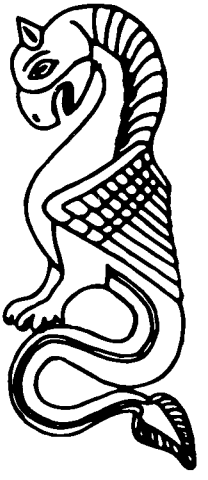
Leeds Studies in English

New Series XXX

1999

Dialectal Variation in English: Proceedings of the Harold Orton Centenary Conference 1998

Edited by
Clive Upton and Katie Wales



Leeds Studies in English
School of English
University of Leeds
1999

Leeds Studies in English

New Series XXX

School of English
University of Leeds
Leeds, England

ISSN 0075-8566

Leeds Studies in English

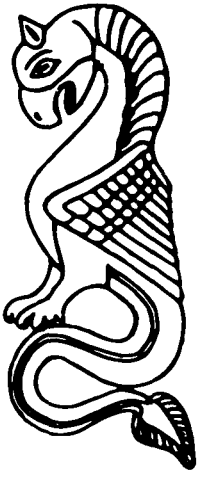
New Series XXX

1999

Dialectal Variation in English: Proceedings of the Harold Orton Centenary Conference 1998

Edited by

Clive Upton and Katie Wales



Leeds Studies in English
School of English
University of Leeds
1999

Leeds Studies in English

Leeds Studies in English is an annual publication from the School of English, University of Leeds, England. It is an international refereed journal publishing articles on Old and Middle English literature, Old Icelandic language and literature, and the historical study of the English language.

Notes for Contributors

Contributors are requested to follow the *Style Book* of the Modern Humanities Research Association, 5th edition (London: 1996). The language of publication is English and translations should normally be supplied for quotations in languages other than English. Each contributor will receive twenty offprints. Contributions from the UK should be accompanied by a large stamped addressed envelope; those from overseas should be accompanied by a large self-addressed envelope and the appropriate value of International Reply Coupons. All contributions should be sent to: The Editor, *Leeds Studies in English*, School of English, University of Leeds, Leeds LS2 9JT, England.

Reviews

Copies of books for review should be sent to Catherine Batt at the above address.

Contents

Editorial Preface

The autumn of 1996 marked a fresh beginning for English language studies in Leeds: a new Chair in English Language was taken up by Professor Katie Wales from the University of London; and Dr Anthea Fraser Gupta from Singapore, and Dr Catriona McPherson from Edinburgh were appointed to lectureships. There is no doubt that the three of them had been attracted to Leeds in no small part by the prestige and fame of Professor Harold Orton and his team and of the *Survey of English Dialects* (published 1962-71). Katie Wales in particular was delighted to find that Stanley Ellis, principal fieldworker, was an 'honorary lecturer'; that Dr Clive Upton, who had worked with Harold Orton, was a visiting lecturer, and about to embark on a new national survey; and that Leverhulme Trust finance was about to be spent on editing the Survey's original surviving recordings in machine-readable form (by Dr Juhani Klemola, temporary lecturer, and Mark Jones).

The very month she was appointed Katie Wales realised the significance of the year 1998: the centenary of Harold Orton's birth (1898-1975); and the (approximate) half-centenary of the institution of the Survey. So the idea of a celebratory international conference took shape, culminating in a 3-day event in March 1998 with over 70 participants from all over Europe and North America, including many of Orton's former colleagues and fieldworkers, as well as Harold's own daughter, Betty Borthwick. As Stanley Ellis reminded us at the conference dinner, the *SED* remains the only systematic nation-wide dialect survey ever to have been completed in England.

The aim of the conference, however, was not only to remember Harold himself, and to take stock of past achievements in English dialectology, but also to highlight current research projects in urban as well as rural dialect studies, English English as well as 'global' English: providing a *rapprochement* between dialectology and sociolinguistics; between diachronic and synchronic perspectives; theory and empirical research. The papers selected from the conference proceedings for this volume hopefully reflect a vigorous variety of perspectives and approaches. What is noteworthy is that time and again the *SED* and its related publications continue to be cited, in ever expanding fields. As Stanley Ellis also reminded us at the conference: 'Harold has a lasting memorial in the existence of the Survey, which, still, 30 years after his death, provides material that scholars will use and enlarge on for further work. He knew he was drawing a firm base line from which to build future studies.'

We wanted also to look forward to the next millennium, to future developments in dialectology. As some of the papers in this volume reveal, there is no doubt that dialectology, like other linguistic disciplines, is benefitting considerably from electronic technology: and how Harold Orton himself would have welcomed that! He would also have delighted in the media interest which the conference attracted: from interviews with almost every local radio station in the country, to the BBC's World Service, television and national broadsheets. At the end of this millennium accents and dialects have lost none of their popular appeal: if anything, non-academic interest has increased. It is even more important, therefore, that dialect studies should continue and flourish, to provide an informed and accurate picture to the community at large of variation in English.

The actual fourteen papers selected for publication here comprise four main groupings. The first group of three papers draw upon the 'Orton corpora': namely, Orton's own Northumbrian Corpus (Kurt Rydland); and the Orton-Dieth *SED* corpus. The papers of both Klemola/Jones and Willy Elmer show just how the *SED* and related materials have adapted well to the electronic age. The second grouping of four papers draws upon existing holdings of well-established dialect data, to provide fresh insights: Merja Black and Stockwell/Minkova from a diachronic perspective, Mark Jones and Davis/Houck/Horvath from a synchronic perspective, the latter looking again at LANE data. The third group, Manfred Görlach and Macafee/McGarrity, are concerned with attitudes to dialects, past and present. The final group comprise studies of urban dialects, non-traditional data (Jane Stuart-Smith; Kerswill/Williams); and also new studies with new methods (Farrar/Grabe/Nolan; Kerswill/Llamas/Upton). The volume concludes with a timely look at the future of dialectology, by William Kretzschmar.

We would like to thank Anthea Fraser Gupta and Catriona McPherson for their invaluable help in preparing this volume for publication.

Katie Wales and Clive Upton

Front Rounded Vowels in Northumbrian English: the Evidence of *The Orton Corpus*

Kurt Rydland

Abstract

This paper examines the phonological properties and the historical development of Northumbrian front rounded vowels of the types [ø(:)] ([ø(:)ə]) and [œ(:)] ([œ(:)ə]), with special reference to the material of the *Orton Corpus*. While studies based exclusively on the *Survey of English Dialects* regard these vowels as two phonemes, or perhaps one, the present paper demonstrates that they constitute three phonemes, /ø:/, /œ:/ and /œ/ (or /ɔ/), two of which have back as well as front realisations (/ø:/ = [ø: ~ o:] etc.; /œ/ = [œ ~ ɔ] etc.). In terms of lexical incidence, /ø:/ occurs typically in the standard lexical set GOAT, while /œ:/ as well as /œ/ (/ɔ/) appears chiefly in LOT and CLOTH. Special features of lexical incidence in the *Orton Corpus* include the use of /ø:/ (in North Tynedale) and /œ:/ (at most localities) in some subsets of THOUGHT.

The paper also offers explanations of the historical development of Northumbrian front rounded vowels. Lass's (1989) suggestion that they may have developed as an approach to the southern GOAT vowel is accepted in general as regards [ø(:)], but is otherwise rejected. Instead, it is shown that all three main types, [ø(:)], [œ:] and [œ], may have developed by various native processes, one of which is contextual, while two are context-free. At the same time, it is recognised that the use of front rounded vowels in the *Orton Corpus* and the *Survey of English Dialects* owes a great deal to the influence of external norms of pronunciation.

1. *Introduction*

1.1 Front rounded vowels (FRVs) are a prominent and well-known feature of

traditional Northumbrian speech. In the Basic Material of the *Survey of English Dialects* (*SED*), they appear typically in such words as *coat* [kø:t, kø'ət, kø^{ə}t, kœ'ət, kœ:^{ə}t] (~[kɷət, kọət]), *fox* [fœks, fœ'ks, fœ:ks] (~[fɔks, fɒks]) and *off* [œf, œ'f, œ:f] (~[ɔf, ɒf]) (Orton and Halliday 1962–63: IV.5.11, VI.14.5–6, IX.2.13), that is, in words belonging to the standard lexical sets GOAT, LOT and CLOTH as defined by Wells (1982). In many words of these sets, FRVs were recorded in most of the Northumbrian dialect area, which may be defined as the area represented by *SED* localities 1.1–1.9 and 3.1–3.2. The area comprises Northumberland, most of Tyne and Wear, and a small part of Durham, and is roughly coextensive with the districts which originally had the Northumbrian burr (/r/ = velar/uvular [ʁ, ʀ]) (see e.g. Påhlsson 1972). FRVs are also attested in words of other lexical sets, mainly STRUT (e.g. *gloves* [glø:vz] VI.14.7; *one* [wœn, wœ'n, wœ:n, wø:n] VII.1.1, VII.2.6, VII.8.18, IX.8.8) and FORCE (e.g. *board* [bø:d, bøəd, bœ:d] I.8.8, I.10.2, V.6.5), but such examples are rare, and seem to be confined to individual items or specific localities. Finally, FRVs occur sporadically as part of various non-centring diphthongs (e.g. [œɪl] *oil* V.2.13, [hœ'ɷl] *howl* VIII.8.11, [mɬøn] *moon* VII.6.3).

This paper examines Northumbrian FRVs of the types [ø(:)], [œ(:)] and [ø(:)ə], [œ(:)ə], that is, front rounded monophthongs and centring diphthongs with a front rounded starting-point. The principal source is the material of the *Orton Corpus*, which was recorded by Harold Orton, or under his direction, between 1928 and 1939 (see Rydland 1998). The corpus contains material from a total of thirty-five localities, thirty-three of which are in the Northumbrian dialect area as defined above. Frequent comparison will be made with the *SED* Basic Material, most of which dates from the early 1950s. The aim of the paper is twofold: first, to analyse and describe Northumbrian FRVs with regard to phonemic status, phonetic realisation and lexical incidence (see section 2); second, to trace and explain their historical origin and development (see section 3).

1.2 The geographical distribution of FRVs is essentially the same in the *Orton Corpus* and the *SED* Basic Material. Both sources record FRVs throughout the Northumbrian dialect area except in the northernmost parts of Northumberland (localities of Berwick and Cornhill) and the districts of Allendale and Knarsdale in south-west Northumberland (localities of Allendale, Allenheads and Coanwood) (see Fig. 1). It may also be mentioned that FRVs are extremely rare in the *SED* material from 1.7 Haltwhistle.

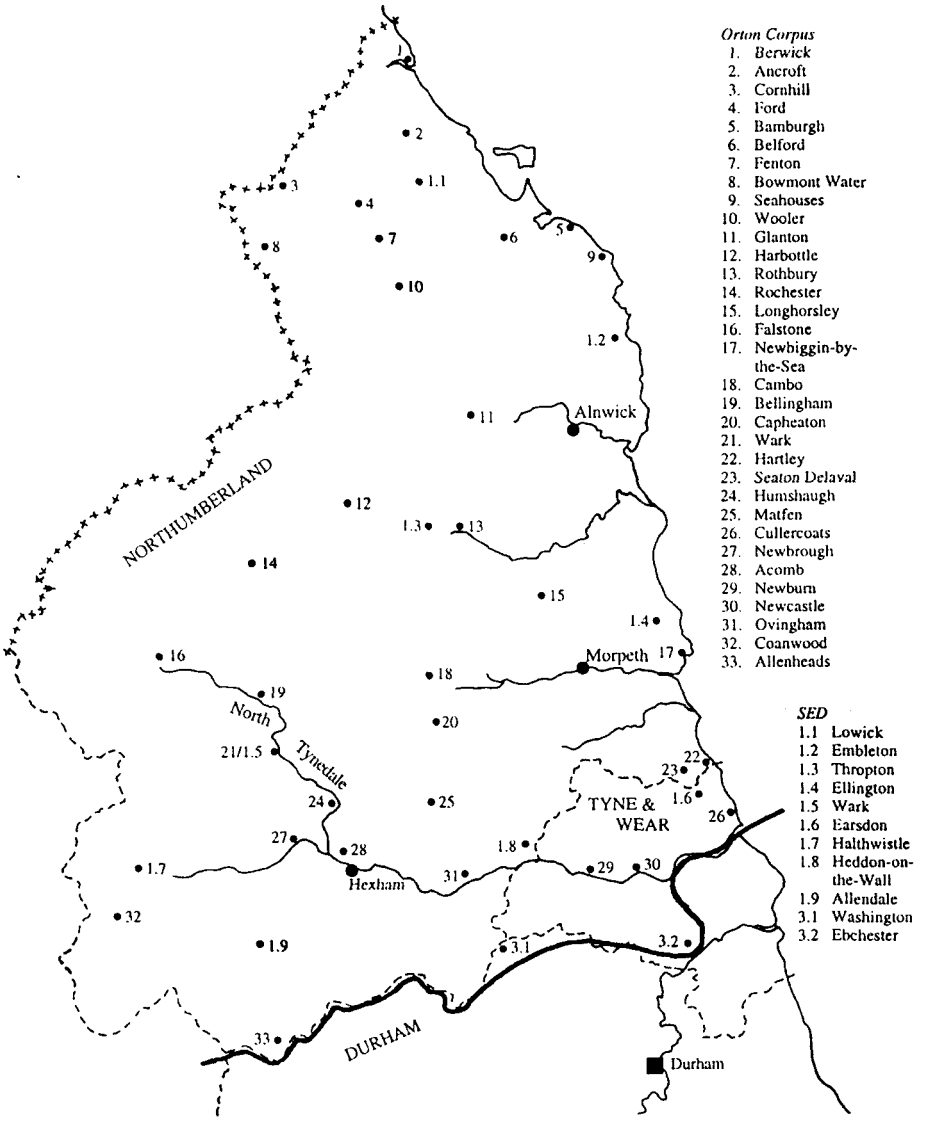


Figure 1: Northumbrian dialect area. Localities in the *Orton Corpus* and the *Survey of English Dialects*

1.3 The phonological properties of Northumbrian FRVs in the *SED* Basic Material have been discussed by Krause (1989) and Lass (1989). Krause presents a detailed analysis of the data from locality 1.1 Lowick, while Lass gives a general overview of all the Northumbrian material. Both studies distinguish two main types of FRVs, namely close-mid [ø(:)] and open-mid [œ(:)] (Lass's symbols); the quantity varies, but is usually long or half-long, less commonly short. Krause analyses them as two phonemes, /ø:/ and /œ:/ (my notation; Krause writes /øø/, /œœ/), with the phonetic variants [ø: (norm), ø', ø:ᵊ, ø] and [œ: (norm), œ', œ:ᵊ, œ]. Further, she demonstrates that /ø:/ and /œ:/ are normally used in different lexical sets: the former occurs mainly in the standard lexical set GOAT, the latter, in LOT and CLOTH, though it is also attested in a handful of GOAT words. Lass appears to take a different view: his account implies that [ø(:)] and [œ(:)] are considered as variants of the same phoneme, with the same lexical incidence. Lass also comments on the origin of Northumbrian FRVs: he suggests that [ø(:)] and [œ(:)] in GOAT may represent an approach to the GOAT vowel of RP, more specifically to the starting-point. Lass makes no attempt to explain the development of FRVs in LOT and CLOTH.

2. *Phonological analysis*

2.1 The *Orton Corpus* records the same basic types of FRVs as the *SED* Basic Material, but presents a different picture of their phonology compared to the descriptions given by Krause (1989) and Lass (1989). In the first place, the *Orton Corpus* invites a different phonemic analysis; second, it gives a wider range of phonetic variants; and third, it shows notable differences in lexical incidence.

Phonemes

2.2 A close study of Northumbrian FRVs in the *Orton Corpus* shows that they represent three different phonemes, which may be written /ø:/, /œ:/ and /œ/ (alternatively, /ɔ/; see Rydland 1995: 568). This analysis differs from Krause's description of FRVs at Lowick in that long and short open-mid [œ:] and [œ] are regarded as separate phonemes. There is no discrepancy in the treatment of the close-mid vowels: as in Krause's account, all such vowels are assigned to the phoneme /ø:/, regardless of quantity.

This phonemic analysis of Northumbrian FRVs was first proposed in Rydland 1995, which was based upon data from two localities in the *Orton Corpus*,

representing about 10% of the total material. Further study has shown that the analysis is valid for most other areas as well. Apparent exceptions are a number of localities where only one or two of the types [ø:], [œ:], [œ] are attested. However, most of these localities are very sparsely represented in the corpus, and few or none of the relevant words were recorded. The absence of one or two FRVs in the material is therefore in all probability due to chance. Open-mid long [œ:] is special in that it is lacking also in the material for two major localities (Capheaton, Glanton), and there is no evidence that the informants concerned ever used it. This suggests that at least some informants for the *Orton Corpus* had only two phonemic FRVs, namely long close-mid /ø:/ and short open-mid /œ/.

The three-phoneme analysis of FRVs in the *Orton Corpus* follows from their phonetic distribution and, above all, from their lexical incidence, as set out in Table 1. It will be observed that there is a marked incidential difference between close-mid [ø:] on the one hand and open-mid [œ:], [œ] on the other: the former occurs mainly in GOAT, the latter two, in LOT and CLOTH. This indicates clearly that [ø:] is phonemically distinct from [œ:] as well as [œ]. Both contrasts are confirmed by numerous minimal pairs such as the following:

/ø:/ ≠ /œ:/: [nø:t] *note* ≠ [nœ:t] *not*; [ø:d] *old* ≠ [œ:d] *odd*; [ø:n] *own* ≠ [œ:n (~ø:n)] *on*; [ʁø:d] *road* ≠ [ʁœ:d] *rod*.

/ø:/ ≠ /œ/: [kø:st] *coast* ≠ [kœst] *cost*; [kø:k] *coke* ≠ [kœk] *cock*; [hø:p] *hope* ≠ [hœp] *hop*; [ø:d] *old* ≠ [œd] *odd*; [sø:k] *soak* ≠ [sœk] *sock*.

Table 1. Incidence and frequency of [ø:], [œ:], [œ] in the *Orton Corpus*

	GOAT	LOT CLOTH	THOUGHT	NORTH FORCE	Misc. items	Total (tokens)	
Vowel type	Tokens %	Tokens %	Tokens %	Tokens %	Tokens %	No (approx)	%
[ø:]	**82**	1.5	14[1]	–	2.5	1540	100
[œ:]	2	**64**	9	18[2]	7	370	100
[œ]	4	**84**	0.5	1.5	10	1360	100

[1] In North Tynedale only.

[2] Only at Newbiggin-by-the-Sea.

It is perhaps less obvious that there should be a phonemic contrast between the two open-mid vowels [œ:] and [œ], both of which occur mainly in LOT and CLOTH. Many items in fact appear with both vowels, in some cases at the same locality, and even with the same informant. This alternation suggests that [œ:] and [œ] are phonemically identical. On the other hand, long [œ:] occurs also in lexical sets which do not normally admit short [œ], notably in THOUGHT. For instance, [œ:] is used in such THOUGHT words as *ball*, *law* and *daughter*. Long [œ:] here belongs to formal style: it is a local equivalent of Standard English /ɔ:/, though its similarity to the latter is not readily apparent. From a phonemic point of view, the use of [œ:] in THOUGHT is important in that it gives rise to minimal and near-minimal pairs between [œ:] and [œ] such as [ʁœ:t] *wrought* ≠ [ʁœt] *rot*, [tœ:t] *taught* ≠ [tœt] *tot* and [bœ:t] *bought* – [pœt] *pot*; [kœ:l] *call* – [kœlt] *colt*; [fœ:n] *fawn* – [fœnd (~fœ:nd)] *fond*; [sœ:s] *sauce* – [tœs] *toss*. These examples demonstrate conclusively that [œ:] and [œ] are phonemically distinct, at least with speakers who use [œ:] in THOUGHT. It may be assumed that they are different phonemes also with other speakers, as the phonemic status of [œ:] is unlikely to depend exclusively on its use in THOUGHT words.

2.3 The FRVs in the *SED* material, too, can be analysed as three phonemes. This is clear from a close examination of all the data (including unpublished incidental material from the original recording-books; see *SED* 1953–55/61) in the light of the material of the *Orton Corpus*. In accordance with their phonemic norms, the phonemes concerned may be written /ø:/, /œ:/, /ɔ/ (=[ɔ, ɒ] ~ [œ]; see further 2.5). As in the *Orton Corpus*, there are marked differences in lexical incidence between the phonetic types [ø:], [œ:] and [œ] (see Table 2), which is a strong indication that they belong to different phonemes. Further, the material turns out to contain a number of minimal and near-minimal pairs for all the contrasts. For instance, the phonemic contrast between [ø:] and [œ:] is clear from such minimal pairs as [tʃø:k] *choke* ≠ [tʃœ:k] *chock*, [kø:m (~kœ:m)] *comb* ≠ [kœ:m] *come* 'came', [nø:t] *note* ≠ [nœ:t] *not* and [ø:n] *own* ≠ [œ:n] *on*. Minimal/near-minimal pairs for the contrast [œ:] ≠ [œ] (phonemically: /œ:/ ≠ /ɔ/) include such examples as [kœ:m] *comb* ≠ [kœm (~kœ:m)] *come* 'came', [fœ:ks] *folks* ≠ [fœks (~fœ:ks)] *fox*, [plœ:t] *ploat* 'to pluck' – [lœt] *lot* and [pœ:k] *poke* 'sack' – [spœk] (~[spœ:k]) *spoke* (preterite).

In general, the three-phoneme analysis of FRVs in the *SED* material is valid for all the relevant localities, though not necessarily for all informants. Possible exceptions are localities 1.7 Haltwhistle and 3.1 Washington, where the distribution of open-mid [œ:] is such that its phonemic status cannot be determined with certainty.

