

per cent in East Anglia) to practically zero occurrences in Ireland and Scotland, the traditional AIN'T-free zone. Again, we can see two distinct groups of dialect areas. For the first group, percentages are at least in double digits, in fact clustering around 20 per cent. For the second group, percentages are considerably lower; the average here is just under 6 per cent – a significant difference. The two groups are displayed in Map 6.3. Not surprisingly, most dialect areas behave similarly for BE and for HAVE, as a comparison of the two maps shows. The most notable differences are the upper southwest, the northeast and London, all of which have rather low percentages for BE, but relatively high ones for HAVE. This is due to the main trend already discovered, that HAVE generally tends to have much higher relative frequencies of AIN'T than BE, and these dialect areas can be regarded as particularly good examples of this tendency.

Ain't *vs* in't

We shall now investigate the differences between the verb forms *ain't* and *in't* which have so far been analysed together as AIN'T. Again, we shall start with uses for BE. Table 6.5 displays the use of AIN'T for BE, subdividing it for *ain't* vs *in't*. Column two, indicating all occurrences of AIN'T

Map 6.3 AIN'T for HAVE in the BNC

Table 6.5 Ain't vs *in't* for BE in the BNC-SpS

BNC code	Total AIN'T for BE	ain't	in't	% in't of total
XEA	75	75	0	0.0
XHC	100	95	5	5.0
XIR	6	6	0	0.0
XLC	99	7	92	92.9
XLO	169	101	68	40.2
XMC	141	34	107	75.9
XMD	12	11	1	8.3
XME	58	58	0	0.0
XMI	46	26	20	43.5
XMS	44	44	0	0.0
XMW	79	18	61	77.2
XNC	134	17	117	87.3
XNE	14	11	3	21.4
XNO	22	14	8	36.4
XSD	4	3	1	25.0
XSL	39	25	14	35.9
XSS	96	89	7	7.3
XSU	3	3	0	0.0
XWA	31	27	4	12.9
Total	1,172	664	508	Ø43.3

for BE, i.e. the sum of all non-standard forms, is identical to Table 6.3.[13] Forms of AIN'T are, for our purposes, divided into *ain't* and *in't*, the two forms that actually occur in the BNC transcriptions. The final column indicates the percentage of *in't* in relation to all non-standard occurrences, i.e. in relation to the sum of *ain't* + *in't*.

If we look at the occurrence of *in't* as opposed to *ain't*, we can see quite generally that there are dialect areas where *in't* does not occur at all, although *ain't* does. For BE, *in't* is not used in five out of nineteen areas (i.e. in over a quarter of all dialect regions).[14] The reverse, however, is never the case: there is no dialect area that employs *in't*, but not *ain't*. This seems to suggest an implicational tendency: if *in't* occurs in a dialect, it is very likely that *ain't* will occur as well. This tendency is displayed in (21).

(21) Implicational tendency I for *in't*:
      in't ⊃ *ain't*

The fact that we do not find a complementary distribution of these two forms is on its own already an indication that the choice of *in't* over *ain't* is not governed purely by the factor of region; it shows that *in't* functions in a much more restricted way than the more widely used *ain't*. In other words, even where it does occur, *in't* does not seem to be a fully functional alternative to *ain't*. Let us now look at the figures for individual

Table 6.6 *Ain't* vs *in't* for BE tags in the BNC-SpS

BNC code	Total AIN'T tags	ain't *tags*	in't *tags*	% in't of total
XEA	14	14	0	0.0
XHC	29	25	4	13.8
XIR	3	3	0	0.0
XLC	74	4	70	94.6
XLO	85	35	50	58.8
XMC	108	13	95	88.0
XMD	11	10	1	9.1
XME	29	29	0	0.0
XMI	25	5	20	80.0
XMS	14	14	0	0.0
XMW	71	13	58	81.7
XNC	116	7	109	94.0
XNE	5	2	3	60.0
XNO	17	9	8	47.1
XSD	2	1	1	50.0
XSL	20	8	12	60.0
XSS	27	21	6	22.2
XSU	1	1	0	0.0
XWA	18	14	4	22.2
Total	669	228	441	Ø65.9

areas. If we compare the average percentages for *in't* in relation to all non-standard strategies, we can see that out of the fourteen areas that employ *in't*, only four use it as the dominant non-standard strategy, i.e. in more than 50 per cent of all cases, as indicated by the averages in the final column. These are Lancashire (at 92.9 per cent), the central Midlands (at 75.9 per cent), the northwest Midlands (at 77.2 per cent) and the central north (at 87.3 per cent), and they form a homogeneous area in the mid-northwest of England. In order to say more about *in't*, both in this dominant *in't*-area and in the rest of the country, we have to investigate the distribution between tag questions and non-tag questions. The first part, the use of *in't* in tag questions, is displayed in Table 6.6.

Table 6.6 shows that *in't* in tag questions is used much more widely (in regional terms) than the first overview might have suggested: here, *in't* is the dominant non-standard strategy for as many as eight dialect areas. These are (in order of decreasing frequency): Lancashire (at 94.6 per cent), the central north (at 94 per cent), the central Midlands (88 per cent), the north-west Midlands (almost 82 per cent), the Midlands (80 per cent), the northeast (at 60 per cent), the lower southwest (60 per cent) and London (at 58.8 per cent). Again, most of these areas are north of the dominant north–south dialect border (e.g. as identified by Trudgill 1990: 63). The only two dialect areas outside this 'northern' area are London and the lower southwest. If we compare these figures to figures for *in't* used in declaratives

(more exactly, outside the environment of tag questions, i.e. in 'non-tags'), we find striking differences.

Table 6.7 shows that only eight areas use *in't* in declaratives at all. Of these eight areas, only one area actually uses *in't* as the dominant strategy in declaratives (more precisely, in non-tags); this is Lancashire at 88 per cent. For all areas, there is a large decrease from the use of *in't* in tag questions to the use in declaratives (except in Lancashire, where, as we have seen, *in't* is equally dominant in both environments). Indeed, in six areas, *in't* occurs exclusively in tags, and not in declaratives at all. These are Merseyside, the Midlands, the northeast, the north, Scotland and Wales. In five more, *in't* occurs in declaratives only very marginally (with occurrences below five): these are the Home Counties, the central Midlands, the northwest Midlands, the lower southwest and the central southwest. These eleven dialect areas account for the majority of regions in which *in't* occurs at all. All these figures point to the fact that *in't* (at least in the data from the BNC) clearly prefers the grammatical environment of a tag question.[15] This is then the answer to the functional differentiation the overall figures already indicated: *in't* seems to have developed into a reduced form for *ain't* that is used mainly in the highly marked environment of tag questions.

If we look at tags more generally, Table 6.8 gives some interesting insights into the overall relation of tag questions to non-tags.

Table 6.7 Ain't vs *in't* for BE, non-tags, in the BNC-SpS

BNC code	Total AIN'T non-tags	Total in't non-tags	% in't of total
XEA	61	0	0.0
XHC	71	1	1.4
XIR	3	0	0.0
XLC	25	22	88.0
XLO	84	18	21.4
XMC	33	12	36.4
XMD	1	0	0.0
XME	29	0	0.0
XMI	21	0	0.0
XMS	30	0	0.0
XMW	8	3	37.5
XNC	18	8	44.4
XNE	9	0	0.0
XNO	5	0	0.0
XSD	2	0	0.0
XSL	19	2	10.5
XSS	69	1	1.4
XSU	2	0	0.0
XWA	13	0	0.0
Total	503	67	Ø13.3

Table 6.8 Non-tags vs tags for BE in the BNC-SpS

Verb	Total	Non-tags	Tags	% of tags
Standard BE	11,851	8,048	3,803	32.1
ain't	664	437	227	34.2
in't	508	67	441	86.8
Total BE	13,023	8,552	4,471	Ø34.3

Whereas overall more than one-third (34.3 per cent) of all occurrences of negative BE (standard and non-standard forms) occur in tags, as the last row shows, this ratio is very similar at 32.1 per cent for the standard English forms of BE; it remains basically the same for *ain't* at 34.3 per cent (the apparent rise is not statistically significant) but increases to an impressive 86.8 per cent for *in't*. The figures from the BNC thus seem to support in every detail Cheshire's finding that *in't* occurs more frequently in tag questions. This is not the case for *ain't*, however: on average, it occurs in tag questions at more or less the same ratio as the standard English forms. If we look at the reverse claim that a non-standard form is preferred in tag questions in general, this is not borne out by the figures from the BNC: the combined figures for any non-standard strategy (i.e. *ain't* plus *in't*) in a tag question only add up to 668/4,471 or 14.9 per cent and are thus in a clear minority when compared to the standard forms of present tense BE. In other words: almost all occurrences of *in't* are in a tag question, but most tag questions are not formed with a non-standard form. The relatively low overall figures for non-standard strategies of course mirror the fact that the BNC contains many standard speakers in its sample.

Let us now investigate a final claim with regard to *ain't* and *in't*, namely the claim that *in't* is sensitive to a distinction of third person vs non-third person subjects. If we indeed want to assume, following Jespersen, that *in't* is not, or perhaps not only, derived from a simplification of, or a variation in, the pronunciation of *ain't*, but from *isn't* in a separate development, then this might still be mirrored in the present day distribution. For those dialect areas that employ *in't*, Table 6.9 shows the figures for third person singular subjects. Table 6.10 completes the picture, detailing the figures for all other persons other than the third person singular.

If the *in't* = *isn't* hypothesis is right, one should expect significant decreases from third person uses to non-third person uses in the ratio of *in't* to AIN'T, i.e. from the figures in Table 6.9 to Table 6.10. As the absolute figures are very low for almost every dialect area, a detailed statistical analysis of the individual regions is unfortunately not possible. The only area that can be tested is the lower southwest. Here, the decrease from *in't* used for the third person singular to *in't* used for all other persons is highly significant (at p < 0.01). The same is true for the overall figures

Table 6.9 AIN'T for 3sg subjects in the BNC-SpS

BNC code	Total AIN'T 3sg	3sg ain't	3sg in't	% in't of total
XHC	41	37	4	9.8
XLC	95	4	91	95.8
XLO	126	61	65	51.6
XMC	122	21	101	82.1
XMD	11	10	1	9.1
XMI	37	17	20	54.1
XMW	75	16	59	78.7
XNC	125	12	113	90.4
XNE	7	4	3	42.9
XNO	19	11	8	42.1
XSD	1	0	1	100.0
XSL	17	8	9	52.9
XSS	51	45	6	11.8
XWA	17	14	3	17.6
Total	744	260	484	∅66.5

Table 6.10 AIN'T for non-3sg subjects in the BNC-SpS

BNC code	Total AIN'T non-3sg	ain't non-3sg	in't non-3sg	% in't of total
XHC	59	58	1	1.7
XLC	4	3	1	25.0
XLO	43	40	3	7.0
XMC	19	13	6	31.6
XMD	1	1	0	0.0
XMI	9	9	0	0.0
XMW	4	2	2	50.0
XNC	9	5	4	44.4
XNE	7	7	0	0.0
XNO	3	3	0	0.0
XSD	3	3	0	0.0
XSL	22	17	5	22.7
XSS	45	44	1	2.2
XWA	14	13	1	7.1
Total	242	218	24	∅9.9

in the last rows. If one looks at the absolute figures for the other regions, it becomes apparent that the general trend that '*in't* forms occur more often with . . . a third person singular subject' (Cheshire 1982: 52) is perceivable for many dialect regions, even though the figures cannot be directly tested. It is also interesting to note that although there are areas where

Table 6.11 Non-tags vs tags for *in't* in the BNC-SpS

	Total	Non-tags	Tags	% tags of total
3sg *in't*	484	64	420	86.8
Non-3sg *in't*	24	3	21	87.5
Total	508	67	441	∅86.8

in't does not occur with a non-third person singular pronoun, although it occurs with a third person singular pronoun, the reverse is never the case. In terms of another implicational tendency, then, if we find *in't* for a non-third person subject, it is very likely that *in't* will be used for a third person subject as well. This implicational tendency for *in't* is displayed in (22).

(22)   Implicational tendency II for *in't*:
       in't for non-3rd sg ⊃ *in't* for 3rd sg

This preference of *in't* for the third person singular must be a reflex of its origin as a reduced form of *isn't* in those areas where it does occur. An average rate of over 66 per cent shows that this reflex of its origin is still strongly present. Nevertheless, also in this subclassification for type of subject, *in't* still prefers the grammatical environment of a tag question, as the summary Table 6.11 shows.

For third person singular subjects, 86.8 per cent of all forms of *in't* occur in a tag question. For all other subjects, the ratio is 87.5 per cent, a very similar figure. (Again, the difference is not significant.) Cheshire's claim that in her data the occurrences of *in't* for third person subjects 'are almost entirely in tag questions' (Cheshire 1982: 52) can therefore be extended to the data from the BNC as well. As the preceding section has shown, however, this is not so much a feature of the third person singular environment (shown by the non-significant difference to other environments), but of the behaviour of *in't* in general, which can be regarded as the reduced alternative of *ain't* that occurs in particular in tag questions. It is therefore as expected that the few occurrences there are of *in't* with a non-third singular subject occur almost exclusively in tag questions.

This final section on the behaviour of AIN'T used for BE takes up Cheshire's claim that in her data, AIN'T occurs 'most often as auxiliary HAVE . . . and least often as auxiliary BE', with copular BE situated somewhere in between (Cheshire 1982: 51). This hierarchy was summarized in (8) and is reproduced below.

(8)   Frequency hierarchy for AIN'T:
       aux HAVE > cop BE > aux BE

This is a claim worth testing. Table 6.12 and Table 6.13 show the use of AIN'T (i.e. *ain't* plus *in't*) for BE, this time subdivided for copular and auxiliary use of BE. The standard present tense forms of BE could not be manually disambiguated for main verb vs auxiliary use, as there were far too many occurrences. However, as one can assume that the ratio of main verb (copular BE) and auxiliary verb use should be quite similar across regions, three random samples were counted instead, each of 200 tokens, of negated present tense BE forms in spoken texts of the BNC. The frequency of auxiliary BE was remarkably stable at 29, 27 and 29 occurrences. An average frequency for auxiliary BE in the spoken part of the BNC of roughly 14 per cent was therefore assumed from these figures, but this procedure of course makes the following results rather tentative.

The figures for AIN'T for copular and auxiliary use do not quite add up to the totals figure given above in Table 6.3 for the simple reason that in some cases it was impossible to decide whether a given form, especially in a tag question, referred back to either auxiliary or copular BE. These unclear examples were included in Table 6.3, as they could usually be assigned to BE (vs HAVE) unambiguously, but were excluded from Table 6.12 and Table 6.13, as assignment to copular or auxiliary BE was impossible. In sum, however, unclear examples only account for seventy-nine occurrences of AIN'T for BE (or 6.7 per cent) – most of which are due to *ain't it/in't it* used as a semi-invariant tag, i.e. following a subject and verb different from *it is*.[16]

Table 6.12 AIN'T for cop BE in the BNC-SpS

BNC code	Copular BE	AIN'T	% AIN'T
XEA	598	49	8.2
XHC	1,313	57	4.3
XIR	170	4	2.4
XLC	557	79	14.2
XLO	2,079	62	3.0
XMC	705	105	14.9
XMD	173	10	5.8
XME	517	38	7.4
XMI	236	31	13.1
XMS	204	25	12.3
XMW	1,072	67	6.3
XNC	659	116	17.6
XNE	393	3	0.8
XNO	160	20	12.5
XSD	261	2	0.8
XSL	340	23	6.8
XSS	796	61	7.7
XSU	128	2	1.6
XWA	642	20	3.1
Total	11,003	774	Ø7.0

Table 6.13 AIN'T for aux BE in the BNC-SpS

BNC code	Aux BE	AIN'T	% AIN'T
XEA	110	19	17.3
XHC	241	33	13.7
XIR	30	2	6.7
XLC	89	10	11.2
XLO	360	37	10.3
XMC	124	25	20.2
XMD	28	1	3.6
XME	97	18	18.6
XMI	48	14	29.2
XMS	48	18	37.5
XMW	171	5	2.9
XNC	99	9	9.1
XNE	73	9	12.3
XNO	25	2	8.0
XSD	45	2	4.4
XSL	65	12	18.5
XSS	153	31	20.3
XSU	22	1	4.5
XWA	113	10	8.8
Total	1,941	258	⌀13.3

A comparison of the 'Total' rows of both tables shows that on average, AIN'T is used more frequently (relatively speaking) for auxiliary BE than for copular BE (we find a rise from 7 per cent to 13.3 per cent). This rise is highly significant at $p < 0.01$. If we look at the individual regions, this trend is borne out practically everywhere. (Again, six areas where AIN'T was used less than five times had to be excluded from statistical tests; these were Ireland, Merseyside, the northeast, the north, Scotland and the upper southwest.) Only one region reverses the trend, this is the central north with a significant fall from 17.6 per cent for copular to 9.1 per cent for auxiliary BE; and in two areas the difference between the two functions is not significant (these are Lancashire and the northwest Midlands). Thus only three areas go against the strong trend indicated by the totals. We can summarize that although (or probably because) auxiliary BE is much less frequent in absolute occurrences,[17] it is much more frequently simplified to AIN'T in relative terms. This distribution is as expected if we presume an underlying general trend of simplification, which seems to be at work here for AIN'T, but it runs counter to Cheshire's proposed hierarchy. How these figures can be related to figures for AIN'T used for HAVE will be investigated in the next section, which at the same time moves us to AIN'T used for HAVE.

AIN'T for HAVE

The general figures for AIN'T used as the present tense negative of HAVE have already been discussed in comparison with the overall figures for BE in Table 6.1. In order to make a comparison with the subcategories of copular and auxiliary BE easier, the following table repeats the percentages for all three verbs. Table 6.14 shows the left-to-right order that Cheshire's scale would predict on the basis of her findings from Reading (aux HAVE > cop BE > aux BE). (However, we have already seen above that BE seems to behave in the reverse order, as AIN'T is much more frequently used, relatively speaking, for auxiliary BE than for the copula.) Columns three and five indicate where the predicted relation is either actually reversed in the BNC material (<) or where a difference is not statistically significant (=).[18] Dialect areas which had to be excluded from statistical testing are marked by (ex.).

From the third and fifth columns it can be seen that Cheshire's ordering of the verbs results in seventeen out of twenty-nine relations that do not fit the predicted order. (The relations excluded from statistical testing have also been excluded from this total.) With ca. 60 per cent of the data showing an unexpected behaviour, this scale is certainly not a good model for the behaviour of AIN'T in British non-standard speech in general. Indeed, nowhere does Cheshire claim that what she found for her Reading

Table 6.14 AIN'T for all three verbs (Cheshire's order)

BNC code	% AIN'T for HAVE	predicted: >	% AIN'T for cop BE	predicted: >	% AIN'T for aux BE
XEA	31.3		8.2	<	17.3
XHC	7.5		4.3	<	13.7
XIR	1.4	ex.	2.4	ex.	6.7
XLC	8.1	<	14.2	=	11.2
XLO	15.0		6.0	<	10.3
XMC	16.6	=	14.9	<	20.2
XMD	4.7	ex.	5.8	ex.	3.6
XME	17.0		6.2	<	18.6
XMI	18.1	=	13.1	<	29.2
XMS	25.4		12.3	<	37.5
XMW	3.2	<	6.3	=	2.9
XNC	9.4	<	17.6		9.1
XNE	15.0	ex.	0.8	ex.	12.3
XNO	13.7		12.5	ex.	8.0
XSD	0.8	ex.	0.7	ex.	4.4
XSL	27.4		6.8	<	18.5
XSS	18.8		7.7	<	20.3
XSU	13.7	ex.	1.6	ex.	4.5
XWA	5.4		3.1	<	8.8
Total	Ø14.1	>	Ø7.6	<	Ø13.3
Absolute fr.	5,021		11,003		1,941

material could have a wider relevance. Her implied scale has purely descrip-
tive power for those Reading peer groups investigated, but is not intended
to support any further claims. As her order of verbs emerges inductively
from her material, it would indeed be difficult to see where any predic-
tive power might come from.

   As we have seen before, AIN'T in general seems to be structured as a
phenomenon of regularization. This phenomenon typically affects infre-
quent contexts first, such that we find the highest relative frequencies of
AIN'T in contexts that are the least frequent in absolute terms. We have
seen that this is the case both when we compare BE and HAVE, and when
we compare auxiliary vs copular BE. As the table above makes clear,
however, a combination of these subdivisions for BE with HAVE does not
lead to very meaningful results. There are indeed dialect areas where the
order is cop BE < HAVE < aux BE, as the idealized regularization distri-
bution would predict, but there are also other areas where this relation
does not hold. To test whether the material fits the regularization scale
any better, Table 6.14 has been rearranged in order of absolute frequen-
cies (cop BE > aux HAVE > aux BE). According to the regularization scale,
the rarest phenomenon should be regularized by use of AIN'T most
frequently. Thus, the expected relative frequency of AIN'T should go in
the opposite direction (cop BE < HAVE < aux BE), provided it makes sense
to subdivide BE in this way. Again, columns three and five indicate where
the expected ratio is contradicted.

   Table 6.15 shows that again only twelve out of thirty relations are as
predicted. In other words, the error rate does not seem to sink. With this
reordering, however, the source of the error becomes a little clearer. What
seems to be responsible for the high error rate is the fifth column: there
is generally no significant difference between AIN'T used for HAVE and
AIN'T used for auxiliary BE. Both uses are significantly more frequent (in
relative terms) than for copular BE, as Table 6.14 shows, but there is no
additional difference between these two verbs. As we have seen in Chapter
2, the distinction auxiliary vs copular BE seems to play only a minor role
in general; there are no features of syntax that pay attention to this distinc-
tion (unlike for HAVE, where auxiliary and full verb uses are distinguished,
for example, by different negation strategies), and this differentiation only
becomes slightly significant when one looks at the figures for BE internally.
On the other hand, there is a huge difference between BE (auxiliary and
copular use taken together) and HAVE, as Table 6.3 and Table 6.4 have
shown. Mixing these (word) external and internal categories, however, does
not result in systematic regularities which could predict ratios for the use
of AIN'T reliably, and a distinction of BE into auxiliary and copular uses
is perhaps not particularly warranted.

   Let us now look at the distinction of *ain't* vs *in't* for HAVE. If the *in't* =
isn't hypothesis is right, there should not be any occurrences of *in't* for a
form of HAVE. On the other hand, it is also possible that *in't* is a reduction

Table 6.15 AIN'T for all three verbs (in order of absolute frequencies)

BNC code	% AIN'T for cop BE	predicted: <	% AIN'T for HAVE	predicted: <	% AIN'T for aux BE
XEA	8.2		31.3	>	17.3
XHC	4.3		7.5		13.7
XIR	2.4	ex.	1.4	ex.	6.7
XLC	14.2	>	8.1	=	11.2
XLO	6.0		15.0	>	10.3
XMC	14.9	=	16.6	=	20.2
XMD	5.8	ex.	4.7	ex.	3.6
XME	6.2		17.0	=	18.6
XMI	13.1	=	18.1	=	29.2
XMS	12.3		25.4	=	37.5
XMW	6.3	>	3.2	=	2.9
XNC	17.6	>	9.4	=	9.1
XNE	0.8	ex.	15.0	=	12.3
XNO	12.5		13.7	ex.	8.0
XSD	0.7	ex.	0.8	ex.	4.4
XSL	6.8		27.4	=	18.5
XSS	7.7		18.8	=	20.3
XSU	1.6	ex.	13.7	ex.	4.5
XWA	3.1		5.4	=	8.8
Total	∅7.6	<	∅14.1	=	∅13.3
Absolute fr.	11,003		5,021		1,941

from *ain't*, or that it has been interpreted as such, in which case we should expect *in't* used for HAVE as well. The figures for those dialects that do have occurrences of *in't* for HAVE are displayed in Table 6.16.

Table 6.16 shows that *in't* is indeed used for HAVE only marginally, which supports Jespersen's hypothesis that *in't* is mainly derived from a contraction of *isn't*. *In't* occurs in only five (out of nineteen) dialect areas for HAVE, i.e. in about one-quarter. Even in those areas where *in't* is actually used, *ain't* is always preferred: the *in't* to AIN'T ratio is never above 10 per cent; in fact, the highest ratio occurs in the central Midlands with 9.4 per cent. The average of 5.8 per cent for all *in't* areas supports this trend. Compared to all non-standard forms from Table 6.4, including those areas where *in't* does not occur at all, the ratio is even lower: *in't* only accounts for 2.4 per cent of all non-standard strategies for HAVE in British English today. We have seen that *in't* for BE is used mainly in tag questions, and this relationship will also be investigated for HAVE.

If we compare the general Table 6.16 with the figures for tag questions in Table 6.17, we see that again of those *in't*-forms that do occur, most can be found in this particular grammatical environment. Indeed, only two instances of *in't* occur outside of tag questions. Although tags are therefore clearly the dominant environment for *in't*, *in't* is never the dominant strategy for a non-standard form in a tag question – the highest ratio

Table 6.16 Ain't vs *in't* for HAVE in the BNC-SpS

BNC code	AIN'T for HAVE	ain't	in't	% in't of AIN'T
XLC	16	15	1	6.3
XLO	121	116	5	4.1
XMC	53	48	5	9.4
XNC	25	24	1	4.0
XSL	76	71	5	6.6
Total	291	274	17	Ø5.8

Table 6.17 Ain't vs *in't* for HAVE tags in the BNC-SpS

BNC code	Total AIN'T tags	ain't	in't	% in't of total
XLC	12	11	1	8.3
XLO	26	21	5	19.2
XMC	22	18	4	18.2
XNC	10	9	1	10.0
XSL	37	33	4	10.8
Total	107	92	15	Ø14.0

can be found in London with 19.2 per cent. All these figures point to the fact that *in't* is generally dispreferred as a non-standard strategy for negating present tense HAVE. Where it does occur, it does so marginally, and mainly in tag questions. This low occurrence of *in't* for HAVE seems indeed to be a remnant of a former exclusive use for BE and speaks against the theory that *in't* might have originated as a contracted form of *ain't*.

We have also seen above that *in't* for BE is clearly preferred for third person singular subjects. On the question of whether *in't* for HAVE is also preferred in this environment, Table 6.18 will shed some light. Because occurrences in individual dialect areas are generally very low, only the total figures are reproduced.

The general picture that emerges from Table 6.18 is similar to the results derived from the respective tables for present tense BE: the relation of *in't* to all non-standard strategies is only 3.3 per cent on average for non-third person subjects, but 12.3 per cent for third person singular subjects (i.e. almost four times as much; this rise is highly significant at $p < 0.01$). There thus seems to be a decided preference for *in't* to occur in a third person singular environment, even when it is used for HAVE. It looks as if this preference of *in't* (which is much more pronounced for BE, as we have seen above) has been taken over when *in't* came to be used for HAVE as well. However, although *in't* prefers to occur with a third person singular subject, if one compares this with figures for standard

Table 6.18 Ain't and *in't* for HAVE in the BNC-SpS

	AIN'T	ain't	in't	% in't of AIN'T
3 sg	81	71	10	12.3
Non-3 sg	210	203	7	3.3
Total	291	274	17	Ø5.8

Table 6.19 3sg HAVE in the BNC-SpS

	Total	has	3sg ain't	3sg in't	% in't of total
Non-tags	233	191	42	0	0.0
Tags	121	82	39	10	8.3
Total	354	273	71	10	Ø2.8

Table 6.20 Tags vs non-tags of 3sg HAVE in the BNC-SpS

Verb	Total	Non-tags	Tags	% tags of total
3sg standard HAVE	273	191	82	30.0
3sg *ain't*	71	42	29	40.8
3sg *in't*	10	0	10	100.0
Total HAVE	354	233	121	Ø34.2

HAVE forms given in Table 6.19, the reverse is certainly not the case: *in't* is never the dominant strategy when a third person subject is negated.

The overall percentages make clear that although *in't* prefers a third person subject, as we have seen in table 6.18, a third person subject certainly does not prefer *in't*: on average, only ten out of 354 or 2.8 per cent of third person subjects occur with *in't* for negative HAVE. As a glance at the last column shows, all occurrences of *in't* are actually in tag questions. If we look at the overall tag question ratio in table 6.20, which rearranges the figures from table 6.19 slightly, again we have a straight increase from standard HAVE, where 30 per cent of all forms occur in tag questions, to *ain't* (at 40.8 per cent) to *in't* (100 per cent). Again, however, the reverse is not the case. Although *in't* occurs in tag questions only, the majority of tag questions do not have *in't*, as table 6.19 has shown: only ten out of 121, or 8.3 per cent of tag questions are formed with *in't*; *ain't* on the other hand rates at twenty-nine out of 121, i.e. 24 per cent, a significant increase.