Table 2. Incidence and frequency of [ø:], [œ:] and [œ] in the *SED* written material

	GOAT	LOT CLOTH	THOUGHT	NORTH FORCE	Misc. items	Total (tokens)	
Vowel type	Tokens %	Tokens %	Tokens %	Tokens %	Tokens %	No (approx)	%
[ø:]	**96**	0.5	0.5	0.5	2.5	495	100
[œ:]	23	**65**	–	4	8	350	100
[œ]	5	**90**	–	–	5	105	100

Phonetic variants

2.4 The *Orton Corpus* records several phonetic variants of each of the three phonemes /ø:/, /œ:/ and /œ/. A survey of the principal types is given in Table 3. The greatest amount of variation is found with /ø:/ and /œ/, both of which include back as well as front variants. In most sources, back variants predominate in the environment of a preceding or following burr (see 1.1), which clearly had a backing effect on front vowels (e.g. [ʁö:d, ʁo:d (~ʁø:d)] *road*; [bʁɔ̈θ, bʁɳθ, bʁɔθ (~bʁœθ)] *broth*; [bɔ̈ʁə, bɳʁə, bɔʁə (~bœʁə)] *borrow*). With /œ/, back variants are fairly common also before a following /k/ (e.g. [nɔ̈k, nɳk (~nœk)] *knock*). Front variants of /ø:/ and /œ/ prevail in all other contexts, where they occur in almost 90% of the tokens.

Table 3. Realisations of /ø:/, /œ:/, /œ/ in the *Orton Corpus* (main types)

Phoneme	Norm	Front variants	Back variants
/ø:/	[ø:]	[ø̈:, ø', ø̈', ø'ə, ø̈'ə] (infrequent) [ø] (very rare)	[ö:, ɳ:, o:, ö'ə, o'ə] (mainly after [ʁ, ʀ])
/œ:/	[œ:]	[œ'] (common); [œ̟:, œ̈:, œ̟'] (rare)	
/œ/	[œ]	[œ̟, œ̈, œ̟] (rare)	[ɔ̈, ɳ, ɔ] (mainly next to [ʁ, ʀ] & before [k])

In contrast to /ø:/ and /œ/, open-mid long /œ:/ does not have back variants. Back rounded counterparts of /œ:/, that is, [ɔ:]-type vowels, are in fact common, but obviously belong to a different phoneme, /ɔ:/. This is clear from minimal pairs such as /kœ:l/ *call* ≠ /kɔ:l/ *curl*, /kœ:d/ *cod* ≠ /kɔ:d/*cord+curd*; /œ:n/ *on* ≠ /ɔ:n/ *earn*.

The phonemic norms of all three front rounded phonemes are typically somewhat retracted from the corresponding cardinal values: the appropriate phonetic label is 'front retracted', or perhaps 'front centralised'. The phonemic norms are illustrated in Fig. 2. Since the front variants are retracted, the realisational range of /øː/ and /œ/ is somewhat less extensive than implied by the symbols in Table 3.

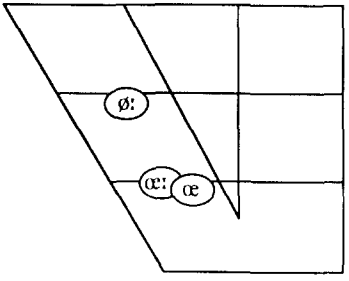


Figure 2. FRVs in the Orton Corpus: phonemic norms

2.5 The principal realisations of /øː/, /œː/ and /ɔ/ (see 2.3) in the *SED* material are set out in Table 4. It will be seen that the *SED* differs from the *Orton Corpus* as regards the frequency of back variants. Thus back variants of /øː/ are extremely rare in the *SED* material. By contrast, back vowels are the regular realisations of /ɔ/ in LOT and CLOTH: they occur in almost 90% of all the tokens, and are essentially context-free, while the corresponding phoneme /œ/ in the *Orton Corpus* shows no more than about 40% of back variants, most of which are context-dependent. The discrepancy, which is large enough to justify the use of different phonemic symbols, is indicative of a gradual shift from [œ] to [ɔ, ɒ] in LOT and CLOTH in the period between the two surveys, probably as the result of approximation to the general northern (and Standard English) vowel of LOT/CLOTH.

Table 4. Realisations of /øː/, /œː/, /ɔ/ in the *SED* material (main types)

Phoneme	Norm	Front variants	Back variants
/øː/	[øː]	[øˑə] (?norm at 3.1) [øˑ], [ø] (rare)	[oː] (very rare)
/œː/	[œː]	[œˑ, œː, œːᵊ]	
/ɔ/	[ɔ, ɒ]	[œ] (infrequent; not at 1.7)	[ɔ̈, ɒ̈, ɒ̜, ɔᵊ] (very rare)

Lexical incidence

2.6 The lexical incidence of FRVs in the *Orton Corpus* is quite similar to that in the *SED* material as described by Krause (see 1.3), but there are also important differences. The principal facts are set out in Tables 1–2. Both corpora give [ø:] chiefly in GOAT, and [œ:] and [œ] in LOT and CLOTH. The main incidential differences involve the use of [ø:] and [œ:] in THOUGHT in the *Orton Corpus*, and the occurrence of [œ:] in GOAT in the *SED*.

2.7 The *Orton Corpus* gives frequent [ø:] in THOUGHT words in North Tynedale in south Northumberland (localities of Acomb, Bellingham, Humshaugh, Newbrough and Wark). The items concerned belong to two historical subsets of THOUGHT, which will be referred to as *DRAW* and *FALL* (keywords for historical sets are written in italicised small capitals). The membership of these sets is indicated in Table 5. In the corpus material from North Tynedale, [ø:] occurs in about two thirds of all tokens of *DRAW* and *FALL*, while it is almost absent in the corresponding material of the *SED*. The *Orton Corpus* here highlights what is clearly a prominent feature of the local dialect.

Table 5. Historical subsets of THOUGHT with North Tynedale [ø:] in the *Orton Corpus*

Keyword	Northern Middle English vowel	Membership
DRAW	/au/	*claw, dawn, draw, law, sauce, saw, thaw ....*
FALL	/a/+/l/(C)	*all, ball, call, fall, hall, salt, stalk, talk, walk ....*

2.8 The *Orton Corpus* also records [œ:] in THOUGHT in many areas. This feature is absent altogether in the *SED* material. Most items with [œ:] belong to *DRAW* and *FALL*, but other historical sets are also represented. Typical examples include *fawn, haws, law, sauce, saw* (=*DRAW*); *all, ball, fall, hall, small, wall* (=*FALL*); *bought, daughter, fought, ought, sought* (words with Northern Middle English /a/ or /ɔ/ before /xt/).

2.9 The use of [œ:] in GOAT is in the main a feature of the *SED* material. Forms in [œ:] do occur in the *Orton Corpus* as well, but are very rare. If the sources are representative of actual usage, the use of [œ:] in GOAT must have increased substantially in the period between the two surveys (from less than 0.5% to almost

7% of all vowels/tokens), at the expense of [ø:] (from 45% to 40%). The change may reflect uncertainty among the *SED* informants about the original lexical incidence of [œ:] and [ø:].

3. *Historical development*

[ø:]

3.1 Close-mid [ø:] occurs almost exclusively in GOAT words, in which it replaces a variety of traditional vowels, in several historical subsets. The most important of these sets and their traditional vowels are listed in Table 6, which also shows the proportion of [ø:] in each set in the *Orton Corpus*. (The corresponding figures for the *SED* material are very similar.) The use of [ø:] for so many different native vowels suggests that it developed by an adaptive change, that is, by adaptation to an outside norm. This lends support to Lass's theory that [ø:] originated as an approach to the GOAT vowel of RP (see 1.3). It may be objected that [ø:] is an unlikely result of such approximation, and there is in fact a possible native source. This is the traditional diphthong [uə] in the historical set *COAL*, as defined in Table 6. [uə] is not very different from [ø:] phonetically, and could have produced an [ø:]-type vowel by internal fronting (probably context-free). There is no direct evidence of such a change, but it is noteworthy that [ø:] is much more frequent in *COAL* than in any other historical subset of GOAT. This suggests that [ø:] first developed in *COAL*, which in turn points to native [uə] as a possible source. Once [ø:] had become established in *COAL*, it was a prime candidate for use in other GOAT words whenever a standard-like pronunciation was needed. The development and use of [ø:] in Northumbrian dialects may thus be seen as involving two different phonological processes: the phonetic type [ø:] arose by a native change in the historical subset *COAL*, but the spread of [ø:] from *COAL* to other subsets of GOAT was the result of adaptation to Standard English.

3.2 North Tynedale [ø:] in the historical sets *DRAW* and *FALL* (see Table 5) is evidently the outcome of a native change, but is closely connected with the development of [ø:] in GOAT. The traditional North Tynedale vowel in *DRAW* and *FALL* was close-mid back [o:], which also appeared in two historical subsets of GOAT, namely *SNOW* and *COLD* as defined in Table 6 (see Ellis 1889: 640, 674–677 (Ellis's (*oo*) = IPA [o:]); compare also Heslop 1892: xix). The change from

Table 6. Main historical subsets of GOAT, traditional Northumbrian vowels and proportion of [øː] in the *Orton Corpus*

Keyword	Northern Middle English vowel	Membership	Traditional vowel	% [øː]
COAL	/ɔː/	*coal, coat, foal, nose, throat*	[uə]	76
BONE	/aː/	*bone, foam, load, most, soap*	[iə, ĭɛ]	31
TOE	/aː/	*no, so, toe*	[ĕiː]	21
ROLL	/ɔu/	*bolt, colt, gold, pole, roll*	[œu, ɔu]	19
SNOW	/au/	*blow, know, mow, snow*	[aː, æː][1], [oː][2]	18
COLD	/a/+/ld/	*cold, fold, hold, old*	[aː, æː][1], [oː][2]	20

[1] General Northumbrian.

[2] North Tynedale.

[oː] to [øː] in *DRAW* and *FALL* may have taken place as outlined in Table 7. The first stage was the emergence of [øː] in *SNOW* and *COLD*, which thus came to have front [øː] as well as back [oː]. The two vowels may have been phonemically distinct to begin with, but then underwent phonemic restructuring: at stage 2, they became variants of the same phoneme, with front [øː] as the norm, while back [oː] was mainly used next to the Northumbrian burr ([ʁ, ʀ]). This restructuring is bound to have affected *DRAW* and *FALL*, since traditional [oː] in these sets was identical with a variant of /øː/ in *SNOW* and *COLD*. The third and final stage was the necessary phonological adjustment, which involved the substitution of [øː] for [oː] in *DRAW* and *FALL*, except in phonetic environments favouring back vowels.

Table 7. Suggested development of North Tynedale [øː] in *DRAW, FALL*

Historical sets	Traditional vowel	Stage 1	Stage 2	Stage 3
DRAW, FALL	[oː]	[oː]	[oː] = /?/	[øː, oː] = /øː/
SNOW, COLD	[oː]	[øː] ~ [oː]	[øː, oː] = /øː/	[øː, oː] = /øː/

[œ]

3.3 The development of open-mid, short [œ] poses few problems. In the *Orton Corpus*, this vowel is the regular Northumbrian reflex of Middle English short back /ɔ/, except in certain phonetic environments. This indicates that [œ] derives from Middle English /ɔ/ by independent (context-free) fronting (see 3.5 for the probable date of this change).

[œː]

3.4 Open-mid long [œː] seems to have originated as a contextual variant of short [œ] in such words as *dog*, *job* and *on*, that is, before voiced consonants in monosyllabic LOT words. This is by far the commonest environment of [œː] in the *Orton Corpus*, and accounts for 95% of all occurrences in LOT. But [œː] was evidently phonemicised at an early date, probably because there were irregularities in its phonetic distribution. In the *SED* material, [œː] is usual in a much wider range of contexts, e.g. before voiceless consonants in LOT and CLOTH, including words of more than one syllable (*drop* [dʁœːpˈ] VII.8.20, *boss* [bœːs] VIII.1.25, *topping* [tœːpən] III.4.8, *offal* [œːfəl] III.11.6; see also 1.1).

The adoption of [œː] in lexical sets other than LOT and CLOTH seems to have been motivated by a number of factors. In THOUGHT, [œː] was apparently introduced as a formal variant, by speakers who considered it as the equivalent of Standard English /ɔː/ (see 2.2). The same explanation may apply to many examples of [œː] in NORTH and FORCE. The spread of [œː] to GOAT is more difficult to account for. The theory put forward by Lass, that [œː] in GOAT is an approach to the GOAT vowel of RP, is hardly tenable. A more likely cause is the native vowel correspondences in the historical sets *DRAW*, *FALL* and *SNOW*, *COLD* (see Tables 6–7). As shown in Table 8, the traditional vowels in these sets were [aː, æː] (=/aː/; see Rydland 1993: 44) in most Northumbrian dialects, while the commonest modern equivalents were [œː] in *DRAW*, *FALL* and [øː] in *SNOW*, *COLD*. The fact that [œː] and [øː] here correspond to, and alternate with, the same traditional vowels, may have created uncertainty about their lexical distribution, and thus paved the way for the use of [œː] in GOAT.

Table 8. Vowel correspondences in *DRAW, FALL* and *SNOW, COLD* (general Northumbrian)

Historical set	Modern vowel	Traditional vowel
DRAW, FALL	[œ:]	[a:, æ:]
SNOW, COLD	[ø:]	[a:, æ:]

Chronology

3.5 The FRVs of modern Northumbrian seem to be a recent phenomenon. Thus there is no trace of such vowels either in Ellis 1889 or in Wright 1905, where GOAT words appear with traditional vowels such as those listed in Table 6, and LOT and CLOTH have back [ɔ]. The closest approximation to FRVs in these sources is Ellis's (œ1ˋ) = IPA [ɐˈ] (central rounded) in *all* and *know* at 1.7 Haltwhistle (1889: 660). These forms are of questionable authenticity (the material from Haltwhistle was transcribed 'conjecturally' (Ellis 1889: 654)), and it is very doubtful if they are connected with the development of Northumbrian FRVs. If the material collected by Ellis and Wright is representative of contemporary usage, FRVs cannot have emerged until after about 1880 or 1890. However, there is reason to suspect that short [œ] may be a good deal older: this vowel is not very different from back [ɔ], and could easily have been mistaken for the latter by Wright's and Ellis's sources, most of whom had no phonetic training. In the absence of evidence to the contrary, it may be conjectured that short [œ] developed at an early date, possibly even in early Modern English. Long [ø:] and [œ:], on the other hand, in all probability did not arise until the end of the 19th century.

4. *Conclusions*

The *Orton Corpus* sheds new light on the modern phonology as well as the historical development of Northumbrian FRVs. The material shows that FRVs constitute three distinct phonemes, not two or one as suggested by other studies (see 1.3, 2.2), and provides essential data about the phonological relationship between FRVs and the corresponding back vowels (see 2.4). This information enables us to give a more satisfactory phonological analysis of FRVs in the *SED* material (see 2.3, 2.5). Further, the *Orton Corpus* reveals important features of lexical incidence not

brought out by the *SED* (see 2.6–8). Finally, the *Orton Corpus* gives significant clues to the historical development of FRVs: they appear to originate in various native developments, but their distribution is partly the result of approximation to outside models of pronunciation (see 3.1–4).

The value of the *Orton Corpus* for the study of Northumbrian FRVs lies in the insight it gives into their original use. This insight provides a clearer understanding of the phonological properties of FRVs, as well as their historical development. By the time the *SED* started fieldwork in the area, the original distribution had been modified, and the resultant patterns tend to obscure synchronic as well as historical relationships. In the matter of Northumbrian FRVs, then, the *Orton Corpus* is an invaluable supplement to the *Survey of English Dialects*.

REFERENCES

Ellis, Alexander J. 1889. *The Existing Phonology of English Dialects* (=*On Early English Pronunciation*, pt. 5). Early English Text Society (Extra Series) 56. London: Trübner & Co.

Heslop, Oliver. 1892. *Northumberland Words. A Glossary of Words Used in the County of Northumberland and on the Tyneside*. London: Kegan Paul, Trench, Trübner & Co.

Krause, Christine B. 1989. A computerized analysis of the stressed vowel systems of Lowick/ Northumberland. MA thesis, Ruprecht-Karls-Universität Heidelberg.

Lass, Roger. 1989. System-shape and the eternal return: front rounded vowels in English. *Folia Linguistica Historica* 10: 163–98.

Orton Corpus. A collection of dialect material from Northumberland, Tyne & Wear and Co. Durham. Card index and raw material (1928–1939). University of Leeds. (Published in Rydland 1998.)

Orton, Harold and Wilfrid J. Halliday. 1962–63. *Survey of English Dialects. The Basic Material*. Vol. 1, *The Six Northern Counties and the Isle of Man*. Leeds: E. J. Arnold.

Påhlsson, Christer. 1972. *The Northumbrian Burr. A Sociolinguistic Study*. Lund Studies in English 41. Lund: CWK Gleerup.

Rydland, Kurt. 1993. Three long vowels in Northumbrian English. The distribution and status of [ɑ:], [a:] and [æ:]. *Studia Neophilologica* 65: 37–55.

Rydland Kurt. 1995. The Orton Corpus and Northumbrian phonology. The material from

Bamburgh and Bellingham (Northumberland). *English Studies* 76: 547–86.

Rydland, Kurt. 1998. *The Orton Corpus. A Dictionary of Northumbrian Pronunciation 1928–1939.* Studia Anglistica Norvegica 10. Oslo: Novus Press.

Survey of English Dialects. 1953–55/61. Field recordings from Northumberland and Durham: localities 1.1–1.9, 3.1–3.2. University of Leeds.

Wells, J.C. 1982. *Accents of English.* Cambridge: Cambridge University Press.

Wright, Joseph. 1905. *The English Dialect Grammar.* Oxford, London, Edinburgh, Glasgow, New York and Toronto: Henry Frowde.

Address

Kurt Rydland
Department of English
University of Bergen
Sydnesplassen 7
N-5007 Bergen
Norway
<kurt.rydland@eng.uib.no>

The Leeds Corpus of English Dialects – Project

Juhani Klemola and Mark J. Jones

Abstract

This paper presents an introduction to the methods employed in the Leverhulme Trust funded *Leeds Corpus of English Dialects* – project. The aim of the project was to transcribe and edit the Survey of English Dialects tape-recordings for publication in machine-readable form. The paper is organised as follows: sections 1 and 2 provide some background information about the Survey of English Dialects (SED) and the SED tape-recordings; section 3 summarises the transcription conventions adopted in our project; section 4 gives some examples of how the transcription conventions were applied; section 5 discusses some examples of data culled from the tape-recordings; and section 6 offers some final remarks on the importance and value of the SED tape-recordings.

1. *Some background information about the SED*

The Survey of English Dialects is the only detailed nation-wide dialect survey which has ever been conducted in England. The idea of a nation-wide survey in England was developed by the Swiss dialectologist, Professor Eugen Dieth and Professor Harold Orton of Leeds. The fieldwork for the SED was undertaken in the 1950s in predominantly rural communities in England by 9 SED fieldworkers. The fieldworkers went through the same detailed questionnaire of over 1300 questions in 313 localities in order to collect comparable information on regional vocabulary, grammar, and pronunciation. The fieldworkers documented the informants' answers, together with other unsolicited information, into fieldworker notebooks in narrow phonetic transcription as the interviews progressed. The collected and edited transcriptions were published between 1962 and 1971 in twelve books – comprising

some 5500 pages – as the *Survey of English Dialects* (B)*: The Basic Material* (Orton et al. 1962-1971). The SED Basic Material have to date provided the data for hundreds of studies on English dialects, including eight different linguistic atlases, a dictionary, and scholarly articles and monographs (see Viereck 1991 for a selective bibliography of studies based on the SED Basic Material). Very many SED-based publications continue to be added to the list every year.

2. *SED tape-recordings*

Soon after the SED fieldwork began in the early 1950s, however, it was decided that the fieldworkers would – in addition to collecting the questionnaire data – also make tape-recordings of casual conversations with the informants. These tape-recordings were made in just under 300 of the 313 SED localities. In the SED publication programme Professor Orton lists a planned volume of phonetic transcriptions 'in both broad and narrow systems' of the tape-recorded material (Orton 1962: 21-22). This part of the SED publication programme was unfortunately never realised, and these unique tape-recordings have remained – stored in the basement of the School of English building at Leeds – for the most part unedited and untranscribed for over 40 years. The material has never been systematically used for research purposes.

The tape-recordings are in the typical dialect interview mould. Harold Orton describes the process of making these tape-recordings in the following terms:

> The material procured was never rehearsed, and, of course, never recited. It was spontaneous, and as a rule consisted of personal reminiscences or opinions, or discussed some task connected with the speaker's occupation, e.g. ploughing, harvesting, hedging, stacking, pig-killing, bread-making. The themes would crop up naturally – so it seemed to the informant – in the course of his conversation with the fieldworker, who, by further skilful management, would ensure that these informal and uninhibited remarks were secured on his tapes for permanent record.
>
> (Orton 1962: 19.)

The original recordings were made with a reel-to-reel tape-recorder. Tape in the 1950s, however, was very expensive, and thus it was not feasible to save the original tape-

recorded interviews in their entirety. In the words of Harold Orton: 'We ourselves felt unable, because of the high cost of tapes and of the lack of the appropriate storage, to preserve the tape-recordings intact. So it was decided to excerpt the best parts only and to re-record these on double sided 12 in. disks' (Orton 1962: 20).

The fact that large sections of the original tape-recordings thus had to be wiped out is, of course, a terrible loss for English Dialectology today. But, to try to find a brighter side to this sorry affair, it is highly likely that the technical quality of the selections from the original interviews transferred onto 78 rpm shellac discs has remained much better than it would have been had the original reel-to-reel tapes been kept instead.

What then are some of the advantages and disadvantages of the SED tape-recordings? In our opinion, the most important thing about the SED tape-recording corpus is that it represents

- **The only systematically collected nation-wide corpus of traditional dialect speech in mid-20th century England.**

We would further like to argue that another strength of the SED tapes is that

- **The informants are mainly NORMs (Non-mobile, older, rural, male)**

To consider this a strength may seem perverse to some, but a good case can be made for arguing that the informant choice that the SED project made is most fortunate, at least for historically orientated dialectology.

Thinking about the disadvantages the recordings have, the most serious drawback of the SED tape-recordings probably is that

- **The surviving individual recordings are relatively short, 8 to 10 minutes on average.**

Another possible disadvantage is that

- **The informants are mainly NORMs (Non-mobile, older, rural, male)**

Obviously, for sociolinguistically orientated research, the type of informant that the

SED aimed at is far from satisfactory.