Going back to Cheshire's claims for Reading, then, it must be said that her finding that 'in tag questions . . . *in't* [is] predominating, for all verbs'

(Cheshire 1982: 56) is certainly not borne out by data from the BNC. Instead, *in't* shows remnants of its derivation from *isn't*, in that it prefers to occur with third person subjects, and this preference is also found for HAVE. Much stronger than this preference, however, is the preference to occur in tag questions, which makes *in't* a functional alternative to *ain't* for precisely this grammatical environment.

AIN'T for full verb HAVE

The final section on AIN'T used for HAVE investigates a phenomenon which linguists are quite sure does not exist: the use of AIN'T for full verb HAVE. However, there are a number of dubious examples from the BNC that should not be discarded right away. If we want to compare the frequency of use of AIN'T for HAVE as a full verb with the standard usage, one of course first has to find out the distribution of full verb negation for the standard English forms of negated HAVE. In order to do this, several random samples of *haven't/hasn't* in spoken texts of the BNC were conducted. They indicate quite consistently that only about 2 per cent of all forms of standard English HAVE + negation are negations of the full verb HAVE. As shown in Chapter 2, the much more usual case for negating full verb HAVE is of course negation with DO-support, even for the more conservative written texts in British English today. It is not surprising that this trend should be even stronger for spoken language.[19] If we conduct searches for full verb HAVE negated without DO-support in the individual dialect areas and restrict the results to the SpS subsample, this ratio is even slightly lower at sixty-four instances out of 4,314 or 1.5 per cent (cf. Table 6.21).

If we consider that in total, 'only' 707 incidents of HAVE are negated by the form AIN'T (cf. Table 6.4), one would expect between 1.5 and 2 per cent of these to be the main verb HAVE (between ten and fourteen occurrences) if AIN'T negates HAVE indiscriminately. As we have seen above, linguists are quite adamant on the other hand that AIN'T can never negate main verb HAVE, so this counterclaim would lead us to an expected occurrence of AIN'T for main verb HAVE of zero.

The data from the BNC-SpS subsample supplies seven instances where AIN'T could be interpreted as being used as the negative of the full verb

Table 6.21 Full verb HAVE of negated HAVE forms

	Total neg HAVE	Full verb	% of total
Spoken texts (random sample)	600	12	2.0
SpS	4,314	64	1.5
AIN'T (expected)	707	ca. 10.5	ca. 1.5

HAVE – but several of them are a little unclear. They are given in (23) to (29).

(23)   Is she, she getting worse? Brain damage *ain't she*? (XHC: KCP 5815)

(24)   *It ain't* nothing to do with my school. (XLO: KPG 4604)

(25)   Mind you *she ain't* nobody to squash. (XLO: KPG 6157)

(26)   Well *ain't you* nothing? ‖ What if I have! (XMS: KCP 7436)

(27)   She said you've a daughter *ain't ya* or summat. (XMW: KSS 2720)

(28)   We've already had one of those. We've already one, *ain't we*, Dave? (XNC: KD2 2636)

(29)   She has only the one home, her little baby *ain't she* the, well he isn't a baby now. ‖ She got three home hasn't she. There's only one at school. (XSS: KBE 7753)

Example (23) hinges on the form *brain damage*. If this is a mispronunciation or mistranscription of the past participle *damaged*, *ain't* is most naturally interpreted as a form of BE (*brain damaged, isn't she*). Examples (24) and (25) are a little dubious because *ain't* here could be substituted by *isn't* as well as *hasn't*, and even an extensive interpretation of the further context could not disambiguate them further. Examples (26) and (27), however, are clear examples of HAVE as a full verb being negated by *ain't*. Example (28) might be due to an ellipsed *had* that was still present in the preceding clause (*We've already [had] one, haven't we, Dave?*), in which case *ain't* would of course be the quite regular negative form of the auxiliary HAVE, rather than the full verb. Example (29) finally is another clear example of *ain't* used for full verb HAVE. The second clause cannot read *her little baby, isn't she*: the 'baby' concerned is a boy, not a girl, as the context makes clear.

The difference between the three unambiguous examples and the expected 10.5 occurrences is of course significant, and one cannot therefore say that AIN'T negates HAVE in all its functions indiscriminately (both as a main verb and as an auxiliary). However, the difference between zero and three is also highly significant, and it would be equally false to say that AIN'T can never function as the negative form of main verb HAVE – this general claim may therefore have to be modified. From the data from the BNC, it rather looks as if AIN'T might be on the way towards behaving like negated HAVE in all its functions. The examples (23) to (26) seem to point to an explanation for this extension of meaning, as ambiguous forms may be interpreted in two different ways. The 'normal' use (AIN'T used for BE or for the auxiliary HAVE) may function as a camouflage for the use of AIN'T as

Table 6.22 AIN'T *got* vs AIN'T in the BNC-SpS

BNC code	Total AIN'T	AIN'T got	% of total
XEA	110	60	54.5
XHC	40	22	55.0
XIR	1	1	100.0
XLC	16	4	25.0
XLO	121	63	52.1
XMC	53	22	41.5
XMD	3	1	33.3
XME	39	24	61.5
XMI	23	6	26.1
XMS	44	24	54.5
XMW	14	8	57.1
XNC	25	7	28.0
XNE	28	21	75.0
XNO	11	3	27.3
XSD	1	1	100.0
XSL	76	34	44.7
XSS	80	40	50.0
XSU	7	3	42.9
XWA	15	7	46.7
Total	707	351	Ø49.6

Table 6.23 AIN'T *got* for HAVE GOT in the BNC-SpS

BNC code	Total	HAVE GOT	AIN'T got	% AIN'T of total
XEA	68	8	60	88.2
XHC	29	7	22	75.9
XIR	1	0	1	100.0
XLC	5	1	4	80.0
XLO	76	13	63	82.9
XMC	35	13	22	62.9
XMD	4	3	1	25.0
XME	25	1	24	96.0
XMI	12	6	6	50.0
XMS	28	4	24	85.7
XMW	16	8	8	50.0
XNC	13	6	7	53.8
XNE	24	3	21	87.5
XNO	4	1	3	75.0
XSD	1	0	1	100.0
XSL	47	13	34	72.3
XSS	66	26	40	60.6
XSU	3	0	3	100.0
XWA	18	11	7	38.9
Total	475	124	351	Ø73.9

a negative form of full verb HAVE.[20] It is of course the case that main verb negation is very rare for AIN'T (it only accounts for about 0.5 per cent of all cases). However, as a look at the standard forms of HAVE has shown, it is almost equally rare here to find main verb negation without DO-support. Indeed, the usual expression for the circumstance of 'having' today is not HAVE on its own any more, but a form of *have got*, in positive as well as negative contexts. And in fact the expected 'regular' non-standard negative *ain't got* accounts for almost 50 per cent of all occurrences of AIN'T, as Table 6.22 shows.

In fact, *ain't got* is far more frequent than the standard *haven't got* – it accounts for 73.9 per cent of all occurrences of negative HAVE GOT, or almost three-quarters of all cases, as Table 6.23 shows. This option is realized in practically every dialect area, even those where *ain't* is traditionally not part of the dialect. The construction *ain't got* therefore may be the carrier that imports this form into new dialect areas.

AIN'T *for other verbs*

AIN'T used for verbs other than BE or HAVE (another impossibility of the dialectology textbooks) is a very marginal phenomenon. Individual examples will probably in most cases be put down to production errors. However, as today's errors might prepare the stage for tomorrow's changes, this section will take a short look at these exceptional forms. All examples are provided in (30) to (42).

(30) Someone had an intruder. I'm not sure | but I think it was Mrs Turner . . . Because it was an old age woman, *they ain't* give anyone's names | old woman was | it was in the paper sort of | met up with an intruder in her house like. (*ain't* for *don't*; XLO: KBF 10905)

(31) I told him if I *I ain't* never get anything else for, when I go on holiday. (*ain't* for *don't*; XMS: KD3 2735)

(32) *I* usually *ain't* er make much fuss. (*ain't* for *don't*; XMW: KC4 2680)

(33) Course I have, been giving it [i.e. the dog] a what we've had, like | when the kids didn't eat their dinner or *we ain't* eat all our dinner | been giving it what's left on the plate, well. (*ain't* for *didn't*; XEA: KB7 9404)

(34) And how long we had 'im | we 'ad 'im a good six months, *ain't we*? (*ain't* for *didn't* (or *hadn't?*); XLO: KB1 406)

(35) *I ain't* see any because I were with Jacqueline weren't I? (*ain't* for *didn't*; XLO: KCX 827)

(36) *I ain't* see their ‖ yeah yeah ‖ Sarah for as long as I can, well as long as I . . . (*ain't* for *didn't*; XME: KC2 5014)

(37) She watched it on video, *we ain't* watch it on Sky. (*ain't* for *didn't*; XSS: KC50)

(38) *I ain't* do nothing against the law. (*ain't* for *didn't*; XSS: KE6 8067)

(39) I was too busy, *ain't I*? (*ain't* for *wasn't*; XMC: KBB 6378)

(40) He said that's it *I ain't* have it no more. (*ain't* for *won't*; XSS: KBE 6470)

(41) You ought to of asked him when he sold it to [name] *ain't ya*? (*ain't* for *oughtn't*? XMC: KBB 5445)

(42) You might as well be off for two, *ain't ya*? (*ain't* for *mightn't*? XMI: KD6 4706)

Excluded from these examples are the seventy-nine almost canonical instances of *ain't it/in't it* as invariant or semi-invariant tag questions; this development has been amply documented elsewhere and is generally accepted as an ongoing change towards a more lexicalized form (e.g. *innit*). In sum, then, *ain't* is used for a relatively wide range of verb forms. These occurrences are summarized in Table 6.24.

Several interesting points emerge from these examples. The examples of *ain't* used in place of *don't/didn't* mostly hinge on the form of the verb following. The BNC transcriptions, disputed as they are, indicate an infinitive rather than a past participle. A past participle would, however, be required for the normal, unmarked and unremarkable reading HAVE. One

Table 6.24 AIN'T used for other verbs in the BNC-SpS

	Total ain't	Tags	% of total
don't	3	0	0.0
didn't	6	1	16.7
wasn't	1	1	100.0
won't	1	0	0.0
oughtn't	1	1	100.0
mightn't	1	1	100.0
Total	13	4	Ø30.8

possible explanation for this phenomenon might be that many dialects have verb paradigms different from the standard English paradigms. However these non-standard paradigms are generally characterized either by a simplification of 'irregular' forms (resulting, for example, in *see*, *seen*, *seen*), or a distribution of simple past and past participle forms reverse of the standard English paradigm (*see*, *seen*, *saw*). It is not usually the case that what looks like an infinitive (as in the examples above) is used as the past participle. We can therefore be quite confident that the instances of *ain't* above do indeed function for *don't* and *didn't*. The contexts also argue for this interpretation and for this reason extended contexts have been provided. *Ain't* used for *don't* and *didn't* make up the majority of all other-verb uses (nine out of thirteen, or 69.2 per cent), and they are the only verb forms for which *ain't* occurs more than once. Also, they are almost the only verbs where *ain't* is used in declaratives (the only other verb is *won't* with just one occurrence). This slight tendency for *ain't* to occur for forms of DO is not very surprising, given the status of DO as the third primary verb – if AIN'T substitutes negated forms for BE and HAVE, why not for DO as well? In addition, AIN'T for DO is precisely what we find in African American Vernacular English, the one dialect that is charac-terized by an almost exclusive use of AIN'T as the all-purpose negator. The fact that AAVE uses AIN'T for the negation of DO, but all other dialects do not, has sometimes been used as a characteristic that sets AAVE apart from all other dialects of English. As data from the BNC shows, however, an extension of AIN'T from the frequent BE and HAVE para-digms to the next frequent DO paradigm seems to be under way in present day British English as well.

Similar arguments as for examples (30) to (38) against transcribers' errors hold for the next two examples where *ain't* is used for *wasn't* and *won't* respectively. Especially in example (40), the sentence is a highly idiomatic expression where only *won't* is usually found (*I won't have it*). Although unstressed *ain't* and *won't* might be quite similar phonetically, in this posi-tion *won't* usually carries primary stress, so that a mix-up on the part of the transcriber does not seem very likely. The most interesting examples, however, are the last two, examples (41) and (42). As the question mark indicates, the verb forms which *ain't* substitutes here are really more or less theoretical. *Oughtn't* as well as *mightn't* are extremely rare in spoken language generally; in tag question position these verbs are virtually non-existent, as we have seen in Chapter 4. This of course opens the question of an alternative: if it is not possible to use *oughtn't* or *mightn't* in a tag, what could be used instead? The obvious solution – at least to the speakers of (41) and (42) – seems to have been to use the one verb form that is already flexible enough to negate two very different verbs (BE and HAVE). This kind of 'emergency' use once the production of the main clause is already under way might indeed point to a perceived structural gap in the system that is emerging from the infrequent use of some peripheral

modals, as, for example, *ought* and *might*. If AIN'T is instinctively used to fill this gap, this is a possible source for a further spread of its use to other verbs.

Summary

We have seen that generally, AIN'T is present throughout England today. Compared with the data from the SED, the use of *ain't* as well as *in't* has spread dramatically. Only Scotland and Ireland still seem to be virtually *ain't*-free dialects, in accordance with most dialect descriptions today. When one looks at the figures in more detail, one can see that non-standard forms (AIN'T) tend to be more frequent (in relative terms) for HAVE than for BE. Also, AIN'T tends to occur more often in tags than in non-tags. This distribution strongly speaks for AIN'T as a process of analogy, which affects infrequent contexts earlier and more strongly than very frequent ones. When we looked at different realizations of AIN'T, we found that dialects which have *in't* also have *ain't* (but not the reverse), which suggests a functional hierarchy (also expressed as an implicational tendency). Characterizing *in't* in more detail, one can say that generally, *in't* shows a preference for occurring as a negative of BE and for occurrence with the third person singular, clearly remnants of its derivation from *isn't*. On the other hand, *in't* seems to have developed into a reduced alternative to *ain't* that today occurs almost exclusively in tag questions. In other words, the regional spread of *ain't* and *in't* seems to have gone hand in hand with a functional differentiation of these two forms.

    We have said that the mechanism of analogy typically proceeds by affecting infrequent items first (Hooper 1976). We have applied this general tendency to the difference between AIN'T for BE and for HAVE, where it might indeed be an argument for a lexical diffusion that began with BE and has only affected HAVE later by analogy. This, of course, does not exclude a natural derivation of AIN'T from both *haven't* and *hasn't*, as presented above. The process of analogy may, however, have promoted the use of AIN'T for HAVE once these forms were established from the contracted and reduced negative forms.

    In typological comparison, *ain't* certainly shows evidence of the trend towards a single negator for a range of verb forms (in particular, auxiliary BE and HAVE and copular BE); its main competitor in the realm of negative clauses is invariant *don't* for main verbs. On the other hand, the stigmatization of *ain't* seems to hinder a rapid spread to new contexts. Nevertheless, in the competition with *don't*, *ain't* seems to be the clear winner: there is no evidence of *don't* being used in place of *ain't*, whereas the reverse holds (in a few rare cases), as we have seen. *Ain't* in Great Britain seems to be making first inroads into the territory of *don't/didn't*, in parallel with AAVE – the only dialect of English which permits *ain't* for *don't* quite regularly.

$$
\left.\begin{array}{l}
\textit{am} \\[16pt]
\textit{is} \\[16pt]
\textit{are}
\end{array}\right\} \quad \textit{ain't}
$$

$$
\left.\begin{array}{l}
\textit{has} \\[16pt]
\textit{have}
\end{array}\right\} \quad \textit{ain't}
$$
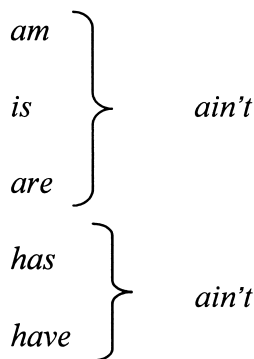
Figure 6.3 Paradigm of AIN'T

The most general, and therefore perhaps the most interesting charac-teristic, however, is the fact that for *ain't*, grammatical distinctions that hold in positive contexts are levelled under negation. For this reason, AIN'T is often described as a regularization or simplification strategy, as it levels the person distinctions that are still present in the positive environment (*am*, *is*, *are*; *has*, *have*) to the one form *ain't/in't* for all persons and numbers in the negative. Hudson, for example, analyses *ain't* (and, incidentally, *don't*) with its partial loss of subject-verb agreement (SVA) under negation as part of a 'steady diachronic shift towards a grammar [of English] where SVA plays no part at all' (Hudson 1999: 204). However, this analysis is slightly problematical as the positive paradigms of BE and HAVE seem to be remarkably robust and show no sign at all of levelling their person/number contrasts.

Instead, we find a very strong trend towards a pattern of asymmetry, displayed in Figure 6.3, that is typical for markedness contexts (in this case, negation), and which we shall encounter again in the following chap-ters. It will be discussed more fully in Chapter 9, when we have investigated more evidence for the pervasiveness of this phenomenon.

7 Third person singular *don't*

Introduction

This chapter investigates the occurrence of *don't* with third person singular subjects – a feature of non-standard dialects which might at first glance look like a failure of subject-verb concord, as examples (1) and (2) show.

(1) She said *she don't* mind us swearing. (KE6 348)

(2) He looks a bit weird, *don't he*? (KC8 1418)

In this chapter, we shall first take a historical look at this form in the traditional dialects of English, in order to determine whether *don't* has a regional origin. We shall then turn our attention to third-singular *don't* in non-standard speech today. *Don't* today will be analysed in its structure, in particular its distribution across various sentence types, and its regional distribution today. In addition, we shall try to answer the question whether third person *don't* has arisen in analogy to *don't* in all other persons, or whether it is an independent development of /z/-deletion. Finally, third-singular *don't* will be argued to be another example of the principle of asymmetry.

The phenomenon

Standard English makes the same distinction for the primary verb DO as for main verbs in its paradigm: *he/she/it* and full NPs in the singular take *does*, all other pronouns and plural NPs are followed by *do*. In other words, the only morphological distinction for DO in standard English is the usual one between third person singular (marked) and all other persons (unmarked). This third person singular vs non-third person singular distinction is neutralized in the past tense – as for all other verbs except BE – in this case to *did* for all subjects. It is therefore not surprising that the common strategy of simplification in non-standard speech should affect the DO paradigm in the present tense as well.

Negatives are formed by the addition of *not*, which can cliticize to *doesn't* and *don't*. As we have seen in Chapter 2, *don't* and *doesn't* do not have a counterpart where the auxiliary would be contracted to the subject (parallel to *I'm not*, *you're not*); this is a feature where standard and non-standard varieties do not differ. The main interest in this chapter will lie on the use of the morphologically unmarked form *don't* being generalized to the third person singular. Further examples of this phenomenon are provided in (3) to (5).

(3)   *He don't* like football, does he? (KR2 690)

(4)   She goes through the book, *don't she*? (KCU 443)

(5)   Well, *it don't* matter, does it? (KCT 9887)

For the sake of comparison, though, we shall also take a look at the generalization of its positive counterpart *do* to the third person singular.

History

Third-singular *don't* is a feature of non-standard dialects that is sometimes noted in the dialect literature, but rarely commented on in great length, let alone investigated in more detail, for reasons that are still unclear. Unlike negative concord and *ain't* in the preceding chapters, third-singular *don't* is certainly not part of a non-standard stereotype and thus does not seem to command either metalinguistic attention by its speakers, or indeed linguistic attention by dialectologists.

Jespersen is one of the few noteworthy exceptions. He has examples for third person *don't* dating back to the seventeenth century, and this is supported by the *OED*, which notes a very early example from Samuel Pepys from 1660 given in (6).

(6)   Sir Arthur Haselrigge do not yet appear in the house.
      (*OED* headword *do*)

Whereas this feature is usually explained as an extension from all other persons (with *don't*) to the third person singular, Jespersen argues in favour of a phonological derivation instead:

> *Don't* for *doesn't* is generally explained from a substitution of some other person for the third person; but as this is not a habitual process – as *do* in the third person sg. is found only in some few dialects, but not in standard English, and as the tendency is rather in the reverse direction of using the verb form in *s* with subjects of the other persons (*says*

> *I, they talks*, etc.), the inference is natural that we have rather a phonetic process, *s* being absorbed before *nt* as in *isn't*, etc.
>
> (1917: 122)

The question of derivation is a very interesting one, and we shall investigate this further in this chapter by means of the distribution of this feature through grammatical environments. First, however, we shall examine the distribution of *don't* and of its positive counterpart *do* in traditional dialects.

As Jespersen notes, positive '*do* in the third person sg. is found only in some few dialects' (Jespersen 1917: 122). In order to investigate the distribution of positive third person *do* in these few traditional English dialects, once again data from the Survey of English Dialects (SED) was examined (Orton, Sanderson and Widdowson 1978). Maps M34 and M35 dealing with *he does* show that positive third person singular *do* for traditional dialect speakers is mainly restricted to the area of East Anglia and the southwest of England. Indeed, in East Anglia the loss of inflection for third person *do* is part of the wider phenomenon of a general loss of third person singular *-s* for all verbs, as shown by other maps and also as investigated, for example, by Trudgill (1974). Trudgill shows that this phenomenon is a typical social marker for speakers in East Anglia, and in 1974 was still extremely widely used in the speech of the lower working class (with frequency indexes of nearly 100 per cent). In the southwest on the other hand, the loss of inflection seems restricted to *do* only.

If we compare this very restricted regional phenomenon with its negative counterpart, as, for example, in Maps M37 and M38 illustrating the use of *he doesn't* (Orton, Sanderson and Widdowson 1978), we see a very different picture. The main four forms, DOESN'T, DUSNO', DON'T and DUNNO can be broadly grouped according to presence vs absence of the third person *-s*. The first two groups of answers, DOESN'T and DUSNO', still contain the inflectional morpheme. These answers are dominant in the north, lying north of a very clear isogloss from the Wash, moving north of Derbyshire and across to the coast at Chester, as the reproduced composite Map 7.1 shows. Answers without the inflectional morpheme dominate south of this isogloss. However, if one investigates the original map more closely, one can see that this southern dialect area is not as homogeneous as its northern counterpart. Throughout the area, with the exception of the lower southwest (i.e. Dorset, Somerset, Devon and Cornwall), there are occurrences of the standard English *doesn't* forms spread across various counties. They do not form a homogeneous sub-area or 'island'; instead, this seems to be an instance of inherent variation throughout the region. One has to remember that the SED is based on questionnaires recording in most cases just one answer per question (per informant). Therefore, it is not possible to record intra-speaker variation for individual items.[1] However, what appears as inter-speaker variation in

3rd person *do* + *don't*

3rd person *don't*
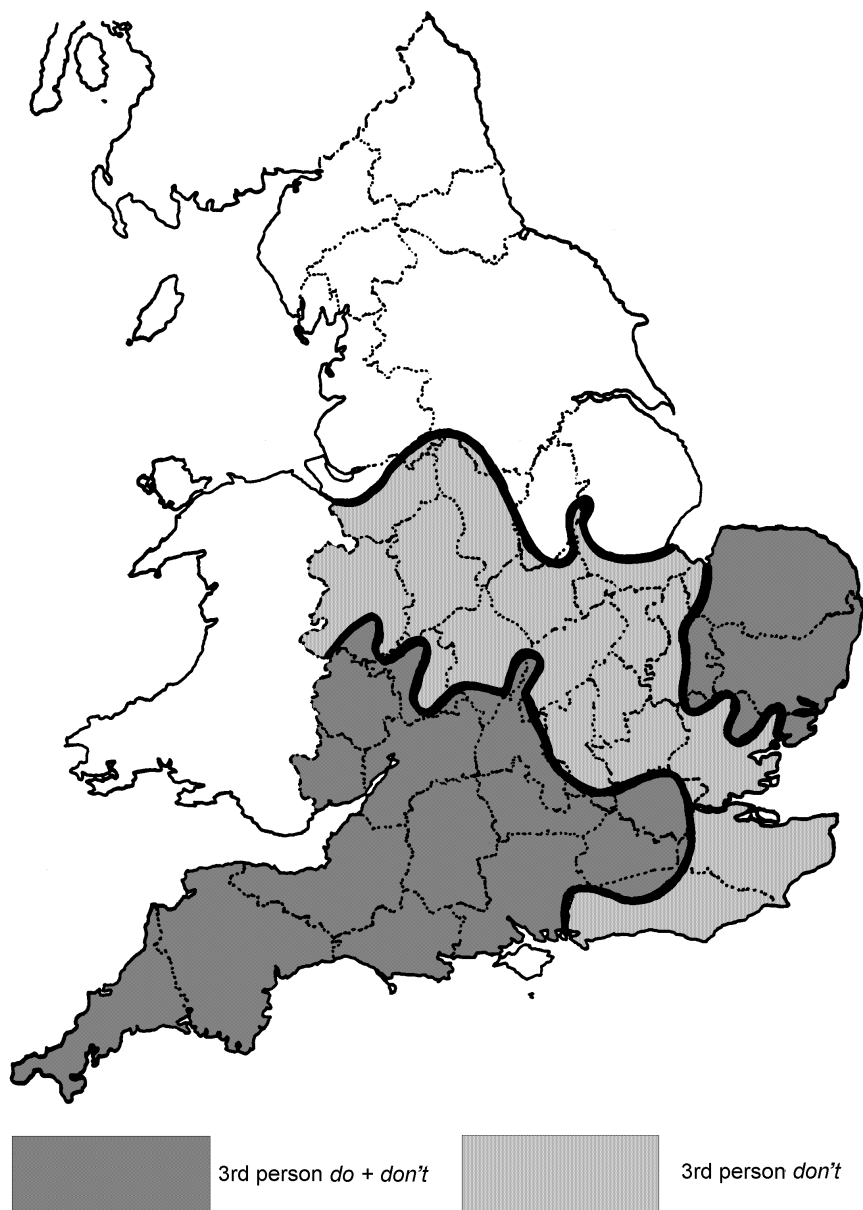
Map 7.1 3sg *do* and *don't* in the SED

Source: Adapted with permission from Orton, H., Sanderson, S. and Widdowson, J. (1978) *The Linguistic Atlas of England*, London: Croom Helm.

the case of this southern dialect area of *don't/doesn't* might in fact be due to several speakers using forms variably. No matter in which of these two ways the map is read, however, it is clear that south of the isogloss identified above, the simplified forms *don't* clearly dominate, although in a small minority of cases the standard English *doesn't* also occurs.

If one compares this result with the use of positive third person singular *do*, one can see that positive *do* is extremely restricted regionally, whereas *don't* is much more widely spread even for traditional dialect speakers: all areas that have positive *do* in the third person also have the negative form *don't*, but the reverse does not hold. Nevertheless, even this more widespread distribution is restricted regionally to the south of England and the Midlands. It is difficult to say whether East Anglia might have been the source for the diffusion of this form; Jespersen's proposed development (loss of the phoneme /z/ through phonological assimilation before /nt/) certainly offers a possible alternative path. With this geographical distribution in mind, however, it should be interesting to compare the traditional distribution with the situation in modern spoken English today.

Data from the BNC

Procedure

Like other high-frequency features of non-standard varieties, third person *don't* can be profitably investigated with the help of the BNC. Although high-frequency items in particular can lead to the most interesting and the most reliable results one can derive from this megacorpus, the high frequency itself presents some problems in handling the data on the interface level. Because it is simply impossible to investigate all occurrences of present tense DO, the searches again had to be limited in a structured way. In order to arrive at results comparable to results for the other phenomena, searches were again restricted to a combination of personal pronouns with *don't/doesn't*. Although full noun phrases were thus excluded from the investigation, a random sample of two times two hundred occurrences from spoken texts suggests that with an occurrence of about 2 per cent for *don't*, about 13 per cent for *doesn't* the number of full NPs is quite negligible in this context. Also in terms of frequent co-occurrences of lexical items, the combination of personal pronoun and *don't/doesn't* is of course of much greater interest, as full NPs are only very rarely repeated.