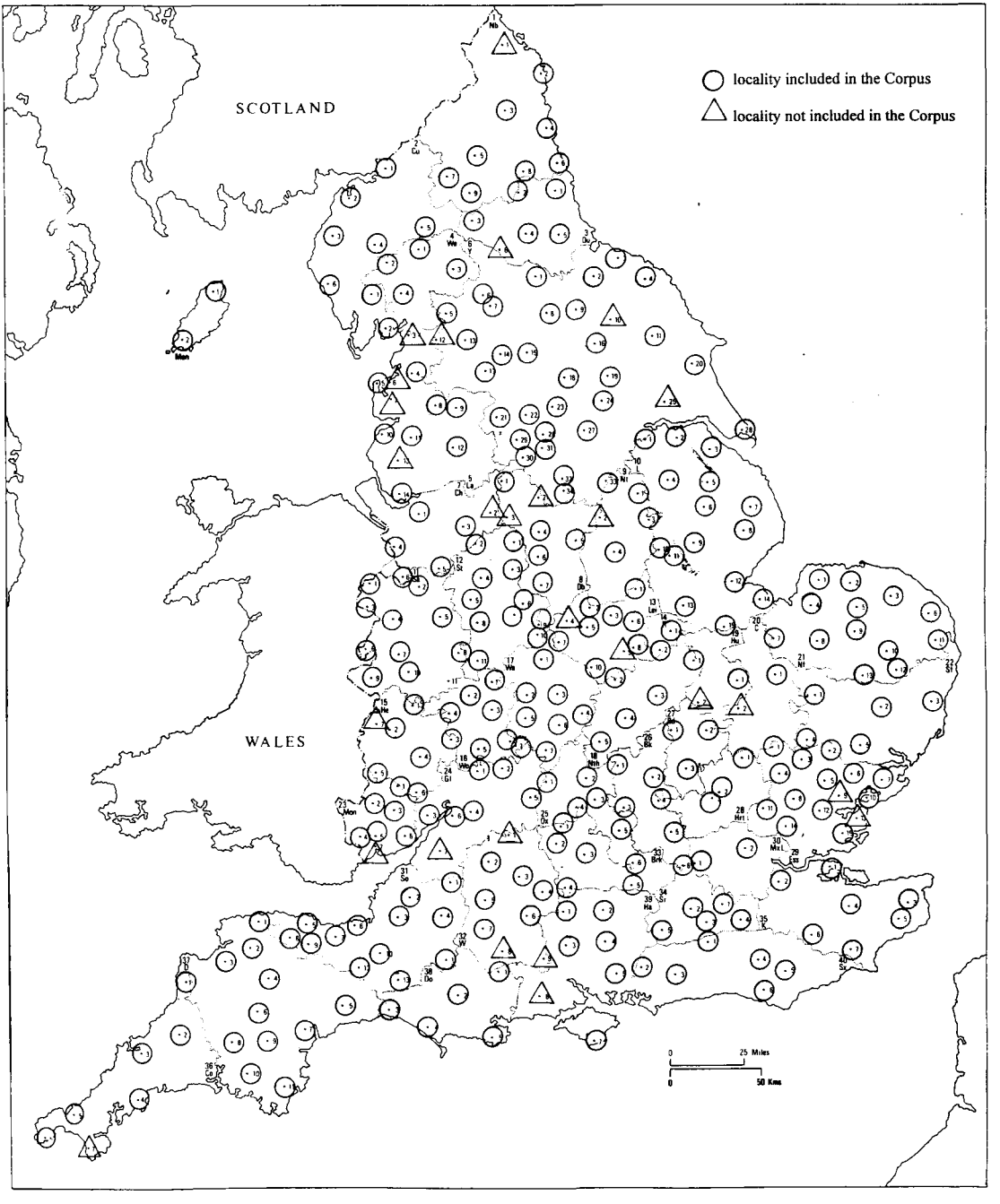
To get back to our present project, the aim of our Leverhulme Trust funded project was to transcribe and edit these tape-recordings for publication in electronic form. The project began in February 1997, when Mark Jones started his 12-month job of transcribing the recordings. The preliminary versions of the transcriptions were completed in December 1997, and the year 1998 was spent on checking the preliminary transcripts and making the necessary corrections. The transcriptions and the tape-recordings will now be published in CD-ROM format as *The SED-CDROM: The Spoken Corpus, recorded in England 1948-1973* (Klemola et al., *forthcoming*).

Map 1 gives an indication of the scope of the Corpus. The 286 circles on this map represent localities where we have a usable recording from; the 27 triangles on the map represent gaps in the recordings. In some localities, no recordings were made, and in some cases the technical quality of the recordings is not good enough for inclusion in the Corpus. The total number of SED localities was 313, the surviving 286 recordings thus represent 91% of the total number of SED localities. And, as *Map 1* indicates, the gaps in the recordings are relatively few and far between.

3. *Transcription conventions*

When transcribing any language, and non-standard language in particular, a decision has to be made as to how to present the data. A phonetic transcription made using the IPA would allow one to represent the material in a form which did justice to its divergence from the standard and presented it in terms of a universally accepted and independently interpretable system. For the purposes of large scale comparison of lexical items and morphological/syntactic features phonetic transcription has very serious drawbacks in that even broad transcriptions make comparison between the same variable in divergent varieties practically impossible. In addition, accurate phonetic transcription takes a long time, and as it was intended that the recordings themselves would accompany any textual representation on CD-ROM – with the orthographic transcriptions aligned with the sound wave – phonetic transcriptions could be made when and where they were required by anyone using the package (cf. Sinclair 1995: 102).

It was decided therefore that the best way to proceed with transcribing the dialect recordings of the SED was to transcribe them in terms of standard English orthographical practices. There are at present no universally agreed conventions for transcribing non-standard speech in this way. A system of transcription conventions



Map 1. Localities included and not included in the *Leeds Corpus of English Dialects*

(CHAT conventions; **C**odes for the **H**uman **A**nalysis of **T**ranscripts) developed for transcribing child speech in the CHILDES project (MacWhinney 1995) was chosen as the starting point for our transcription conventions on the basis of its widespread use and relative unobtrusiveness. The CHAT conventions allow the transcriber to mark pauses and other discourse phenomena in a relatively easy way, which represents the sample of speech but does not cause undue difficulty to the reader. Standard punctuation is used to broadly reflect intonational units in the recording rather than morphosyntactic ones.

We have modified the CHAT conventions for use in the *Leeds Corpus of English Dialects* – project, but even with a set of transcriptional practices as a guideline, the task is not an easy one and ongoing revisions were made. The aim of the transcription has been to present the recorded material in as unadulterated a form as possible. Over the course of the transcription, many of the original conventions have been altered, especially those which it was felt imposed an undue amount of interpretation of the data. For the ease of analysis, several innovations have been made beyond the basic CHAT system which allow dialect material to be found as easily as possible. This has not been limited to dialect lexical items, but also involves standard English words used in a particular dialectal sense (e.g. *come* as a past tense form), where these would otherwise remain undistinguished and cause users of the system many hours of unnecessary labour.

When contracted forms appear, such as *he'd* for *he would*, the contracted section is preceded by a space before the apostrophe. This allows search routines to identify all the contracted forms en masse, without separate searches having to be conducted for *he'd*, *she'd* etc. The contraction is not glossed. Genitive case markings have no space preceding the apostrophe. The following are the transcription conventions used in the *Leeds Corpus*:

+... Indicates one of three possibilities

a) The speaker has trailed off and left a sentence unfinished before another speaker speaks.
b) as made a false start and leaves the sentence unfinished, continuing with another sentence
c) speech interrupted but continues after the interruption.

+/. The speaker has been forcibly interrupted.

+"	Indicates the start of a quote (direct speech).
"+	Indicates the end of a quote (direct speech).
[/] word [\]	The word/phrase bracketed has been unintentionally repeated.
[: word]	Replace by 'word'. Marks (a) the use of a standard English form in a non-standard way, e.g. *he come* [: came] *yesterday*, etc. or (b) the use of a non-standard form with the standard equivalent in brackets e.g. *they goed* [: went] *there*.
xxx	Speech unintelligible.
[*word*]	a) Marks a dialect term, e.g. *I* [*kenn*] *him*. b) Marks a technical term, one which is used across large parts of the country with reference to a particular field of activity, e.g. the thatching term [*yelm*], a layer of cut straw, which occurs frequently across the country when thatching is discussed. Particular instances may not be synonymous.
[+word+]	Represents a dialect pronunciation of a particular word, when this is contrasted with a pronunciation reflecting the standard more closely, e.g. *No, I don't mean water, I mean* [+water+].
spl-	Phonetically identifiable section of a word uttered. Occurs mostly when speaker stutters. Not followed in these instances by repetition marker [/].
[!=a]	Indicates non-linguistic sounds affecting the conversation. Includes [!= clears throat], [!= coughs], [!= laughs], [!= sneezes], [!= sniffs], [!= spits], [!= yawns].
‖	Indicates a discontinuity in the tape-recording (The 78 rpm disks only contain extracts from the original tape-recordings, and a surviving recording may consist of several extracts from the original tape-recording).

4. *Transcription examples*

We have selected two short extracts from geographical extremes of the SED recordings to illustrate how the conventions work in practice. The first extract is from a recording from Heddon-on-the-Wall in Northumberland (SED 1 Nb 8). The recording was made by Stanley Ellis on March 21, 1953:

(1) <TM The blacksmith hoops that as +...
after you get all your fellies and things on.
You allow about # three quarters of an inch gape +...
hole # uh # [*atween*] # all the points of the [/] the [\] fellies here,
all the cuts,
you know,
[*atween*] the spokes # where the # felly joints are,
you leave them open.
xxx them fellies is all pinned and nailed you know,
and to keep them together.
Well,
that 's what they call # the allowance for the +...
when the ho- hoop 's put on,
hot,
it contracts you know. TM>
<SE Hmm. SE>
<TM They keep dousing it with cold water,
and draws all of your joints,
all your # [/] your [\] felly in. TM>

The second extract comes from a tape-recording from South Zeal in Devon (SED 37 D 6). The recording was made by Stanley Ellis on April 1, 1963:

(2) <TW You see,
[*they 'm*] all in a # clique up there.
Choose ever anybody starts to go up # on Dartmoor,
you 've had it.
Off they go.
There 's one fellow now,
at the present moment,

I can # tell you a man,
(a) young fellow,
he been trying to keep up sheep up there,
and [*they 'm*] drive 'em # all over the place.
A chap # doesn't know what to do with them.
And he 's trying to # put sheep on the moor,
but # they won't let [*en*] go there,
and they ain't going to neither. TW>
<SE And is he really entitled to do it? SE>
<TW He 's entitled so much as I am and uh everybody as in the village. TW>

5. *Some examples of data*

In order to give some indication of the range of interesting linguistic features that can be found in the tape-recordings, we would finally like to discuss some examples of data culled from the corpus.

Examples (3) to (7) give an idea of the occurrence of a typically Northern feature of verb syntax in the corpus. The construction type exemplified under (3) to (7) was coined the 'Northern subject rule' by Ihalainen (1994). Other names for the phenomenon include 'the personal pronoun rule' (McIntosh 1983) and 'Northern Present-tense Rule' (Montgomery 1994).

(3) <BE uh a lot of fellows is +...
objects to them things because they 're
the beasts when they lig down sometimes catches their knees on them and you get beasts with big knees.
But I don't that big knees are altogether # bl- +...
should all be +...
altogether blamed on that because they # [/] they [\] run hand in hand with abortions # chiefly. BE>
(We1 Great Strickland, Westmorland)

(4) <RM and uh # a lot of people thinks a # [/] a [\] pit man 's just a bloody duck egg.
But he 's not.
Oh no.

He knows his uh [/] he knows his [\] job. RM>
(Nb6 Earsdon, Northumberland)

(5) <FS But he had to pay for these chickens.
But [/] but [\] he 'll sell them him at uh +... FS>
<HS A shilling a week. HS>
<FS At four week old.
and they think they 've +...
because they don't eat much the first four weeks. FS>
(Y21 Heptonstall, Yorkshire)

(6) <DD Oh,
some takes four or five crops.
some three or four.
Some of them breeding spots.
Where they just has breeding ewes and breeds their own,
they 'll nobbut take three crops off 'em.
(Y13 Horton-in-Ribblesdale, Yorkshire)

(7) +" Keep your money in your pocket, "+
I said,
+" we don't breed 'em for you,
we breed 'em for ourselves. "+
(Y19, York, Yorkshire)

The Northern subject rule states essentially that, in the present tense, the verb takes the *-s* ending in all persons, singular and plural, unless it is adjacent to a personal pronoun subject (except for the third person singular, where the *–s* ending is used regardless of the type and proximity of the subject NP). Thus in (3) you find the inflected forms ***objects*** and ***catches*** on the second and third lines of the example, when the verb is not adjacent to a personal pronoun subject. On the other hand, when the verb **IS** adjacent to a personal pronoun subject, you find forms without the ***-s*** marker, as in ***they lig*** (line 3), ***you get*** (line 4), or ***they run*** (line 7).

Similarly, in (4), where the subject is a full NP, it is followed by a verb form in ***-s***, ***a lot of people thinks***, whereas in (5), with an adjacent personal pronoun subject, no ***-s*** marker is used in ***they think***. Examples (6) and (7), with the verb ***breed***, give further indication of the Northern subject rule operating in the Leeds

Corpus.

The phenomenon of the Northern subject rule is extremely interesting in terms of its geographical distribution, its history, and its typological rarity. The examples given here are just meant to give an indication of the fact that the SED tape-recordings provide a good source of data for the study of dialect syntax; for a more detailed discussion of the geographical distribution and history of the Northern subject rule construction, see Klemola *(in press)*.

The final set of examples, examples (8) to (10), are given here as an indication of the potential the SED tapes may have also for the study of the geographical distribution of various discourse markers in English dialects. Before we started transcribing the tapes, we were unaware of the geographical distribution of ***man*** as a discourse marker, associating it mainly with American English, especially African American Vernacular English. Therefore it was somewhat surprising to find the form being used in very traditional Northumberland speech, pronounced as /man/ rather than /mæːn/, however.

(8) as I say,
the [/] the [\] working man was just a bloody slave,
man.
Aye.
(Nb6 Earsdon, Northumberland)

(9) The women never had no time to gan in,
oh no,
the women's work was never done,
man,
poor buggers.
(Nb6 Earsdon, Northumberland)

(10) Badger.
Oh,
there 's no dog can kill a badger,
man.
(Nb3 Thropton, Northumberland)

A preliminary analysis of the geographical distribution of the use of *man* as a discourse marker in the recordings indicates that this phenomenon is very definitely a

feature of the upper North. The 75 instances of *man* in this function found in the Corpus were confined to the Northern counties of Northumberland, Cumberland, Durham, and Westmorland.

6. *Final remarks*

We believe that the completion of the *Leeds Corpus of English Dialect* project will open up an exciting field of further research. The availability of the *The SED-CDROM: The Spoken Corpus, recorded in England 1948-1961*(Klemola et al., *forthcoming*) will make it possible for the first time to analyse an extensive corpus of spontaneous, comparable tape-recorded (English English) dialect speech and thus to study the regional variation in English dialect morphosyntax on a nation-wide scale. Access to the corpus of SED-tape-recordings will make it possible to study such areas as the syntax of negation, relative clauses, pronoun systems, etc., and thus to make important new discoveries about the grammatical structure of traditional dialects of English English. Combined with other historical corpora available, such as the *Helsinki Corpus of English Texts* and the *Corpus of Early English Correspondence*, it will also be possible to use the results of the synchronic analysis as the basis of a diachronic study of the morphosyntactic properties of non-standard vernacular varieties of early Modern English.

The SED-CDROM: The Spoken Corpus, recorded in England 1948-1961 will be an extremely valuable source of data not only for dialectological research, but also for sociolinguists, historical linguists, and phoneticians interested in the study of language variation and change. Furthermore, it is expected that the *Corpus* will also be of interest to oral historians, to schools as source material for English language A-level courses, and to the general public interested in the richness of the English language.

ACKNOWLEDGEMENT

We gratefully acknowledge the support of the Leverhulme Trust (grant no: F/122/AT) which made possible the work reported here.

REFERENCES

Ihalainen, Ossi. 1994. The dialects of England since 1776. In Robert Burchfield (ed), *The Cambridge History of the English Language. Volume 5: English in Britain and Overseas. Origins and Development,* pp. 197-274. Cambridge: Cambridge University Press.

Klemola, Juhani. in press. The origins of the Northern Subject Rule: a case of early contact? In Hildegard Tristram (ed), *Celtic Englishes II.* Heidelberg: Carl Winter Verlag.

Klemola, Juhani et al. forthcoming. *The Survey of English Dialects on CD-ROM: The Spoken Corpus Recorded in England 1948-1961.* London: Routledge.

McIntosh, Angus. 1983. Present indicative plural forms in the later Middle English of the North Midlands. In: Angus McIntosh, Michael L. Samuels, and Margaret Laing (eds.), *Middle English Dialectology: Essays on Some Principles and Problems*, pp. 116-22. Aberdeen: Aberdeen University Press.

MacWhinney, Brian. 1995. *The CHILDES Project: Tools for Analyzing Talk.* Second Edition. Mahwah, NJ: Lawrence Erlbaum.

Montgomery, Michael B. 1994. The evolution of verb concord in Scots. In Fenton, A. and D.A. MacDonald (eds), *Studies in Scots and Gaelic: Proceedings of the Third International Conference on the Languages of Scotland*, pp. 81-95. Edinburgh: Edinburgh University Press.

Orton, Harold. 1962. *Survey of English Dialects* (A): *Introduction.* Leeds: E.J. Arnold; repr. Routledge, London, 1998.

Orton, Harold, Michael V. Barry, Wilfrid J. Halliday, Philip M. Tilling, and Martyn F. Wakelin (eds). 1962-1971. *Survey of English Dialects* (B): *The Basic Material* (4 vols in 12 parts). Leeds: E.J. Arnold; repr. Routledge, London, 1998.

Sinclair, John. 1995. From theory to practice. In Geoffrey Leech, Greg Myers, and Jenny Thomas (eds), *Spoken English on Computer: Transcription, Mark-up and Application*, pp. 99-109. Harlow: Longman.

Viereck, Wolfgang (in collaboration with Heinrich Ramisch). 1991. *The Computer Developed Linguistic Atlas of England 1.* Tübingen: Max Niemeyer Verlag.

Addresses
Juhani Klemola
Department of English
University of Helsinki
PO Box 4 (Yliopistonkatu 3)
FIN-00014 University of Helsinki
Finland
<juhani.klemola@helsinki.fi>

Mark J. Jones
51 Lydgate Hall Crescent
Crosspool
Sheffield S10 5NE
England
<markjjones@hotmail.com>

The Phonetic Database Project (PDP) – A New Tool for the Dialectologist

William Elmer

Abstract

The *Basic Material* of the *Survey of English Dialects* is rightly admired as one of the richest sources for the study of dialect variation, represented most visibly in a number of linguistic atlases. Despite their undisputed value, these atlases find natural limits where the *generalization* of linguistic features and processes is concerned; also, they do not lend themselves easily to interaction with the researcher. These reasons (among others) stand behind our attempt at converting the complete *Basic Material* (BM) into electronic form, in order to allow fast access to its more than 450'000 phonetic entries, through a small number of search routines, with the results presented in the form of lists and maps. The *Phonetic Database Project* (PDP) hopes to have come some way towards this aim. Its format, potential and limits are illustrated, resulting in the conclusion that, while it relieves the dialectologist of much laborious manual work, the imaginative part still has to be played by the linguist, in a yet to be fully established type of interaction with the computer.

One of the clearest signs of the substance of a scholar's work is seen when future generations turn to it again with new questions and expectations, generated by the development of the field. The unbroken interest in the *Survey of English Dialects* (SED) – the joint project of Harold Orton and Eugen Dieth – is proof of the fruitful nature of this work. My contribution towards a conference dedicated to the memory of Harold Orton intends to suggest one way in which the SED could inspire dialectologists to continue work in a domain shaped by more than thirty years of study as well as to chart new directions for future research.

As to the legacy of work based on the SED, it is hardly necessary here to

mention its varied but still focussed nature, determined quite significantly by the format in which the data are presented in the BM. Insofar as the representation of SED data in linguistic atlases is concerned, I believe that, while there is now quite a number of them, representing different orientations, the limits of atlas-based areal dialectology seem to have been reached. The main reason for this situation has to do with the generalization of linguisitc features or processes. The most obvious initial mode of representation which the arrangement of the BM entries (according to phonological, morphological, syntactic and lexical keywords, counties and localities) seems to suggest, is the single map. This is the basic format of the *Linguistic Atlas of England (LAE)* (Orton et al. 1978), the *Atlas of English Sounds (AES)* (Kolb et al. 1979) and practically all of their companions, and it is unnecessary to repeat here that work based on this type of evidence has yielded rich results. Still, it is true that this orientation has a tendency to favour 'interesting' individual maps (even if arranged in groups) – a fact which may derive from the very format of an atlas: despite the ingenuity of editors, the problem of selection remains a question of interests and priorities. The situation is reflected by the widespread experience by dialectologists of often being 'one map short'. Generalizations towards the overall behaviour of sounds and forms are not easily made in this context – both for reasons of data accumulation and presentation on a map. The search for more general or even 'global' patterns – based on the *complete* data offered by the BM – is not practically feasible at present. The main reason for this is the time-consuming nature of work in this area. The BM contains about 450'000 phonetic entries, and many of them occur not only under their respective keywords, but also elsewhere ('hidden forms'), which makes them difficult to find. (If you intend to study the areal occurrence properties of palatalization, velarization, vocalization, (un)rounding etc., the amount of laborious manual work is multiplied: short of going through the complete material, such aspects cannot be investigated systematically in the BM). In addition to this, another restriction conditions our work: this is the necessarily reductive selection that has to be made at the outset of any investigation. Once a group of sounds or items is defined, it cannot be changed without considerable additional toil. (Imagine a context-dependent phonological investigation of the relative effect of different consonant clusters on the lengthening or shortening of Middle English vowels, where it may be important to be able to extend or narrow the database in the light of intermediate results). All of us who have ever tried to draw maps of seemingly clearly patterned and suggestive data have experienced that at different stages in our work we wanted to modify the pathway of inquiry, a situation which reflects the inherently interactive nature of dialectological work; however, under the constraints of manual work this is again just not possible.

The answer to the problem of accessing and handling BM data efficiently and accurately lies in their conversion into electronic form. This obviously appealing idea has so far failed to realize because of the narrow phonetic transcription used in the BM as well as of its format. When we decided at the English Seminar of the University of Basel to tackle this problem, our aims were therefore relatively modest: we wanted to make the complete SED material available on computer, in its original form (i.e. without any loss of phonetic detail). Reasonably simple access should be combined with a number of basic search routines, and the results represented in the form of lists and maps. We wanted to offer a tool for the working dialectologist, to relieve him/her from much mundane work and in particular make possible the kind of flexible research sketched before.