In order to arrive at comparable figures for the very different dialect areas, the usual frequency index was calculated. It indicates the percentage of actual occurrences in relation to all possible occurrences, i.e. to the sum of standard English *doesn't* plus non-standard *don't*. The results are displayed in Table 7.1.

Table 7.1 3sg *don't* in the BNC-SpS

BNC code	Dialect area	Total	3sg don't	% of total
XMS	South Midlands	51	33	64.7
XEA	East Anglia	183	113	61.7
XSS	Central southwest England	236	123	52.1
XMC	Central Midlands	162	76	46.9
XSL	Lower southwest England	79	34	43.0
XNO	Northern England	22	8	36.4
XNC	Central northern England	151	53	35.1
XLO	London	543	184	33.9
XMI	Midlands	70	22	31.4
XWA	Wales	180	45	25.0
XSU	Upper southwest England	32	7	21.9
XME	Northeast Midlands	111	24	21.6
XNE	Northeast England	96	18	18.8
XLC	Lancashire	163	25	15.3
XHC	Home Counties	294	43	14.6
XSD	Scotland	49	50	10.2
XMW	Northwest Midlands	272	27	9.9
XMD	Merseyside	56	5	8.9
XIR	Ireland	35	2	5.7
XHM	Humberside	210	0	0.0
Total		2,813	847	30.1

Table 7.1 gives a first overview of the phenomenon of *don't* in the third person. To facilitate a comparison across dialects, the rows have been arranged in order of (relative) frequencies of *don't* so that those dialect areas where *don't* is used most frequently appear at the top of the table. Table 7.1 already shows that third person *don't* is clearly present in all dialects;[2] what is more, we find a very high overall rate for this phenomenon (more than 30 per cent on average). Cheshire, Edwards and Whittle report auxiliary third-singular *do* as more frequent in the south (Cheshire, Edwards and Whittle 1993: 78), and indeed those areas in the BNC where third-singular *do* accounts for over 50 per cent of all forms are all situated in the south (East Anglia, the south Midlands and the central southwest). A possible north–south divide is of course what would be expected on the basis of the historical distribution of these two forms as discussed above. A closer look at Table 7.1 shows, however, that although the three dialect areas with the highest proportion of *don't* belong to the south, high frequencies can also be found in the Midlands (especially the central Midlands at 46.9 per cent), and the north (northern England with 36.4 per cent, and the central north with 35.1 per cent). On the other hand, not all areas in the south have a high proportion of third person *don't* (for example, in the Home Counties it occurs with a frequency of 'only' 14.6 per cent). To determine whether, overall, the impression of a south–north divide holds, the same procedure applied to neg concord will

Table 7.2 South–north grouping of 3sg *don't* in the BNC-SpS

	Total	*3sg* don't	*% of total*
North (8 areas)	782	161	Ø20.6
Midlands (5 areas)	795	194	Ø24.4
South (7 areas)	1,418	538	Ø37.9
Total	2,813	847	Ø30.1

be applied here, and we shall investigate a very broad regional grouping of dialect areas in relation to third person singular *don't*.[3]

Table 7.2 shows that the south does indeed behave significantly differently from the rest of the country; as expected, third person singular *don't* is used much more frequently here than in the Midlands or in the north (both differences are statistically highly significant at p < 0.01). The difference between the north and the Midlands is not statistically significant; we seem to have a relatively homogeneous Midlands/north area for this phenomenon. Nevertheless it has to be stressed that averages even in the north of over 20 per cent indicate a highly frequent phenomenon, and the next sections will shed some light on the structural properties of this pervasive feature.

Tags vs non-tags

As it has been suggested that third person *don't* is particularly a feature of the syntactic environment of tag questions, occurrences here have been investigated separately. Table 7.3 provides a first overview.

Table 7.3 shows that third person singular *don't* in tag questions is equally present practically everywhere, and that it is as frequent on average in tags as in general (cf. Table 7.1 above). It is of course possible that the subgroup of tags investigated in more detail here has influenced the general figures. We shall therefore compare *don't* in tags more carefully with their complementary environment of mainly declaratives (full interrogatives are also included here, but they are very rare overall), more precisely here called 'non-tags'.

Table 7.4 compares third person *don't* in tags with the complementary figures in non-tags, and the averages in the last column already indicate that although there is a slight increase in percentages from 29.8 per cent for non-tags to 30.8 per cent for tags, the difference between these two figures is only minimal, and statistical tests support this impression; these figures do not differ significantly.[4] For the moment, this is, however, only tested for the overall averages. It might be possible after all, that these averages cover up significant regional differences. In order to investigate this in more detail, the figures for all individual areas have been compared between non-tags and tags in Table 7.5.

Table 7.3 3sg *don't* in tag questions in the BNC-SpS

BNC code	Total tags	don't	% of total
XEA	52	34	65.4
XHC	79	17	21.5
XHM	9	0	0.0
XIR	3	0	0.0
XLC	46	7	15.2
XLO	118	23	19.5
XMC	51	16	31.4
XMD	11	2	18.2
XME	31	5	16.1
XMI	26	9	34.6
XMS	19	13	68.4
XMW	81	12	14.8
XNC	54	29	53.7
XNE	16	1	6.3
XNO	6	3	50.0
XSD	8		0.0
XSL	18	9	50.0
XSS	77	44	57.1
XSU	14	2	14.3
XWA	53	12	22.6
Total	772	238	30.8

Table 7.4 3sg *don't* in tags vs non-tags in the BNC-SpS

	Total	don't	% of total
Non-tags	2,041	609	29.8
Tags	772	238	30.8
Total	2,813	847	Ø30.1

For the statistical tests, seven areas had to be excluded because occurrences of *don't*, either overall or in tags, were below a threshold of five, and so a statistical analysis was not possible. These excluded areas are not marked in the last column of Table 7.5. They are Humberside, Ireland, Merseyside, the northeast, the north, Scotland and the upper southwest. Because of this relatively large portion of exclusions (35 per cent or over a third of all areas), the following results can at best be tentative and can only point towards certain trends. For the rest of the dialect areas, the following pattern emerges: for four areas, there is a significant difference between the use of *don't* in tags and the use of *don't* in declaratives. In two cases, *don't* is actually used significantly less frequently in tags than in non-tags, namely in London and the central Midlands. For the other two areas, *don't* in tags is used significantly more frequently than in declaratives; these

Table 7.5 *Don't* in tags vs non-tags in the BNC-SpS

BNC code	Total non-tags	don't	%	Total tags	don't	%	Statistical tests
XEA	131	79	60.3	52	34	65.4	non-tags = tags
XHC	215	26	12.1	79	17	21.5	non-tags < tags
XHM	19	0	0.0	9	0	0.0	
XIR	32	2	6.3	3	0	0.0	
XLC	117	18	15.4	46	7	15.2	non-tags = tags
XLO	425	161	37.9	118	23	19.5	non-tags > tags
XMC	111	60	54.1	51	16	31.4	non-tags > tags
XMD	45	3	6.7	11	2	18.2	
XME	80	19	23.8	31	5	16.1	non-tags = tags
XMI	44	13	29.5	26	9	34.6	non-tags = tags
XMS	32	20	62.5	19	13	68.4	non-tags = tags
XMW	191	15	7.9	81	12	14.8	non-tags = tags
XNC	97	24	24.7	54	29	53.7	non-tags < tags
XNE	80	17	21.3	16	1	6.3	
XNO	16	5	31.3	6	3	50.0	
XSD	41	5	12.2	8	0	0.0	
XSL	61	25	41.0	18	9	50.0	non-tags = tags
XSS	159	79	49.7	77	44	57.1	non-tags = tags
XSU	18	5	27.8	14	2	14.3	
XWA	127	33	26.0	53	12	22.6	non-tags = tags
Total	2,041	609	29.8	772	238	30.8	non-tags = tags

are the Home Counties and the central north. The remaining nine areas behave in the same way as the overall averages: there are no significant differences in the occurrence of *don't* in tags or outside of tags. These nine areas are East Anglia, Lancashire, the east Midlands, the Midlands, the south Midlands, the northwest Midlands, the lower southwest and Wales.

No clear regional distribution emerges for these two types. Those areas where *don't* is generalized less frequently in tags than outside of tags seem to concentrate in the middle, with London and the central Midlands. The reverse phenomenon, however, is found in areas as far away as the Home Counties and the central north. However, the areas where generalized *don't* occurs at more or less the same rate in all grammatical environments (tags and non-tags) seem to concentrate in the Midlands, reaching down as far as East Anglia and the lower southwest, and also including Wales, as Map 7.2 shows.

The large area of undifferentiated *don't* corresponds in a rather striking way to the traditional *don't* area as displayed in Map 7.1. If we take a look at those areas which had to be excluded from statistical tests, we can see that only in the upper southwest, the figures also seem to indicate a decrease from *don't* in non-tags to *don't* in tags. No difference at all between these two grammatical environments, however, seems to be the rule, and most areas excluded from the tests also seem to follow this pattern, as far as
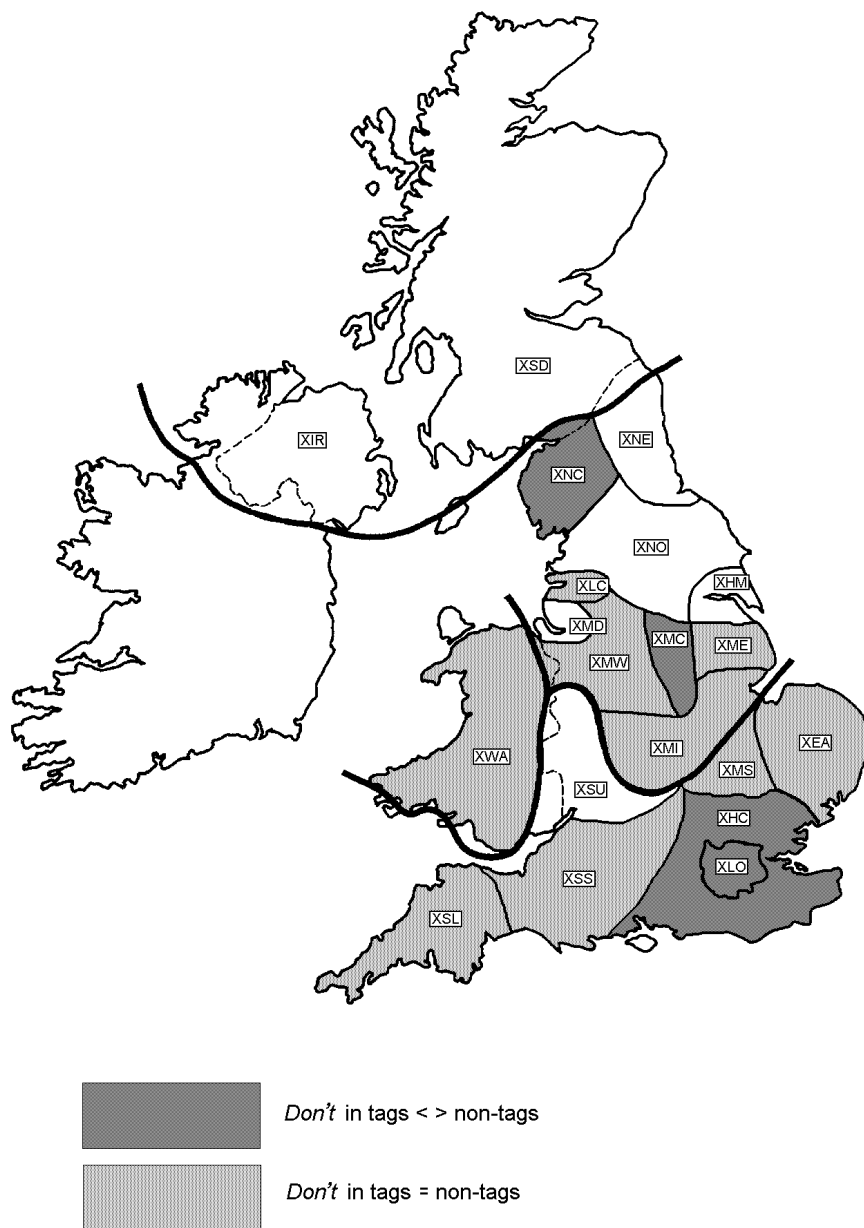
Map 7.2 3sg *don't* in tags vs declaratives in the BNC

one can judge on the basis of these very low occurrences. The dominant type in most areas is therefore an undifferentiated *don't* which does not behave significantly differently in tags or non-tags. What we do not find is a (possible) functional differentiation for DO in the majority of dialect areas today.

Subjects and sentence types

An investigation into the different kind of personal pronouns that function as antecedents of third person *don't* might tell us something else about this phenomenon. As the pronouns occur with very different absolute frequencies (*it* with a form of negated DO is roughly as frequent as *he* and *she* taken together), it might be expected that the use of invariant *don't* varies as well. Table 7.5 gives the figures for all dialect areas together (with the exception of Humberside, as there are no occurrences of third person *don't*)[5] divided for the type of third person singular subject pronoun (*he*, *she* or *it*). Table 7.6 also differentiates syntactic environments into declaratives, interrogatives and tag questions, which incidentally validates the approach taken above of grouping declaratives and interrogatives together as non-tags: full interrogatives are very rare – the eighty-seven occurrences only amount to 3 per cent of all cases – and they do not pattern significantly differently from the much more usual declaratives.

   The very similar percentages in the 'Total' row already make it clear that there are no significant differences between *don't* used in declaratives, interrogatives or tags; they roughly mirror the less detailed subdivisions (of tags and non-tags) in Table 7.4. If we look at the different kinds of subjects and put them in order of increasing absolute frequency (*it* > *he* > *she*), we see that this does not correlate with the relative frequencies of third person *don't*. Although *it* by itself is the most frequent pronoun overall, the combination *it don't* (or *don't it* for interrogatives and tags) is the least frequently used form. Relative and absolute frequencies are obviously not in a relation of proportion. However, the reverse is not the case either: the relative frequencies cannot be ordered in a scale *she* > *he* > *it* (this would be the exact reverse of the absolute frequencies). If we look more closely at the difference between *he* and *she* in the different syntactic environments, we can see that in only one case is there a significant difference (namely for declaratives). In the two other environments, both interrogatives as well as tag questions, the difference between *he* and *she* occurring with *don't* is statistically not significant. The difference between the use of *don't* with *it*, and with either *he* or *she*, on the other hand, is significant in both declaratives and tag questions. Although there is no significant difference found with interrogatives, occurrences are so low in this environment that they do not influence the overall figures, which therefore also indicate this significant difference between *he* or *she* on the one hand and *it* on the other hand. If we group *he* and *she* together as personal antecedents,

Table 7.6 3sg *don't* by subject and sentence type in the BNC-SpS

	Total	%	Decl.	%	Tag	%	Int.	%
it	334/1,379	24.2	201/870	23.1	127/485	26.2	6/24	25.0
he	314/813	38.6	248/628	39.5	57/158	36.1	9/29	31.0
she	199/593	33.6	135/429	31.7	54/120	45.0	10/34	29.4
Total	847/2,785	30.4	584/1,927	30.3	238/763	31.2	25/87	28.7

Table 7.7 3sg *don't* by subject per dialect area in the BNC-SpS

BNC code	Total	it	he	she
XEA	113/183	36/79	52/69	25/35
XHC	43/294	15/144	20/77	8/73
XIR	2/35	0/17	1/14	1/4
XLC	25/163	9/86	10/41	6/36
XLO	184/543	68/249	70/166	46/128
XMC	76/162	45/100	20/37	11/25
XMD	5/56	3/31	2/16	0/9
XME	24/111	9/60	10/31	5/20
XMI	22/70	11/39	9/21	2/10
XMS	33/51	8/20	18/20	7/11
XMW	27/272	9/133	10/96	8/43
XNC	53/151	28/80	14/36	11/35
XNE	18/96	10/64	4/18	4/14
XNO	8/22	3/7	5/11	0/4
XSD	5/49	4/30	1/16	0/3
XSL	34/79	9/31	16/27	9/21
XSS	123/236	50/113	32/57	41/66
XSU	7/32	3/14	4/10	0/8
XWA	45/180	14/82	16/50	15/48
Total	847/2,785	334/1,379	314/813	199/593

there emerges a clear statistical difference compared to the use of *don't* with *it*. *Don't* tends to be used with personal antecedents (*he*, *she*) rather than impersonal *it*, but does not differentiate between tag questions and declaratives in terms of relative frequency.

Although *do* is semantically bleached, in that its meaning is no longer restricted to animate agents, the non-standard form *don't* seems to mirror the very old semantic heritage of *do*, as the differences in distribution show. On the other hand, even for *it* rates of non-standard occurrences are very high (at 24.2 per cent on average), so that the type of subject does not seem to be a very important factor for this phenomenon. Nevertheless, we shall investigate the different subject areas for the different subjects as well as sentence types.

When the figures for the individual dialect areas are subdivided for the type of subject and for the sentence type respectively, they will become

very small in most cases; a quantitative approach is therefore not possible any longer. Because of this, the following section will simply look at overall occurrence vs non-occurrence, in order to find a system of 'gaps'. The question is: if a form stops occurring, which ones fall out of the system first? Table 7.7 and Table 7.8 therefore indicate the use of third person *don't* subdivided by subject type and by sentence type respectively. The columns are ordered in terms of increasing absolute frequencies; non-occurrences are marked as grey cells.

Table 7.7 shows that in all dialect areas that have third person *don't*, *don't* occurs with *he*. Although *he* is not the most frequent personal pronoun, this is clearly the preferred combination in structural terms. As we have seen above in Table 7.6, this is also mirrored by the fact that this is the most frequent combination (in relative frequency). Of the other two pronouns, only one dialect area (Ireland) does not have the form *it don't* (or: *don't it*), but in four, *don't* does not occur with *she*. This is then clearly the option that is realized most rarely. As *she* is the least frequent personal pronoun of the three pronouns investigated (in absolute terms), it is not surprising that it is the first pronoun to have gaps in this matrix. There is no obvious hierarchical relation between *she* and *it*, however (in Ireland, where *it don't* does not occur, *she don't* occurs nevertheless). This might be due to the fact that Ireland shows a highly irregular system, a system different from the rest, or – more likely – that figures are simply too small to be relied on; after all, Ireland has just two occurrences of third person *don't*, and the distribution might simply be due to chance. A similar point can be made for Scotland, where the difference between no occurrences for *she* and just one occurrence for *he* can also not be the basis for positing a principled difference. It is well known for other phenomena that Scotland and Ireland often pattern differently, so that a situation where these two dialect areas would completely fall out of the system would therefore not be very surprising. As Chapter 3 has already indicated, the one other English dialect area with a purported Celtic substrate, Wales, generally does not pattern with its northern Celtic sister dialects, and this is clearly also the case for this phenomenon, where Wales patterns with the rest of the English dialects, as shown by the distribution in Table 7.7.

We can summarize so far that if a dialect has *she don't*, it is very likely that it will also have *he don't*, as shown in (7).

(7) Implicational tendency for third person *don't*:
       she don't ⊃ *he don't*

This is equivalent to the hierarchy in (8).

(8) Hierarchy for third person *don't*:
       he don't > *she don't*

Table 7.8 3sg *don't* by sentence type per dialect area in the BNC-SpS

BNC code	Total	Decl.	Tag	Int.
XEA	113/183	76/125	34/52	3/6
XHC	43/294	26/199	17/79	0/16
XIR	2/35	2/29	0/3	0/3
XLC	25/163	16/115	7/46	2/2
XLO	184/543	157/409	23/118	4/16
XMC	76/162	55/106	16/51	5/5
XMD	5/56	3/42	2/11	0/3
XME	24/111	14/73	5/31	5/7
XMI	22/70	13/43	9/26	0/1
XMS	33/51	20/32	13/19	—
XMW	27/272	15/190	12/81	0/3
XNC	53/151	24/95	29/54	0/2
XNE	18/96	16/75	1/16	1/5
XNO	8/22	5/15	3/6	0/1
XSD	5/49	5/40	0/8	0/1
XSL	34/79	25/61	9/18	—
XSS	123/236	76/138	44/77	3/11
XSU	7/32	5/18	2/14	—
XWA	45/180	31/122	12/53	2/5
Total	847/2,785	548/1,927	238/763	25/87

How *it don't* fits into this system is not very clear, as all dialect areas (except Ireland) have both *it don't* and *he don't*. This does not contradict a possible scale *it > he > she* – an order that would suggest itself from the absolute frequency distribution – but does not offer very convincing support for it either. In general, however, the difference in pronoun subjects does not seem to be a major factor for determining the distribution of third person *don't*. Let us now look at the different syntactic environments in the same manner, as detailed in Table 7.8.

Again in Table 7.8, all non-occurrences are indicated by a darker background for easier reference. The mark '—' indicates that the syntactic environment in question does not occur at all. These 'gaps' are therefore neither an argument for nor against the hierarchy proposed below. Compared with Table 7.7, *don't* does not occur in many more cases. In two areas, third person singular *don't* does not occur in tag questions. But in a total of eight areas, *don't* does not occur in interrogatives (other than tag questions). Here, a non-occurrence for tag questions coincides with a non-occurrence in interrogatives in both cases (Ireland and Scotland). The hierarchy for sentence types is displayed in (9).

(9) Hierarchy for third person *don't*:
     don't in declaratives > *don't* in tag questions > *don't* in full
     interrogatives

Again, one can state the same relations as an implicational tendency:

(10)   Implicational tendency for third person *don't*:
        interrogative *don't* ⊃ tag *don't* ⊃ declarative *don't*

In other words: if a dialect has third person singular *don't* in an interrogative, it is very likely to have *don't* in tags as well; and if a dialect has *don't* in tags, it will have third person *don't* in a declarative environment, too. If we combine the two Tables 7.7 and 7.8, only indicating occurrence (no mark) or non-occurrence (marked by 0)[6] of third person singular *don't*, we get an ordering of combined subcategories that is displayed in Table 7.9.

   The two final columns indicate the number of structures with occurrence (the first figure) and non-occurrence (the second figure) of third person singular *don't* for each dialect area. The dialect areas have been ordered according to increasing figures of occurrence of *don't*. Where *don't* occurs in the same number of possible positions for more than one dialect area, this subgroup is internally ordered in decreasing frequency of *non*-occurrence. In this way, the dialect areas with the most 'gaps' in the system occur at the top of the table, those with no gaps at all at the very bottom. The columns

Table 7.9 Non-occurrence of 3sg *don't* by sentence type, subject and dialect area in the BNC-SpS

Subject	he	it	she	she	she	he	it	it	he	Total occ.	Non-occ.
Sentence	*Int.*	*Int.*	*Int.*	*Tag*	*Decl.*	*Tag*	*Tag*	*Decl.*	*Decl.*		
XIR	0	0	0	—		—	0	0		2	5
XMD	0	0	0	0	0	0				3	6
XSD	0	—	—	0	0	0				3	4
XNO	0	—	0	0	0					4	4
XSU	—	—	—	0	0					4	2
XMI	0	—	—	0						5	2
XNE	0		—	0		—				5	2
XMW	0	—	0							6	2
XHC	—	0	0							6	2
XNC	0	—	—							6	1
XMS	—	—	—							6	
XSL	—	—	—							6	
XEA		0	0							7	2
XWA	0	—								7	1
XME				0						8	1
XSS		0								8	1
XLC	0									8	1
XLO	0									8	1
XMC										9	
Total occ.	4	5	6	11	15	15	18	18	19		
Non-occ.	11	5	6	8	4	2	1	1			

have been ordered in terms of increasing frequency of occurrence from left to right. Again, if several positions were filled in the same number of dialect areas, the internal structuring is in order of decreasing frequency of non-occurrence. In other words, those pronouns and syntactic structures least often occurring with *don't* appear at the left, those with the most frequent occurrences of *don't* at the right-hand side of the table.

This ordering results in a surprisingly homogeneous matrix. There are only five dialect areas that behave slightly irregularly (with positive occurrences to the left of non-occurrences), but with the exception of the northeast Midlands, these irregular occurrences are restricted to a single position. If we look at the ordering of pronouns and syntactic environments in the very first row that has resulted in this homogeneity, several features spring to mind: for all personal pronouns, the interrogative environment has the most non-occurrences. There is a big gap between only six dialect areas using *don't* in interrogatives with *she*, and the neighbouring column, with eleven dialect areas using *don't* in tag questions with *she*. Thus, all interrogative environments cluster at the left of the table. Declaratives and tags on the other hand do not show the same kind of overall clustering. However, for each of the three pronouns, *don't* occurs in many more slots for declaratives than for the respective tags. For each pronoun, then, the hierarchy posited in (9) (declarative > tag > interrogative) is mirrored in the matrix.

If we look at the ordering of pronouns inside the various syntactic environments, however, no consistent order emerges. For interrogatives, we find the order *he* < *it* < *she*; for tag questions, the order *she* < *he* < *it*; and for declaratives, *she* < *it* < *he*. This matrix indicates, in other words, that the overriding constraint on the use of *don't* with third person singular subjects is the syntactic environment, not the personal pronoun of the subject. This order of syntactic environments exactly corresponds to the absolute frequencies, with interrogatives being the rarest structure, and declaratives the most frequent. Although the relative frequencies of the use of *don't* in these various structures is very consistently the same at around 30 per cent, as we have seen in Table 7.6, the structural diversity is highly differentiated, as Table 7.9 has shown, and it corresponds closely to the absolute frequencies: the rarer a syntactic environment (e.g. interrogatives), the fewer structures have third person singular *don't*. Conversely, the more frequent a syntactic environment (e.g. declaratives), the more structures have third person singular *don't*.

This is then basically a phenomenon that is in a proportional relation: the more frequent a syntactic environment is (in absolute terms), the more diverse is its internal structure in terms of 'possible slots' for the use of *don't* instead of *doesn't*. And this distribution can perhaps answer our initial question of whether *don't* originated through analogy, spreading from all other persons to the third person singular, or whether *don't*, as Jespersen argued, arose through /z/-deletion.

A direct frequency relation has been recognized as an indication for the spread of a lexeme by phonetic simplification – a change that is basically phonetically motivated (cf. Hooper 1976). Other well-known phenomena that can be quoted as instances of this proportional frequency relation are contracted forms: the more frequent a word or a combination of words, the shorter they will get (cf. Krug 1998 for the concept of *string frequency*; and this also conforms to Zipf's law, cf. Zipf 1949: 20ff.).

The competing analysis, a spread of *don't* from all other grammatical environments to the third person, would be a spread based on analogy. Again, this phenomenon can of course be frequently observed in language change. An example would be the regularization of irregular past tense forms in analogy with the more regular paradigms. In contrast to the first hypothesis of phonological shortening, this kind of change is cognitively motivated. As Hooper (1976) has shown, however, this kind of change behaves in the opposite way from shortening processes: the *less* frequent a phenomenon is in absolute terms, the more often it will undergo a regularization process in relative terms. Thus, very infrequent forms are typically affected first by regularization processes; conversely, highly frequent forms can preserve the most irregularities (see, for example, the present and past tense paradigms of BE). In principle, it should therefore be possible to distinguish between these two types of change simply by their frequency distributions. Idealized diagrams are provided below and they indicate again that although both procedures are frequency related, absolute frequencies play a diametrically opposed role: for shortening processes, the higher the absolute frequency, the higher the relative frequency. For regularization processes, however, the higher the absolute frequency, the lower the relative frequency, as displayed in Figure 7.1.