In this, I am privileged to have had three collaborators in my department who complemented each other ideally and who together shaped the project which I am presenting here: Michael Gasser scanned the complete BM, Ernst Rudin designed the computer format (the code into which the phonetic transcription has to be converted and reconverted again, among other things), and wrote all the programs, Guy Schiltz contributed the cartographic module. All of this together combines into PDP, the *Phonetic Database Project*.

So we now have the entire BM as a database (in ASCII format) on our computer. This work took some time, and it presented us with an exciting mixture of problems, not least that of coding the phonetic transcription. The lasting experience, however, was not shaped by such difficulties, but by another aspect: as work progressed, we developed the highest respect for the fieldworkers who actually heard and transcribed what a sophisticated computer would often just not swallow. Soon we realized that that fieldworkers like Peter Wright and Stanley Ellis naturally did not have a digitalized version of their magnificent transcriptions in mind. (Should I admit that sometimes we asked ourselves whether they would have heard the same wonderful dialect sounds had they been faced with devising a system for their encoding). Once we had a grip on the narrow phonetics and the wide variety of diacritics, we had to make accessible the sophisticated ways in which editorial information is given in the BM. It took some time to convince the computer that an unfilled circle means 'incidental material', that anything in square brackets is 'illustrative material' and – most confusingly for a device based on the binary principle – that some signs are used with variable functions, a fact which may have led Harold Orton to consider such information at times 'to be somewhat empirical' (Orton 1962:23).

PDP cannot replace the BM; it differs from the real thing in that you cannot browse in it but have to ask it questions, exact questions in a format prescribed by the

program. Searches are essentially for the occurrence of sounds or forms (phonetic or orthographic), context-dependent or context-free. The results are then shown as lists or maps. PDP is not an atlas! Our maps are not stored, but generated from the results of a search. An important aspect is speed: PDP checks any search unit through the complete BM in a matter of a few minutes. Here lies its potential as a working tool: it allows fast and flexible access to the SED in the domains mentioned. It allows us to *ask* the BM questions. Let us therefore investigate some aspects of *ask* (and phonologically related words) in English dialects with the aim of illustrating what PDP can do – its potential and its limits. The point here is to show the type and range of applications of our program, not to perform a systematic study of the areal variation properties of *ask*.

Ask is a lexical (not a phonological) SED keyword, and the best way of getting an initial idea of its occurrence is of course to look it up in the four BM volumes. However, should the green books not be available, PDP offers us two approaches for a first look at the data: we can either consult the keyword itself (ASK HIM, IX.2.4) or we can search for *ask* in the responses section. While the former search is limited to responses to IX.2.4, the latter also finds occurrences of *ask* elsewhere, including the *incidental material*. Let us thus enter the lexical searchstring *ask*. PDP answers our question as in Figure 1.

What PDP tells us, then, (not surprisingly) is that *ask* occurs as an answer to question IX.2.4, and it also tells us that the word is not found elsewhere. The program now creates a file of the responses to this question and this is the database for our searches. We move now to the phonetics section of the BM; for any search, PDP looks at the phonetic entries listed under the keyword *ask* for occurrences of the searchstring in question. In the present case, we expect the two most interesting parameters of variation to be quantity [ask/a:sk] and quality [ɑ:sk/a:sk]. Let us start with the short vowel. Our question then is: where is *ask* attested as [ask], i.e. with a short low front vowel? – In a matter of less than a minute, the **ask** map (short vowel) occurs on the computer screen: see Figure 2, Appendix.

The form [ask] is attested in all the filled cells. This view of England is produced from the following base map, made by Guy Schiltz: see Figure 3, Appendix.

What Guy did was to take the original map of localities and have the computer figure out the equi-distance from one point to all the surrounding ones (so-called Thyssen polygons, a cartographic method common in dialectometry), so that now we have these 313 cells. If a searchstring is attested in a locality, the corresponding cell is filled. (Note that on maps of this type, PDP shows the occurrence properties of search units, in our case a phonetic variant of the SED keyword *ask*.)

```
(File created by QRScan on 23. 1.1999 at 12.47.
                              Copyright e.r. 1997)

PHONETIC DATABASE PROJECT          English Seminar
      UNIVERSITY OF BASEL

LIST OF FINDINGS in the RESPONSES SECTION.
Original Filename = FL.Q2

Keyword Pattern: *

SEARCHSTRING: ask



N92-4.EXT; IX.2.4 ASK HIM
      ASK

E92-4.EXT; IX.2.4 ASK HIM
      ASK
      ASK

W92-4.EXT; IX.2.4 ASK HIM
      ASK

S92-4.EXT; IX.2.4 ASK HIM
      ASK
           .
           .
           .
```

Fig. 1 List of findings (ask)

* stands for a search through the complete material (in all four regions). If we wanted to limit a search to the north, we would have an N* here. N92-4 is the filename IX.2.4 for North, East, West and South, ASK HIM is the keyword, and then follow the responses.

PDP does not produce symbol maps of the type we know – showing a keyword and its range of variation. No computer program can do this – or will, for some time. This part of our work still has to be done by the dialectologist (fortunately, some would say), but I hope to show that it can now be done in a well-informed and efficient way. For the *ask*-map, PDP has checked the occurrences of [ask] in the four files and found it 73 times in 73 localities. The overall area in which [ask] is attested agrees only in part with what we expect. There is an abundance of forms in the East, but the Western half of the North is still underdetermined: the unfilled cells only mean that [ask] does not occur. We should therefore search for other variants, notably the two other front vowels [æ] and [ɛ]. The search here finds 8 occurrences of [æsk] and 4

occurrences of [ɛsk]. Instead of adding these forms to the first map, we can combine the three searches, either by entering the three searchstrings or (more elegantly in our case) by searching for the occurrence of short vowels, including the variants [as] and [ast] (for 'asked'). Such cases are quite comfortably catered for by our program, by the possibility of using wildcard characters ('placeholders') in the searchstring.

Our search for long vowel occurrences of *ask* proceeds in the same way: first, the long low front vowels. As we do not at this stage want to distinguish between them, we do a combined search. See Figure 4 **ask** (long front vowel).

Figure 4 shows that there are 66 findings of these two vowels; this is an interesting intermediate working map which will find its fuller interpretation in a wider context. Having made sure that there are no occurrences of [ɛ:], we turn back to the variant [ɑ:sk]. Figure 5 is the map, again awaiting fuller evidence (long back vowel).

All of this is exactly what you see on the screen. We could now superimpose this onto the long front vowel map to show the distribution of lengthening in the context of *ask*. But we could also ask PDP for the same map directly, entering the searchstring *long A-type vowel* (i.e. front and back variants); note that half-long forms are not included, although this would have easily been possible. Equally, we could contrast this map with that for the short vowel in the same context by superimposing the two pictures.

We now have a first impression of the main distributional patterns of *ask*, and we have also experienced the kind of flexible work which PDP allows. We can now focus on some interesting aspect or we can generalize our investigation to other contexts. We are still in the domain of the single word *ask*. This means that as dialectologists and historical linguists we expect to find metathesis here. Let's see what PDP says. The answer to our searchstring [aks/æks] and [ɛks] (respectively *short low vowel+sk*) is Figure 6, **aks** (metathesis). If anything, this map answers the question about the West, which remained largely blank on the short vowel map. Combining these two pictures we can see in Figure 7 (**aks** vs. **ask**) how they complement each other.

All of the searches illustrated so far are efficient and accurate, and they show how PDP works. But as mentioned, its real interest lies in generalizing work. Our next step then is to find 'all occurrences of the (orthographical) sequence -ask-'. Again we first establish the complete list of relevant lexical items. PDP scans all the responses sections in the BM for occurrences of *ask*-words, and what we get on the screen is Figure 8 (a list of findings ***ask***).

1 – again the filename corresponds to the questionnaire number.

2 – II.2.10a is the SED keyword 'cowslip'.
3 – III.5.4 shows the keyword 'basket'.
4 – IV.9.8 gives *ask* as the word for 'newt', which we delete from the list.
5 – V.9.7 'clothes-basket': here the answer 'wisket' seems to interfere with our search for short *ask*-forms; in order to exclude findings of 'wisket', we just limit our search to *A-type vowel*, as shown before.
6 – IX.2.4 finally is our keyword *ask*. Thus, in addition to this, we have found 6 further *ask*-words, 3 of which are keywords themselves. Then follow the files for East, West and South. Filtering down the list to the *ask*-words proper establishes our new database for the searches we have already applied to the single keyword *ask*. The database now comprises 26 files, compared to 4 files for the single word. Of course not all words are evenly attested; we have completeness here in terms of inventory, not geographical coverage. Let us again look first at the short front vowels [a/æ/ɛ]. We are given their occurrence by a single search: see Figure 9 ***ask*** (short vowel).

Comparing this map with the same map for the single word *ask*, we note a considerable extension of the short vowel area, Figure 10: ***ask*** vs. **ask**.

The short front vowels are attested in 138 localities. We again note the unfilled regions in the North-West, but this time they have a different status: this evidence is based on all *ask*-words in the BM, which means that there are no more short forms of this type to be found. Next, the long variants [a:] and [æ:], in Figure 11: ***ask*** (long front vowel).

Again the picture is now considerably more complete, including the unfilled area in the South-East. The search for back [ɑ:] completes the picture: the South-East is now filled, and there is quite a coherent [ɑ:]-area in Norfolk: see Figure 12, ***ask*** long back vowel).

And Figure 13 is a first view of the complementary pattern formed by the occurrence of long front and back variants: ***ask*** (front vs. back long vowel).

In agreement with the evidence we find in the LAE and elsewhere, the South-East stands out as a compact [ɑ:]-region. Notice the hint of variation in East Anglia. We can now again superimpose the short and long vowel areas: see Figure 14 ***ask*** (long vs short vowel).

Note the astonishingly clear-cut pattern, with metathesis in the North-West and isolated examples of variation in the Southern area. (We can study its particular nature by looking at these localities individually, cf. below). At this point PDP has done its job (for the moment). The maps are now ready for comparison, scrutinizing and interpretation – as well as for flexible re-arrangement in the light of new hypotheses. We may find that we want to modify our *lexical* database, e.g. by introducing

subcategories for morphological classes – 2-syllable words, ing-forms etc. Or we may want to broaden or narrow our *phonetic* database (so far we have disregarded diacritics, but our searches can be made sensitive to any diacritic symbol). PDP allows us to react immediately to any new question which our cumulative results generate; we experience the thrill of interactive work with the SED map. And in all of this we are working with the *complete* material. Even if the data are not exhaustive or geographically balanced enough to represent a phenomenon fully – here is all the evidence there is.

We conclude this presentation with a further look at possible generalizations which involve extensions of the database (the domain of cumbersome manual work again). One interesting generalization would be towards the *class* of final plosives: to include *ast* and *asp* besides *ask*.

Let us have a glimpse at some *ast*-maps and illustrate just one aspect of the generalization process: the question is 'should *pasture* be included in the *ast*-class'? A comparison of the two maps shows that adding the word *pasture* to our data brings 5 more forms to the short vowel, but 25 additional occurrences to the long vowel map: Figure 15, ***ast*** (long vowel).

Again we notice the almost variation-less transition from short to long vowel. This is evidence which complements the *ask*-maps in interpretable ways. Looking at the relevant lists which contain these sequences, we are faced with the kind of decisions typical of PDP work: decisions concerned with the definition of a representative data set. This is an empirical matter which can at least be plausibly dealt with if we can look at the phonetic and areal properties of different subclasses side by side. One of the promises of PDP is exactly that it lets us define comparable evidence (based on etymological, morphological or other criteria) in a very flexible manner. I expect PDP to contribute to greater awareness of the problem of representativity in the context of a gradual generalization of our data. Just for completeness' sake, Figure 16 presents the combined maps for long and short realizations of the *ask*-, *ast*- and *asp*-class together (***as+C*** (short vowel)).

The short vowel map shows thinly but evenly attested forms throughout the southern lengthening area – an invitation to have a closer look at variation. The generalized long vowel map is in Figure 17 (***as+C*** (long vowel)).

Again a combined view of the short and long vowel maps highlights the respective compact regions as well as the variation areas. And of course at this point we wonder what the situation is with the final fricatives, as in the *pass/path/laugh*-class of words, and this naturally leads on to the question of the class of *o*-words in the same context (*cross* etc.) – and so on.

Investigations like these involve huge data sets. PDP offers the necessary help to tackle such tasks and – due to the creative potential of interactive work – opens up new dimensions for rich and imaginative questions to be put to the SED.

Our presentation has so far focussed on the cartographic aspect. However, the *Phonetic Database Project* also allows the systematic phonetic study of English dialect sounds in detail. As mentioned briefly, PDP produces with each search a phonetic list of the data found. The BM entries which we scanned into the computer were coded for the purpose of the search procedures, but they can be re-converted into exactly their original printed form. On the screen it looks like Figure 18.

These are the data from which our *ask*-maps are produced; diacritics are fully preserved, as are informants' remarks and editorial notes, as well as the incidental material (i.m.). (BN (Ch 6) stands for 'basic note', pp. is the conversion of the symbol used in the BM into its meaning 'present participle').

Next is the list of the broader search, (the sequence *any vowel (short, half-long, long) followed by -sk).* Note again that the list (of which we only see a part) is complete, i.e. it includes 'hidden' occurrences and incidental material (i.m.), repetitions are also indicated (Figure 19).

The fact that PDP searches the entire BM accurately for any phonetically defined unit in a very short time makes it attractive for the study of general phenomena and processes. In this context the study of diacritics is especially interesting: if we are interested in the occurrence – e.g. in 'transition regions' – of vowel variants, we can in addition to these search for diacritic symbols. In the same way, the different types of /r/, the r-colouring of vowels and other typological phenomena can be investigated rather comfortably.

This step leads us from the preceding binary to (mildly) quantitative maps and to questions of method and principles such as the definition of our database, mentioned already in the context of *pasture.* Should it be restricted according to historical, synchronic-systematic or other linguistic criteria? How do results gained in this way compare with 'global' patterns based on other rationales? These are fundamental issues; they are inspiring issues, too, dealing with the question of representativity in dialect studies. PDP does not solve such problems for us, but they can now at least be approached in a systematic and efficient manner.

I have tried to give you an impression of the type and range of applications which PDP offers, and I have tried to make it clear that PDP is but a tool in the hands of the dialectologist, a flexible and useful tool, designed to relieve us from some aspects (not all) of the manual work associated with areal dialectology. We have seen that PDP is applicable to traditional dialectological tasks but that its real promise lies

in quantitative work; besides producing the kind of maps and lists shown here, it is now possible to study the general phonetic processes mentioned as well as phonotactic properties of English. If we can add genuine statistical possibilities, e.g. by calculating occurrence percentages for any sound or relative similarity values for all 313 localities (in a manner similar to dialectometrical work), and if we can add possibilities for studying areal variation on the basis of the Middle English sound system, then PDP will be a really powerful tool. We are now starting work on both aspects. Awaiting the final version of PDP, I hope to have given you an idea of what it could contribute to dialect study based on the SED and to have shown that the creative and imaginative part will always have to come from the dialectologist, although it will be inspired by the constant feedback provided by flexible interaction with intermediate evidence. If this invitation results in a new and more widespread realization of the scientific interest – and the cultural value – of the SED, the work of our team will have achieved its purpose.

REFERENCES

Kolb, Eduard, Beat Glauser, Willy Elmer and Renate Stamm. 1979. *Atlas of English Sounds*. Bern: Francke.

Orton, Harold, Wilfrid Halliday, Michael, Barry, Philip Tilling and Martyn Wakelin (eds). 1962-71. *Survey of English Dialects: The Basic Materials*. 4 vols. Leeds: E.J. Arnold, (repr. 1998 London: Routledge).

Orton, Harold, Stewart Sanderson and John Widdowson (eds). 1978. *The Linguistic Atlas of English*. London: Croom Helm.

Address
William Elmer
English Seminar
University of Basel
Nadelberg 6
CH-4051 Basel
Switzerland
<elmerw@ubaclu.unibas.ch>

APPENDIX

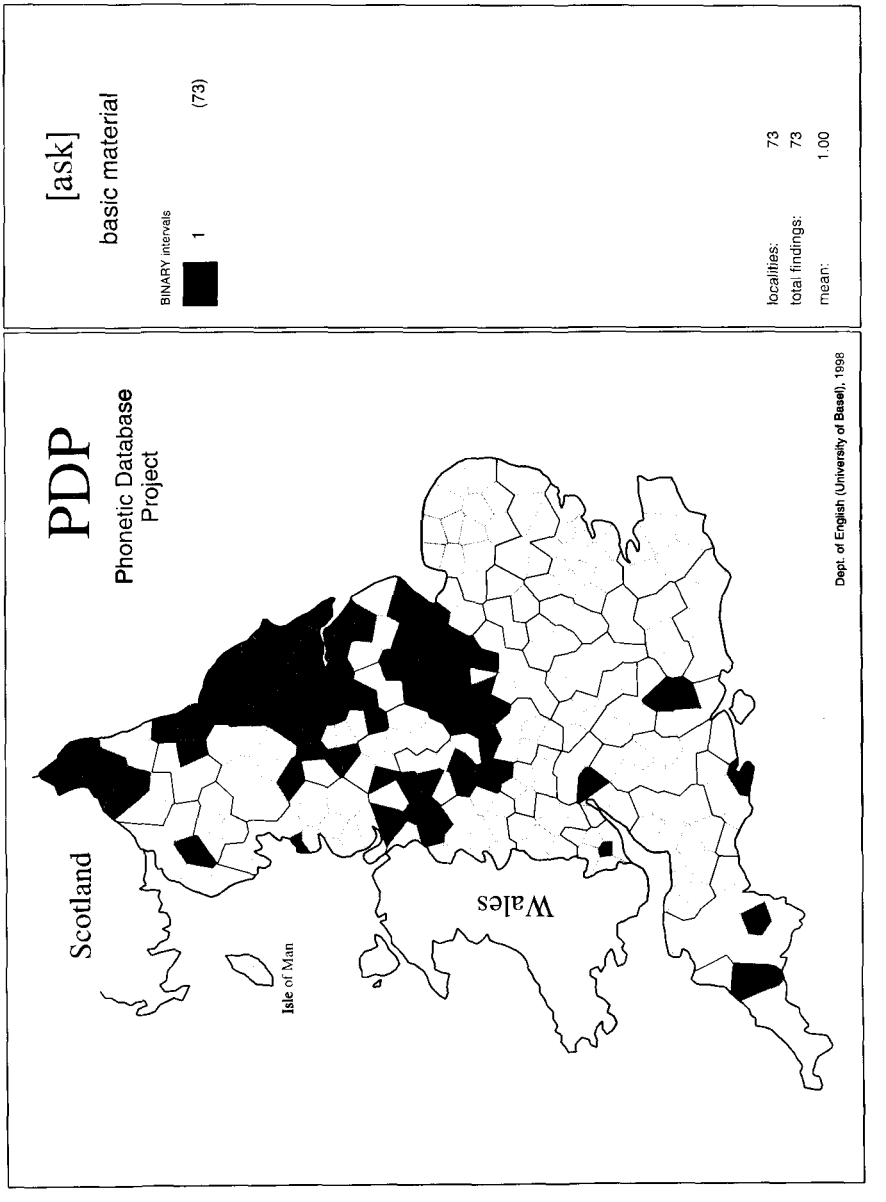


Figure 2: **ask** (short vowel)