As already indicated, we can use these two different sets of distributional relations as an analytical tool to determine the nature of the process of third person *don't*. We have found a proportional distribution that is typically characteristic of shortening phenomena. This distribution supports
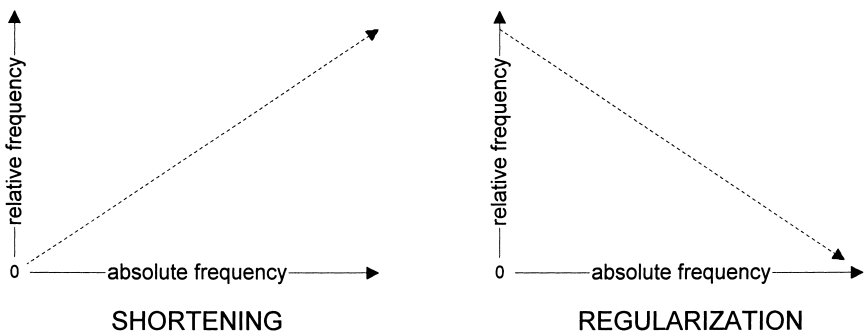
Figure 7.1 Idealized frequency distributions
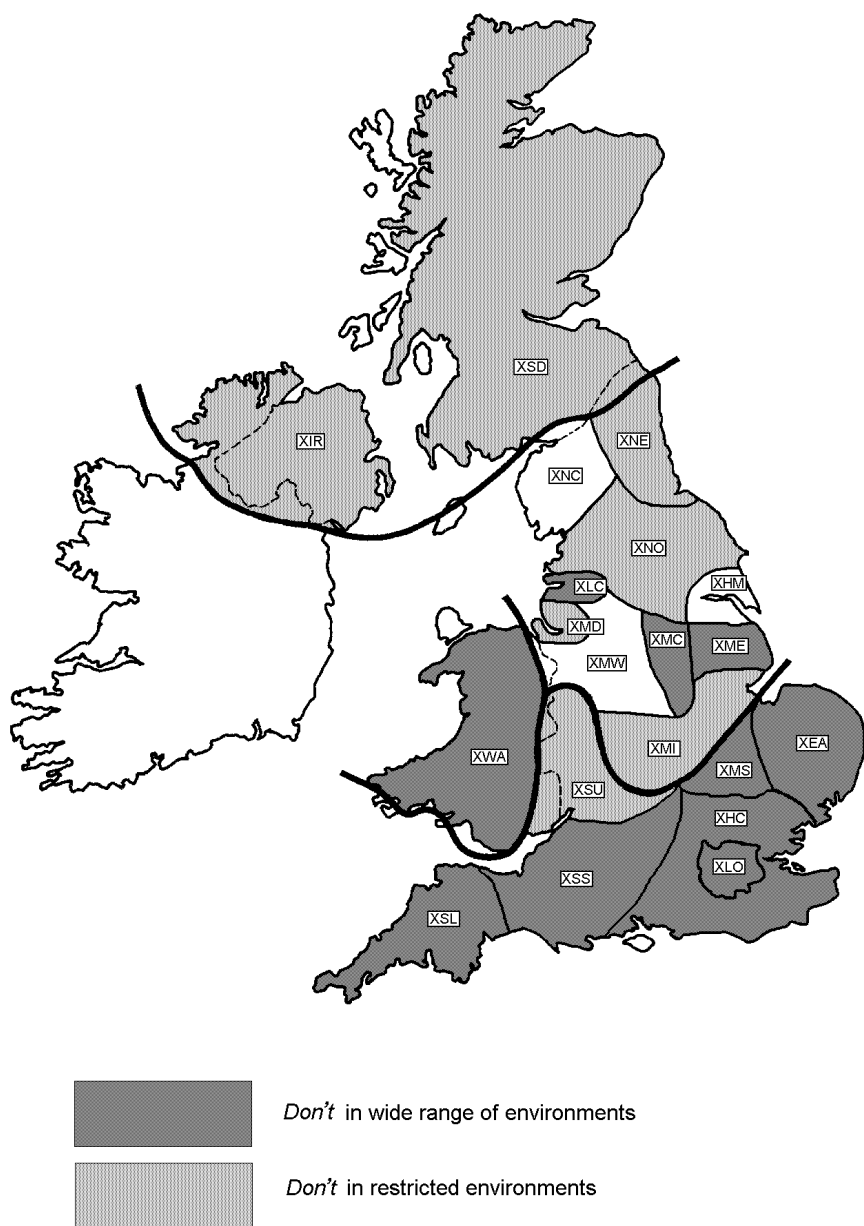
Map 7.3 3sg *don't* in the BNC

Source: Adapted with permission from Trudgill, P. (1990) *The Dialects of England*, London:
Edward Arnold.

the hypothesis of a phonetically motivated process. We should therefore regard third person *don't* as a phenomenon of coalescence or /z/-deletion before the contracted negator {nt} as proposed by Jespersen. Third person *don't* therefore does not seem to be primarily the result of a regularization strategy, spreading from all other persons to the third person singular in analogy, as we have seen that infrequent environments are by no means affected first.

If we now consider the distribution of dialect areas over the matrix in Table 7.8, it is clear that northern areas and in particular the (northern) Celtic Englishes cluster at the top of the table (e.g. Ireland, Merseyside, Scotland, the north and the northeast), i.e. here, *don't* occurs in only very restricted environments. Southern and Midlands dialect areas are found towards the bottom of the table, in particular the Home Counties, the south Midlands, the lower southwest, East Anglia, Wales, the northeast Midlands, the central southwest, London and the central Midlands. Here, *don't* occurs in a much wider range of environments. Of course there are some exceptions, notably the upper southwest and the Midlands among the more northern group, whereas Lancashire patterns with the more southern ones, but the overall distribution of filled slots, in addition to the frequency approach chosen above, supports the impression that the occurrence of third person singular *don't* still shows a regional distribution along north–south lines. This distribution is displayed in Map 7.3.

Summary

Jespersen already notes a parallel between third person singular *don't* and the paradigm of *ain't*: 'as with *ain't*, the distinction of person and number has been obliterated in the negative forms' (Jespersen 1940: 435), and data from the BNC have shown that we find *don't* used for all persons practically everywhere throughout Great Britain today. Positive *do* on the other hand is regionally extremely restricted, if not disappearing altogether. No matter what the origin of this phenomenon is, that is, whether *don't* has arisen out of loss of the phoneme /z/ before the negative clitic {nt} – as proposed by Jespersen and as our analysis has confirmed to a degree – or whether it is the result of a transfer from the plural form, again we have a system of asymmetries which has its exact parallel in the paradigm for AIN'T (cf. Chapter 6). The asymmetries in the paradigm for non-standard present tense DO are displayed in Figure 7.2.

Another tendency can be observed that results from this process. With the obliteration of the third person singular *doesn't*, *don't* is becoming a kind of invariant neg marker for the present tense for all persons. As we have seen in Chapter 2, DO-support is of course only used for those sentences where no other auxiliary is present in the positive form. It has already been pointed out in Chapter 2 that the English system of negation is extremely rare cross-linguistically: an obligatory auxiliary has to be inserted which is

$$\left. \begin{array}{l} \textit{do} \\[2.5em] \textit{does} \end{array} \right\} \quad \textit{don't}$$
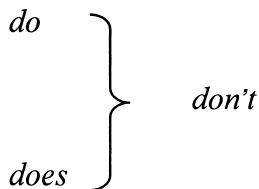
Figure 7.2 Paradigm for DO

then negated by a negative particle. As has also been shown, English in this respect resembles most closely some non-Indo-European languages which employ a proper negative auxiliary, like Finnish or Estonian. As Payne (1985) has pointed out, in languages with negative auxiliaries, there is a strong diachronic tendency to generalize one form, usually the most frequent third person singular, to all other persons and thus to achieve a kind of invariant neg marker rather than an inflected negative auxiliary. The Estonian negative marker *ei*, for example, 'is an invariant particle . . . which derives historically from the third person singular of the present form of the negative verb' (Payne 1985: 215). Although negative verbs are more common than the exceptional structure of obligatory auxiliaries that English uses, they are still relatively rare in the world's languages. Much more common typologically is an invariant neg marker (they occur in 45 per cent of languages investigated by Dahl 1979); this strategy for negation ranges at the very top together with morphological negation (as in Turkish), at ca. 41 per cent. Proper negative auxiliaries on the other hand are only employed by about 16 per cent of the languages investigated by Dahl (1979). An invariant neg marker is thus a much more natural 'strategy' for negation cross-linguistically than negative verbs, and it is therefore not surprising that languages with negative verbs should show a trend to simplify this relatively complicated system by substituting it with just one invariant form. It seems to be due to phonological reasons, as pointed out by Jespersen, that in the case of English, the form chosen for this invariant neg marker is not the very frequent third person singular form *doesn't*, as one might expect from the cross-linguistic data, but the unmarked form *don't*, which has arisen out of *doesn't* through coalescence. On the other hand, if *don't* is evolving into a general negation marker, it certainly has a very long way to go. Its main contender, *ain't*, is alive and well, and the remaining auxiliaries also show no sign of being replaced by *don't*. As we have seen, the reverse seems to be the case: there are instances of AIN'T used for DO, but the reverse is certainly not the case. If *don't* should really be on the way to becoming an invariant neg marker, this development can therefore only be described as being in its very initial stages.

8 Past tense BE

Introduction

The regularization of the past tense BE paradigm is a phenomenon that has long been neglected by dialectologists and sociolinguists, although it is extremely widespread as well as frequently used. Recently this feature has commanded the attention of several renowned linguists across the English-speaking world, and yet comparative studies on a wider scale are still missing. In particular, this chapter deals with the use of *was* for *were*, as well as the reverse use of *were* for *was*, for regularization occurs both ways.

(1) Did she see what *you was* doing? (KBK 4110)

(2) So you were lucky to be in today really *wasn't you*? (KDW 1781)

(3) *She were* in earlier but *she weren't* feeling very well. (KCX 3306)

(4) *It were* Julie's birthday yesterday. (KB1 3970)

In contrast to preceding chapters, the positive forms are also included in this investigation, as they form interesting patterns with their negative counterparts. We shall look at the situation in historical dialects, and then proceed to present day data from the BNC. A functional explanation shall be advanced for the pervasiveness of this phenomenon.

The phenomenon

The past tense paradigm for BE is the only irregular paradigm in the whole past tense system in standard English. There are two forms (in contrast to just one form for all other verbs), *was* and *were*, which are distributed as follows: *I*, *he*, *she*, *it* and all singular NPs take *was*; *we*, *you*, *they* and all plural NPs take *were*. Singular and plural meanings of *you* are not distinguished in this respect. Indeed, it can be argued that *you* as the

historical plural form, which was generalized to the singular with the demise of *thou/thee*, still behaves morphologically like a plural form. The *was/were* distinction in standard English can therefore be regarded as a distinction according to number (singular vs plural forms). Exceptions to this rule of grammatical concord are collective noun phrases, which are formally singular but denote a group of people, for example *the police* or *the government*. With these cognitively 'plural' nouns, continuation with plural pronouns becomes increasingly possible, and these collective NPs are themselves regularly used with plural verb forms in standard British English according to notional, rather than grammatical, concord (cf. Quirk *et al.* 1985: 758f.): *The audience were enjoying every minute of it.*

The negative of *was/were* is formed quite regularly by adding the sentence negator *not* or by cliticization, resulting in *wasn't/weren't*. Non-negative contraction is not possible for *was/were*, as shown in Chapter 2. The past tense BE system for standard English can be displayed as in Figure 8.1. The single line indicates morphemic difference (as the first and third persons singular clearly employ a different morpheme from the remaining persons).

As this is the only paradigm that is different from all other past tense paradigms, it is a possible candidate for regularization in non-standard speech; all other verbs, auxiliary or main ones, have only one past tense form even in the standard (e.g. *has/have – had*; *finds/find – found*), and a distinction is clearly not perceived as cognitively necessary.

Surprisingly, the regularization of past tense BE used to receive relatively little attention even among scrupulous charters of dialect differences. Even in 1990, Trudgill still notes in all of five lines that:

> the verb <u>to be</u> in English is highly irregular, and most dialects have forms which differ from Standard English at least in some respects. Some, for instance, have <u>was</u> throughout the past tense – <u>you was</u>, <u>we was</u>, <u>they was</u> – while others have generalized <u>were</u> – <u>I were</u>, <u>she were</u>. Yet others distinguish between positive <u>was</u> and negative <u>were</u>:
>> He was there, weren't he?
>> You was there, weren't you?
>
> (Trudgill 1990: 98)

He laconically continues that 'even more interesting is the amount of variation found in the present tense' (Trudgill 1990: 98), and he gives no indication as to whether there might be a regional distribution. Cheshire (1982) goes into a little more detail, although she also only notes exemplary

was : wasn't
———————
were : weren't

Figure 8.1 Past tense BE (standard English)

dialect regions: 'in Bradford, for example, the form *were* occurs, with all persons . . . In Reading English the form *was* is used not only with first and third person subjects . . . but also with other subjects' (Cheshire 1982: 44). She gives the relative frequencies of non-standard *was* and *were* for her sample of adolescent speech in Reading; however, as she does not supply any absolute figures, it is very difficult to interpret her results. Nevertheless, Cheshire already notes that 'the occurrence of the negative particle *not* or *-n't* has an interesting effect on the non-standard forms' (Cheshire 1982: 45).

Recently, the phenomenon of *was/were*-variation has commanded considerably more attention, especially among US scholars. Tagliamonte has studied this phenomenon in a range of US and British dialects (cf. Tagliamonte 1998; Tagliamonte and Smith 1999), and Wolfram (with various co-authors) has concentrated on isolated and native American communities (cf. Christian, Wolfram and Dube 1988; Wolfram and Sellers 1999; Wolfram and Dannenberg 1999). Exemplary, is the study by Wolfram and Schilling-Estes (1996), which will be the basis for much of the following discussion. Wolfram and Schilling-Estes apply the relatively scarce information on British dialects to their analysis of the dialect of Ocracoke, North Carolina. Ocracoke is characterized as a 'post-insular island' in the Outer Banks in the southern United States (for details cf. also Wolfram 1996, Wolfram, Cheek and Hammond 1996 and especially Wolfram, Hazen and Schilling-Estes 1999) – the isolation of this island community has only recently ended, and increasing contact with the mainland has led to an interesting mix in the dialect system. The very distinctive historical Ocracoke 'brogue' is dying out, but instead of straightforward assimilation to the mainland southern US dialect, a new mixed dialect is emerging. One of the distinctive features described by Wolfram and Schilling-Estes of the historical dialect is the generalization of *were* (in both positive and negative contexts) for personal pronouns.[1] Mainland southern states American English on the other hand is characterized by an overall generalization of *was*. In this situation of dialect contact, a new type seems to be emerging, in parallel to Cheshire's and, in part, Trudgill's observations: on Ocracoke there is a tendency to generalize *was* to all persons in positive contexts, and to generalize *weren't* in negative contexts.

Data from the BNC

Procedure

Past tense BE is another highly frequent phenomenon where an investigation with the help of the BNC promises interesting results. The little interest irregular past tense forms of BE have found in the academic literature so far might be an indication that this feature is not as stigmatized as, for example, the use of *ain't*.[2] In order to make the huge amounts of data more manageable, and to obtain results that are comparable to the

Table 8.1 Was with plural subjects in the BNC-SpS

BNC code	Dialect area	Total plural	was	% of total
XEA	East Anglia	214	88	41.1
XHC	Home Counties	522	35	6.7
XHM	Humberside	76	2	2.6
XIR	Ireland	97	7	7.2
XLC	Lancashire	276	21	7.6
XLO	London	874	88	10.1
XMC	Central Midlands	358	37	10.3
XMD	Merseyside	46	2	4.3
XME	Northeast Midlands	230	21	9.1
XMI	Midlands	77	6	7.8
XMS	South Midlands	134	7	5.2
XMW	Northwest Midlands	421	36	8.6
XNC	Central northern England	337	32	9.5
XNE	Northeast England	158	19	12.0
XNO	Northern England	75	13	17.3
XSD	Scotland	231	35	15.2
XSL	Lower southwest England	106	8	7.5
XSS	Central southwest England	282	68	24.1
XSU	Upper southwest England	55	13	23.6
XWA	Wales	354	64	18.1
Total		4,923	602	∅12.2

investigations of third person *don't* and of *ain't* in the preceding chapters, searches were restricted to combinations of a personal pronoun or existential *there* with *was/wasn't/were/weren't* in the SpS subsample of the BNC per dialect region. Full NPs will therefore have to be neglected in this investigation.[3] Occurrences of hypothetical *was/were* (and, where it occurred, *wasn't/weren't*) – defined as occurrence after *if* – were excluded. Also excluded from the following counts was the set phrase *as it were*.

Besides looking for pronoun + verb combinations, searches for the inverted word order were also conducted to find the tag questions for a comparison of non-tags and tags. In order to find a measurement for non-standard generalization, the percentage (for every dialect area) of actual generalized forms was calculated in relation to all possible forms. In practice, this means – for the singular pronouns – adding all instances of generalized *were* forms (*I were, he were*, etc.) to the corresponding StE forms *I was, he was*, etc., and the reverse for *was*-generalization for the plural forms. We shall look at each verb form individually, starting with *was* used with plural subjects.

Table 8.1 shows that in positive clauses, generalized *was* is used in every dialect region, albeit at different ratios. The distribution ranges from 2.6 per cent (in Humberside) to a maximum of 41.1 per cent in East Anglia. The average at 12.2 per cent is not notably higher than for some of the other non-standard phenomena, notably negative concord and *ain't*, as we

Table 8.2 *Was* with plural subjects in tag questions in the
BNC-SpS

BNC code	Total tags	was tags	% of total
XHC	7	1	14.3
XLO	12	1	8.3
XME	14	3	21.4
XMW	3	1	33.3
XSS	4	1	25.0
XSU	2	1	50.0
Total	42	8	Ø19.0

have seen in the preceding chapters. Generalization to *was* in tags also
occurs, but this seems to be a minority option. We find generalized *was*
in tag questions in only six dialect areas (out of twenty), or in under one-
third of all cases, as Table 8.2 shows.

Although in those dialect areas where *was*-generalization does occur, it
is slightly more frequent than in non-tags (the only exception being London
with 8.3 per cent in tags, compared to 10.1 per cent overall), it must be
remembered that in all areas with the exception of the northeast Midlands,
was in tag questions occurs only once, and percentages are therefore at
best misleading. Nevertheless, it is interesting to see that all dialect areas
where the use of non-standard *was* is also extended to tag questions are
situated in the south and the Midlands. However, the table also shows
that positive tags in general are relatively rare. According to standard
grammar, as reverse polarity tags they occur only after negative main
clauses. As we have already seen, negative main clauses are in general
much rarer than their positive counterparts. In addition, only a small
minority of all main clauses carry a question tag, so that these grammat-
ical constraints explain the relative overall infrequency of positive tag
questions, as well as the overall infrequency of *was*-generalization in this
environment.

Table 8.3 shows that generalized *were* in positive sentences occurs signif-
icantly less frequently – only about half as frequently – than its plural
counterpart *was*: on average in only 6.7 per cent of all possible occur-
rences, whereas we found 12.2 per cent for *was*. The regional distribution
is more uneven, and there are two dialect areas (the small Merseyside and,
not surprisingly, Humberside) where this non-standard form is not found
at all. Again, generalization does occur in tag questions as well, as Table
8.4 shows.

Again, only six dialect areas out of twenty have non-standard *were* in tag
questions, but this time these dialect areas are concentrated in the Midlands,
with extensions to London and to the north. Whereas *was*-generalization
seemed to be more concentrated in the south, *were*-generalization seems to
be more a phenomenon of the Midlands. The average ratio of non-

Table 8.3 Were with singular subjects in the BNC-SpS

BNC code	Total singular	were	% of total
XEA	851	6	0.7
XHC	1,759	27	1.5
XHM	230	0	0.0
XIR	411	1	0.2
XLC	817	59	7.2
XLO	2,964	451	15.2
XMC	1,042	81	7.8
XMD	252	0	0.0
XME	765	31	4.1
XMI	278	8	2.9
XMS	385	59	15.3
XMW	1,703	26	1.5
XNC	888	206	23.2
XNE	886	137	15.5
XNO	231	28	12.1
XSD	904	5	0.6
XSL	390	1	0.3
XSS	1,147	17	1.5
XSU	265	3	1.1
XWA	1,152	19	1.6
Total	17,320	1,165	Ø6.7

Table 8.4 Were with singular subjects in tag questions
in the BNC-SpS

BNC code	Total tags	were *tags*	% of total
XLO	78	3	3.8
XMC	18	1	5.6
XME	28	4	14.3
XMS	17	1	5.9
XMW	44	1	2.3
XNC	23	1	4.3
Total	208	11	Ø5.3

standard forms of 5.3 per cent in tag questions is similar to the 6.8 per cent overall. Although the absolute figures are much higher for generalized *were* than they were for generalized *was* (singular pronouns are after all much more frequent than plural ones), again generalization in tag questions is almost restricted to single occurrences. The one main exception is again the northeast Midlands, where the maximum of four occurrences is found. The very fact that generalized *were* does practically not occur in a tag question speaks for itself, although the individual figures could of course not be tested for statistical significance. Let us now look at the respective figures for negative environments.

Table 8.5 Wasn't with plural subjects in the BNC-SpS

BNC code	Total singular	wasn't	% of total
XEA	16	0	0.0
XHC	64	1	1.6
XHM	8	0	0.0
XIR	6	2	33.3
XLC	41	3	7.3
XLO	114	6	5.3
XMC	49	2	4.1
XMD	8	0	0.0
XME	27	3	11.1
XMI	6	0	0.0
XMS	23	2	8.7
XMW	47	6	12.8
XNC	50	0	0.0
XNE	14	0	0.0
XNO	10	0	0.0
XSD	12	1	8.3
XSL	18	0	0.0
XSS	39	1	2.6
XSU	4	0	0.0
XWA	39	3	7.7
Total	595	30	Ø5.0

Table 8.6 Wasn't with plural subjects in tag questions in the BNC-SpS

BNC code	Total tags	wasn't	% of total
XLC	16	2	12.5
XLO	40	1	2.5
XMW	25	1	4.0
Total	81	4	Ø4.9

The generalization of *was* in negative contexts is significantly less frequent than its positive counterpart (5 per cent on average vs 12.2 per cent for positive *was*-generalization). Nine dialect areas out of twenty, i.e. almost half, show no occurrences at all. Where generalized *wasn't* does occur, it does so only minimally in most dialect areas. We can conclude that *wasn't* is clearly not a favoured generalization strategy. Again, the figures are even lower for tag questions; here, generalization occurs in only three dialect areas with a total occurrence of four, as Table 8.6 shows.

If we look at generalized *weren't* in Table 8.7, we see a very different picture.

Table 8.7 Weren't with singular subjects in the BNC-SpS

BNC code	Total plural	weren't	% of total
XEA	114	61	53.5
XHC	231	39	16.9
XHM	30	0	0.0
XIR	49	1	2.0
XLC	214	63	29.4
XLO	409	133	32.5
XMC	136	36	26.5
XMD	24	4	16.7
XME	101	24	23.8
XMI	49	26	53.1
XMS	63	26	41.3
XMW	226	13	5.8
XNC	134	59	44.0
XNE	79	19	24.1
XNO	41	9	22.0
XSD	77	5	6.5
XSL	67	33	49.3
XSS	211	110	52.1
XSU	46	5	10.9
XWA	165	37	22.4
Total	2,466	703	∅28.5

Table 8.7 shows that non-standard *weren't* is used virtually everywhere. Also, at an average ratio of 28.5 per cent it occurs most frequently of all generalization strategies of past tense BE. Generalized *weren't* is almost six times as frequent as its negative competitor *wasn't*, and more than four times as frequent as its positive counterpart *were*; these differences are statistically highly significant (at $p > 0.01$). In tag questions, generalized *weren't* similarly occurs very frequently, as Table 8.8 shows.

Non-standard *weren't* at last occurs frequently enough to allow statistical significance tests for individual dialect areas (although some areas still had to be excluded; these are Ireland, Merseyside, Scotland and the upper southwest, where occurrences are below five). For the remaining areas, in six the difference between generalized *weren't* in tags and generalized *weren't* in non-tag environments is not statistically significant. These areas are Lancashire, London, the northeast Midlands, the south Midlands, the northeast and the central southwest. The remaining nine dialect areas generalize *weren't* much more frequently in tag questions than in non-tags and thus conform to the trend already indicated by the averages. In fact, in one area (Merseyside) all occurrences of *weren't* with a singular subject are inside tag questions. The striking differences between the four generalization strategies in question are summarized in Table 8.9 which details the overall ratio of a non-standard form in tags vs non-tags. (All figures

Table 8.8 Weren't with singular subjects in tag questions
in the BNC-SpS

BNC code	Total tags	weren't	% of total
XEA	55	39	70.9
XHC	79	25	31.6
XIR	13	1	7.7
XLC	54	17	31.5
XLO	156	56	35.9
XMC	55	20	36.4
XMD	12	4	33.3
XME	44	13	29.5
XMI	22	15	68.2
XMS	33	14	42.4
XMW	98	10	10.2
XNC	50	34	68.0
XNE	27	8	29.6
XNO	29	8	27.6
XSD	16	4	25.0
XSL	45	32	71.1
XSS	100	49	49.0
XSU	14	2	14.3
XWA	87	33	37.9
Total	989	384	Ø38.8

Table 8.9 Non-standard forms in tag questions vs non-tags in the BNC-SpS

	Total	Non-tags	Tags	% tags of total
was	261	253	8	3.1
were	854	843	11	1.3
wasn't	15	11	4	26.7
weren't	703	319	384	54.6
Total	1,833	1,426	407	Ø22.2

are restricted to those dialect areas where the non-standard form occurs
in tag questions at all.)

Table 8.9 shows very clearly that non-standard forms prefer the environ-
ment of negative tag questions. While we find only a minute minority of
the two positive forms, non-standard *was* and *were*, in tag questions, this
figure rises to over a quarter of all occurrences for the negative *wasn't*,
and over half of all occurrences for negative *weren't*. We can see that these
regularization strategies do not apply indiscriminately, but seem to be
structured internally in very specific ways. The fact that the negative forms
prefer to occur in tag questions hints at a functional differentiation, the

pattern of which shall be investigated in more detail in the remainder of this chapter. From this introductory overview, however, we can already say that for present day British English, we find a preference for *was*-generalization over *were* in positive contexts, while we can see an even stronger preference for *weren't* over *wasn't* in negative contexts.

Positive vs negative contexts

The next section compares the two generalization strategies in positive clauses with generalization in negative clauses in more detail. From the overview we have already seen that the *percentage* of generalization in negative clauses is always much higher than in positive clauses. Taken together, *was*- and *were*-generalization in positive clauses average at 1,767/22,243 or almost 8 per cent, generalization in negative clauses on the other hand at 733/3,061 or 23.9 per cent, a significant difference. Generalization in negative clauses is therefore three times as likely as in positive clauses. If we look at this ratio for the individual areas, we find a very similar picture.

Although some areas again had to be excluded from statistical testing because despite the conflation of categories, the occurrences were still below five (Humberside, Ireland and Merseyside), of the rest only three

Table 8.10 Generalization in positive contexts in the BNC-SpS

BNC code	Total positive	Generalized forms	% of total
XEA	1,065	94	8.8
XHC	2,281	62	2.7
XHM	306	2	0.7
XIR	508	8	1.6
XLC	1,093	80	7.3
XLO	3,838	539	14.0
XMC	1,400	118	8.4
XMD	298	2	0.7
XME	995	52	5.2
XMI	355	14	3.9
XMS	519	66	12.7
XMW	2,124	62	2.9
XNC	1,225	238	19.4
XNE	1,044	156	14.9
XNO	306	41	13.4
XSD	1,135	40	3.5
XSL	496	9	1.8
XSS	1,429	85	5.9
XSU	320	16	5.0
XWA	1,506	83	5.5
Total	22,243	1,767	Ø7.9

Table 8.11 Generalization in negative contexts in the
        BNC-SpS

BNC code	Total negative	Generalized forms	% of total
XEA	130	61	46.9
XHC	295	40	13.6
XHM	38	0	0.0
XIR	55	3	5.5
XLC	255	66	25.9
XLO	523	139	26.6
XMC	185	38	20.5
XMD	32	4	12.5
XME	128	27	21.1
XMI	55	26	47.3
XMS	86	28	32.6
XMW	273	19	7.0
XNC	184	59	32.1
XNE	93	19	20.4
XNO	51	9	17.6
XSD	89	6	6.7
XSL	85	33	38.8
XSS	250	111	44.4
XSU	50	5	10.0
XWA	204	40	19.6
Total	3,061	733	Ø23.9

Table 8.12 Statistical comparison of positive and negative
        generalization

BNC code	% positive gen.	Statistical relation	% negative gen.
XEA	8.8	<	46.9
XHC	2.7	<	13.6
XLC	7.3	<	25.9
XLO	14.0	<	26.6
XMC	8.4	<	20.5
XME	5.2	<	21.1
XMI	3.9	<	47.3
XMS	12.7	<	32.6
XMW	2.9	<	7.0
XNC	19.4	<	32.1
XNE	14.9	=	20.4
XNO	13.4	=	17.6
XSD	3.5	<	6.7
XSL	1.8	<	38.8
XSS	5.9	<	44.4
XSU	5.0	=	10.0
XWA	5.5	<	19.6
Total	Ø7.9	<	Ø23.9

behaved differently from the average. In the northeast, the north and the upper southwest, generalization in a positive clause and generalization in a negative clause do not differ significantly. The upper southwest still has very low occurrences, which may be a reason for this unexpected behaviour. For the northeast and the north, however, it looks as if here the old system of Northern Subject Rule might still play a role. (As mentioned before, this is something that still awaits further analysis. First investigations show, however, that all full noun phrases follow the standard English system rather than a system of the Northern Subject Rule. Nevertheless, a frequent verb like past tense BE might preserve a reflex of this rule in its distribution of *was* and *were*.) Although the differences are not significant, the trend even in these three areas is clear: generalization is more frequent in negative contexts than in positive ones. In general, however, for fourteen out of seventeen areas that were admitted to statistical testing, there is a highly significant difference between the two environments, as Table 8.12 shows.