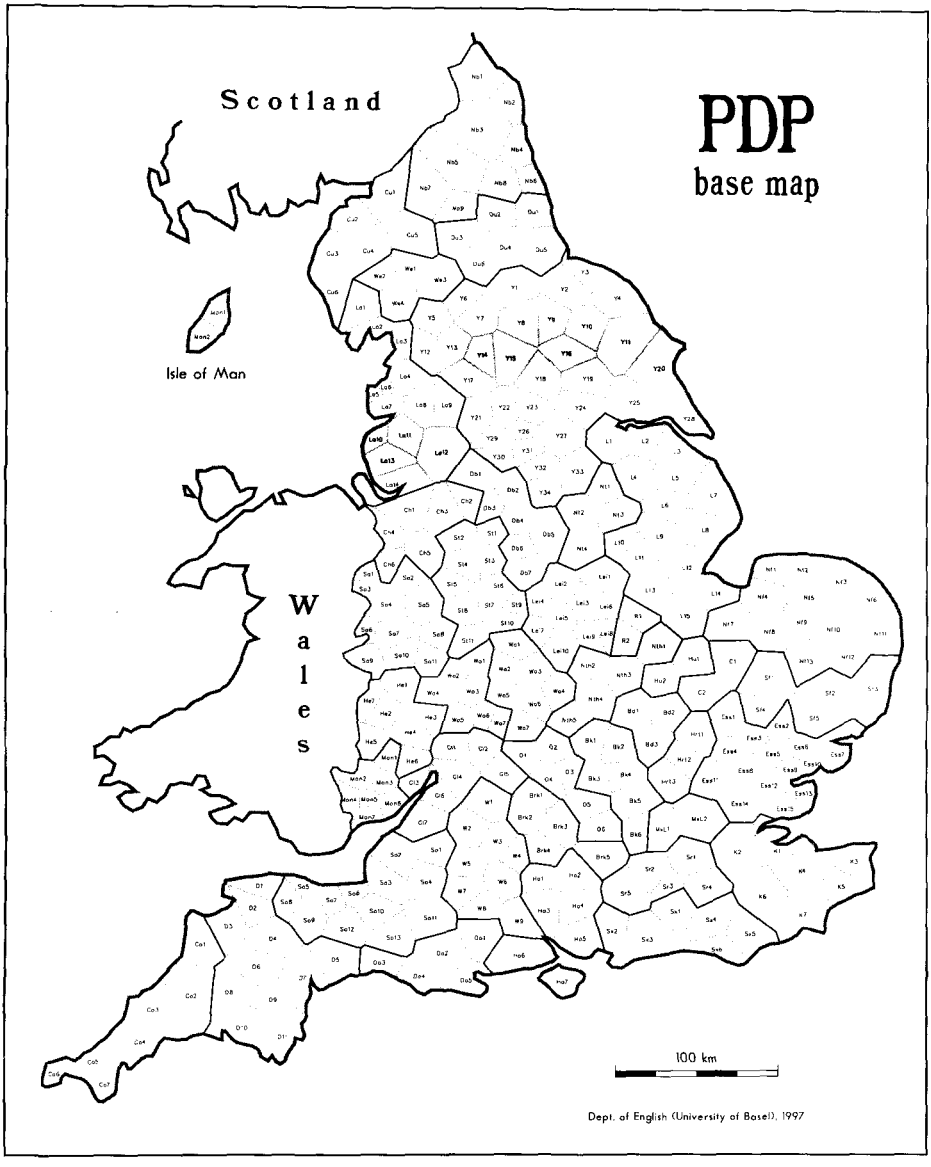


Figure 3: base map

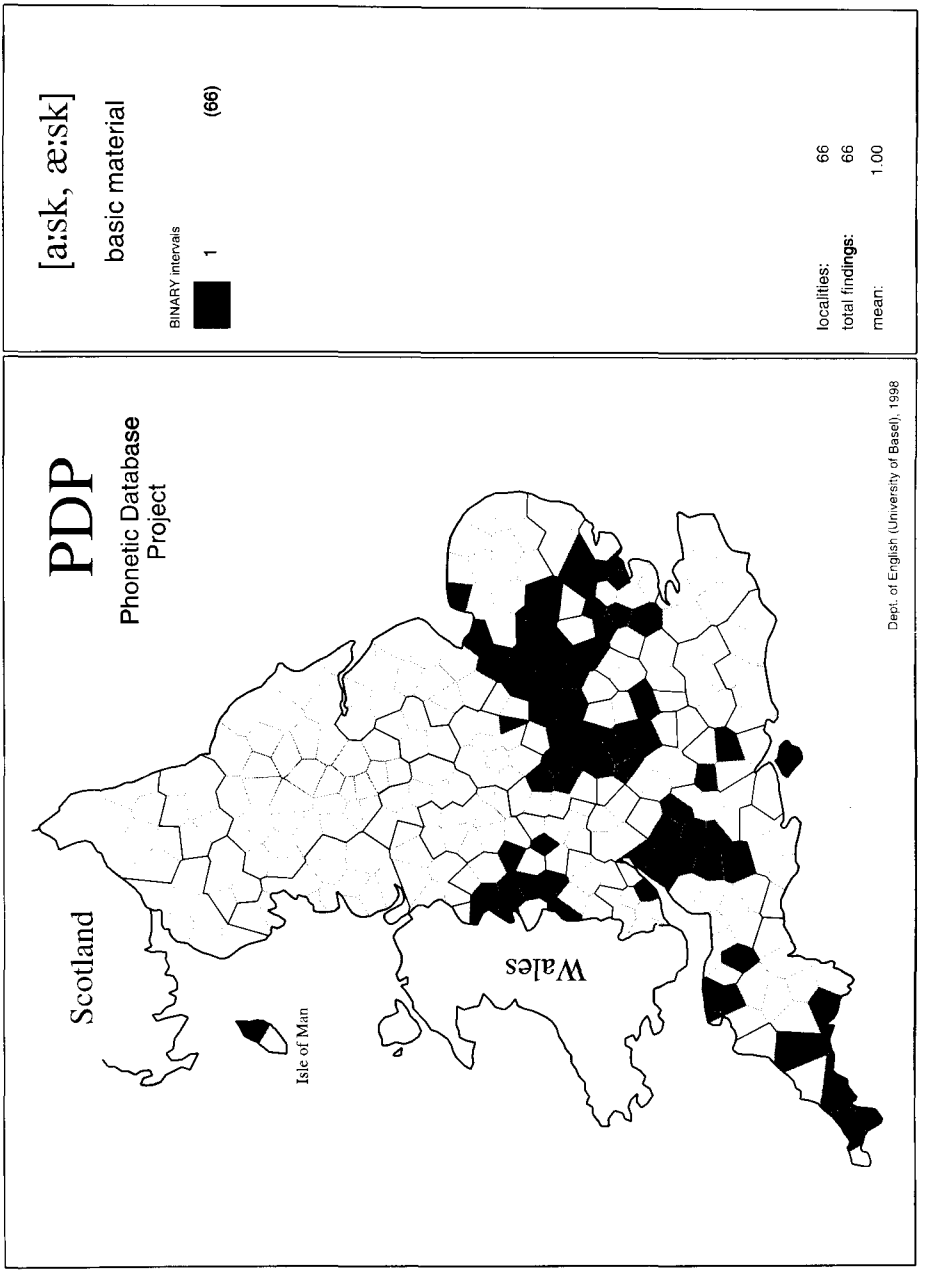


Figure 4: **ask** (long front vowel)

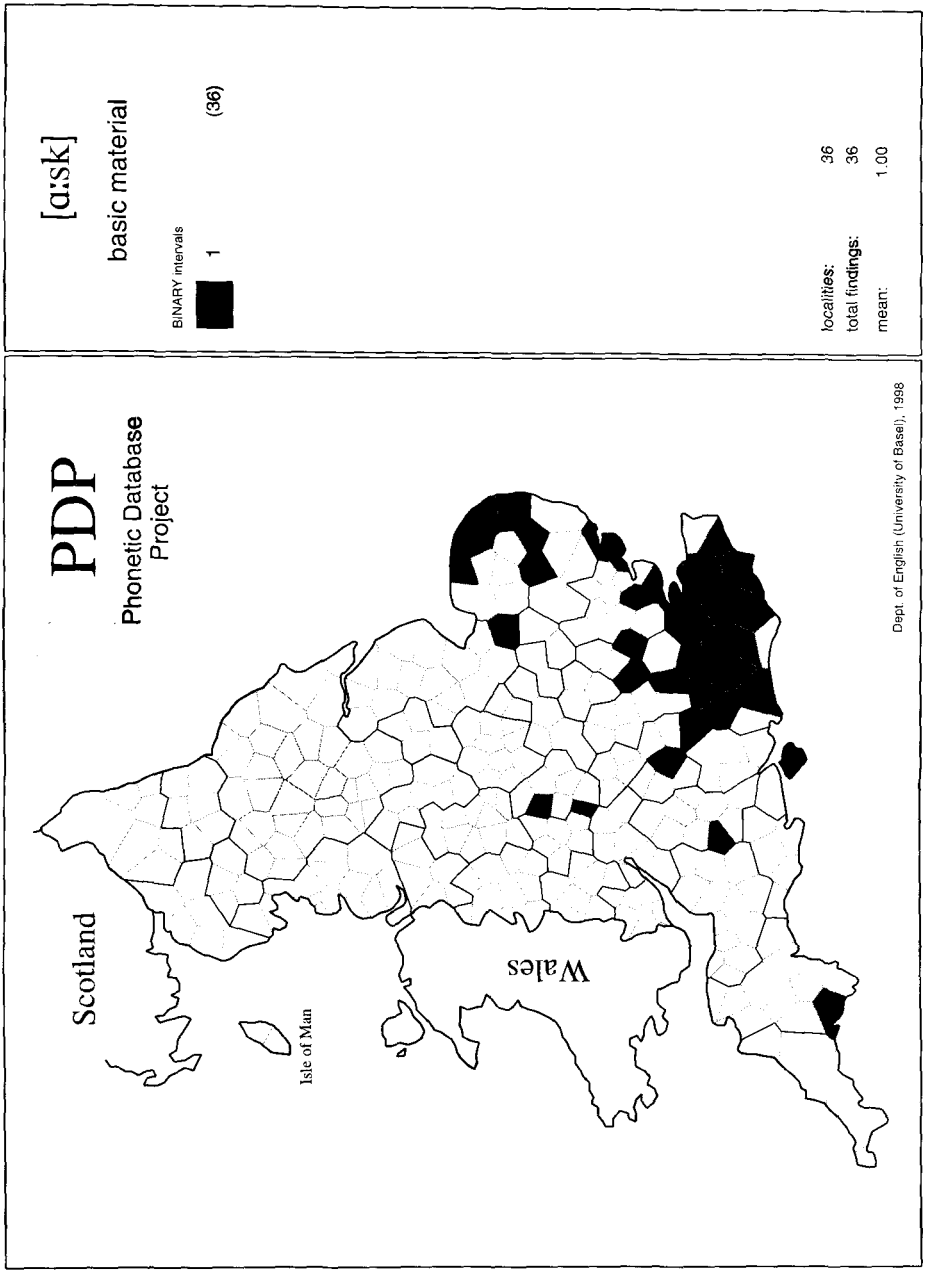


Figure 5: **ask** (long back vowel)

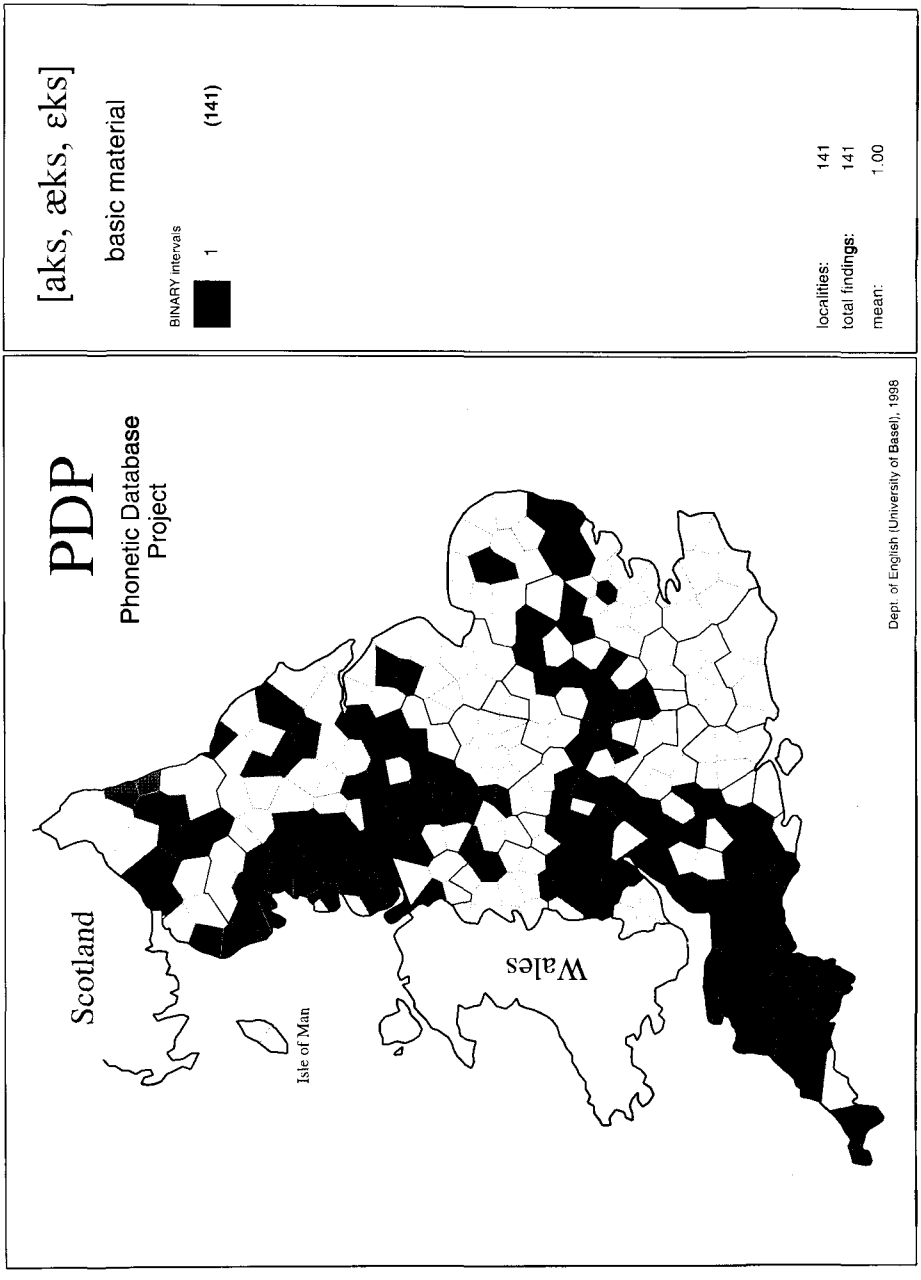


Figure 6: **aks** (metathesis)

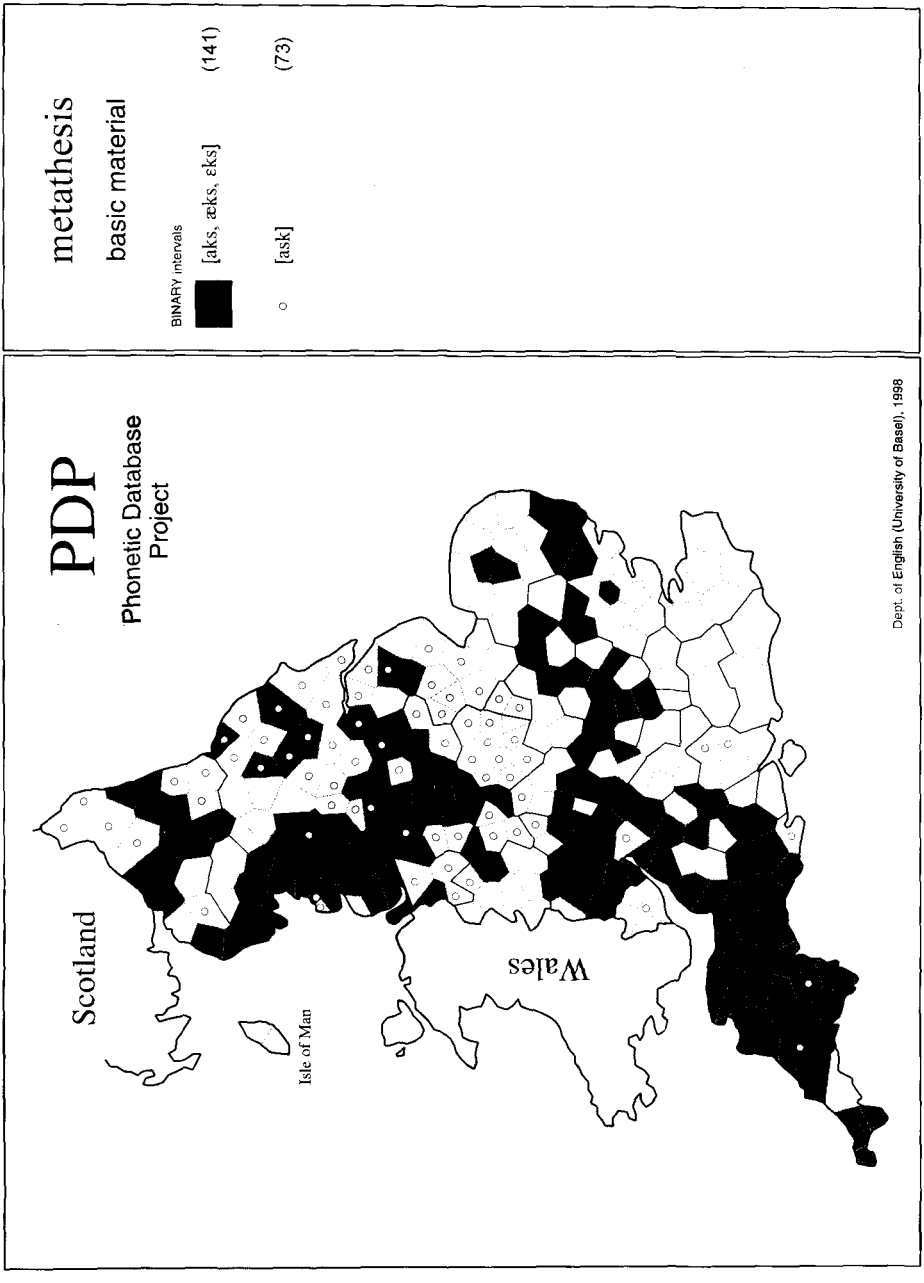


Figure 7: **aks** vs. **ask**

```
(File created by QRScan on 23. 1.1999 at 13. 6.
 Copyright e.r. 1997)


PHONETIC DATABASE PROJECT        English Seminar
UNIVERSITY OF BASEL

LIST OF FINDINGS in the RESPONSES SECTION.
Original Filename = FL.Q2

Keyword Pattern: *

SEARCHSTRING: *ask*

          ①          ②

N22-AA.EXT; II.2. 10 (a) COWSLIP (DAISY DANDELION)
      BASKETWEED,

N23-6.EXT; II.3.6 SOWING-BASKET
      BASKET,

N35-4.EXT; III.5.4 BASKET
      BASKET,              ③

N45-A.EXT; IV.5. 10 HAREph.
      JERK/MASKER

N48-8B.EXT; IV.8.8 (b) SKEP
      (BASKET/SKEP/STRAW-)HIVE,

N49-8.EXT; IV.9.8 NEWT
      (DRY-)ASK,  ④
      WATER-ASK,
      ASKER(T),  ⑤

N59-7.EXT; V.9.7 CLOTHES-BASKET
      CLOTHES-BASKET/WISKET,

N91-3.EXT; IX. 1 .3 ASKEW
      ASKEW,       ⑥

N92-4.EXT; IX.2.4 ASK HIM
      ASK

N92-7.EXT; IX.2.7 AJAR
      ASKEW
          .
          .
          .
```

Figure 8: list of findings (***ask***)

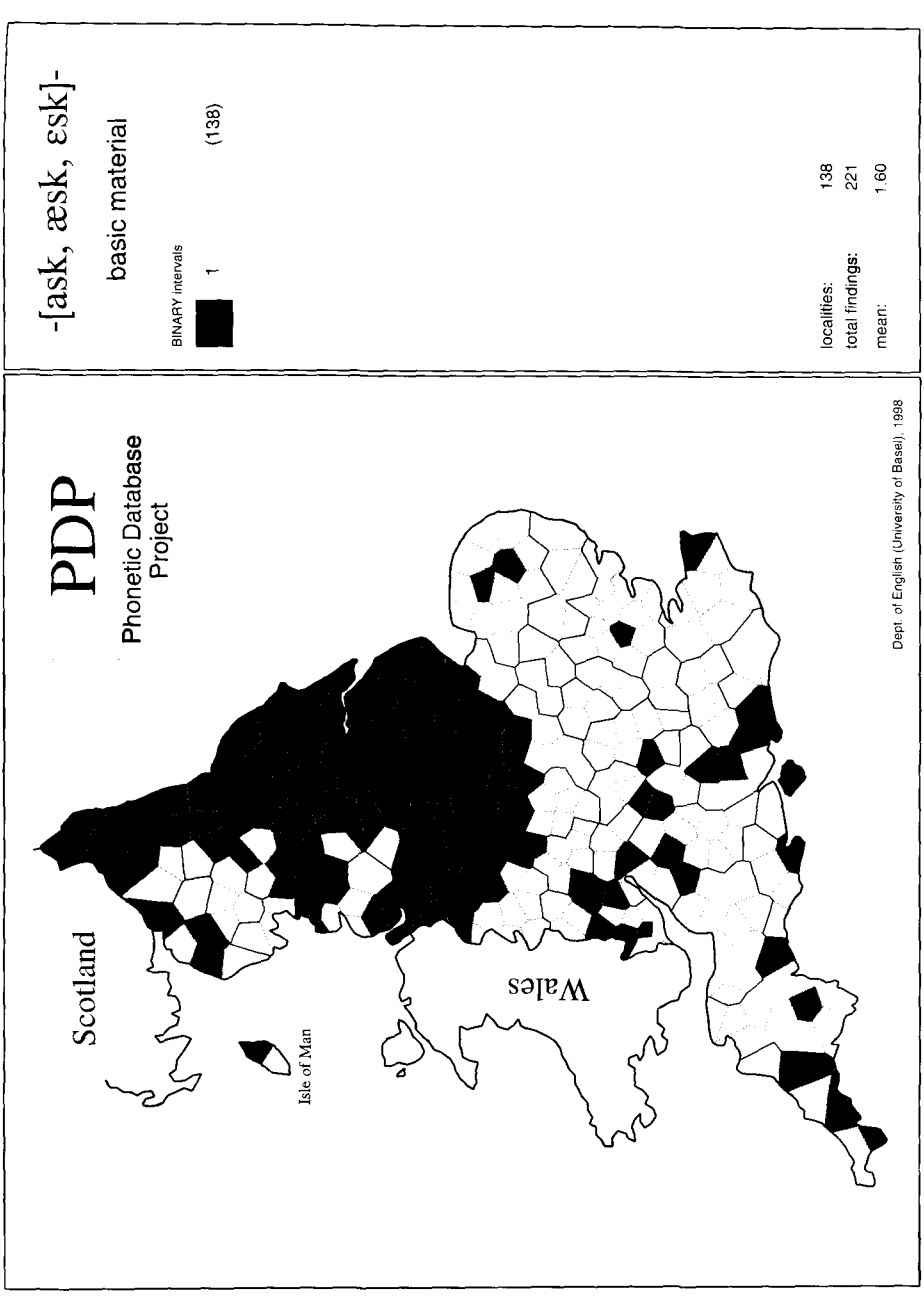


Figure 9: ***ask*** (short vowel)