The striking difference between generalization in positive and negative contexts in most dialect areas points to the fact that any generalization strategy is strongly preferred in a negative context over positive contexts. This tendency accords well with a strong cross-linguistic trend that many grammatical distinctions which hold in positive clauses are levelled under negation. Negative clauses in general therefore tend to be less differentiated than their positive counterparts. This is clearly also the case for the past tense forms of BE in non-standard British English, where the tendency to generalize (and thus simplify) an irregular form is strongest under negation. The different generalization tendencies in non-standard English thus conform much more consistently to a general typological pattern than the standard, where – as shown above – the morphological singular–plural distinction holds equally in positive and negative contexts.

Generalization types

As the initial tables have already indicated, most generalized forms occur in all dialects, albeit with different frequencies. In a next step, the combination of *was-* and *were*-generalization for each dialect region is investigated, as well as the relation between these two strategies. The question to be answered is: which generalization strategy is dominant for a particular dialect region? In comparing one non-standard with another non-standard generalization strategy, the question of possible non-representativeness – which so importantly influenced the discussion of, for example, negative concord (see Chapter 5) – can be disregarded, as the relation of standard to non-standard forms, detailed in the four tables above, no longer plays a role here. What we are concerned with here is the internal relationship between two non-standard strategies. In principle, four (2^2) combinations
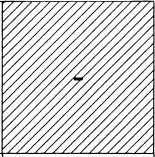
Figure 8.2 Possible and actual combinations of generalization strategies

of these two strategies are possible, but as a glance at the tables above shows, only three of these are realized. The one combination that is not attested (a combination of *were* and *wasn't*) is marked grey in Figure 8.2.

The three strategies that have emerged as possible can be described in more detail as follows:

(1) A dialect could choose to generalize *was* in positive as well as negative contexts. This is probably the most predictable generalization strategy, as *was* occurs so much more frequently even in the standard than *were*: the first and third persons singular account for over 67 per cent of all personal pronoun occurrences in the whole of the BNC, i.e. more than two-thirds. This most predictable generalization strategy also seems to be the type that is favoured in other Englishes. It is reported to be widespread in the United States and is also the dominant strategy in African American Vernacular English (AAVE). Wolfram and Sellers (1999: 109) speculate that this generalization type occurs spontaneously. It results in a simplified past tense paradigm for BE with just one verb form (*was*) for all persons in both positive and negative contexts, thus neutralizing the singular–plural distinction of the standard English system and simplifying it, bringing it in line with all other past tense paradigms.

(2) A second possibility is the exact counterpart: *were/weren't*-generalization to the first and third persons singular. Maybe due to the lower overall occurrence in the standard (about 33 per cent – just under a third) this option might not appear as likely, but as the tables show, it certainly does occur. Again, this generalization type results in a simplified past tense paradigm for BE with just one verb form (this time *were*) for all persons in both positive and negative contexts, which regularizes the irregularity of the standard English by simplification.

Were/weren't-generalization is much rarer worldwide, and Wolfram and Sellers (1999) claim it is characteristic of relic areas. *Were/weren't* does not tend to arise spontaneously.

(3)   The third option is linguistically the most interesting. It is a mix of the other two systems. As shown in the matrix above, the two generalization strategies cannot be combined freely: the combination of generalized *were* (dominant in positive clauses) with generalized *wasn't* (dominant in negative clauses) – though logically possible – does not occur in the BNC sample, and has not been documented in the literature for any variety of English worldwide. Any combined system must therefore consist of *was* generalized in positive clauses and *weren't* generalized in negative clauses. The gap in the system suggests an implicational tendency for British English dialects: if *were*-generalization dominates in positive clauses, *weren't*-generalization will dominate in negative clauses as well.

(1)   Implicational tendency for non-standard past tense BE:
       non-standard *were* ⊃ non-standard *weren't*

A combination of generalized *were* with generalized *wasn't* does not seem to be a possible option.

From this overview, we can now compare the generalization strategies for every single dialect area. The combination of generalization preference in the positive contexts with generalization in negative contexts gives us the generalization types, which can be WAS (the preferred strategy is *was* and *wasn't* respectively), WERE (the preferred strategy is *were* and *weren't*, respectively), or 'mixed' (the preferred strategy is *was* in positive contexts, but *weren't* in negative ones, as detailed above). The individual areas are presented in Table 8.13. A simple dash (—) indicates that statistical testing was not possible because occurrences were too low in absolute frequency. An equation sign (=) indicates that statistical testing did not indicate a significant difference between the two strategies. Results in brackets are tentative, because one of the figures was too low for testing, but the underlying trend might still be of interest.

If we compare generalized *was* and generalized *were* for each dialect area, we find that in most cases, the difference between the two strategies is statistically significant (at $p < 0.02$). Two areas had to be excluded from statistical tests because occurrences for both forms were below five (Humberside and Merseyside), in three other cases (Scotland, the lower southwest and the upper southwest) only the figures for generalized *were* were below this mark and they thus seemed to exemplify the trend towards generalized *was* for positive contexts *in extremis*, as it were. The statistical analysis revealed that for three areas, the difference between *was*- and *were*-generalization in positive contexts was not statistically significant; these areas are Lancashire, the north and the lower southwest. (However, as the occur-

Table 8.13 Positive and negative generalization strategies in the
BNC-SpS

BNC code	Positive	Negative	Generalization type
XEA	was	— (weren't)	(mixed)
XHC	was	— (weren't)	(mixed)
XHM	— (was)	—	?
XIR	was	— (wasn't)	(WAS)
XLC	=	— (weren't)	?
XLO	were	weren't	WERE
XMC	was	— (weren't)	(mixed)
XMD	— (was)	— (weren't)	(mixed)
XME	was	— (weren't)	(mixed)
XMI	was	— (weren't)	(mixed)
XMS	were	— (weren't)	(WERE)
XMW	was	wasn't	WAS
XNC	were	— (weren't)	(WERE)
XNE	were	— (weren't)	(WERE)
XNO	=	— (weren't)	?
XSD	— (was)	— (wasn't?)	(WAS)
XSL	(=)	(weren't)	?
XSS	was	(weren't)	(mixed)
XSU	— (was)	— (were)	(mixed)
XWA	was	(weren't)	(mixed)
Total	was	weren't	mixed

rences for the lower southwest were so low, this area may have to be
excluded from consideration after all. This leaves only two areas with no
clear differences in positive context.) In four areas, *were*-generalization is
actually preferred to *was*; these are London, the south Midlands, the central
north and the northeast. For all other areas (eleven out of eighteen tested,
or over 60 per cent), *was*-generalization dominates in positive contexts.

If we look at the figures for negative contexts, the very low figures in
almost every case for generalized *wasn't* make a statistical analysis not fea-
sible. On the other hand, the very strong trend for generalized *wasn't* NOT
to occur in negative contexts in most dialect areas of course speaks for itself.
Where statistical analyses could be conducted (only in the case of the largest
dialect areas – London and the northwest Midlands), the difference between
generalized *wasn't* and *weren't* is highly signifi-cant. As the figures for *weren't*
in some cases were also below five, for five areas no general trend for neg-
ative contexts could be established; these were Humberside, Ireland,
Merseyside, Scotland and the upper southwest. In all other cases, i.e. in
fourteen out of fifteen areas, *weren't*-generalization is much more frequent
than *wasn't* (which, as we have seen, in many cases does not occur at all).
The only exception is the northwest Midlands, where *wasn't*-generalization
is significantly more frequent than *weren't*-generalization.

*Table 8.14 Was/were-*generalization in the NITCS

	Generalized of total	%	Dominant pattern
Generalized *was*	157/1,608	9.8	WAS
Generalized *were*	71/3,204	2.2	
Generalized *wasn't*	22/98	22.4	WASN'T
Generalized *weren't*	11/222	5.0	

If we combine the information on positive and negative contexts, we can say WAS-generalization in positive as well as negative contexts (generalization strategy (1)) occurs at the most in three dialect areas: possibly in Scotland and Ireland, but the figures are too low to say anything definite, but notably also in the northwest Midlands. For Northern Ireland this finding can be corroborated by results from the NITCS that are displayed in Table 8.14.

It also fits well with the observation that non-standard American English seems to prefer a WAS-generalization strategy especially in those areas that have historically been strongly influenced by Scotch-Irish immigrants, as, for example, the southern United States.

WERE-generalization in both positive and negative contexts (generalization strategy (2)) occurs in four cases, although the areal distribution is not as coherent. These four dialect areas are the contiguous regions of northeast England and the central north, as well as the south Midlands and London.

In two cases, no clear pattern emerges because in positive contexts, no significant difference between the two strategies could be found, as shown above, although both prefer *weren't* in negative contexts. These unclear areas are Lancashire and the north. Humberside again must be disregarded, as there are no instances of generalization in negative contexts. The rest of the country basically shows the mixed type generalization pattern. With seven out of twelve included dialect areas, this generalization type accounts for over 58 per cent of all dialect regions where statistical tests could be applied at least partly. This generalization type certainly seems to be the preferred strategy for non-standard British English. The regional distribution of these three generalization strategies is displayed in Map 8.1. The overall averages for the figures from the BNC also support this impression. On the other hand, this investigation has shown that even for a frequent phenomenon like past tense BE, the figures are too low in many cases to make statistical tests of cross-dialectal comparison possible; only for the largest dialect areas London and the northwest Midlands could it be statistically established that they follow one generalization strategy rather than another.
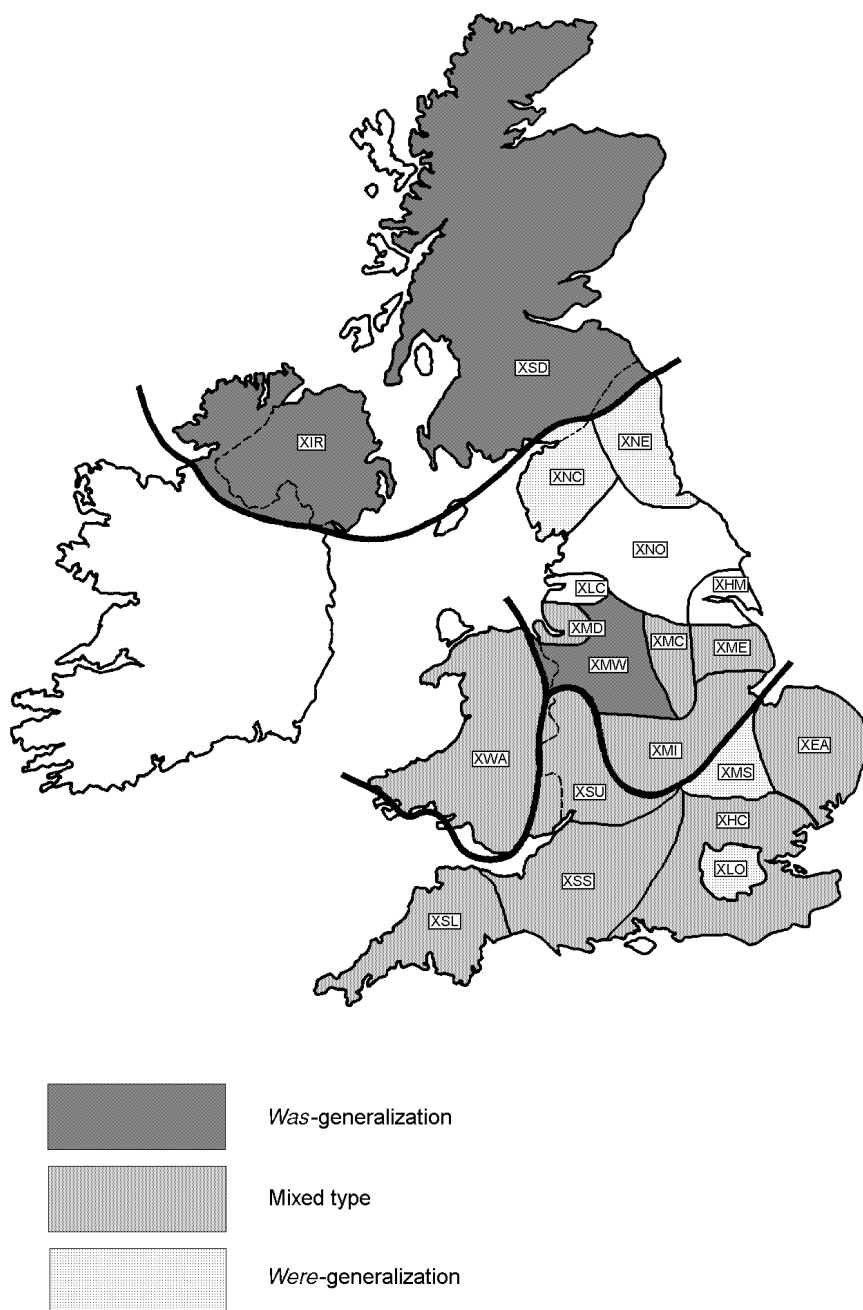
Map 8.1 *Was*/*were*-generalization types in the BNC

Source: Adapted with permission from Trudgill, P. (1990) *The Dialects of England*, London: Edward Arnold.

Historical and functional explanations

As has been mentioned above, Wolfram and Schilling-Estes (1996) suggest that the mixed generalization type they found on Ocracoke is a relatively recent phenomenon and has emerged from the contact of a former *were*-generalizing dialect with the surrounding southern States AmE which employs the more frequent *was*-generalization. This clearly amounts to a historical claim. A second suggestion they make is that the new system has evolved in part in a parallel development to *ain't*, which is also extremely frequently used on Ocracoke.[4] As shown in Chapter 5, *ain't* is a negative marker which does not have a direct positive counterpart **ai* and must probably be analysed as a single morpheme. It is therefore maximally distinct from the regular positive forms of present tense BE and another good example that grammatical (in this case person) distinctions which hold in the positive clauses are levelled under negation. There are then a range of parallels between present tense BE – *ain't* and past tense BE *was* – *weren't*.

   In order to investigate these two claims for those areas that employ the mixed system in Great Britain with the help of the BNC material, in a first step the possibility of temporal variation was investigated. As the BNC contains purely synchronic material, the only approach possible is an apparent time approach. This is theoretically (and practically) possible, but as in the discussion on a combination of dialect region and social class shown above (see Chapter 5), a problem arises when the two categories of 'dialect region' and 'age group' are combined. The lack of representativeness for a combination of these two categories leads to empty cells in the matrix for practically all but the largest dialect areas, notably London, and this makes any regional investigation into apparent time variation a very dubious affair and, for most of the dialects, gives few results, which could at best be a hint in a certain direction. Nevertheless these particular searches were conducted for London and the results are displayed in Table 8.15.

   Table 8.15 shows that the younger speakers (speakers under 35) clearly prefer a mixed system, in accordance with the rest of the nation, whereas older speakers (age groups 3 and 4, aged 35 to 59) have a clear preference for WERE-generalization. The over 60-year-olds are only poorly

Table 8.15 Was/were-generalization per age groups in XLO

Age group	% was	% were	% wasn't	% weren't	Gen. type
0 (0–14)	8.6	1.2	4.3	32.1	mixed
1 (15–24)	8.3	6.5	12.5	26.3	mixed
2 (25–34)	31.0	3.3	11.7	28.9	mixed
3 (35–44)	10.7	44.1	4.5	67.1	WERE
4 (45–59)	5.3	21.3	0.0	24.1	WERE
5 (60+)	6.3	5.7	5.0	11.5	mixed?

represented and the results are not a clear indication of either strategy. Very tentatively, one might conclude from these figures that some kind of change towards a mixed system could indeed be under way. Comparisons with other dialect areas are, however, not uniformly possible and results must therefore remain at best a hypothesis. The historical claim therefore awaits more investigation.

A comparison of the BNC data with the older, traditional data collected in the SED, is, unfortunately, not possible for *was/were*-generalization. The questions eliciting these items only contain positive contexts. Thus we do find material for *was-* and *were*-generalization in positive contexts which we can compare diachronically with the BNC material. The interesting systematic question, however, cannot be answered as negative contexts were not elicited with the SED questionnaire.

The second claim – a development parallel to AIN'T – can be investigated by comparing the BNC data for AIN'T (used for BE), as detailed in Chapter 6, with the regional distribution of the mixed generalization type for past tense BE. In order to show the regional distribution of AIN'T, the data from Chapter 6 are repeated here. The rows have again been arranged in order of decreasing (relative) frequency of AIN'T. The dialect area of Humberside has been excluded as there are no occurrences of AIN'T.

As we have already seen in Chapter 6, AIN'T for BE shows a striking regional distribution, such that we find the majority of dialect areas where the frequency of AIN'T is clustering around 17 per cent, whereas for the

Table 8.16 AIN'T for BE in the BNC-SpS

BNC code	Dialect area	Total	AIN'T for BE	% of total
XNC	Central northern England	767	134	17.5
XMS	South Midlands	253	44	17.4
XMC	Central Midlands	840	141	16.8
XMI	Midlands	285	46	16.1
XLC	Lancashire	656	99	15.1
XNO	Northern England	186	22	11.8
XEA	East Anglia	715	75	10.5
XSS	Central southwest England	953	96	10.1
XSL	Lower southwest England	409	39	9.5
XME	Northeast Midlands	616	58	9.4
XLO	London	2,447	169	6.9
XHC	Home Counties	1,564	100	6.4
XMW	Northwest Midlands	1,250	79	6.3
XMD	Merseyside	202	12	5.9
XWA	Wales	756	31	4.1
XNE	Northeast England	468	14	3.0
XIR	Ireland	200	6	3.0
XSU	Upper southwest England	150	3	2.0
XSD	Scotland	306	4	1.3
Total		13,195	1,172	∅8.9

Table 8.17 Correlation of AIN'T and generalization
types

BNC code	% AIN'T for BE	Generalization type
XNC	17.5	WERE
XMS	17.4	WERE
XMC	16.8	mixed
XMI	16.1	mixed
XLC	15.1	
XNO	11.9	
XEA	10.5	mixed
XSS	10.1	mixed
XSL	9.5	(mixed)
XME	9.4	(mixed)
XLO	6.9	WERE
XHC	6.4	(mixed)
XMW	6.3	WAS
XMD	5.9	(WAS)
XWA	4.1	mixed
XNE	3.0	WERE
XIR	3.0	(WAS)
XSU	2.0	(mixed)
XSD	1.3	(WAS)

second group of dialect areas, AIN'T only occurs with a frequency of around 5 per cent, with the 'Celtic' Englishes at the very bottom of the table. If we correlate this regional distribution of AIN'T with the distribution of *was/were-* generalization types (where these could be determined), the following picture emerges as in Table 8.17.

Table 8.17 is of course very speculative for two reasons: as the preceding sections have shown, the data base in many cases is too small to determine a statistically significant preference for one of the three non-standard generalization strategies in many dialect areas. In other cases, especially in negative environments the preference for *weren't* over *wasn't* is so strong that generalized *wasn't* does not occur at all, and again these relations cannot be tested statistically for variability (zero occurrences are obviously not variable and for this reason usually represent a knockout constraint). Nevertheless, the differences between these two strategies can give some indications, under careful observation, even if the differences cannot be tested statistically. The generalization types that seem to emerge from the data have therefore been added to the final column in brackets to indicate their tentative status.

Table 8.17 shows that there is no very clear correlation between the two phenomena (AIN'T and past tense BE generalization) such that all dialect areas with high frequency AIN'T could automatically be assigned to the

mixed generalization type: the two dialect areas with the highest frequency of AIN'T, the central north and the south Midlands, both generalize WERE in positive as well as negative contexts. The reverse is also not the case: not all dialect areas with the mixed generalization type use AIN'T with a particularly high frequency. Thus, the mixed type also seems to occur in Wales, in the upper southwest and in the Home Counties, three areas that can be found in the lower frequency group for AIN'T. Although a very strong correlation therefore clearly does not hold, one can nevertheless see that mixed generalization in the past tense paradigm is much less frequent in the bottom (low frequency) group than in the top group. Here, only three out of nine areas employ the mixed generalization system in the past tense. In the high frequency AIN'T group, on the other hand, the ratio is six out of eight (for two areas, no strategy could be determined as both generalized *was* and *were* were equally frequent in positive contexts) and thus the mixed generalization type is clearly the dominant strategy. There does seem to be a weak correlation, then, between the frequency of AIN'T and the employed generalization strategy. However, this correlation is not strong enough to warrant the postulation of an implicational tendency. Possibly, AIN'T is just one of several factors that come together to determine the direction of non-standard generalization of past tense BE.

The mixed type generalization system might seem an unlikely strategy, as it does not result in a simplified system like straightforward *was/wasn't*- or *were/weren't*-generalization which can easily be cognitively motivated. What we see instead is that the singular–plural distinction we find in the standard system is eroded in those dialect areas that employ the mixed system. The standard distinction singular–plural is abandoned in favour of a morphological distinction between positive and negative contexts. This makes much more sense cognitively. As we have seen, a singular–plural distinction is clearly not necessary cognitively for past tense verbs in English: the past tense forms of BE constitute the only verb paradigm where this distinction is still in place. A positive–negative distinction on the other hand is extremely valuable, as the information carried by the negator is very important – after all, the negator turns the meaning of a clause into its exact opposite. A hearer who misunderstands the negator will fundamentally misunderstand the whole utterance. Especially with the contracted forms *wasn't* and *weren't*, which tend to undergo further contraction to *n'*, any additional information about the actual negative character of a clause is cognitively very important. The clear acoustic differences between *was* and *weren'(t)* – absence of a sibilant, difference in vowel quality and vowel quantity – maximize this distinctiveness. In these contexts, the negator itself may become redundant where the positive–negative distinction is transferred to the morphological level completely. Even in the case of complete loss of *-n't*, the intended message would then still be unambiguous, with *was* indicating positive clauses, *were* indicating negative ones.

Summary

The three non-standard systems of past tense BE that have emerged from our investigation can be clearly divided into two strategies. The first is, again, regularization. Straightforward *was/wasn't*-generalization results in a regularized past tense paradigm that abolishes the irregular third person singular–non-third person singular distinction of the standard that does not generally exist in the past tense. *Were/weren't*-generalization, although the much rarer strategy, essentially does the same, only employing different lexical means. To illustrate, we can compare their simplified paradigms with the standard English paradigm (reproduced here as Figure 8.3).

In both cases, the simplification becomes easily apparent: instead of four forms, these two strategies can employ just two, abolishing the morphemic differences in addition. The mixed system, however, cannot be as easily described as a regularization strategy. It is displayed schematically in Figure 8.6.

Here, the morphemic difference of standard English is not abolished, as in the other two cases, but reallocated to indicate the difference in polarity. The morphological marking has been transferred from a distinction of person and number (as in the standard system) to the distinction of positive and negative. This strategy is therefore perhaps best described as remorphologization.

It is clear that, in the face of variability, these descriptions only summarize idealized endpoints of an ongoing process. In the meantime, what we find for present tense BE – just as we found it for *ain't* and for *don't* – is a highly asymmetrical system where person-distinctions are kept relatively intact in the positive contexts but are increasingly levelled under negation. (In contrast to *ain't* and *don't*, however, levelling does occur in the posi-

was : wasn't
————————
were : weren't

Figure 8.3 Past tense BE (standard English)

was : wasn't

Figure 8.4 Past tense BE (generalization strategy 1)

were : weren't

Figure 8.5 Past tense BE (generalization strategy 2)

was || weren't

Figure 8.6 Past tense BE (generalization strategy 3)

am

is   ain't

are

was

were   wasn't/weren't
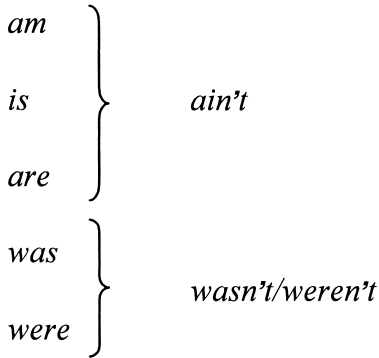
Figure 8.7 Non-standard BE (present and past tense)

tive paradigm as well – if only at very reduced ratios.) This final result can again be displayed graphically and Figure 8.7 stresses the parallelism that we find in the present and past tense paradigms of non-standard BE.

The motivating factor for these pervasive asymmetries as we have encountered them in the present chapter and in the preceding ones will be discussed more fully in Chapter 9.

9 Conclusion

Standard English vs non-standard varieties

What has been discussed in the preceding chapters will now be summarized. In this first section, the description of standard English in Chapter 2 will serve as the foil for comparing the investigated standard and non-standard features of the negative system of present day (spoken) British English. It must be stressed again that non-standard speech is not understood as comprising traditional dialect speakers – these have always constituted a minority, and especially today certainly do not represent the average non-standard speaker. Instead, our main interest has focused on the majority of speakers today. As we have seen, newer, non-standard constructions are used with varying frequencies throughout the population across all dialect areas, albeit with sometimes striking regional differences.

We have seen in Chapter 3 that the unmarked sentence negator of standard English, *not/-n't*, is also the unmarked negator in most non-standard English dialects today. Negators like *noan* or the clitics *-no/-na/-nae* which go back to the lexicalization of a different negative polarity item in Old English are today regionally very restricted and are mainly found in the northwest Midlands and in Scotland and some neighbouring northern English dialect areas. Furthermore, there are no differences between standard English and modern non-standard dialects in elementary characteristics like basic word order. It is SVO in all cases, and many criteria from the more traditionally typological studies on negation could therefore not be applied directly. Equally, a comparison with the historically earlier stages of English shows that all dialect areas seem to have adopted the same system as standard English early on, and there are practically no remnants of an older system in the dialects today (the exception of neg concord is discussed below). Thus, the negator in all varieties has to take a position after the first auxiliary (or form of BE) of the clause. The identity of these very basic underlying characteristics in both the standard and the non-standard varieties is of course a prime argument for considering them varieties of the same language (English) in the first place. They can be considered part of the common core of English grammar.

The greatest number of morphological differences could be established for Scots. Again, it is therefore not surprising that Scots is the only serious candidate that could be considered a sister language of English, on linguistic grounds, rather than a dialect of the standard southern variety.[1]

The negation of the three primary verbs could be ordered on a scale of irregularity for standard English, with the full verb HAVE taking an intermediate position between the irregular paradigm of BE (which cannot take DO-support) and the full verb DO (which obligatorily takes DO-support). For most modern dialects today, we have seen in Chapter 4 that in the majority of dialect areas, full verb HAVE behaves in a much more regular way than in standard English, and one can imagine this as a shift of HAVE along the irregularity scale towards the more regular pole, i.e. in the direction of DO. In this respect, it seems to be the (written) standard that preserves an earlier stage of the English language, whereas the spoken language is more progressive and certainly more regular. The paradigm of BE on the other hand is just as irregular in never allowing DO-support for negation in non-standard dialects as in the standard variety; this is therefore another candidate for a common 'core' grammar.

Gaps and irregularities

The first implicational tendency that could be proposed for standard English relates to contraction strategies. We have seen that in standard English, whenever auxiliary contraction was possible in the system, negative contraction also had to be possible (but not vice versa). The only exception to this strong tendency was *(I) am not*: here, only the auxiliary, but not the negative can be contracted. The lack of a negative contracted form **amn't* thus constituted a first systematic gap of the standard. The detailed investigations in Chapter 4 have shown that indeed in all regions, auxiliary contraction is much less frequent than negative contracted forms for almost all verbs, thus confirming the implicational tendency that was proposed for standard English, not only in terms of systematic possibilities, but also in terms of frequency relations. Chapter 4 also showed that the gap in the system of standard English (**amn't*) is not closed in most varieties of non-standard English, contrary to what might have been expected. Again, historical dialects that contain the form *amn't* are restricted to the northwest Midlands, Scotland and Ireland. However, the frequency distribution of the two competing contraction strategies showed that for all varieties of English, the whole paradigm of the verb BE behaves 'irregularly' or in exception to the implicational tendency: auxiliary contraction rather than negative contraction is the preferred strategy for all verb forms of the present tense: *he's not*, *we're not*, *I'm not* are always preferred to *he isn't*, *we aren't* and the non-existent *I amn't*. A cognitive explanation was advanced for the exceptional behaviour of BE: it has an exceptional status among the auxiliary and primary verbs in that its semantic load is particularly small. Because of this, it is the one verb

that is most prone to contraction and even deletion. This tendency is supported by cross-dialectal and cross-linguistic evidence that copula deletion (or simply copula absence)[2] is indeed a frequently observed phenomenon. The pattern of a markedness reversal for the verb BE could thus be explained in terms of a different prototypical function of the verb BE compared to the other frequent verbs.

Although this markedness reversal does not fill a gap that exists in the standard English system, but instead seems to widen it from one present tense form to the whole present tense paradigm, nevertheless, it is still an effective regularization strategy, just as the filling of the gap would have been: overall, it leads to a more regular system of the language. Although a whole, more irregular, paradigm is the price that is paid, this irregularity itself can be explained as being functionally motivated.

In the realm of modal verbs, Chapter 4 has also shown that one more 'gap' of standard English, namely the lack of epistemic *mustn't* is indeed closed in most non-standard varieties. This anomaly of the standard system can thus in retrospect be confirmed as such, a gap.

Another clear case of a regularization strategy could be found in the non-standard past tense paradigm of BE, as we have seen in Chapter 8. Both *was/wasn't*- and *were/weren't*-generalizations lead to a past tense paradigm for BE that is in line with all other past tense paradigms of English: namely one without person distinctions.

The generalization of *was* in positive contexts and *weren't* in negative contexts (generalization strategy (3)) on the surface again looks like a more complicated, rather than a simpler or more regular, system. It could be shown, however, that this solution is ultimately even more economical and leads to a different kind of regularity: the morphological distinction of the standard (singular vs plural) is transferred to the difference between positive and negative contexts. This remorphologization ultimately makes the negative clitic -*n't* redundant, and thus ensures (through a kind of doubling effect) that the speaker's intended message has twice as good a chance of being transmitted and received.

Gap-filling, remorphologization and other regularization methods therefore are strategies employed by non-standard varieties of English that clearly reduce irregularities and inconsistencies of the standard variety and that can thus be easily motivated by cognitive, and even language internal, factors.

Cross-linguistic patterns

There are other phenomena of non-standard English (in the area of negation) that cannot so easily be interpreted as resulting from regularization, and that do not necessarily result in a more coherent system. Nevertheless, they can be cognitively motivated, because a comparison with dominant

patterns in the languages of the world shows that here, the non-standard varieties of English comply with typological patterns where the standard does not. Although in these cases, explanations move beyond reasons of internal consistencies, typological patterns suggest that there are under-lying regularities which are the results of cognitive factors.

When we looked at the interplay of negative quantifiers with the sentence negator in Chapter 5, for example, we noticed striking differences between standard English and the non-standard dialects. In practically all dialects, co-occurrence of these two negating strategies is permitted. In no case is this obligatory but, equally, in no case – except the standard – is it impos-sible to find either the occurrence of a negative quantifier with the sentence negator, or the co-occurrence of two negative quantifiers in one clause with an overall negative meaning. This points to a fundamental typological difference between the system of standard English and the non-standard varieties: standard English, as pointed out in Chapter 2, patterns typo-logically with its standard Germanic sister languages in being neg-impermeable, whereas all non-standard varieties have preserved the older status of being neg-permeable. The widespread systematic possibility for the use of negative concord does not constitute a regularization phenom-enon as such. However, it brings the non-standard varieties in line with a cross-linguistic trend – both with the majority of the languages of Europe, and with the majority of languages worldwide. The cross-linguistic pref-erence for multiple expression of negation in one clause is, as we have seen, cognitively motivated: the semantic load that the negator carries is particularly large, and redundant marking of this information is one formal means for ensuring the correct interpretation of the speaker's intended meaning. In the terms of natural morphology, negation as the (function-ally) more marked member has to employ more morphological features on the expression level to ease the processing load.[3] Semantic markedness and formal marking typically go together in what Mayerthaler calls *construc-tional iconicity*, so that semantically more marked categories 'are more often coded with features than are base categories. ... The intuitive basis for this is that what is semantically "more", should also be constructionally "more"' (Mayerthaler 1988: 18, cf. also Wurzel 1984: 203). Conversely, it is not surprising that during the process of standardization this multiple expression of negatives was particularly frowned upon, as it runs counter to what Stein has termed the standardization principle 'No double surface realization' (Stein 1998: 38).[4]

A second phenomenon can also be explained by directly referring to the typological characterization of standard English. We have seen that standard English possesses an extraordinary strategy for negating clauses, in that a negative clause requires the presence of an auxiliary, so that clauses with a main verb have to add the semantically empty 'dummy' auxiliary DO for the purposes of negation. This is extremely marked and typologically very unusual. We have compared this system to 'proper'

negative auxiliaries that exist in non-Indo-European languages like Finnish and Estonian. As Payne (1985) has noted, there are strong tendencies in these languages with negative auxiliaries to change them into invariant neg markers, the most usual negation strategy worldwide. Against this background, it is therefore not surprising to find very high ratios of invariant *don't* in the third person singular in all non-standard varieties of English – another feature where non-standard English is more progressive than the standard, and where the development still seems to be going on. Because of the history of the English language, this regularization has (so far?) only extended to the present tense of DO – that is, only positive clauses with a main verb in the present tense are negated in these varieties with the same (and thus invariant) negative marker *don't*. Clauses with a modal still make use of the traditional negative pattern, while BE and auxiliary HAVE are negated by *ain't*. A trend towards a truly invariant negative marker can thus, at the most, be said to be at an incipient stage for non-standard varieties of English.[5]

Asymmetries

As noted throughout this book, some phenomena of the non-standard negative systems that are investigated here can be explained as the filling of systematic gaps in standard English. A grammatical system that has filled a gap can be considered as more regular; in this way gap-filling is clearly a regularization strategy. Gap-filling phenomena and (other) regularization strategies, however, cannot account for another pattern that we have encountered, the highly pervasive asymmetries.[6]

The use of the form *ain't*, for example, clearly levels several person distinctions, and, moreover – probably due to a historical coincidence – the distinction between the two verbs BE and HAVE, as Chapter 6 has shown. However, this levelling only takes place under negation, as *ain't* functions as the negative verb form for all persons for both verbs. In other words, as this levelling is specific to negative contexts, it therefore leads to *more* surface irregularities rather than fewer: the person distinctions have to be (and indeed are) still observed in positive contexts, whereas they are regularized under negation. The outcome is a highly asymmetrical system where many more distinctions are present in the positive paradigms than in the negative paradigms.

In this respect, the use of invariant *don't* follows the same trend, as shown in Chapter 7. Invariant *don't* is also not particularly well described as a regularization strategy. Although the use of *don't* for the third person singular levels the person distinction of the third person singular vs all other persons, on the surface this again leads to *more* irregularity in the system rather than less, because the person distinction as such is not abolished. Speakers will now have to remember to use *he does*, *she does*, *it does* in positive contexts, but *he don't*, *she don't*, *it don't* in negative ones. Again,

am

is

are } *ain't*

has

have
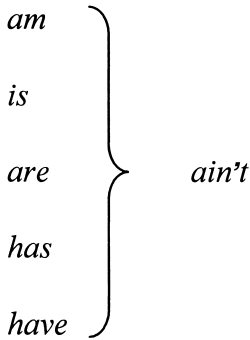
Figure 9.1 Asymmetrical AIN'T paradigm

do
} *don't*
does
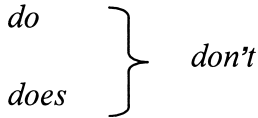
Figure 9.2 Asymmetrical DO paradigm

was

} *wasn't/weren't*

were
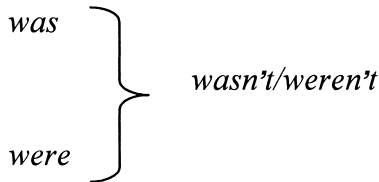
Figure 9.3 Asymmetrical past tense BE paradigm

practically all non-standard varieties could be shown to be moving in the direction of this highly asymmetrical system, displayed in Figure 9.2, and it remains to be discussed why.

Past tense BE, finally, shows that in the process towards a regularized or a remorphologized system (depending on the dominant generalization strategy), levelling is again strongest under negation, so that again we have this pervasive trend towards an asymmetrical system, displayed in Figure 9.3.

In sum, it can be said that there is a very strong tendency in non-standard varieties to distinguish negative clauses from their positive counterparts by way of strongly asymmetrical paradigms. The reasons and motivations shall be discussed now with the help of the very general markedness criteria that were introduced in Chapter 1. We have already seen that very generally, negative contexts are considered marked in contrast to their positive counterparts. Typological studies have discovered

a trend that many grammatical distinctions of positive clauses are levelled under negation. Greenberg, for example, mentions that for negative forms cross-linguistically, 'there is sometimes neutralization of categories which are distinct in the positive form. Thus in Shilluk there is syncretization of the present and future in the negative' (1966: 50). Payne states that 'neutralization of tense distinctions in negative sentences is a not infrequent phenomenon' (Payne 1985: 230); for example, in some languages, morphological futures exist 'in positive verb forms only: the non-past negative serves as a negative for both the present and future tenses' (219). Givón and Miestamo add information from the Bantu language Bemba, where the distinction between 'tomorrow future' and 'after tomorrow future' is lost in the negative; only the 'tomorrow future' is used there (cf. Miestamo 2000). Drawing on a wide range of materials, Hagège finds neutralization of tenses in negative contexts in 42 per cent of his cross-linguistic sample of 754 languages (1982: 85). Givón also cites further evidence from the languages of Rwanda and Zulu, which allow the use of only a subset of focus particles under negation (Givón 1978: 86–7, 92).

One could say then that negative sentences in general tend to be less differentiated than their positive counterparts. In order to see how this pattern can help us understand the pervasive patterns of asymmetries, we shall return to very basic underlying typological principles that surface in the markedness criteria formulated in Greenberg (1966) and reformulated in Croft (1990) which we already encountered in Chapter 1. We established that affirmative statements in standard English are clearly unmarked, in opposition to their negative counterparts, in that they conform to two markedness criteria: they are typically realized by zero, and they are far more frequent than negative statements.

Non-standard English on the other hand also marks negative clauses by the morpheme {-n't}, often by the addition of a form of DO as well, and thus conforms to criterion S1. Equally, negative clauses are much less frequent than their positive counterparts in spoken language. Although it has been shown that negation is far more frequent in conversation than in written texts,[7] negative clauses are still much rarer than positive ones

Table 9.1 Markedness criteria

	Name	*Explanation*
S1	Zero value	The unmarked value is typically realized by zero
S2	Syncretization	The unmarked value has more distinct forms in the paradigm
S5	Irregularity	The unmarked value has more allomorphs or is more irregular
S8	Frequency	The unmarked value is more frequent in text counts

Source: Adapted from Croft (1990: 70–94).

Table 9.2 Fulfilled markedness criteria

	Name	Standard English	Non-standard English
S1	Zero value	√	√
S2	Syncretization	—	√
S5	Irregularity	—	√
S8	Frequency	√	√

also in spoken language: Chapter 8 suggests a ratio of roughly 1:7 for informal spoken English (in all, there were over 22,000 instances of positive *was/were*, against just over 3,000 of negative *wasn't/weren't*).[8] Criterion S8 is therefore also clearly fulfilled. The pervasive pattern of asymmetry, however, can now also be explained with a view to the remaining criteria in Table 9.1. Criterion S2 is related precisely to patterns that fuse (or 'syncretize') grammatical distinctions in the marked environment, and this is exactly the phenomenon we have found again and again, for *ain't, don't* as well as *wasn't/weren't.* These patterns of asymmetries seem to be directly related to the marked character of negation worldwide. Miestamo gives an ontological reason for asymmetrical paradigms as they can be found cross-linguistically between affirmative and negative contexts: 'In the asymmetric paradigms there is a "vertical" analogy (or iconicity): the ontology of non-fact is less differentiated than the ontology of fact, and linguistic structure reflects this distinction' (Miestamo 2000: 78). These asymmetries can thus again be iconically motivated. And finally, in abolishing partly irregular person distinctions under negation, this makes the negative paradigms also more regular than their positive counterparts. In other words, non-standard English even shows a trend of complying with criterion S5, where the standard is simply neutral. Summing up, standard and non-standard English differ in how strongly negation is marked morphologically, as Table 9.2 shows.

   In other words, the patterns of asymmetry show that non-standard English is changing towards a morphological system which overall enhances the marked character of negation. Again, there are good functional reasons for this pattern which, on the surface, appears to be rather complicated. Enhancing the marked character of negation serves to make negative clauses more different from positive clauses. Ultimately redundant marking ensures that speakers have a better chance of making themselves understood, and hearers have an easier task of identifying the intended meaning. Non-standard varieties simply seem to go a step further in this direction than the standard. This functional motivation seems to be the underlying reason for the pervasive asymmetrical patterns we have encountered in this study.

Use of the BNC

This final section reconsiders the use of the BNC for purposes of cross-dialectal comparison. As has been noted throughout the preceding chapters, there are many problems involved with these kinds of investigations, apart from the purely practical – that, due to limitations in the soft- and hardware, searches have to be restricted to certain combinations as demonstrated in the individual sections, and that therefore many questions cannot be answered satisfactorily yet. Even if we assume that the greatest part of the SpS subsample is relatively faithfully transcribed, or that possible mistakes even themselves out in a sample of this size, it is clear that the choice of non-linguists as transcribers and, especially, the sometimes unusual 'normalization' practices, mean that some phenomena that might be of interest to linguists are lost to objective investigations. Even some morphosyntactic phenomena like the shape of the negator suffer from this limitation and cannot be examined reliably with the help of the BNC – contrary to claims by the compilers. Apart from this transcription problem, we have encountered the problem of sheer size and inadequate software – a superficial problem that could of course easily be remedied in the future. Finally, we have seen that for several dialect areas, subcategorization leads to empty categories and thus becomes meaningless. For the investigation of non-standard features, the presence of non-standard speakers of lower socio-economic status is of course an essential prerequisite. This is therefore a serious limitation, and the white spots on various maps and figures show that often a detailed regional conclusion had to be postponed to future studies.

On the other hand, these unrepresentative areas are spread relatively evenly over Great Britain. Although small-scale comparisons will not be possible with the help of the BNC – after all, the BNC is not a dialect corpus, in fact not even a sociolinguistic corpus – large-scale comparisons can indeed lead to meaningful results. In particular, the investigation of neg concord has shown that a comparison of northern, Midland and southern dialect areas can confirm the regional differentiation of a phenomenon that is widely assumed to have no regional differentiation, or where first results (the study by Cheshire, Edwards and Whittle 1993, for example) have been doubted because of prevailing opinion.

In sum, however, it should be remembered that the main arguments of this book relate to a typological, and thus ultimately functional, explanation of the negation patterns found in non-standard dialects, and that any geographical distribution is of only secondary importance. Nevertheless, the immense wealth of material contained in the BNC and especially the fact that all of it is unobserved natural speech outside any constrained interview format is argument enough to try to use this new source to the greatest advantage, even if this sometimes entails purposes that were perhaps not originally envisaged. In this way I hope to have shown that large-scale investigations of relatively high-frequency phenomena can indeed lead to meaningful and, in some cases, new results for non-standard spoken British English today.

Appendix
Shared grammatical features of British urban dialects

Difference to StE	Example
Differences in the NP	
Absence of plural marking	'To make a cake you need two pound of flour'
What as subject relative pronoun	'The film what was on last night was good'
Them as demonstrative adjective	'Look at them big spiders'
Differences in the VP	
Non-standard *was*	'We was singing'
There was with plural subject	'There was some singers here a minute ago'
There's with plural subject	'There's cars outside the church'
Past participle *sat* following BE	'She was sat over there looking at her car'
Past participle *stood* following BE	'And he was stood in the corner looking at it'
Should of	'You should of left half an hour ago!'
Differences in the AdvP	
Adverbial *quick*	'I like pasta. It cooks really quick'
Differences in the field of negation	
Never as past tense negator	'No, I never broke that'
Ain't/in't	'That ain't working/that in't working'
Differences in syntax/word order	
Give me it	'Give me it, please'

Note: Regrouped after Cheshire, Edwards and Whittle (1993: 64–5).

Notes

1 Introduction

1 Incidentally, this is not only true for grammatical phenomena, which have always been suspected of being much more uniform than accents, but is increasingly also the case for matters of pronunciation. The extremely rapid spread of /t/-glottalization is one example (cf. contributions in Foulkes and Docherty 1999); other examples include the spread of what has been termed 'Estuary English' from the southeast of England to large parts of the country (cf. Altendorf 1999).

2 Cf., for example, Kerswill and Williams (2000); Cheshire, Edwards and Whittle (1993). A similar process must have accompanied the rise of the standard variety in Early Modern English; Nevalainen (2000) terms this process 'supralocalization'.

3 Ramat (1987: 48).

4 This is probably true of artificial as well as of natural languages: negation is one of the basic logical operators that can define every other higher operation.

5 Cf. Wierzbicka (1998: 144) for a recent list of conceptual primitives.

6 The – both western and eastern – history of this study is conveniently summarized and discussed in Horn (1989), who takes the reader up to current semantic and pragmatic views on negation. However, as Horn has little to say on varieties, his otherwise excellent book is unfortunately of only little practical value for the present study.

7 For a programmatic sketch of this approach cf. also Kortmann (1999) and Anderwald and Kortmann (2002).

8 Cf. van der Auwera (1998: 814) for results from several studies of adverbial constructions, and for the use of this term.

9 Four other criteria have been left out and will not play a role in this discussion: Croft shows that two criteria applying to neutralization (his S3 and S4) do not belong to markedness criteria in general, as neutralization is generally a different phenomenon from markedness; criterion S6 (relating to defectivation) even for Greenberg 'can be considered a form of syncretism' (Greenberg 1966: 29), and can thus be subsumed under S2, and the last criterion, S7, is specific to the categories of number and gender only and thus does not apply to our case of polarity. For a critique of these criteria for markedness cf. Battistella (1990: 25ff.).

10 This is not as obvious as it may sound; there are indeed languages (for example, Welsh) that possess a marker of affirmation (cf. Bernini and Ramat 1996: 110), as we shall see in Chapter 3.

11 For details cf. Mair and Hundt (1997).

12 For some first comparative studies as well as a discussion of the term *Celtic Englishes* itself cf. the contributions in Tristram (1997, 2000).

13 As an anonymous reader helpfully pointed out, the usual terms are *North*, *Central*, and *South* (cf. also Trudgill 1990: 33, 63). As we will have to reserve an abbreviation for the 'Celtic' varieties, however, the term *Midlands* (or *Mid*) rather than *Central* has been chosen, in order to have unambiguous one-letter shorthands for all areas.

14 Although the length of the questionnaire meant that many interviews are incomplete, or that interviews were begun with one informant and completed with another, cf. Chambers and Trudgill (1998: 23).

15 Cf. Kirk (1990, 1992) for details.

16 Although the interest in Northern Ireland as a dialect area is rising, as the recent publication of McCafferty's (2000) study on Londonderry shows exemplarily.

17 The possibility of utilizing the wealth of material that lies hidden in folklore archives and museums is only just beginning to be investigated. It opens up particularly rich avenues of investigation into dialect grammar, and is the basis of a regionally representative dialect corpus that is at the moment under construction at Freiburg University in Germany (the Freiburg English Dialect corpus FRED).

18 A similar method has been employed by sociolinguists with the explicit aim of avoiding the observer's paradox, and this has been described as 'a very successful method of collecting data' (Janet Holmes on the Wellington Corpus of Spoken New Zealand English, p.c.).

19 The division of the United Kingdom into dialect regions seems to follow Trudgill's division of modern dialects (cf. Trudgill 1990: 66), although this is not made explicit anywhere in the material surrounding the BNC. It is, however, confirmed by the fact that other dialectologists have noticed this correspondence independently (Klemola p.c.).

20 Obviously not included were utterances, for example, by German or French native speakers, or by speakers of Indian or Jamaican English.

21 However, the authors of the *Handbook* promise that 'corrigible errors . . . will be corrected, as resources permit, in later versions of the BNC' (Aston and Burnard 1998: 37). This second version was not yet available at the time these studies were conducted (first announcements were not made in the BNC discussion list until December 2000), but subsequent checks have revealed that retranscriptions on a large scale have not been made.

22 This corresponds to the 95 per cent confidence level usually employed in the social sciences. Cf. Kretzschmar and Schneider (1996: 37ff.) for a justification of different – equally arbitrary – cut-off points in sociolinguistic and dialectological statistical analyses.

2   Negation in standard English

1 In the typological literature, Dahl also employs similar terminology; he notes that 'in morphological neg constructions, neg is an inflectional category of the verb' (Dahl 1979: 81). Payne on the other hand curiously states that 'morphological negation occurs whenever the negative morpheme must be considered to form part of the *derivational* morphology of the verb' (Payne 1985: 226, my emphasis). It is possible that Payne uses *derivation* here in the generative sense of the creation of a surface structure, and not in contrast to *inflection*. Van Schaaik in his treatment of negation in Turkish unfortunately only refers to the Turkish negation marker as a 'suffix' (van Schaaik 1994: 38f.), keeping the morphological status of this suffix open to debate. Nevertheless, in cases where the negative morpheme has become internal to the verb morphology, I would maintain that it makes more sense to consider it an inflectional marker. Dryer also supports

this analysis when he calls the English prefix *un-* a derivational affix and bound morphemes expressing the same meaning as English *not* 'inflectional affixes' (Dryer 1988: 93). Bernini and Ramat describe the Turkish construction as 'a bound morpheme [which] occurs within the verbal inflection' (Bernini and Ramat 1996: 9). Zwicky and Pullum even argue for the status of English *-n't* as an inflectional category (Zwicky and Pullum 1981).

 2 Besides, any analysis of a prefix like *un-* as 'negative' is at best imprecise. The meaning of *unhappy*, for example, is quite distinct from 'not happy', because *happy – unhappy* are contrary opposites: some middle ground, where neither would apply, is perfectly possible. *Happy – not happy* on the other hand are complementary opposites: where one is false, the other is of necessity true. In other words, *unhappy* is a subclass of *not happy* at the extreme end of the scale. For contradictory terms, on the other hand, incorporated and *not*-forms are roughly synonymous. Consider also the subtle difference between *un-* and *non-*: *un-* designates the contrary, *non-* the complementary opposite.

 3 On the status of *-n't* as an inflectional morpheme cf. Zwicky and Pullum (1981).

 4 I am grateful to Christian Mair for stressing this point.

 5 For details on the corpora cf. the Introduction.

 6 Cf. also Zipf's law, according to which the shortest and most frequent words are the most irregular and the oldest (Zipf 1949: 20ff.). A prime example for a highly frequent irregularity is of course the morphological paradigm of present tense BE with person distinctions that are unique in the morphology of English.

 7 Von der Gabelentz's concept of a spiral, rather than a cycle, as a model for language change might indeed be more appropriate here (von der Gabelentz 1901: 256).

 8 The finite verb is usually the basis for deciding the V component typologically; Dryer (1988), however, takes the verb stem for deciding the position of V.

 9 And indeed this is a universal phenomenon which finds expression in English in a whole range of minimizers (negative intensifiers) in object position, as, for example, relatively set phrases like *he didn't give* a damn, *he didn't see* a thing, *he didn't care* a hoot. See Spitzbardt (1957) for more examples.

10 Zwicky and Pullum (1981) regard *-n't* as an inflectional marker rather than a clitic. If we accept Zwicky and Pullum's assumed development, this would enlarge the standard English morphological inventory to nine inflectional categories – at least for some verb forms.

11 On this lengthening process cf. Lass (2000).

12 See the simplification in words like *damn* /dæm/, *damned* /dæmd/, *hymn* /hɪm/, *column* /ˈkɒləm/.

13 For further discussion of the several possible developments of *ain't* cf. also Chapter 6.

14 But see also Chapter 4 for forms of present tense BE.

15 The term is Bernini and Ramat's (1996: 117). It is equivalent to Quirk *et al.*'s *negative items* (1985: 782) and Haspelmath's *negative indefinite pronouns* (1997: 192).

16 On the difference between *-one* and *-body* forms see below. Cheshire (1999) excludes *never* from this list of negative quantifiers on the basis of its different interactional behaviour, claiming that *never* functions as an emphatic negator, not as a negative quantifier of time.

17 From the BNC.

18 More precisely, *S[ome]*-forms are glossed as 'specific existential with presupposition of existence, even in negative contexts, e.g. "John didn't see something" ⊃ "there exists some thing that John didn't see"'. *A[ny]*-forms in English under negation 'receive an interpretation of non-existence and therefore their use constitutes an alternative to negative quantifiers' (Bernini and Ramat 1996: 119).

Haspelmath notes that cross-linguistically, these quantifiers (in his terminology: indefinite pronouns) typically occur in parallel series like the English series displayed above (Haspelmath 1997: 21).

19 See the Introduction for more details on the corpora. I have added the figures for *nobody* and *no one* which are not considered by Quirk *et al.* (1985).

20 Biber *et al.* stress that this kind of logical multiple negation 'is a complex choice which requires deliberate planning. It . . . is found particularly in writing' (Biber *et al.* 1999: 179).

21 In the terminology of Bernini and Ramat (1996). Haspelmath calls this language type 'V-NI', which has 'negative indefinites that never co-occur with verbal negation, e.g. the standard English *no*-series' (Haspelmath 1997: 201), as opposed to 'NV-NI', obligatory co-occurrence of verbal negation and negative quantifier, and '(N)V-NI', where co-occurrence is not obligatory in all cases.

22 This is true for all N-quantifiers with the exception of *never*. Especially in spoken language, combination of sentence negator and *ever* are extremely rare, and *never* on its own is the more idiomatic expression.

23 Cf. Givón (1978: 82ff.) who argues that the subject should be regarded as outside the scope of the negator. This is also explicit in the analyses of Quirk *et al.*, who say that 'the scope of negation normally extends from the negative item itself to the end of the clause' (Quirk *et al.* 1985: 787).

24 We can gloss an example *x is not y* as 'it is predicated of x that y is not the case'.

25 This is a further indication that *not* is fully grammaticalized and is indeed the normal unmarked negator today. As Bernini and Ramat (1996: 100f.) make clear, this is by no means as obvious as it may sound, as nine of the investigated forty-five languages of Europe use different morphs for these constructions, among them also two Germanic ones closely related to English (German and Frisian).

26 A careful analysis of some BNC texts produces the following counterexample:

> I had *not* mentioned Sister Island, *nor* its old name of Murder Cay.
> (CCW 105)

Here, it is obviously the noun phrase in object position that is negated, but this negation is raised to sentence negation, in exact parallel to the purportedly 'ungrammatical' sentence above.

27 The term 'verb phrase' VP is used here in the narrow sense (following Quirk *et al.* 1985: 96), not in the wider sense as in generative grammar, i.e. only encompassing one or more verb constituents, not including everything that follows the verb.

In some very rare cases, different phrases, functioning as the same sentence constituent, can be co-ordinated, as in the following examples:

> Being in a concert party whose members are buying shares is *neither wrong nor against the law*. (co-ordination of AP and PP, both functioning as the subject complement; A85 35)

> *Neither today nor indeed at the Mansion House*. (co-ordination of AdvP and PP; A55 452)

> In this last case, *neither . . . nor* can even be used to co-ordinate adverbials of different functions, one of time (*today*) and one of place
> (*at the Mansion House*). This unusual co-ordination, however, results in a very marked stylistic effect.