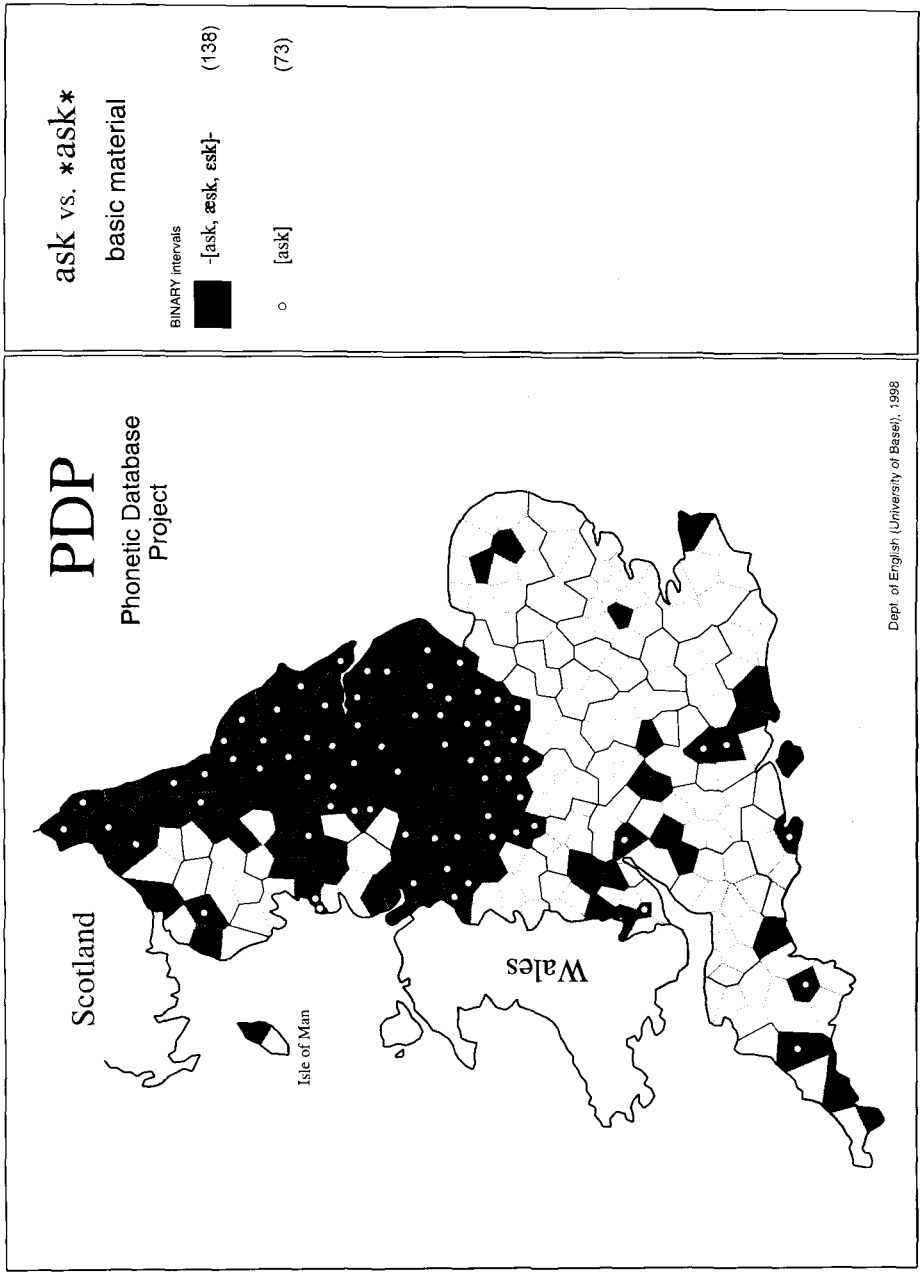


Figure 10: ***ask*** vs. **ask**

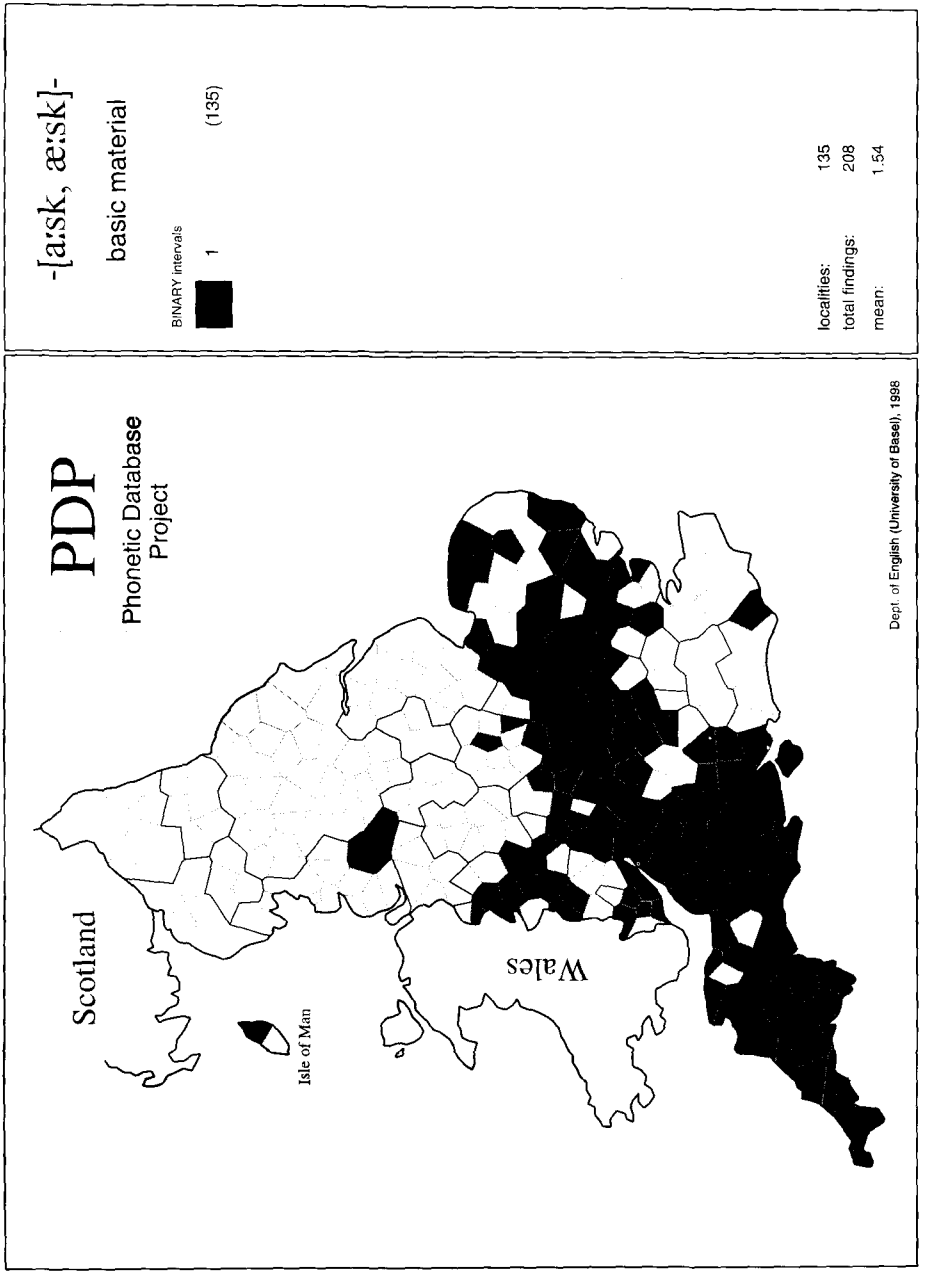


Figure 11: ***ask*** (long front vowel)

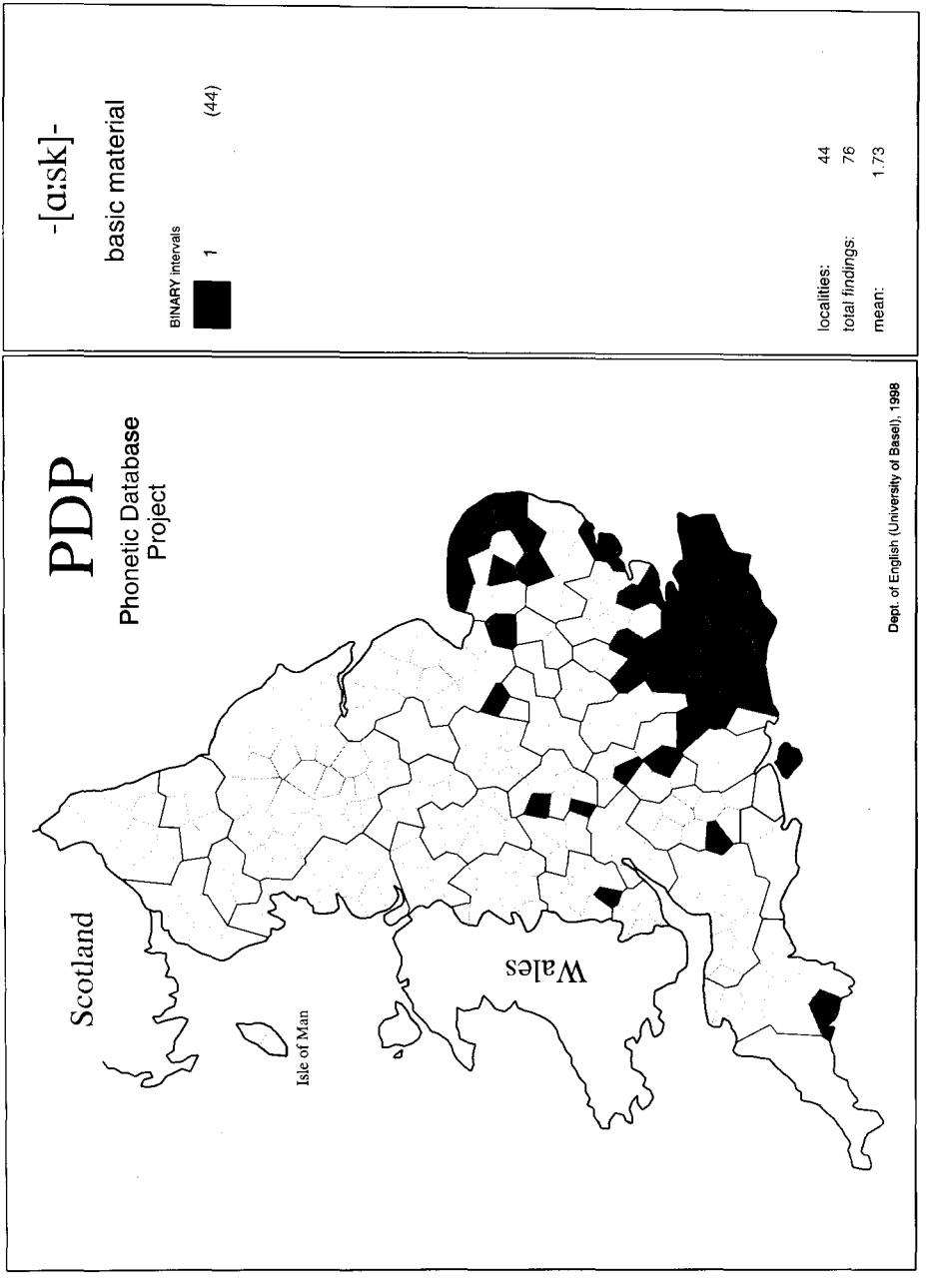


Figure 12: ***ask*** (long back vowel)

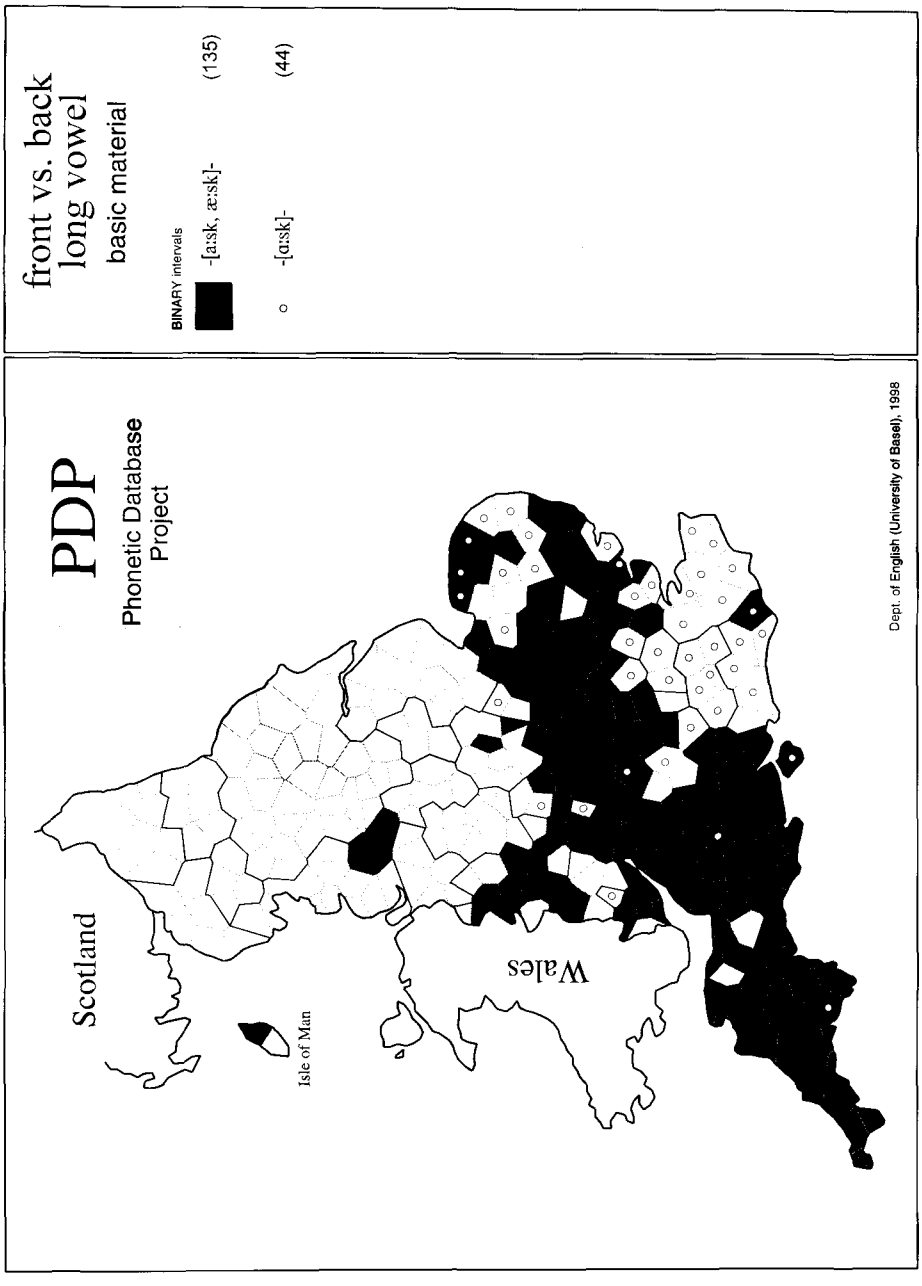
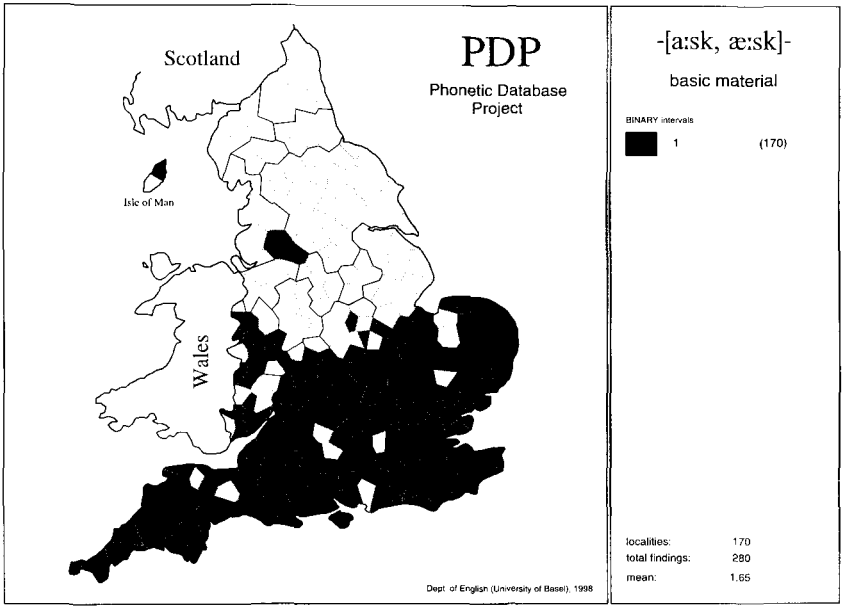


Figure 13: ***ask*** (front vs. back long vowel)



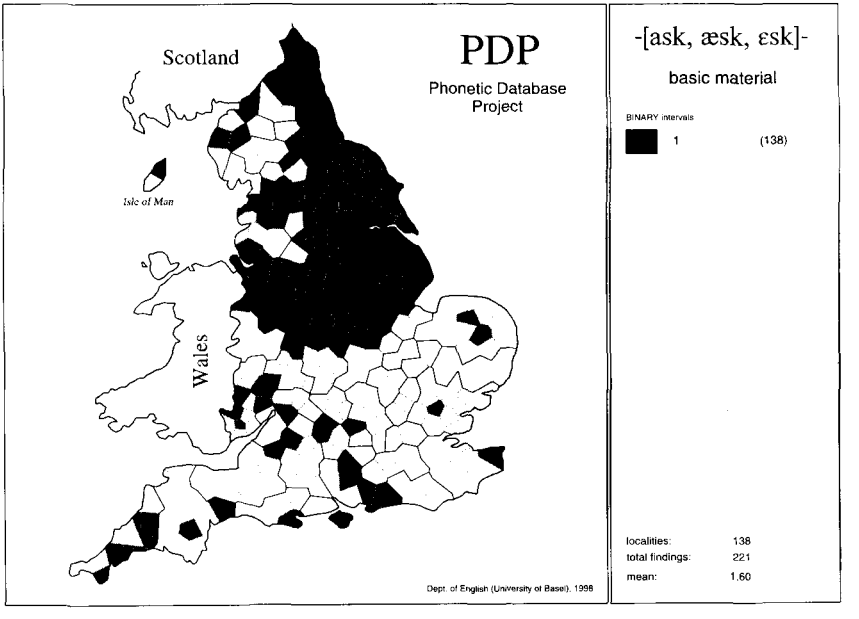


Figure 14: ***ask*** (long vs. short vowel)

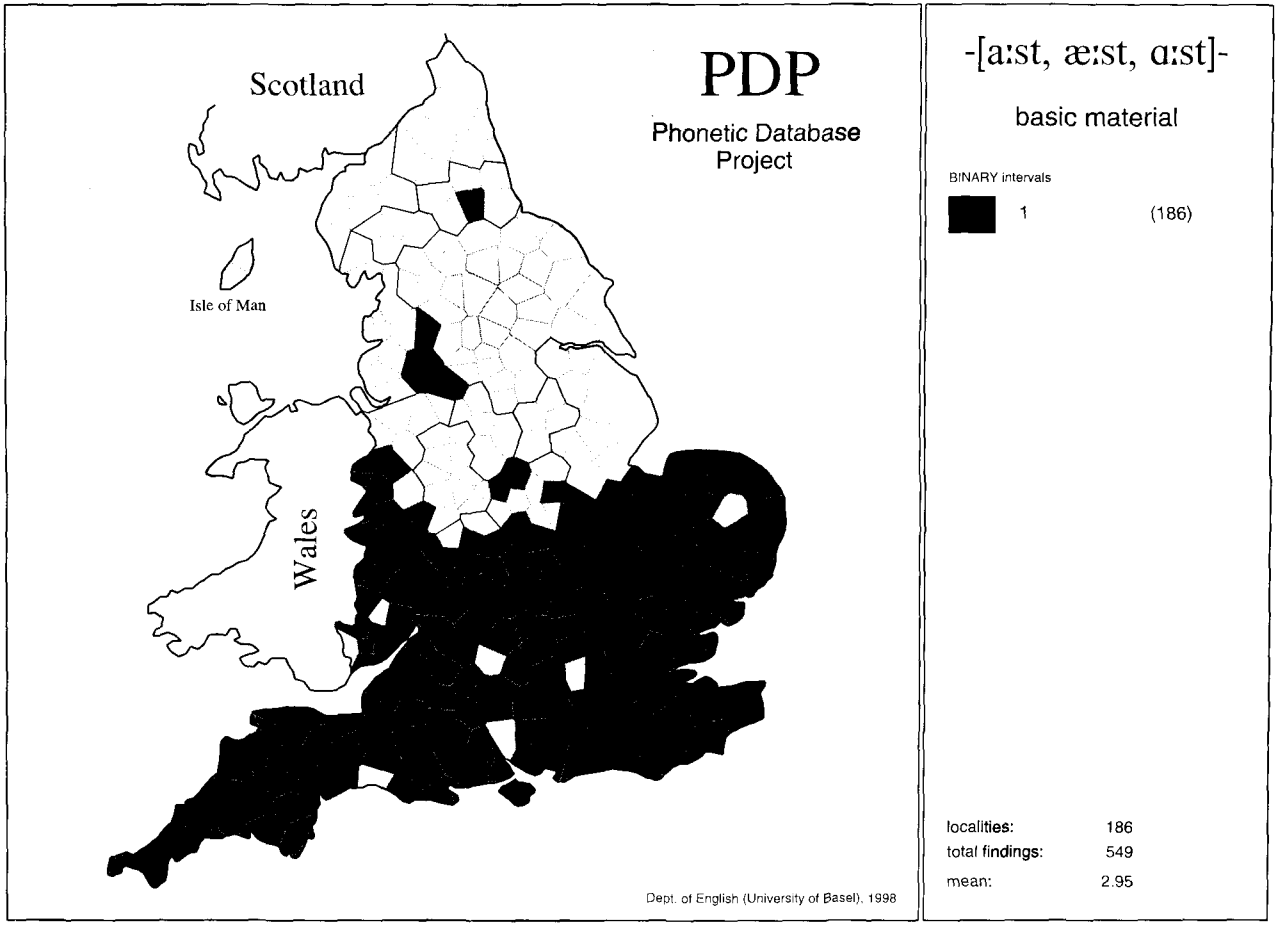


Figure 15: ***ask*** (long vowel)