28 Quirk *et al.* (1985: 127f.).
29 Cf. Palmer (1986) for detailed analyses.
30 Quirk *et al.* classify 'ability' as a kind of epistemic meaning, saying that ' "ability" is best considered a special case of possibility' (Quirk *et al.* 1985: 221):

> For the 'ability' sense, *can/could* may be paraphrased by use of the *be able to* construction, or in some cases by *be capable of* or *know how to*. However, the same meaning can also be approximately captured by the *be possible* construction . . . For this reason, the 'ability' meaning of *can/could* can be considered a special case of the 'possibility' meaning, *viz.* one in which the possibility of an action is due to some skill or capability on the part of the subject referent.
>
> (Quirk *et al.* 1985: 222)

This neatly makes a three-fold classification unnecessary. Palmer (1986) similarly does not consider the 'ability' or 'disposition' senses as a separate category: 'It is doubtful whether this should be included within modality at all' (Palmer 1986: 12).

31 After Horn (1989: 10). The Scholastic square of opposition was first devised for antonyms (semantic opposites) and their negated forms, and especially for the quantifiers *all – none – some*. It has been shown, however, that it can also very usefully be applied to modal meanings. The 'names' for the corners (A I E O) derive from the vowels in Latin *affirmo* (for the left-hand side) and *nego* (for the negated, right-hand side).
   The relations that hold between the various corners of this square can be informally defined as follows:

> If one of a pair of *contradictory* terms is true, then the other term *must* be false. (And vice versa: If one of the terms is false, the other must by necessity be true.)

> In a pair of *contrary* terms, both might be false at the same time, but they can never both be true. (Something cannot be at the same time obligatory and forbidden.)

> *Subcontraries* might both be true at the same time, but they can never be both false. (I can be permitted to go to the cinema, but also be permitted not to go at the same time – if my parents don't care what I do in my spare time, for example.)

> *Entailment*, finally, is also a logical relation: x *entails* y if in all cases where x is true, y is necessarily true. (Typically, therefore, y is a kind of subclass of x.)

For a more sophisticated account that also includes answers to various criticisms that have been voiced against this square cf. van der Auwera (1996). As we will not go into much detail here the simple Aristotelean square will suffice for our purposes.

32 As Horn (1978, 1989) has pointed out, the O-corner is universally not lexicalized. He gives a pragmatic explanation for this curious systematic lack.
33 Examples adapted from A4X 209, A0L 314, ADM 1909, B03 1085, AB7 2778 and AYK 719 respectively.
34 The present table is only concerned with standard English. Epistemic *mustn't* is documented for American English and indeed is quite regular in that dialect. The extent to which epistemic *mustn't* is possible and present in non-standard British English is investigated further in Chapter 4.

35 Figures are based on Dahl's sample of just over 240 languages; although he claims that 'the sample is not satisfactory for statistical generalizations' (1979: 76), they seem to be a first valid indication of very general trends. Addition of the figures might result in more than 100 per cent because some languages employ more than one negation strategy.

36 Compare this to the situation in French, where the negator *ne . . . pas*, today analysed as a discontinuous morpheme, is only just beginning to be reduced to postverbal *pas*. Standard French is thus still at stage 3, whereas spoken and informal French is at stage 4, moving towards stage 5 (Ashby 1981).

37 Indeed, all SVO-languages investigated by Dryer (1988) which permit postverbal sentence negation belong to this one language family, and Haspelmath notes that in Europe, 'all and only the V-NI languages [those languages not permitting the co-occurrence of verbal negation and negative quantifiers] have postverbal negation elements that arose by Jespersen's Cycle' (1997: 203).

38 Examples from Schlag (1997: 215).

39 Indeed, the widespread possibility of DO-periphrasis in all Germanic languages has been cited as one possible path of explanation for its emergence in English, see Rissanen (1991).

40 These figures do not add up to the total of forty-five, because this criterion does not apply to the missing six languages, which do not possess an N-quantifier. These are the Goidelic Celtic languages (Irish and Scottish Celtic), Basque, and the Finnic languages. See also Chapter 3 on the status of the Celtic languages.

3   Regional variation

1 For an outsider position cf. Kallen (1997) who claims that this early English did not die out, as is usually believed. (Kallen's position has the elegant consequence that English did not have to be reintroduced in the sixteenth century and OE survivals are therefore more of a possibility.)

2 For the situation of English in Scotland see below.

3 There is considerable disagreement over the terminology used to characterize these varieties. *Ulster Scots* and its speakers have also been classified as the 'Anglo-Irish', 'Hiberno-English' (although this term is usually reserved for English spoken in southern Ireland), 'Ulster English', 'Planter English', and 'Scotch-Irish' (especially in American terminology), and there are also voices in favour of Ulster Scots as a language in its own right. I would argue with Montgomery that Ulster Scots is a direct descendant of mainland Scots, but in the close contact with English in Ulster has developed to form one end of a continuum in Ulster today (cf. Montgomery and Gregg 1997).

4 For a systematic analysis of these terms cf. Bernini and Ramat (1996: 117ff.).

5 From Acquaviva (1996: 284, 288).

6 From Bernini and Ramat (1996: 90).

7 From Harris (1984: 305).

8 Gregg (1972) uses this clitic negative as one of the determining factors for demarcating his core Ulster Scots area.

9 The obligatoriness of negattrac is only suspended if the quantifier is stressed, or in hypotheticals, or if a subordinate clause is c-commanded by a negative or a hypothetical.

10 Numbering of the examples as in the original.

11 For a very different, grammar-internal explanation see Alison Henry's analysis in a minimalist framework. She argues that the underlying logical form for a Belfast English sentence like *Any country couldn't stand it* should correctly look something like $[_{TP}neg[_{TP}Any\ country[_{T}could[_{VP}stand\ it]]]]$; in other words, although

the NPI is not c-commanded by a negator at the sentence level, it is so in the logical form (LF). The only difference to standard English is the subject that can take a position at the sentence level different from standard English. Thus, she claims, there is 'a natural way to account for the possibilities of NPIs in subject position in Belfast English, and their impossibility in standard English. NPIs are possible in Belfast English because the subject can be in SPEC/TP at LF' (Henry 1995: 29). However, this explanation remains purely grammar internal; Henry does not give an explanation as to what might have caused this striking difference in grammar.

12 Filppula compares his Irish corpus with a corpus of educated spoken British English, where 'the average frequency of clefting was . . . just about half of the corresponding figure for Dublin and much less than a third of the figures for Clare and Kerry' (Filppula 1999: 248).

13 Example from Harris (1993: 171).

14 Cf. Chapter 1 for details of this corpus.

15 Cf. MacKinnon (2000: 44).

16 On Norn cf. Barnes (2000).

17 On Scottish Gaelic cf. in particular MacKinnon (2000).

18 Cf. Smith (2000). The status of Scots, however, is still (or again) open to debate. Claims vary between the poles of considering Scots a dialect of English to granting it status as a separate language (a sister language of English). As in the even more complex case of Ulster Scots discussed above, the decision seems to be mainly a political question. In the following descriptions, I will use the terms *Scots* and *Scottish English* interchangeably.

19 Although there are also scholars who hold that the Picts were not even of Indo-European origin; cf. Price (2000b) for a summary of the most recent (and not so recent) discussion.

20 For a graphic representation of the decline of Gaelic as a wave process cf. Withers (1979).

21 Cf. MacKinnon (2000: 45).

22 This is only to be expected if we consider Scots to be a sister language of English, even if a very closely related one, rather than a dialect of the same system.

23 Cf. Brown and Millar (1980) for all verb forms.

24 However, if the position of the sentence negator is taken up (and it is by the clitic *nae*), then the only position left for the second standard negator *no* (as these are not examples for negative concord) is not clause but constituent negation. Constituent negation can take almost any position, as Chapter 1 has shown, so that these purported differences in scope are by no means specific to Scottish English.

25 From Millar and Brown (1979: 28).

26 Also from Millar and Brown (1979: 28).

27 It should not be confused with the informal English particle *eh*, which can always be substituted by other discourse markers like *huh*, *m*, etc.

28 From Millar and Brown (1979: 33).

29 From Miller (1993: 116).

30 Also Montgomery (1998). As Montgomery has noted, in the southern United States double modals occur mostly towards a female interlocutor. Bailey and Tillery (1999), however, have shown that this is due to the Rutledge effect – the fieldworker Barbara Rutledge who conducted most of the interviews unfortunately skewed the atlas data considerably.

31 From Miller and Brown (1982: 13).

32 From Miller and Brown (1982: 12).

33 On the other hand, Wales has been called 'the first colony of an expanding English state' (Williams 1990: 19).

34 Cf., for example, Greenberg (1966). Nevertheless, Welsh is not a counter-example, as positive and negative contexts are merely equally marked.

35 Bernini and Ramat even claim that Welsh (and Breton) has 'only neg2', i.e. discontinuous negation by two negative particles (Bernini and Ramat 1996: 48). However, their own examples show that a negative sentence can be negated by the use of *nid* alone (Bernini and Ramat 1996: 12). Dahl (1979) more accurately classifies Welsh as belonging to two groups, those where negation is effected by a preverbal particle, and those where it is effected pre- as well as postverbally, i.e. by a discontinuous construction.

36 From Bernini and Ramat (1996: 33).

37 Cf., for example, Awbery (1990).

38 From Bernini and Ramat (1996: 131). The same authors detail that 'Welsh . . . possesses only the negative pronouns *neb* 'nobody' and *dim (byd)* 'nothing' and for the adverbial categories makes use of A items' (Bernini and Ramat 1996: 164). This characteristic makes Welsh rather similar to Irish again.

39 The transcribed forms are of course based on the dialect, not on the standard English phonological forms.

40 From Beal (1993: 198), who herself quotes McDonald (1985).

41 Cf. the following chance example from a linguistics (!) textbook: *All languages do not have a standard variety* (Downes 1998: 35). Intuitively, the use of this kind of construction seems to function as topic-comment structuring.

42 It has to be remembered that the SED stopped at the borders of Wales and Scotland, so that all SED maps only contain reliable information for England.

43 Cf. Chapter 8 for more details on *was* and *were*.

44 An examination of the kinds of subjects these /dn/ forms take confirms that they are indeed forms of *isn't* and *wasn't*, respectively.

45 Taylor (1997) analyses this substitution of /d/ for /z/ in a generative framework as 'stopping': 'There is a tendency among some Deep South speakers to change /z/ to [d] when followed by /n/ . . . When schwa-deletion is applied to *business* and *reasonable*, a /zn/ sequence occurs in which /z/ also becomes [d]. The derivation of *business* as [bɪːdnɪs] leads us to conclude that schwa-deletion precedes stopping' (214).

46 In the framework of Taylor (cf. the preceding note) one would perhaps have to analyse the Somerset and Devon data as resulting from a different ordering of the rules of schwa-deletion and stopping, so that words where a /zn/ sequence results from schwa-deletion are not affected by 'stopping'. This would result in a correct prediction for *idn't* and *wadn't*, and *business* as well as *reasonable*. 'Analogous' forms like *weredn't*, however, could not be accounted for, as they do not have an underlying /zn/ sequence.

4   Filling the gaps?

1 It has to be stressed that this is a feature more of general spoken English than of *non-standard* English – on the other hand, we have to assume a continuum from more standard to more non-standard speech, and a regional differentiation for the standard speakers should apply all the more to the non-standard speakers as well.

2 However, as this is mainly a software problem, the near future may see more comprehensive studies of similar phenomena. On the other hand, full noun phrases are relatively rare in spoken language anyway. Personal pronouns and existential *there* make up the bulk of subjects, and it is hoped that the major developments could in this way be documented.

3 This applies in particular to *'s not*, which was assigned to *is not* or *has not*; and *'d not*, which was disambiguated for *had not* or *would not*. From sample studies it

was judged that a contraction of *should not* to *'d not* does not occur; *'d not* was therefore never assigned to *should not*. Equally, it is impossible to assign *'ll not* with future meaning unambiguously to *shall not* as opposed to *will not*. As *shall* in general is very rare anyway and *shan't* in particular only occurs with the first person (singular and plural), it was decided to classify all instances of *'ll not* with *will not* instead.

4 All significances at p < 0.05; cf. Chapter 1.

5 The following dialect areas could be tested for statistical significance: Home Counties, Ireland, Lancashire, London, the central Midlands, the northeast and Scotland.

6 A final explanation for these unexpected results is the possibility that the BNC might not be reliably transcribed. Although the compilers of the BNC seem quite adamant that this is the case whenever studies are conducted that have not been originally envisaged, several pilot studies have shown that the BNC is much more reliable than its reputation (cf. Krug 1998). Besides, the hypothesis that the transcribers might have mistaken *he isn't* for *he's not* is not very plausible. Nevertheless, possible flaws in the basic data of course have to be taken into consideration.

7 However, the most highly ranked constraint lex ruling out the use of *amn't* ultimately rests on the unpronounceability of **amn't*: 'the absence of a pronunciation for [*amn't*] will filter it out from the candidate set' (Bresnan 2001: 37). Although this correctly describes the situation in standard English today, this is unsatisfactory from an explanatory point of view; as Hudson rightly criticizes, the usual procedure would be to fix the pronunciation, not to do without the word altogether (Hudson 2000: 298).

8 Incidentally, this procedure would also lead to wrong positive forms for all irregular negative contracted forms, e.g. *I* /wʊ/ instead of *I will*, *I* /dəʊ/ for *I do*, etc. To prevent these forms, Hudson stipulates the actual forms 'to reflect exceptional positive evidence' (Hudson 2000: 315). Contracted *I'm* can also not be derived regularly from *am* but has to be stipulated (Hudson 2000: 315).

9 Hazen's analysis is based on Noske (1982). Syllables are assigned numerical values for their syllable structure (according to a 'markedness scale' for English: the higher the figure, the more marked):

Onset	Rime	Value
C	V	0
Ø	VC	1
CC	VCC	2
CCC	VCC	3

Each syllable in addition receives an extra weighting of 1. *Isn't* in this count thus has a weighting of 6; *is not* is phonotactically better formed with a weighting of 5. *Ain't* finally is phonotactically the least marked with a value of just 3, and therefore preferred.

10 For the sake of simplicity, we will not be concerned here with other categories of number, like the dual or the paucal.

11 Croft gives examples for morphologically more complex singulars like Russian *gorox* 'peas' vs *gorošina* 'one pea', Syrian Arabic *xass* 'lettuce', *xass-e* 'one lettuce' and Turkana *ŋa-kì* 'ears' vs *a-k-it* 'one ear', amongst others (Croft 1990: 145).

12 Cf. Ferguson (1996) for a succinct summary of varieties which (stereo?)typically do without the copula: 'Normal speech, baby talk, foreigner talk and pidgins.'

It is also true that there are languages that do not possess a lexeme for 'have'. This other case is different, however. The concept of 'having' in these languages is expressed by a polymorphemic construction including the lexeme BE, for example in constructions like 'it is at me' (for example, again, in Russian, Finnish, Irish Gaelic). To my knowledge, the concept of 'having' cannot be expressed by zero. However, it is true that the copula BE in these constructions can then again be deleted, as in

> Russian *u menja kljutš*
> at me key

'I have (the) key'. This still leaves the preposition and the pronoun as possession markers.

13 It should be noted just as an aside that *copula absence* typically also includes the absence of forms of *be* used as an auxiliary, as is made explicit by Rickford *et al.* (1991: 103).

14 For a detailed discussion of copula absence in AAVE cf. the contributions in Mufwene *et al.* (1998).

15 The *OED* notes for BE that 'the primary sense appears to have been . . ., "to occupy a place" (i.e. *to sit, stand, lie*, etc.) in some specified place; thence the more abstract [sense] was derived by abstracting the notion of particular place, so as to emphasize that of actual existence, "to be somewhere, no matter where, to be in the universe, or realm of fact, to have a place among things to exist."' (*OED* headword *be* v. B).

16 This is of course the reason why we still find postverbal negation with HAVE (and even more so with BE) and this idea was our starting point in Chapter 2.

17 Example from Quirk *et al.* (1985: 225).

5   Negative concord

1 For recent analyses of OE multiple negation in a generative framework c.f., for example, Beukema and Tomić (1995), Haeberli and Haegeman (1995, 1999), van Kemenade (1999) and Ukaji (1999).

2 Jack's analysis of Middle English prose texts indicates that in early Middle English, *any*-forms are not found in negative clauses. This changed until late Middle English, where the strict rule was relaxed – at least for *any* itself, as Jack notes: 'instances in which negative concord has failed to apply to forms other than *any* (and conjunctions) are very rare, but not entirely unattested' (Jack 1978a: 70, see also Jack 1978c). Indeed, Iyeiri argues that multiple negation only reached its peak in Middle English times, and cites evidence from a corpus of verse that at that time multiple negation was a feature of formal, southern style (Iyeiri 1999).

3 Cf. the lexicalized examples for 'cancelling', i.e. logical, double negation in Latin in such forms as *non-nemo* (lit. not-nobody, 'someone'), *non-nulli* (not-none, 'some'), *non nunquam* (not-never, 'sometimes'), etc.

4 Although Jespersen claims that he has only two instances from Shakespeare, Singh's study indicates that multiple negation even in Shakespeare's texts is still relatively frequent (Singh 1973).

5 Cf. the Introduction for details on this subcorpus.

6 Cf. also Nevalainen (1998) for a justification of this narrow definition.

7 Unlike the other examples, this example looks very much like either a production or a transcription error. The second negator should in all probability be *no*, rather than *not*. The postverbal position of this word alone makes it highly unlikely that *not* could really be systematically possible here. However, this question cannot be resolved without having recourse to the original tapes.

8 Cf. note 7 above.
9 In the absence of access to the original tapes, however, any final decision depending on intonation and stress pattern has to be left open.
10 A third possibility, which is often mentioned by the compilers of the BNC to discourage studies not originally envisaged, can in my opinion be discarded for this phenomenon: transcription errors that the 'skilled audio-keyboarders' may have made, due to insufficient linguistic training. In this case I believe it is highly unlikely that a sentence like *I didn't have none* would be transcribed as *I didn't have any*.
11 Unfortunately, Cheshire, Edwards and Whittle do not investigate how far the differences they found are statistically significant.
12 The south Midlands are here included in the south. This seems justified because the south Midlands pattern so clearly with their southern neighbours East Anglia and the central southwest. Trudgill (1990: 63) also draws the main north–south dialect border north of the south Midlands. Lancashire and Merseyside are traditionally and uncontroversially classified as 'northern' areas, cf. Edwards and Weltens (1985).

6 AIN'T

1 It seems that this has not always been the case: the *OED* interestingly notes that *ain't* 'is also found as a (somewhat outmoded) upper class colloquialism' (cf. *ain't* v.1).
2 From Hughes and Trudgill (1979: 14).
3 From Hughes and Trudgill (1979: 36).
4 From Trudgill (1990: 97).
5 From Trudgill (1990: 96).
6 From Hughes and Trudgill (1979: 36).
7 The same process can be quoted for the third person form *has/hasn't*.
8 Cf. Chapter 2.
9 Although – as an aside – it has to be noted that the two areas that are reported to be virtually *ain't* free, Scotland and Ireland, are the two areas where the form *amn't* exists at all. To my knowledge, the historical (and regional) link between these two phenomena has not been examined in any satisfactory manner yet.
10 The term *simplification* or *regularization* in this context refers to a simplification of the negative paradigm only. It will be argued in the summary that overall, a more irregular system is evolving.
11 For numerous examples from phonetics and morphology cf. Greenberg (1966) and Croft (1990).
12 The number of declaratives + full questions, i.e. everything but tag questions, can be easily calculated by subtracting the tag-question figures in Table 6.1 from the totals. For rough estimates, this 'rest' of non-tags was also estimated to be largely indicative of declaratives only, as full questions are extremely rare and can therefore be disregarded in the following discussion.
13 Only Humberside is excluded from this and the following tables because it shows no occurrence of either *ain't* or *in't*. This only marginally changes the overall totals.
14 Again, Humberside has been excluded.
15 Because of the very low figures of occurrence, most areas could not be tested for statistical significance. Where figures where large enough, however (over five), i.e. in London and the central north as well as for the totals, the decrease from tag questions to declaratives was highly significant at $p < 0.01$.
16 This is a well-known phenomenon that has been investigated elsewhere – I shall not go into it in detail here, as it also entails a further reduction to *innit* which

is not conditioned by negation any more (for in-depth treatments of the current development of *innit* cf. Krug 1998 and Stenström and Andersen 1996).

17 It has to be remembered that the absolute figures for standard BE for every dialect area are based on the average frequencies of auxiliary vs copular BE, as detailed above. It is basically a conjecture that for all areas, auxiliary BE is used significantly less often than copular BE, although this is warranted by the random samples mentioned above.

18 All significance tests have of course been conducted on the absolute, not the relative, frequencies.

19 Despite this comparison, the two phenomena should not be confused. In Chapter 2 and in Chapter 4, the question was: how is HAVE as a full verb negated with or without DO? Here, the question is: of negated HAVE forms (i.e. *hasn't, haven't*), how many are full verbs?

20 On the concept of camouflaging cf. Spears (1982), Wolfram (1995).

7   Third person singular *don't*

1 And indeed the aim of the SED was to record the most traditional form, not, for example, knowledge of standard English forms.

2 Again, there are no non-standard occurrences in Humberside, which will therefore be discounted in the following discussions.

3 The assignment of dialect areas is the same as for neg concord, cf. section 'Data from the BNC', Chapter 5.

4 Although Table 7.1 and Table 7.3 give the percentages in the last columns for quick reference, all statistical tests have of course been conducted on the absolute figures. All significances at $p < 0.05$.

5 This slightly increases the percentages from the preceding tables, although not significantly.

6 Again, the sign — indicates that the construction in question does not occur at all.

8   Past tense BE

1 More precisely, the old system seems to have been one that followed the Northern Subject Rule, where both the type of subject and distance between subject and verb had an influence on the verb form. Full noun phrases as well as more distant subjects tended to have verb forms in *-s*, directly adjacent personal pronouns had non-*s* forms. (Cf. Montgomery 1994 on these two constraints for historical Scots and their – late – extension to the BE paradigm.) As my investigation, however, is limited to directly adjacent pronouns, it is correct to say that for pronouns, Ocracoke historically was a *were*-generalizing dialect.

2 Wolfram and Schilling-Estes' informant interviews incidentally also support this impression (cf. Wolfram and Schilling-Estes 1996: 140ff.).

3 The very interesting question of whether the Northern Subject Rule – or a remnant of it – still plays a role for the following patterns must be left open, as the combination of noun phrases with *was/were* can simply not be investigated systematically with the BNC at the moment. Sample studies of the smaller northern dialect areas indicate, however, that today instances of the Northern Subject Rule are extremely rare in the BNC material even in the northern area where most occurrences should be expected.

4 See also Hazen (1996) on *ain't* (cf. the discussion in Chapters 4 and 6).

9 Conclusion

1 Again, however, it must be pointed out that a decision in favour of the desig-
nation *language* rather than *dialect* is always (also) a political one; purely linguistic
criteria are not always enough to decide between the two terms. Emerging
Scottish nationalism and the process of devolution from the central English
government are political movements that point in this direction.

2 It has to be remembered that this term is also applied to the deletion of auxil-
iary *be* for English (cf. exemplarily for AAVE Rickford *et al.* 1991; Rickford
1998).

3 Mayerthaler translates his German distinction of *markiert* (marked in the func-
tional sense) vs *merkmalhaft* (morphological marking on the expression level) as
marked vs *featured* (Mayerthaler 1981, 1988; cf. also Dressler *et al.* 1987: 28ff.),
although the more usual English terms seem to be *semantic markedness* vs *formal
marking* (cf. Battistella 1990: 35ff.).

4 The absence of neg concord from the standardized language varieties at least
of Western Europe might be considered a typical feature of *Ausbau*-languages.
The historical development of the Slavic languages, however, shows that this is
not necessarily so; for the Eastern Slavic languages the reverse has been the
case: originally N1-languages were standardized as neg concord languages. How
far the model of Ancient Greek (a neg concord language) in the area of influ-
ence of the Orthodox church vs Latin (a neg-impermeable language) in the area
of influence of the Catholic church was indirectly responsible for these devel-
opments cannot be discussed in the context of this book, but looks like a tempting
hypothesis. (For areal distributions that suggest similar spheres of influence on
adverbial subordinators cf. Kortmann 1997: 267ff.)

5 Some English-based creole languages have taken this development to a much
more natural extreme with the generalization of *duon/don* to a proper invariant
negative marker for all persons and tenses (Schneider 2000: 216–17) – although
the much more widely used negator is of course *no* (Holm 1988: 171–4). This
might be stating the obvious, but the negator *don* < *don't*, where it does occur,
should of course not be confused with the marker of perfectivity *don* < *done*
which is widespread in creoles and pidgins, as described, for example, by Holm
(1988: 407).

6 It has to be stressed that the term *asymmetries* in this context always refers to
asymmetric paradigms. For a different kind of asymmetry (between positive and
negative clauses) and their motivation cf. Miestamo (2000).

7 First indications came from Tottie (1991), who found that in her material, 'the
frequency of negative expressions was more than twice as high in the spoken
texts as in the written texts' (1991: 17). In fact, she quotes absolute frequencies
of 27.6/1,000 words for spoken texts, 12.8/1,000 words for written texts. Based
on much larger corpora, Biber *et al.* basically confirm this ratio; they cite a text
frequency for negators of over 20,000 per million words for conversational texts
vs only around 5,000 to 6,000 per million words in written texts, i.e. a differ-
ence of up to factor four, cf. Biber *et al.* (1999: 159). This corresponds roughly
to more than one negator every 30 seconds in spoken language (cf. Biber *et al.*
1999: 39).

8 Extrapolating from the figures in Chapter 8 and Biber *et al.*'s observations (cf.
preceding note) one might assume a ratio of between 1:20 and 1:30 for written
texts (one negative clause for twenty to thirty positive ones).

Bibliography

Acquaviva, Paolo (1996) 'Negation in Irish and the representation of monotone decreasing quantifiers', in Robert Borsley and Ian Roberts (eds) *The Syntax of the Celtic Languages*, Cambridge: Cambridge University Press.

Altendorf, Ulrike (1999) 'Estuary English: Is English going cockney?', *Moderna Språk* 93: 2–11.

Anderwald, Lieselotte and Bernd Kortmann (2002) 'Typology and dialectology: A programmatic sketch', in Jaap van Marle and Jan Berns (eds) *Present Day Dialectology*, Vol. 1: *Problems and Discussions*, Berlin: Mouton de Gruyter.

Ashby, William (1981) 'The loss of the negative particle *ne* in French: A syntactic change in progress', *Language* 57: 675–87.

Aston, Guy and Lou Burnard (1998) *The BNC Handbook: Exploring the British National Corpus with SARA*, Edinburgh: Edinburgh University Press.

Austin, Frances (1984) 'Double negatives and the eighteenth century', in N.F. Blake and Charles Jones (eds) *English Historical Linguistics: Studies in Development*, Sheffield: Centre for English Cultural Tradition and Language (CECTAL).

Awbery, Gwenllian M. (1990) 'The geographical distribution of Pembrokeshire negatives', in Martin J. Ball, James Fife, Erich Poppe and Jenny Rowland (eds) *Celtic Linguistics*, Amsterdam and Philadelphia: John Benjamins.

Awbery, Gwenllian M. (1997) 'The English language in Wales', in Hildegard L.C. Tristram (ed.) *The Celtic Englishes*, Heidelberg: C. Winter.

Bailey, Guy and Jan Tillery (1999) 'The Rutledge Effect: The impact of interviewers on survey results in linguistics', *American Speech* 74: 389–402.

Barnes, Michael P. (2000) 'Norse and Norn', in Glanville Price (ed.) *Languages in Britain and Ireland*, Oxford: Blackwell.

Barry, Michael V. (1996) 'Historical introduction to the dialects of Ulster', in Caroline I. Macafee (ed.) *Concise Ulster Dictionary*, Oxford: Oxford University Press.

Battistella, Edwin L. (1990) *Markedness: The Evaluative Superstructure of Language*, Albany, NY: State University of New York Press.

Beal, Joan (1993) 'The grammar of Tyneside and Northumbrian English', in Jim Milroy and Lesley Milroy (eds) *Real English: The Grammar of English Dialects in the British Isles*, London: Longman.

Beal, Joan (1997) 'Syntax and morphology', in Charles Jones (ed.) *The Edinburgh History of the Scots Language*, Edinburgh: Edinburgh University Press.

Bernini, Giuliano and Paolo Ramat (1996) *Negative Sentences in the Languages of Europe: A Typological Approach*, Berlin and New York: Mouton de Gruyter.

Beukema, Frits (1999) 'Five ways of saying no: The development of sentential nega-
    tion in English in a Government and Binding perspective', in Ingrid Tieken-Boon
    van Ostade, Gunnel Tottie and Wim van der Wurff (eds) *Negation in the History of
    English*, Berlin: Mouton de Gruyter.
Beukema, Frits and Olga Mišeska Tomić (1995) 'Negation in English: A diachronic
    view', *Folia Linguistica Historica* 16: 123–35.
Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan
    (1999) *Longman Grammar of Spoken and Written English*, Harlow: Longman.
Bresnan, Joan (2001) 'Explaining morphosyntactic competition', in Mark Baltin and
    Christ Collins (eds) *Handbook of Contemporary Syntactic Theory*, Oxford: Blackwell.
Britton, Derek (1992) 'Secondary contraction and deletion of negative enclitics in
    English dialects', *Zeitschrift für Dialektologie und Linguistik* 59: 38–49.
Brown, Keith and Martin Millar (1980) 'Auxiliary verbs in Edinburgh speech',
    Transactions of the Philological Society 78: 81–133.
Chambers, J.K. (1995) *Sociolinguistic Theory: Linguistic Variation and its Social Significance*,
    Oxford: Blackwell.
Chambers, J.K. and Peter Trudgill (1998) *Dialectology*, 2nd edn, Cambridge:
    Cambridge University Press.
Cheshire, Jenny (1982) *Variation in an English Dialect: A Sociolinguistic Study*, Cambridge:
    Cambridge University Press.
Cheshire, Jenny (1999) 'English negation from an interactional perspective', in Ingrid
    Tieken-Boon van Ostade, Gunnel Tottie and Wim van der Wurff (eds) *Negation
    in the History of English*, Berlin: Mouton de Gruyter.
Cheshire, Jenny, Viv Edwards and Pamela Whittle (1993) 'Non-standard English
    and dialect levelling', in Jim Milroy and Lesley Milroy (eds) *Real English: The
    Grammar of English Dialects in the British Isles*, London: Longman.
Christian, Donna, Walt Wolfram and Nanjo Dube (1988) *Variation and Change in
    Geographically Isolated Communities: Appalachian English and Ozark English*, Tuscaloosa,
    AL: University of Alabama Press.
Croft, William (1990) *Typology and Universals*, Cambridge: Cambridge University Press.
Dahl, Östen (1979) 'Typology of sentence negation', *Linguistics* 17: 79–106.
Downes, William (1998) *Language and Society*, 2nd edn, Cambridge: Cambridge
    University Press.
Dressler, Wolfgang U., Willi Mayerthaler, Oswald Panagl and Wolfgang U. Wurzel
    (1987) *Leitmotifs in Natural Morphology*, Amsterdam and Philadelphia: John Benjamins.
Dryer, Matthew S. (1988) 'Universals of negative position', in Michael Hammond,
    Edith A. Moravcsik and Jessica R. Wirth (eds) *Studies in Syntactic Typology*,
    Amsterdam and Philadelphia: John Benjamins.
Edwards, Viv (1993) 'The grammar of southern British English', in Jim Milroy and
    Lesley Milroy (eds) *Real English: The Grammar of English Dialects in the British Isles*,
    London: Longman.
Edwards, Viv and Bert Weltens (1985) 'Research on non-standard dialects of British
    English: Progress and prospects (1)', in Wolfgang Viereck (ed.) *Focus on: England
    and Wales*, Amsterdam and Philadelphia: John Benjamins.
Elworthy, Frederic Thomas (1877) 'The grammar of the dialect of West Somerset',
    Transactions of the Philological Society 1877–8–9: 143–253.
Fasold, Ralph (1976) 'One hundred years from syntax to phonology', in Stanford
    B. Steever, Carol A. Walker and Salikoko S. Mufwene (eds) *Papers From the
    Parasession on Diachronic Phonology*, Chicago: Chicago Linguistic Society.

Fennell, Barbara A. and Ronald R. Butters (1996) 'Historical and contemporary distribution of double modals in English', in Edgar W. Schneider (ed.) *Focus on the USA*, Amsterdam and Philadelphia: John Benjamins.

Ferguson, Charles A. (1996) 'Absence of copula and the notion of simplicity: A study of normal speech, baby talk, foreigner talk and pidgins', [Original from 1971] in Thom Huebner (ed.) *Sociolinguistic Perspectives: Papers on Language in Society, 1959–1994*, New York and Oxford: Oxford University Press.

Filppula, Markku (1991) 'Subordinating *and* in Hiberno-English syntax: Irish or English origin?', in P. Sture Ureland and George Broderick (eds) *Language Contact in the British Isles: Proceedings of the Eighth International Symposium on Language Contact in Europe, Douglas, Isle of Man, 1988*, Tübingen: Niemeyer.

Filppula, Markku (1999) *The Grammar of Irish English: Language in Hibernian Style*, London and New York: Routledge.

Foulkes, Paul and Gerard Docherty (1999) *Urban Voices: Accent Studies in the British Isles*, London: Edward Arnold.

Givón, Talmy (1984) *Syntax: A Functional-Typological Introduction*, Amsterdam and Philadelphia: John Benjamins.

Glauser, Beat (2000) 'The Scottish/English border in hindsight', *International Journal of the Sociology of Language* 145: 65–78.

Görlach, Manfred (1997) 'Celtic Englishes?', in Hildegard L.C. Tristram (ed.) *The Celtic Englishes*, Heidelberg: C. Winter.

Greenberg, Joseph (1966) *Language Universals*, The Hague and Paris: Mouton.

Gregg, Robert J. (1972) 'The Scotch-Irish dialect boundaries in Ulster', in Martyn F. Wakelin (ed.) *Patterns of the Folk Speech of the British Isles*, London: Athlone Press.

Haeberli, Eric and Liliane Haegeman (1995) 'Clause structure in Old English: Evidence from negative concord', *Journal of Linguistics* 31: 81–108.

Haeberli, Eric and Liliane Haegeman (1999) 'Negative Concord and Verb Projection Raising in Old English and West Flemish', in Ingrid Tieken-Boon van Ostade, Gunnel Tottie and Wim van der Wurff (eds) *Negation in the History of English*, Berlin: Mouton de Gruyter.

Hagège, Claude (1982) *La structure des langues*, Paris: PUF.

Hansen, Klaus, Uwe Carls and Peter Lucko (1996) *Die Differenzierung des Englischen in nationale Varianten*, Berlin: Erich Schmidt.

Harris, John (1984) 'Syntactic variation and dialect divergence', *Journal of Linguistics* 20: 303–27.

Harris, John (1993) 'The grammar of Irish English', in Jim Milroy and Lesley Milroy (eds) *Real English: The Grammar of English Dialects in the British Isles*, London: Longman.

Haspelmath, Martin (1997) *Indefinite Pronouns*, Oxford: Clarendon.

Hawkins, John (1986) *A Comparative Typology of English and German: Unifying the Contrasts*, London: Croom Helm.

Hazen, Kirk (1996) 'Linguistic preference and prescriptive dictum: On the phono-logical and morphological justification of *ain't*', in Jennifer Arnold, Renée Blake, Brad Davidson, Scott Schwenter and Julie Solomon (eds) *Sociolinguistic Variation: Data, Theory, and Analysis. Selected Papers from NWAV 23 at Stanford*, Stanford, CA: CSLI Publications.

Henry, Alison (1995) *Belfast English and Standard English: Dialect Variation and Parameter Setting*, New York and Oxford: Oxford University Press.

Henry, Alison (1997) 'The syntax of Belfast English', in Jeffrey Kallen (ed.) *Focus on Ireland*, Amsterdam and Philadelphia: John Benjamins.

Holm, John (1988) *Pidgins and Creoles*, Cambridge: Cambridge University Press.

Hooper, Joan B. (1976) 'Word frequency in lexical diffusion and the source of morphophonological change', in William M. Christie Jr (ed.) *Current Progress in Historical Linguistics: Proceedings of the Second International Conference on Historical Linguistics, Tucson, Arizona, 12–16 January 1976*, Amsterdam, New York and Oxford: North Holland.

Horn, Laurence R. (1978) 'Some aspects of negation', in Joseph H. Greenberg, Charles A. Ferguson and Edith Moravcsik (eds) *Universals of Human Language*, Vol. 4: *Syntax*, Stanford, CA: Stanford University Press.

Horn, Laurence R. (1989) *A Natural History of Negation*, Chicago: University of Chicago Press.

Horn, Laurence R. and Yasuhiko Kato (2000) 'Introduction: Negation and polarity at the Millennium', in Laurence R. Horn and Yasuhiko Kato (eds) *Negation and Polarity: Syntactic and Semantic Principles*, Oxford: Oxford University Press.

Hudson, Richard (1999) 'Subject-verb agreement in English', *English Language and Linguistics* 3: 173–207.

Hudson, Richard (2000) '**I amn't*', *Language* 76: 297–323.

Hughes, Arthur and Peter Trudgill (1979) *English Accents and Dialects*, London: Edward Arnold.

Ihalainen, Ossi (1991a) 'On grammatical diffusion in Somerset folk speech', in Peter Trudgill and J.K. Chambers (eds), *Dialects of English: Studies in Grammatical Variation*, London and New York: Longman.

Ihalainen, Ossi (1991b) 'Periphrastic *do* in affirmative sentences in the dialect of East Somerset', in Peter Trudgill and J.K. Chambers (eds) *Dialects of English: Studies in Grammatical Variation*, London and New York: Longman.

Iyeiri, Yoko (1999) 'Multiple negation in Middle English verse', in Ingrid Tieken-Boon van Ostade, Gunnel Tottie and Wim van der Wurff (eds) *Negation in the History of English*, Berlin: Mouton de Gruyter.

Jack, George B. (1978a) 'Negation in later Middle English prose', *Archivum Linguisticum* 9: 58–72.

Jack, George B. (1978b) 'Negative adverbs in early Middle English', *English Studies* 59: 295–309.

Jack, George B. (1978c) 'Negative concord in early Middle English', *Studia Neophilologica* 50: 20–39.

Jespersen, Otto (1917) *Negation in English and Other Languages*, København: Det Kgl. Danske Videnskabernes Selskab. Historisk-filologiske Meddelelser.

Jespersen, Otto (1940) *A Modern English Grammar on Historical Principles*, Vol. 5: *Syntax*, London: Allen and Unwin.

Kahrel, Peter and René van den Berg (eds) (1994) *Typological Studies in Negation*, Amsterdam and Philadelphia: John Benjamins.

Kallen, Jeffrey (1997) 'Irish English: Context and contacts', in Jeffrey Kallen (ed.) *Focus on Ireland*, Amsterdam and Philadelphia: John Benjamins.

Keenan, Edward L. and Bernard Comrie (1977) 'Noun phrase accessibility and universal grammar', *Linguistic Inquiry* 8: 63–99.

Kerswill, Paul and Ann Williams (2000) 'Creating a New Town koiné: Children and language change in Milton Keynes', *Language in Society* 29: 65–115.

Kirk, John M. (1990) *Northern Ireland Transcribed Corpus of Speech*, Colchester: Economic and Social Research Council Data Archive, University of Essex.

Kirk, John M. (1992) 'The Northern Ireland Transcribed Corpus of Speech', in Gerhard Leitner (ed.) *New Directions in English Language Corpora: Methodology, Results, Software Developments*, Berlin and New York: Mouton de Gruyter.

Klima, Edward S. (1964) 'Negation in English', in Jerry A. Fodor and Jerrold J. Katz (eds) *The Structure of Language: Readings in the Philosophy of Language*, Englewood Cliffs, NJ: Prentice-Hall.

Kortmann, Bernd (1997) *Adverbial Subordination: A Typology and History of Adverbial Subordinators Based on European Languages*, Berlin: Mouton de Gruyter.

Kortmann, Bernd (1999) 'Typology and dialectology', in B. Caron (ed.) *Proceedings of the 16th International Congress of Linguists, Paris 1997*. CD-ROM. Amsterdam: Elsevier Science.

Kretzschmar, William A. and Edgar W. Schneider (1996) *Introduction to Quantitative Analysis of Linguistic Survey Data: An Atlas by the Numbers*, Thousand Oaks, CA: Sage Publications.

Krug, Manfred (1998) 'String frequency: A cognitive motivating factor in coalescence, language processing and linguistic change', *Journal of English Linguistics* 26: 286–320.

Labov, William (1966) *The Social Stratification of English in New York City*, Washington, DC: Center for Applied Linguistics.

Labov, William (1972) 'Negative attraction and negative concord in English grammar', *Language* 48: 773–818.

Lass, Roger (2000) 'A branching *path*: Low vowel lengthening and its friends in the emerging standard', in Laura Wright (ed.) *The Development of Standard English 1300–1800*, Cambridge: Cambridge University Press.

Macafee, Caroline I. and Colm Ó Baoill (1997) 'Why Scots is not a Celtic English', in Hildegard L.C. Tristram (ed.) *The Celtic Englishes*, Heidelberg: C. Winter.

McCafferty, Kevin (2000) *Ethnicity and Language Change*, Amsterdam: John Benjamins.

McDonald, Christine (1981) 'Variation in the use of modal verbs with special reference to Tyneside English', University of Newcastle: PhD thesis.

McDonald, Christine (1985) 'Notes on local syntax', Mimeo.

MacKinnon, Kenneth (2000) 'Scottish Gaelic', in Glanville Price (ed.) *Languages in Britain and Ireland*, Oxford: Blackwell.

Mair, Christian, and Marianne Hundt (1997) 'The corpus-based approach to language change in progress', in Uwe Böker and Hans Sauer (eds) *Anglistentag 1996 Dresden*, Trier: WVT.

Mayerthaler, Willi (1981) *Morphologische Natürlichkeit*, Wiesbaden: Athenaion.

Mayerthaler, Willi (1988) *Morphological Naturalness*, Ann Arbor, MI: Karoma.

Mazzon, Gabriella (1994) 'OE and ME multiple negation: Some syntactic and stylistic remarks', in Francisco Fernández, Miguel Fuster and Juan José Calvo (eds) *English Historical Linguistics 1992: Papers from the 7th International Conference on English Historical Linguistics. Valencia 22–26 September 1992*, Amsterdam and Philadelphia: John Benjamins.

Miestamo, Matti (2000) 'Towards a typology of standard negation', *Nordic Journal of Linguistics* 23: 65–88.

Millar, Martin and Keith Brown (1979) 'Tag questions in Edinburgh speech', *Linguistische Berichte* 60: 24–45.

Miller, Jim (no date) 'Negatives in Scottish English', Working paper.

Miller, Jim (1993) 'The grammar of Scottish English', in Jim Milroy and Lesley Milroy (eds) *Real English: The Grammar of English Dialects in the British Isles*, London: Longman.

Miller, Jim and Keith Brown (1982) 'Aspects of Scottish English syntax', *English World-Wide* 3: 3–17.

Milroy, Jim and Lesley Milroy (eds) (1993) *Real English: The Grammar of English Dialects in the British Isles*, London: Longman.

Milroy, Lesley (1980) *Language and Social Networks*, Oxford: Blackwell.

Montgomery, Michael (1994) 'The evolution of verb concord in Scots', in Alexander Fenton and Donald A. MacDonald (eds) *Studies in Scots and Gaelic: Proceedings of the Third International Conference on the Languages of Scotland*, Edinburgh: Canongate Academic.

Montgomery, Michael (1998) 'Multiple Modals in LAGS and LAMSAS', in Michael Montgomery and Thomas Nunnally (eds) *From the Gulf States and Beyond: The Legacy of Lee Pederson and LAGS*, University of Alabama Press.

Montgomery, Michael and Robert J. Gregg (1997) 'The Scots language in Ulster', in Charles Jones (ed.) *The Edinburgh History of the Scots Language*, Edinburgh: Edinburgh University Press.

Mufwene, Salikoko S., John R. Rickford, Guy Bailey and John Baugh (eds) (1998) *African-American English: Structure, History and Use*, London and New York: Routledge.

Nevalainen, Terttu (1998) 'Social mobility and the decline of multiple negation in Early Modern English', in M. Krygier and Jacek Fisiak (eds) *English Historical Linguistics 1996*, Berlin and New York: Mouton de Gruyter.

Nevalainen, Terttu (2000) 'Supralocalization and the rise of Standard English', in Ricardo Bermúdez-Otero, David Denison, Richard M. Hogg and C.B. McCully (eds) *Generative Theory and Corpus Studies: A Dialogue from 10 ICEHL*, Berlin and New York: Mouton de Gruyter.

Noske, Roland (1982) 'Syllabification and syllable changing rules in French', in Harry van der Hulst and Norval Smith (eds) *The Structure of Phonological Representations*, Dordrecht: Foris.

Ó Dochartaigh, Cathair (2000) 'Irish in Ireland', in Glanville Price (ed.) *Languages in Britain and Ireland*, Oxford: Blackwell.

OED. Oxford English Dictionary (1994) 2nd edn CD-ROM, Oxford: Oxford University Press.

Orton, Harold, Steward Sanderson and John Widdowson (1978) *The Linguistic Atlas of England*, London: Croom Helm.

Palmer, Frank R. (1986) *Mood and Modality*, Cambridge: Cambridge University Press.

Payne, John (1985) 'Negation', in Timothy Shopen (ed.) *Language Typology and Syntactic Description: Clause Structure*, Cambridge: Cambridge University Press.

Payton, Philip (1997) 'Identity, ideology and language in modern Cornwall', in Hildegard L.C. Tristram (ed.) *The Celtic Englishes*, Heidelberg: C. Winter.

Payton, Philip (2000) 'Cornish', in Glanville Price (ed.) *Languages in Britain and Ireland*, Oxford: Blackwell.

Petyt, K.M. (1978) 'Secondary contractions in West Yorkshire negatives', in Peter Trudgill (ed.) *Sociolinguistic Patterns in British English*, London: Arnold.

Pitts, Walter (1986) 'Contrastive use of verbal -*z* in slave narratives', in David Sankoff (ed.) *Diversity and Diachrony*, Amsterdam and Philadelphia: John Benjamins.

Pollock, Jean-Yves (1989) 'Verb movement, UG and the structure of IP', *Linguistic Inquiry* 20: 365–424.

Price, Glanville (2000a) 'Introduction', in Glanville Price (ed.) *Languages in Britain and Ireland*, Oxford: Blackwell.

Price, Glanville (2000b) 'Pictish', in Glanville Price (ed.) *Languages in Britain and Ireland*, Oxford: Blackwell.

Pryce, W.T.R. (1990) 'Language shift in Gwent, c. 1770–1981', in Nikolas Coupland (ed.) *English in Wales: Diversity, Conflict, and Change*, Clevedon: Multilingual Matters.

Pust, Lieselotte (1998) ' "I cannae see it": Negation in Scottish English and dialect data from the British National Corpus', *Arbeiten aus Anglistik und Amerikanistik* 23: 17–30.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik (1985) *A Comprehensive Grammar of the English Language*, Harlow: Longman.

Ramat, Paolo (1987) *Linguistic Typology*, Berlin and New York: Mouton de Gruyter.

Rawick, George P. (ed.) (1977) *The American Slave: A Composite Autobiography*, Westport, CT: Greenwood.

Rickford, John R. (1998). 'The Creole origins of African-American Vernacular English: Evidence from copula absence', in Salikoko S. Mufwene, John R. Rickford, Guy Bailey and John Baugh (eds) *African American English: Structure, History and Use*, London and New York: Routledge.

Rickford, John R., Arnetha Ball, Renée Blake, Raina Jackson and Nomi Martin (1991) 'Rappin on the copula coffin: Theoretical and methodological issues in the analysis of copula variation in African American Vernacular English', *Language Variation and Change* 3: 103–32.

Rissanen, Matti (1991) 'Spoken language and the history of *do*-periphrasis', in Dieter Kastovsky (ed.) *Historical English Syntax*, Berlin and New York: Mouton de Gruyter.

Schilling, Karl (1906) *A Grammar of the Dialect of Oldham (Lancashire)*, Darmstadt: G. Otto.

Schlag, Aila (1997) *PONS Reisewörterbuch Finnisch*, Stuttgart, etc.: Klett.

Schneider, Edgar W. (1989) *American Earlier Black English: Morphological and Syntactic Variables*, Tuscaloosa: University of Alabama Press.

Schneider, Edgar W. (2000) 'Feature diffusion vs contact effects in the evolution of New Englishes: A typological case study of negation patterns', *English World-Wide* 21: 33–63.

Shorrocks, Graham (1980) 'A grammar of the dialect of Farnworth and district', Sheffield: University of Sheffield.

Shorrocks, Graham (1996) 'The second person singular interrogative in the traditional vernacular of the Bolton metropolitan area', in James E. Black and Virginia Motapanyane (eds) *Microparametric Syntax and Dialect Variation*, Amsterdam and Philadelphia: John Benjamins.

Shorrocks, Graham (1999) *A Grammar of the Dialect of the Bolton Area. Part II: Morphology and Syntax*, Frankfurt, etc.: Lang.

Singh, Rajendra (1973) 'Multiple negation in Shakespeare', *Journal of English Linguistics* 7: 50–6.

Smith, Jeremy J. (2000) 'Scots', in Glanville Price (ed.) *Languages in Britain and Ireland*, Oxford: Blackwell.

Spears, Arthur (1982) 'The semi-auxiliary *come* in Black-English Vernacular', *Language* 58: 850–72.

Spitzbardt, Harry (1957) 'Negationsverstärkungen im Englischen', *Die neueren Sprachen* 56: 73–83.

Stein, Dieter (1998) 'Syntax and varieties', in Jenny Cheshire and Dieter Stein (eds) *Taming the Vernacular: From Dialect to Written Standard Language*, London: Longman.

Stenström, Anna-Brita and Gisle Andersen (1996) 'More trends in teenage talk: A corpus-based investigation of the discourse items *cos* and *innit*', in Carol E. Percy, Charles F. Meyer and Ian Lancashire (eds) *Synchronic Corpus Linguistics: Papers from the 16th International Conference on English Language Research on Computerised Corpora (ICAME 16)*, Amsterdam and Atlanta, GA: Rodopi.

Strang, Barbara (1970) *A History of the English Language*, London and New York: Methuen.

Tagliamonte, Sali (1998) '*Was/were* variation across the generations: View from the city of York', *Language Change and Variation* 10: 153–91.

Tagliamonte, Sali and Jennifer Smith (1999) 'Analogical leveling in Samaná English', *Journal of English Linguistics* 27: 8–26.

Taylor, William C. (1997) 'Rule ordering in the phonology of Alabama-Georgia consonants', in Cynthia Bernstein, Thomas Nunally and Robin Sapiro (eds) *Language Variety in the South Revisited*, Tuscaloosa and London: University of Alabama Press.

Thomas, Alan (1997) 'The Welshness of Welsh English', in Hildegard L.C. Tristram (ed.) *The Celtic Englishes*, Heidelberg: C. Winter.

Tieken-Boon van Ostade, Ingrid (1982) 'Double negation and eighteenth-century English grammars', *Neophilologus* 66: 278–85.

Tottie, Gunnel (1991) *Negation in English Speech and Writing*, San Diego, CA: Academic Press.

Tristram, Hildegard L.C. (ed.) (1997) *The Celtic Englishes*, Heidelberg: C. Winter.

Tristram, Hildegard L.C. (ed.) (2000) *Celtic Englishes II*, Heidelberg: C. Winter.

Trudgill, Peter (1974) *The Social Differentiation of English in Norwich*, Cambridge: Cambridge University Press.

Trudgill, Peter (1983) *On Dialect: Social and Geographical Perspectives*, Oxford: Blackwell.

Trudgill, Peter (1984) 'Standard English in England', in Peter Trudgill (ed.) *Language in the British Isles*, Cambridge: Cambridge University Press.

Trudgill, Peter (1990) *The Dialects of England*, Oxford: Blackwell.

Trudgill, Peter and J.K. Chambers (eds) (1991) *Dialects of English: Studies in Grammatical Variation*, London and New York: Longman.

Ukaji, Masatomo (1993) ' "I not say": Bridge phenomenon in syntactic change', in Matti Rissanen *et al.* (eds) *History of Englishes: New Methods and Interpretations in Historical Linguistics*, Berlin and New York: Mouton de Gruyter.

Ukaji, Masatomo (1999) 'On the scope of Negative Concord', in Ingrid Tieken-Boon van Ostade, Gunnel Tottie and Wim van der Wurff (eds) *Negation in the History of English*, Berlin: Mouton de Gruyter.

van der Auwera, Johan (1996) 'Modality: The three-layered scalar square', *Journal of Semantics* 13: 181–95.

van der Auwera, Johan (1998) 'Conclusion', in Johan van der Auwera (ed.) *Adverbial Constructions in the Languages of Europe*, Berlin and New York: Mouton de Gruyter.

van Kemenade, Ans (1999) 'Sentential negation and clause structure in Old English', in Ingrid Tieken-Boon van Ostade, Gunnel Tottie and Wim van der Wurff (eds) *Negation in the History of English*, Berlin: Mouton de Gruyter.

van Schaaik, Gerjan (1994) 'Turkish', in Peter Kahrel and René van den Berg (eds) *Typological Studies in Negation*, Amsterdam and Philadelphia: John Benjamins.

Viereck, Wolfgang (1989) 'In need of more evidence on American Black English: The ex-slave narratives revisited', *Neuphilologische Mitteilungen* 92: 247–62.

von der Gabelentz, Georg (1901) *Die Sprachwissenschaft: Ihre Aufgaben, Methoden und bisherige Ergebnisse*, Leipzig: Tauchnitz.

Wells, John C. (1990) *Longman Pronunciation Dictionary*, Harlow: Longman.

Wierzbicka, Anna (1998) 'Anchoring linguistic typology in universal semantic primes', *Linguistic Typology* 2: 141–94.

Williams, Colin H. (1990) 'The Anglicization of Wales', in Nikolas Coupland (ed.) *English in Wales: Diversity, Conflict, and Change*, Clevedon: Multilingual Matters.

Wilson, James (1926) *The Dialects of Central Scotland*, Oxford: Oxford University Press.

Withers, C.W.J. (1979) 'The language geography of Scottish Gaelic', *Scottish Literary Journal* 9: 41–54.

Wolfram, Walt (1995) 'On the sociolinguistic significance of obscure dialect structures: NP_i *call* NP_i Ving in African American Vernacular English', *American Speech* 69: 339–60.

Wolfram, Walt (1996) 'Endangered dialects: Sociolinguistic opportunity and obligation', in James E. Alatis, Carolyn A. Straehle, Maggie Ronkin and Brent Gallenberger (eds) *Linguistics, Language Acquisition, and Language Variation: Current Trends and Future Prospects. GURT 1996*, Washington, DC: Georgetown University Press.

Wolfram, Walt and Clare Dannenberg (1999) 'Dialect identity in a tri-ethnic context: The case of Lumbee American Indian English', *English World-Wide* 20: 179–216.

Wolfram, Walt and Natalie Schilling-Estes (1996) 'Dialect change and maintenance in a post-insular island community', in Edgar W. Schneider (ed.) *Focus on the USA*, Amsterdam and Philadelphia: John Benjamins.

Wolfram, Walt and Natalie Schilling-Estes (1998) *American English*, Oxford: Blackwell.

Wolfram, Walt and Jason Sellers (1999) 'Ethnolinguistic marking of past *be* in Lumbee Vernacular English', *Journal of English Linguistics* 27: 94–114.

Wolfram, Walt, Adrianne Cheek and Hal Hammond (1996) 'Competing norms and selective assimilation: Mixing Outer Banks and Southern', in Jennifer Arnold, Renée Blake, Brad Davidson, Scott Schwenter and Julie Solomon (eds) *Sociolinguistic Variation: Data, Theory, and Analysis. Selected Papers from NWAV 23 at Stanford*, Stanford, CA: CSLI Publications.

Wolfram, Walt, Kirk Hazen and Natalie Schilling-Estes (1999) *Dialect Change and Maintenance in the Outer Banks*, Tuscaloosa: University of Alabama Press.

Wright, Joseph (ed.) (1898) *The English Dialect Dictionary*, Oxford: Oxford University Press.

Wurzel, Wolfgang U. (1984) *Flexionsmorphologie und Natürlichkeit: Ein Beitrag zur morphologischen Theoriebildung*, Berlin: Akademie-Verlag.

Zipf, George K. (1949) *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*, Cambridge: Addison Wesley.

Zwicky, Arnold M. and Geoffrey K. Pullum (1981) 'Cliticization vs inflection: English *n't*', *Language* 59: 502–13.

Index

All parts of speech are in *italics*. Page numbers in *italics* refer to maps and figures.