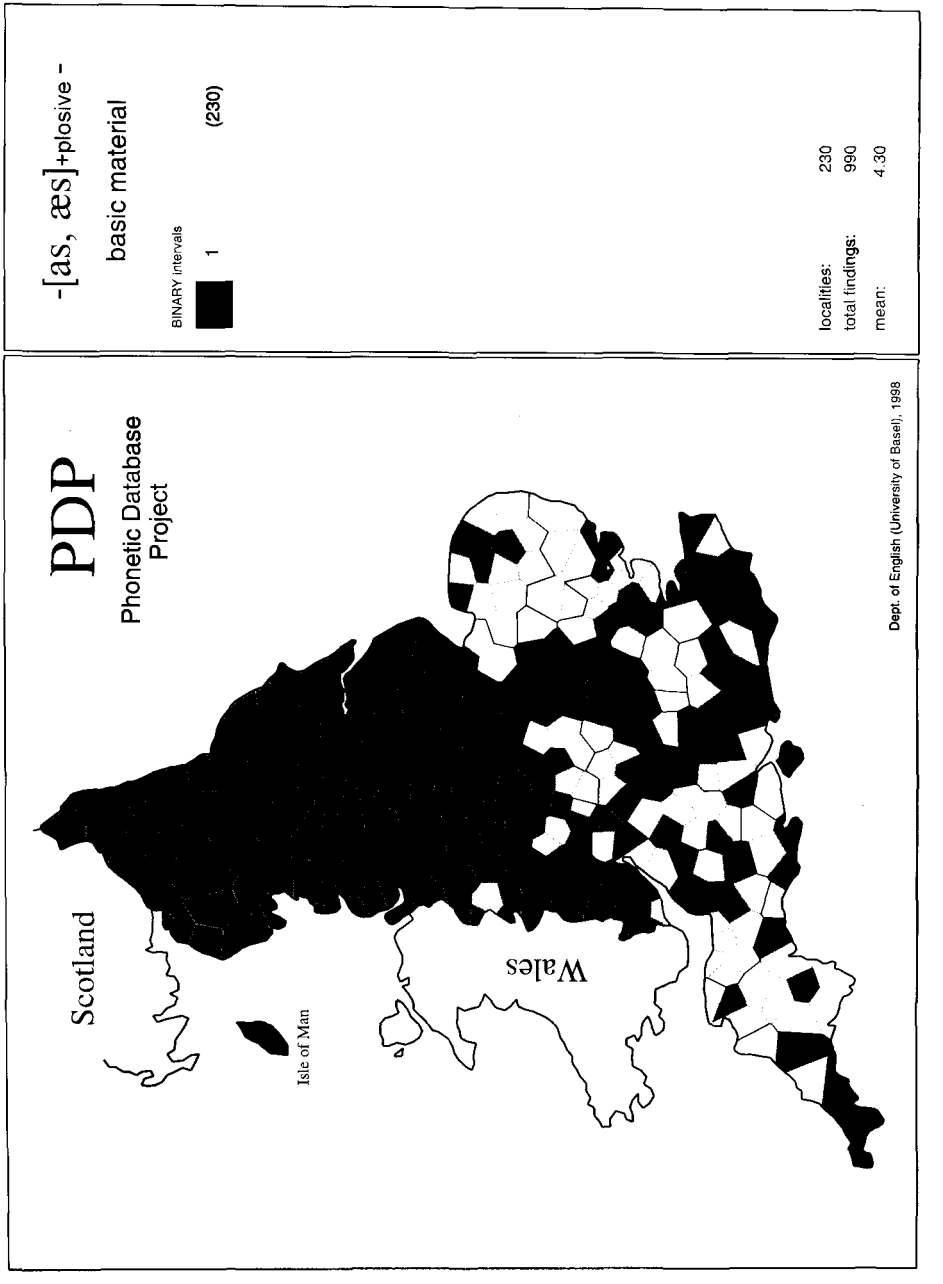


Figure 16: ***as+C*** (short vowel)

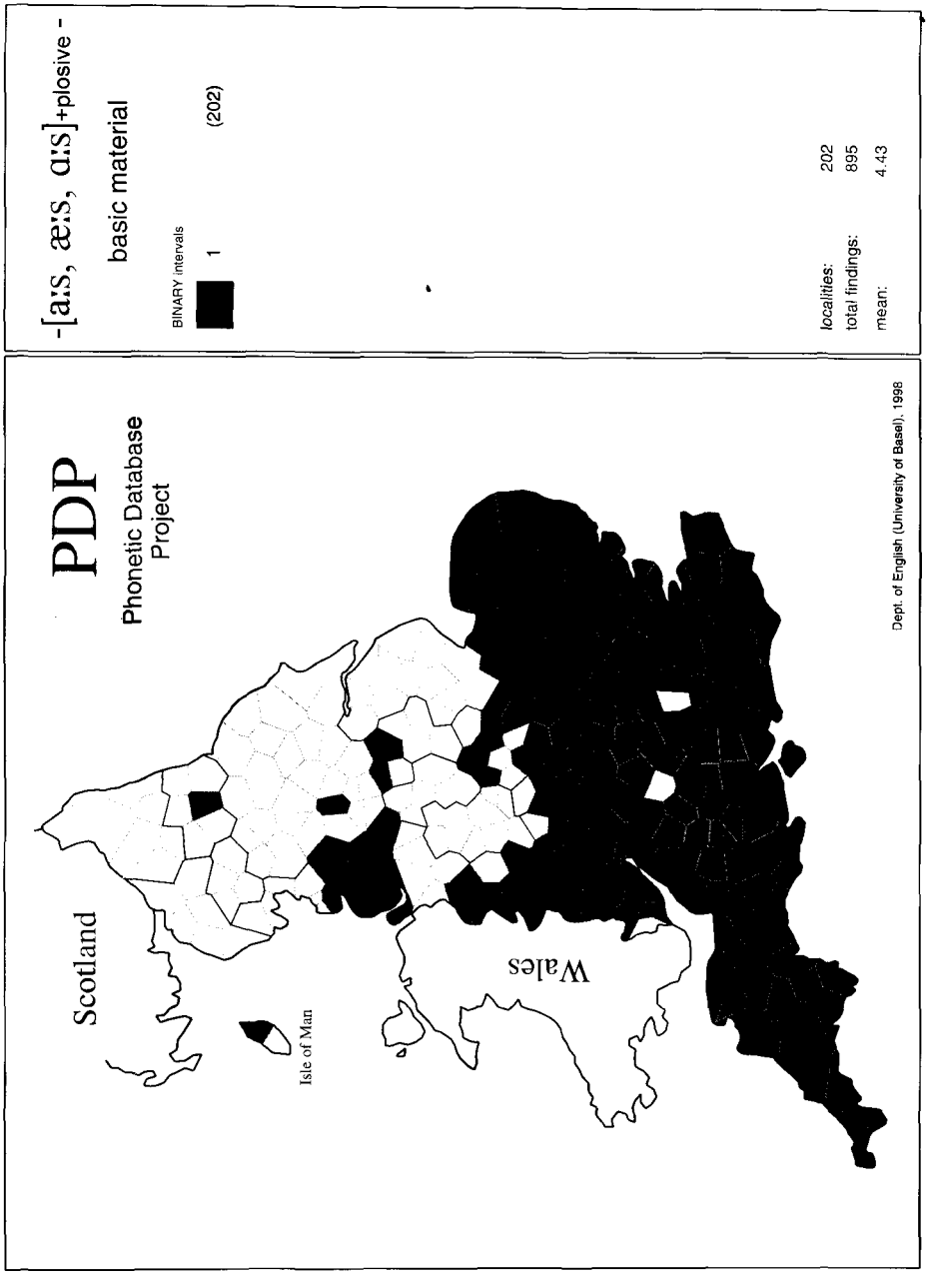


Figure 17: ***as+C*** (long vowel)

(File created by QScan on 23. 1.1999 at 14.13. Copyright e.r. 1997)

PHONETIC DATABASE PROJECT English Seminar UNIVERSITY OF BASEL

LIST OF FINDINGS. Original Filename: RS.Q2. For STATISTICS, see file ST.Q2

Keyword Pattern: *92-4

SEARCHSTRING: %???#46?#07?* or %???:#46?#07?* or %???.#46?#07?*

N92-4.EXT; IX.2.4 ASK HIM

Nb 1-3: ask
Nb 5: ask
Cu 4: ask,
Du 1: ask
Du 4: ask,
Du 5: ask,
La 4: ɛsk,
La 5: ask
Y 2: ask
Y 4: ask
Y 9: ask
Y 10: ask,
Y 11: ask
Y 16: ask
.
.
.
Ess 9: a̱:sk
Ess 10: ä:sk
Ess 11: ɑ:sk
Ess 12: ɑ:sk
Ess 14: ä:sk
Ess 15: ɑ:sk
.
.
.
Ch 1: ask
Ch 2: ask,
Ch 5: ask
Ch 6: asks BN [pp. askɪn3]
Sa 2: ask
.
.
.

Figure 18: phonetic list: IX.2.4 *ask (him)*

Ch 6: tlo:zbaskɪt
Db 1: tlʏ·əsba:skɪt
Db 2: klɷəzbaskɪt,
Db 2: klo:zbaskɪt REP
Db 3: tlü:zbaskɪt
Db 4: tl^{ɷ}ü:zbaskɪt
Db 5: tlɷəzbaskɪt
Db 6: tlü:zbaskɪt
Db 7: tl^{ɷ}u:zbaskɪt
Sa 1-2: klo:zbaskɪt
Sa 3-4: klo:zba·skɪt
Sa 5: klo:zbaskɪt
Sa 6: klo:zba:skɪt
Sa 7: wɪskət
Sa 8-9: wɪskɪt
Sa 10: klo:zwɪskɪt
Sa 11: wɪskɪt,
Sa 11: wɔ:ʃɪnba:skɪt REP
St 1: klozbaskɪt
St 2: kloɷ:zbaskɪt
St 3: klɷ:zbaskɪt,
St 3: klɷ:zwɪskɪt REP
St 4: kloɷzbaskɪt
St 5-6: klɷ:zbaskɪt

St 7: klɔɷzbaskɪt
St 8: baskɪt
St 9: flaskɪt
St 10-11: kloɷzbaskɪt
He 1: klo:zwɪskɪt
He 2: klo:zba:skɪt
He 3-6: klo:zbæskɪt
He 7: kloɷzbạ:skɪt
Wo 1: klɒɷzbaskɪt
Wo 2: wɪskɪt
Wo 3: kloɷzba:skɪt
Wo 4: kloɷzbæ·skɪt
Wo 5: klo:zba:skɪt
Wo 6: wɛʃbæ:skɪt
Wo 7: kloɷzbɑ:skɪt
Wa 1: baskɪt
Wa 2: klɔɷzbɑ:skɪt
Wa 3: klo̜ɷzba:skɪt
Wa 4: kloɷzba:skɪt
Wa 5-6: klo̜ɷzba:skɪt
Wa 7: wɒʃɪnba:skɪt,
Wa 7: kloɷzba:skɪt REP BN ["older"]
Mon 1: bɑ:skɪt
Mon 2-3: klo:zba:skɪt

Figure 19: phonetic list: ***V(:)sk***

Parallel Lines Through Time: Speech, Writing and the Confusing Case of *She*

Merja Black

Abstract

The approach to Middle English dialectology that was developed in connection with the *Linguistic Atlas of Late Mediaeval English* treats the written text, rather than its spoken correlate, as the primary object of study. Such an approach allows for the direct analysis of surviving data rather than the indirect approach demanded by the reconstruction of speech. For interpretative purposes, however, the interaction between the written and spoken modes cannot be ignored, and is of fundamental importance for the comparison of medieval and modern dialect materials.

This is illustrated by the history of certain regional variants of the pronoun 'she'. A geographical correspondence between the survival of some conservative written forms in the Middle English materials and a partially gender-free pronoun system in modern varieties is of particular interest, in that it appears to reflect a significant divergence between the medieval written and spoken systems. It is here suggested that the divergence has its basis in the different communicative strategies required by the two media, in particular the greater reliance of the written mode on such syntagmatic tracking devices as gender and case. The actual selection of written variants does not seem simply to mirror regional spoken usage, but rather to reflect social and textual factors. This example highlights the importance, in interpreting Middle English written forms, of an integrated approach that combines a corpus-based analysis with a close study of the individual texts.

1. *Setting the question: the spoken mode in the study of Middle English*

1.1. The study of past states of language differs in two important respects from the study of present-day varieties: the evidence survives in the written mode only, and is limited by chance survival. A central problem in historical linguistics has, accordingly, involved the interpretation of written data: how does one make assumptions about the spoken language in a period from which only written data survive? In the study of Middle English (henceforth ME), the centrality of this question has varied with changing approaches to the material. Traditional historical linguistics tended to treat the written form merely as (inadequate) evidence for the spoken language, the latter being the primary object of study. More recently, with the methodology developed in connection with the *Linguistic Atlas of Late Mediaeval English* (henceforth LALME) (McIntosh, Samuels and Benskin, 1986), the focus has shifted to the written language as an autonomous system, to be studied in its own right before any extrapolations to the spoken mode are made.

Following the LALME tradition, the question of the relationship between the spoken and written modes has in recent years remained relatively marginal in ME studies. The present paper relates to preliminary work for a new history of ME orthography and phonology, based on the LALME framework but necessarily facing different challenges.[1] Using as an example the late survival of traditional spellings such as *heo* for the feminine pronoun in the Southwest Midland area, it will suggest that these challenges can only be met by taking fully into account the interaction between the spoken and written systems.

1.2. As is well known, the historical circumstances during the early part of the ME period gave rise to a situation where English writing came to display linguistic variation to an extent usually only found in the spoken mode. This was a direct consequence of the temporary displacement, after the Norman Conquest, of English by Latin and French as the written language for most functions. As English gradually regained its status as a written language from the thirteenth century onwards, it lacked a centralized model, and the texts show considerable variation on all levels of language. It is not until the middle of the fifteenth century that the written dialects gradually begin to be replaced by a standard usage.

Accordingly, the surviving ME materials form a large corpus that reflects the natural variation and the changes in English during a very dynamic period. Traditionally, only a small fraction of the available material was held to be of value as

linguistic evidence: the great majority of texts, which survive as scribal copies only, tended to be dismissed as containing mixed or 'corrupt' language, while only authorial holographs or original documents could be expected to provide a 'pure' dialect. Views on the purity of dialects have since changed; it is now generally acknowledged that variation, not uniformity, is characteristic of natural language, and that this variation is of an orderly kind and can be analysed and explained.

While these insights mainly derive from studies of present-day states of language, a similar idea, applied to ME studies, was originally formulated by McIntosh (1963 [1989], 1974 [1989], 1975 [1989]): there is no reason to assume that ME scribes would typically produce a random mixture of dialectal forms. A scribe would generally translate a text written in a dialect different from his own, and would tend to replace alien forms but retain ones that belonged within his own repertoire. The inventory of forms in any ME text is thus not random but can be explained, and is often definable in regional terms. By detailed dialectal analysis, it is possible to determine whether a text contains a regionally consistent language; such texts can then be localized by fitting them in relation to each other. This insight formed the basis of the LALME methodology and made possible the construction of a framework of several hundreds of localized texts and, as a consequence, a very large amount of mapped-out data.

To achieve this, it was necessary to work with finer distinctions than those available by traditional methods. It was shown by McIntosh (1963 [1989], 1974 [1989], 1975 [1989]) that distinctions within the written mode, whether corresponding to spoken-language features or not, form patterns that are in themselves regionally significant. As the orthographic data can be studied directly, it has been possible to build up a much more sophisticated picture of regional variation in ME than would ever have been considered possible before.

The strict separation of the written and spoken modes, and the study of the former in isolation, have thus been (and continue to be) necessary methodological steps in building up a typology of ME scribal usage. However, as soon as our main interest is no longer purely typological – as soon as the data are to be interpreted or commented upon in some way _ a consideration of the relationship between the spoken and written modes becomes unavoidable. As has been pointed out by Smith (1996: 6) in a similar context, history presupposes an attempt to make sense of things; if we wish to make sense of ME spelling it is impossible to ignore the spoken mode. In other words, we cannot approach ME as though 'its users were deaf and dumb' in the words of McIntosh (1956 [1989: 11]), as they were not, and the fact is in itself highly significant for the dynamics of the language.

2. *The spoken and written systems in ME*

2.1. As indicated in the previous section, speech and writing form separate systems, and variation in one does not necessarily relate to variation in the other. At the same time, both are manifestations of the same 'language' – the same lexis and grammar – and some correspondence must be assumed between them. It has frequently been pointed out that alphabetic spelling systems, when newly devised, ideally tend to reflect the phonemic distinctions of the spoken language. Because of the greater conservatism of the written medium, the correspondences between the two systems are likely to diminish with time, so that the spelling system, as in present-day English, may end up being far from an accurate reflection of the distinctions in speech. However, there are limits as to how far the written system can be distanced from the spoken one without becoming uneconomical, that is, accessible only to specialists with the time and motivation to learn it (cf Smith 1996: 15-17).

The spoken and written media, then, can be said to work autonomously in that choices within each are governed by different factors; at the same time, they interact constantly, each influencing the other. Figure 1 shows diagrammatically the relationship between speech and writing, using the basic model given by Samuels (1972: 6).[2] The two media are represented by parallel lines running along the time axis. The arrows between the lines show different kinds of interaction, including spelling pronunciations and adjustments in the written mode to changes in speech. Samuels (1972: 6) comments on the model as follows:

> (a) a majority of linguistic changes arise in the spoken language, and may or may not ultimately spread to the written medium; (b) certain (though fewer) changes originate in the written language, and may or may not spread to the spoken medium; and (c) the main influence of the written language is a conservative one – it acts as a brake, inhibiting the general acceptance of many changes that arise in the spoken language.

These points may be taken as axiomatic for most periods of the language. In general, we expect written language to be more conservative, and to reflect a somewhat earlier stage compared with the spoken mode. A certain level of fixity, or standardisation, may, moreover, be seen as a typical (if not universal) characteristic of written language: compared with the spoken mode, it is likely to be much more closely focused on a particular set of conventions, permitting less variation. This situation

could be depicted by drawing a section of the parallel lines, where the variable spoken mode is represented by a whirl, suggesting a centre of focus with fuzzy edges, and the fixed written mode by a dot (see Figure 2). Most changes take place within the spoken medium, and the influence of the written mode on the spoken is on the whole a restrictive one, limiting change.

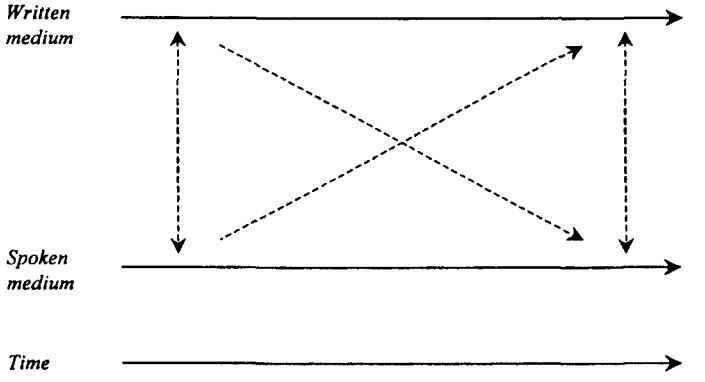


Figure 1. The interaction of the written and spoken media over time (after Samuels 1972:6)

However, such a model cannot be applied directly to the ME situation, which is characterized by a high degree of variability in the written mode. Smith has compared this variability to that of present-day spoken usage: '[m]edieval written standards seem to have worked in the same way as spoken Received Pronunciation does in Present-Day British English: they are a sort of mean towards which scribes tend' (Smith 1996: 67). In ME, then, we must assume a model where both the written and spoken modes are characterized by more or less loosely focused variation; such a situation could be depicted by drawing the sections of both lines in Figure 1 as whirls (see Figure 3). The interaction between the two modes will, then, to some extent differ from that described by Samuels. Rather than an asymmetrical relationship, where the influence of the spoken language tends to speed up change in the written mode, and the written mode acts as a brake on the spoken, we might assume two parallel variable systems in a more fluid and dynamic process of interaction.

2.2. To bring the preceding discussion into focus, the copying behaviour of the ME scribe should be considered. While it makes sense to say that, in ME, written language reflects regional (and other) variation, we can assume no *direct*

correspondence between the variants used by a scribe in writing and in speech. Firstly, much orthographic variation has no counterpart in speech (e.g. *myn* ~ *min* 'mine'). Secondly, even such variation that can be assumed to be connected with features of the spoken language, such as the occurrence in a scribe's written usage of variants such as *man* and *mon* for 'man', does not necessarily imply an equivalent variation in his speech, but simply that both forms were familiar and acceptable to him in writing. The scribe's choice in a given text of either form, or a particular mixture of both, will in the main depend on two factors: his general copying strategy and the linguistic usage of the exemplar.

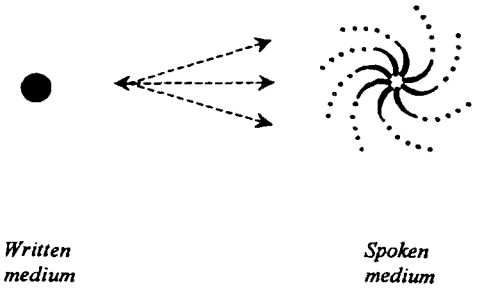


Figure 2. The written and spoken media: section of the arrows in Figure 1

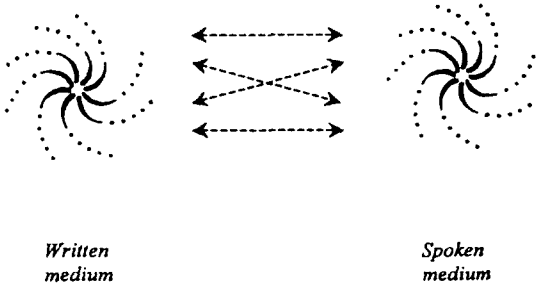


Figure 3. The written and spoken media: the ME situation

In a seminal article, Benskin and Laing (1981) outlined the copying strategies available to the ME scribe. The typology follows the original, much-quoted tripartite distinction made by McIntosh (1973 [1989: 92]), dividing the scribes into A) translators, B) *literatim* scribes and C) those who do something in between. The original premise of the LALME project was that a large number of scribes could be

treated as A types, so that their output could be used as direct evidence for a single regional usage.

In reality, however, it appears that most, perhaps all, ME texts reflect to a greater or lesser extent the scribal behaviour known as constrained selection, described by Benskin and Laing (1981: 72-75) as a variant of group C, but shading into both A and B. Constrained selection works within passive repertoires. A scribe translating a text written in a dialect different from his own will replace alien forms with familiar ones; however, when he encounters forms familiar and acceptable to him, whether part of his own active repertoire or not, he copies them as they stand. The extent of scribal passive repertoires, or constraints, must be assumed to vary immensely, from very wide to very narrow. At one extreme, a scribe with wide constraints, copying from a text in a familiar dialect would be likely to produce, in essentials, a literatim (or B type) copy, while a scribe with narrow constraints, copying from a very different dialect, would produce something very close to a thoroughgoing translation.

Scribal constraints cannot be assumed to reflect regional patterns only, even though these must have been of primary importance. The familiarity of a scribe with a given form may reflect geographical vicinity, but it may also reflect his previous copying experience For example, a Midland scribe used to copying large quantities of northern texts will almost certainly have a raised tolerance of northern forms compared with scribes of the same area who lack such experience. Towards the end of the ME period, the increased mobility and circulation of texts may be assumed to have raised considerably the familiarity and tolerance of scribes towards 'alien' forms, especially in large urban centres, resulting in a process whereby originally strictly regional forms become generalised. Such forms were compared by Samuels (1981[1988: 91]) to 'the coins when two currencies are combined', that 'pass from writer to writer . . . and their regional significance is lost'; at least some of the many northern forms that came to form part of Standard English must have begun their spread as such loose change.

2.3. Even in Middle English, we cannot, then, assume a one-to-one relationship between written and spoken forms: to make inferences about the spoken mode, a more sophisticated approach is needed. Firstly, in order to observe and validate correspondences between patterns in the written material and postulated features of the spoken language, a reasonably large amount of data and contextual information is required. Secondly, the correspondences should, as far as possible, be related to data that allow direct inferences about spoken usage. Traditional sources for such information include the following:

- Rhyming and alliterative usage
- Scribal patterns (back spellings, confusion of graphs, etc.)
- Contemporary comments on pronunciation / usage
- Imitation of features of other dialects
- Texts written using the orthographic conventions of another language
- Comparison with modern spoken dialects

All these sources are potentially useful as bases for deductions about the spoken system, even if they each have their limitations. Apart from the obvious restriction of applying only to verse, the use of rhymes and alliteration as evidence for spoken usage involves two major problems. Firstly, it presupposes that medieval writers were consistent in their usage, and aimed at producing perfect rhymes or alliterative patterns. Secondly, the rhyming or alliterative usage of a text generally reflects the usage of its original author, and may be very different from the usage of the scribal text in which it appears (see Benskin and Laing 1981: 69-71).

Certain recurring patterns within a single scribal usage may be used to make inferences about the relationship between the written and spoken systems. Most notably, confusion between two forms (e.g. the occasional use of <a> for expected <e> and vice versa within a single text) generally suggests hyperadaptation of some kind, reflecting a merger in the scribe's dialect, or at least a considerable conflict between two systems. On the other hand, the use of exceptional or isolated spellings confined to a single lexeme or word class may simply illustrate that *chaque mot a son histoire,* but might also signal other factors and should not be accepted uncritically as evidence for a distinct pronunciation.

Contemporary comments on pronunciation are central to the study of Early Modern English, during which period they are plentiful; however, such comments are exceedingly rare in Middle English. More useful evidence is provided by various occasional biproducts of language contact: imitations of dialectal speech occur in several Middle English texts and can provide valuable data (see e.g. Smith 1995 for Chaucer's use of Northern dialect in the *Reeve's Tale).* Furthermore, there are some rare but important examples of Middle English spelt according to the conventions of other languages, notably Welsh (see Black 1998a). Texts such as these are of particular interest for the light they can throw upon such problematic phenomena as changes in vowel quality during the ME period.

Finally, comparison with modern regional speech is a potentially fruitful method of making deductions about past speech-patterns, including their relationship to the

written mode. Some of the possibilities involved in comparing the collected ME data with those derived from modern dialect surveys – notably the Survey of English Dialects (SED) – were discussed in the Introduction to LALME (I: 27); so far, however, relatively little work has been done along these lines. In the following, an attempt will be made to relate some of the SED data to the LALME material for the Southwest Midland area. Apart from the data in LALME itself, the discussion is based on a close study of 24 texts, dated to the fourteenth and fifteenth centuries and localized in the Herefordshire area.[3]

3. *Diachronic comparison: the case of she*

3.1. Because of its historical orientation, SED is particularly useful for diachronic comparison. There are, naturally, limitations as to the comparability of the data collected for the SED with that in LALME: the overall set of dialectally significant features has undergone great changes over time, and the two surveys aimed at collecting different types of material.[4] The most fundamental difference, however, involves the medium: the SED records variation in the spoken mode, LALME in the written. Accordingly, even where data for the 'same' item have been collected for both surveys, they are not directly comparable: the comparison can only be indirect, dependent on our interpretation of the medieval written forms and, ultimately, on our theory of the relationship between the two media.

In some cases, correlations can be fairly straightforward. Figure 4 shows a well-known correspondence between ME spellings and twentieth-century pronunciations of the word 'man', as recorded by LALME and the SED respectively. The two patterns are, of course, not directly equivalent. While the SED pattern reflects an actual pronunciation of 'man' as something like [mɒn], the LALME pattern does not show how speakers in a particular area pronounced the word, but rather the area where scribes were in the habit of spelling 'man' with <o>. The two are clearly different matters; still, it would seem unreasonable to doubt that the distributions show a significant correspondence of some kind (see Wakelin 1982).

At first sight, a set of similar correspondences seems to emerge in the maps showing the medieval and modern distributions of forms of the feminine personal pronoun 'she' (Figure 5). However, on closer consideration these patterns turn out to be rather less straightforward to interpret than is the case with the *man/mon* example.

The distributions of the northwestern *ho/oo* type and the dominant *she* type may fairly safely be assumed to reflect some kind of continuity, even if the derivation of

these forms may be controversial. The southwestern patterns, on the other hand, present a problem. The distributions of ME *heo* and ModE *her* (or, more precisely, *er*) are remarkably similar, and a direct derivation of the latter from the former has been tentatively suggested (see Duncan 1972: 190 and references there cited). Less controversially, *her* is usually derived from the object form, making the correspondence on the map more or less coincidental. It will here be suggested that *heo* and *(h)er* are indeed connected, but that the connection is of a more complex kind than direct derivation, and one which demonstrates well the interaction between writing and speech.

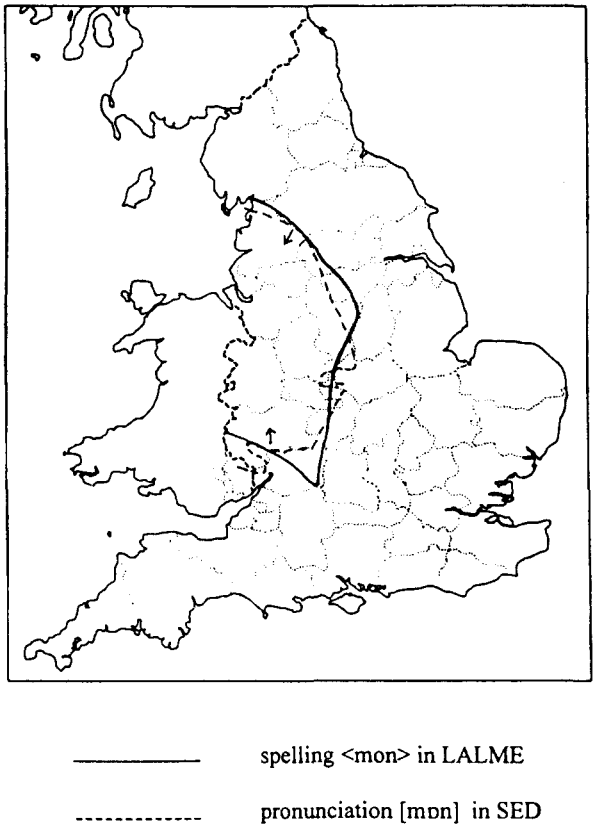
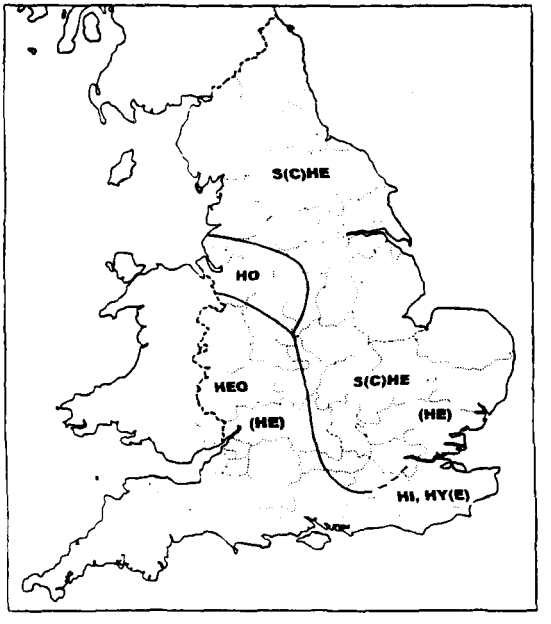
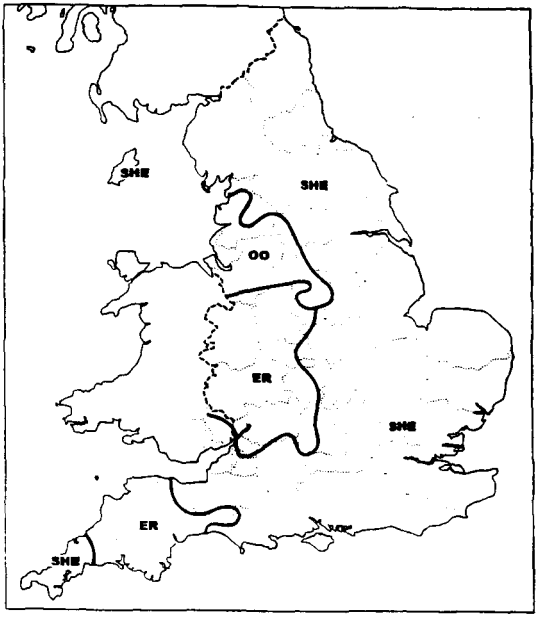


Figure 4. The distribution of MON type forms for 'man' in LALME and SED



LALME



SED (from Upton and Widdowson.1996)

Figure 5. The distribution of forms for 'she' in LALME and SED respectively

3.2. The Old English third-person pronoun system distinguished between masculine *he* and feminine *heo*, the contrasting vowels of which may fairly safely be assumed to have been a long front mid vowel and some kind of mid-height diphthong (see Table 1). In the late OE period, *eo* appears to have been monophthongized to a rounded front mid vowel, usually symbolized /ø:/; as a result, the distinction between 'he' and 'she' was maintained by a contrast between an unrounded front mid vowel and a rounded one. In ME spelling, the rounded vowel is usually represented by <eo>, sometimes <oe>, <ue> or <u>. Subsequently, /ø:/ became unrounded and fell together with /e:/. The latter, systemic change seems to have begun in the eastern and northern dialects, and took place at markedly different rates in different areas.[5] The regular result in the third-person pronoun system would have entailed the loss of formal gender distinction; in those (non-southern) dialects where OE *heo* was also the form for 'they', formal distinction of number would likewise be lost. The resulting system would have been markedly less functionally efficient; however, as is well known, the situation was remedied in most dialects by the adoption of new forms of the present-day 'she' and 'they' types, which gradually spread out from the northeastern Scandinavian-influenced areas.[6]

Period / material	**Masculine**	**Feminine**	**Plural**
Old English (OE)	he	heo	hie / heo
ME reflexes of OE forms	he	he	hi / he
ME – Present-day English (adjusted system)	he	she	they
LALME Herefordshire texts (14th and 15th centuries; main types only)	he	heo (11 texts) she (10 texts) he (2 texts)	hi, they

Table 1. Simplified summary of the development of the English third-person pronoun system

Text	(LALME code, repository and manuscript, contents)	Attested forms of 'she'
7260	London, BL Royal 17 B xliii *(Mandeville's Travels)*	sche (schee)
7280	London, BL Harley 2281 *(Prick of Conscience)*	heo
7301	Cambridge, CCC 293, hand A *(Piers Plowman* C-text)	sche (hue) ((he she))
7302	Cambridge, CCC 293, hand B *(Piers Plowman* C-text)	hue 6 sche 4 he 3
7310	Oxford, Bodleian Laud Misc 553, hand A *(Agnus Castus)*	he (((hee)))
7320	London, BL Harley 2376 (*Piers Plowman* C-text)	he (((sche)))
7330	Oxford, Bodleian Digby 171 (*Piers Plowman* C-text)	scheo ((sche 3heo s3heo)) > heo ((he))
7340	Cambridge, CUL Dd.VI.29, fols 110-124v (medica)	sche 2
7350	Oxford, Bodleian Douce 78 *(Titus and Vespasian)*	sche ((schee))
7361	London, BL Sloane 5, hand A (medica)	she
7370	Oxford, Bodleian Rawlinson B 171 *(Prose Brut)*	she
7380	Princeton, Garrett 138 *(Prick of Conscience)*	scheo 2 heo 1 hue 1
7391	Oxford, Bodleian Tanner 201, hand A *(Mem. Credencium)*	hue 3 > heo 10+
7392	Oxford, Bodleian Tanner 201, hand B *(Mem. Credencium)*	heo 5
7410	London, BL Add. 46919 (William Herebert, poems)	hoe 3 a 1
7420	Cambridge, CUL Kk.I.12 (*Prose Brut*)	sche (((he schee she)))
7430	London, BL Cotton Cleopatra D ix, hand B *(Southern English Legendary: Gregorius)*	heo (((hij)))
7450	Cambridge, St John's College B.12 *(Confessio Amantis)*	heo (((sche he)))
7460	Cambridge, Selwyn College 108 L.1 (New Testament)	heo ((he))
7481	Oxford, Bodleian Rawlinson B 173, hand A *(Prose Brut)*	sche ((scheo))
7500	London, BL Harley 201, hand A *(Robert of Gloucester)*	heo ((he)) (((ho)))
7510	Oxford, St John's College 6 (Lydgate, *Troy Book)*	sche ((scheo)) (((heo)))
7520	Longleat, Marquess of Bath's MS 5, fols 1-35.	scheo (sche) ((heo))
9260	London, BL Harley 2253 (miscellany)	heo hue ((he)) (((ho hy)))

Table 2. The attested forms of 'she' in the LALME texts localized in Herefordshire

For the pronoun 'she', the Herefordshire material from the fourteenth and fifteenth centuries shows a large number of individual forms (see Table 2), which may be grouped to form three main types: *heo, he and she*.[7] The *heo* type (including the variant spellings *hoe, hue*) is most common, appearing as the dominant form in eleven of the 24 texts, with *she* (including *sche(e)*) a close second, being dominant in ten texts. *He* (including *hee*) is least common, occurring as a dominant form only in two texts; however, it is present as a minor variant in most of the texts.

It might be assumed that this pattern reflects, in a direct way, an ongoing change within the spoken mode, in accordance with the general historical development outlined above: the *heo* type gradually disappears, leaving a dysfunctional *he,* which is immediately replaced by the new *she* type. On closer inspection, however, this assumption involves several problems.

Firstly, although the *heo* forms are numerous, their orthographic form is exceptional in the Herefordshire material: *heo* is, in fact, the only lexical item where the spelling <eo> (or <oe>, <ue>) is regularly retained in texts dated after the mid-fourteenth century. On the basis of the evidence both of rhymes and back spellings, it seems fairly certain that the distinction between /ø:/ and /e:/ must in general have disappeared from all or most spoken systems at least by the second half of the fourteenth century.[8] However, in the single form *heo,* <eo> spellings appear even as late as the mid-fifteenth century.

Because of the isolation of these forms, it should be asked whether it is plausible to assume a correspondence between them and any distinctive spoken-language feature. On the one hand, each word having its own history, *heo* might simply have retained a rounded vowel longer than other words. A connection between a heavy functional load and the late retention of rounding has been made by Sundby (1963: 141 ff.) and Kristensson (1977), with reference to the entire class of /ø:/ (i.e. the long variety of the rounded front mid vowel, as opposed to the short one). A comparatively late retention of the rounded vowel in *heo* would certainly make sense from a functional point of view. On the other hand, the period of time involved would seem exceedingly long for a phoneme to survive marginally, in a single word. Moreover, the late survival of /ø:/ in *heo* also seems unlikely in light of the numerous back spellings of <eo> for historical /e/ in many of the texts in which *heo* is dominant.

The similarly frequent *she* type also involves a problem with regard to correspondences with spoken features. Comparison with the traditional dialects of the present century gives little positive evidence for *she* ever having formed part of the spoken system. The dialects of the Southwest Midland area are, in fact, unique among traditional varieties of English in that they, at least in certain contexts, lack formal

gender distinction in the personal pronoun system; according to a local saying, 'everything in Herefordshire is he, except the tomcat who is she' (Leeds 1985: 22). The normal spoken-language form both for 'he' and 'she' is *er,* although the object form *'im* may be used for the masculine when required for clarity, as in *'im went out but er didn't'* (cf. Leeds 1985: 23).

The form *er* 'she' is usually derived from the object form *her,* and assumed to represent the general tendency for pronoun exchange found in the Southwest. The masculine pronoun has been explained differently: Ihalainen (1994: 216) derived it from a low-stress form that came to be identical to the feminine form from pronoun exchange: 'Hyper-rhoticity probably accounts for the merger . . . The feminine pronoun derives from *her* whereas the masculine pronoun comes from the weak form *a* [ə], which introduces an *r* in final position.' Ihalainen's derivation is supported by 'the fact that in some varieties of south-western English, most notably in Devon, the masculine enclitic is *a,* whereas the feminine pronoun is *er.*'

It is, however, doubtful whether pronoun exchange alone can account for the form of the feminine pronoun in the Southwest Midland area. It has been pointed out (Duncan 1972: 190) that the area of *er* 'she' is clearly different from the southwestern area of general pronoun exchange, and that *(h)er,* unlike all other pronoun-exchange forms – including *(h)im* – is used consistently as a subject, both in stressed and unstressed position. Moreover, it might be argued that the coincidental loss of a major formal grammatical distinction as the result of two pronouns undergoing entirely different developments is not an intrinsically attractive explanation.

A more plausible explanation might be to derive both *er* 'he' and *er* 'she' from ME *he,* used without gender distinction. The phonological development causes no problems: in modern dialects, 'aitch-dropping' is a general feature in the Southwest Midlands, as in most parts of England, and the speech especially of the Hereford and Worcester area is marked by extensive hyper-rhoticity, that is, the addition of (usually retroflex) *r* after final schwa. Apart from well-known examples such as *yeller, feller,* there are recorded present-day forms *mer* 'me' and *ther* 'thee', which form an exact parallel to *er* 'he' (Leeds 1985: 15).

3.3. The main problem with the suggested derivation of *er* 'he/she' from ME *he* 'he/she' is that the latter type of system occurs relatively infrequently in the ME written materials. If a direct correspondence is assumed between the written and the spoken forms, as in the case of *mon,* there are no very strong grounds for postulating a widespread system without formal gender distinction, from which the modern one could be derived. Even though the *heo* and *she* forms are in themselves problematic,

the fact remains that they are dominant in the surviving texts.

This dominance is, however, based on the assumption that all informants are of equal evidential value. Bearing in mind the linguistic character of ME scribal texts – the likelihood that they reflect constrained selection rather than active inventories – it would seem inevitable that some texts provide better evidence about a particular dialect than others. Accordingly, it may not be enough to consider the overall figures relating to the Herefordshire texts: in order to make sense of the distribution of forms, we should look more closely at the individual texts, taking into account their textual background and relative status as dialectal evidence.

A clear difference emerges, first of all, between texts with a northern or eastern dialectal background, and those with a more local textual history. Virtually all the texts showing dominant *she* can be shown to represent scribal translation from a northern or eastern dialect. These include: BL MS Royal 17 B xliii of *Mandeville's Travels;* three manuscripts of the Prose *Brut* (Bodleian Rawlinson B 171 and B 173; Cambridge University Library Kk.I.12); two medical manuscripts (Cambridge University Library Dd.VI.29 and BL Sloane 5); Oxford, St John's College MS 6 of Lydgate's *Troy Book,* and Longleat, Marquess of Bath's MS 5 of a Wycliffite sermon handbook. One very late manuscript, Bodleian Douce 78 (*Titus and Vespasian*) already shows the influence of standardisation.

In contrast, virtually all texts with a local, or at any rate southern/western, background show *heo* as the dominant form. These include: BL MS Harley 2281 of the 'southern version' of the *Prick of Conscience,* a text with a geographically restricted spread centred on the Southwest Midland area; Bodleian MS Tanner 201 of the *Memoriale Credentium,* a text surviving in dialectally closely related copies and probably originating in Gloucestershire; the poems of William Herebert (BL MS Add. 46919), an authorial holograph by a writer of Herefordshire origin, and BL MS Harley 201 of Robert of Gloucester's *Chronicle. Heo* is also the dominant form in Cambridge, St John's College MS B.12 (34) of Gower's *Confessio Amantis,* in which it has been shown to represent a fairly thorough scribal translation into Herefordshire usage.[9] Finally, two of the three manuscripts of the C-text of *Piers Plowman* in the Herefordshire material (Cambridge, Corpus Christi College 293 and Bodleian Digby 171) show a mixture of *she* and *heo* type forms; as has been shown by Samuels (1985 [1988]: 80), such mixtures go back to Langland's own usage.[10]

Only two texts in the material show *he* as the dominant form for 'she', even though *he* appears as a minority form in most of the texts. However, these two texts were on close analysis judged to provide exceptionally good evidence for the Herefordshire dialect. The text of the herbal *Agnus Castus* in Bodleian MS Laud Misc

553 shows a regular and strongly regional usage; as it probably reflects a date fairly early in the fourteenth century, its use of *he* rather than *heo* is all the more significant (see Black 1996: 94, 239). The other text, BL MS Harley 2376, is a version of the C-text of *Piers Plowman,* which shows an unusually thorough translation, involving all levels of language, into the scribe's own usage. In comparison with other manuscripts of *Piers Plowman,* this usage would, moreover, appear to correspond relatively closely to a spoken system (see Black 1998b and forthcoming).

It would, then, appear that the *she* type is mainly restricted to scribal translations of northern and eastern texts, and may simply have belonged to the passive repertoires of Herefordshire scribes, without forming part of spoken usage. The question remains, however, how the dominance of the *heo* type forms should be interpreted. It was already noted in the previous section that the possibility that the *heo* forms reflect an actual survival of /ø:/ seems very unlikely: why, then, should <eo> be retained in *heo* long after its use had been discontinued elsewhere?

The most reasonable explanation, it would seem, has to do with the different pragmatic requirements of the written and spoken media. Compared to speech, writing requires a much higher degree of explicitness, due to the lack of immediate speaker interaction: accordingly, certain syntagmatic tracking devices, such as the distinction of gender and number, are inherently of greater functional importance in the written mode than in the spoken. The situation in modern Finnish might be used as an example. The Finnish system of personal pronouns does not distinguish gender: *hän* corresponds to both 'he' and 'she'. This causes few communicative problems in the spoken mode, as the immediate context makes it clear (if necessary) whether 'he' or 'she' is meant, and any misunderstandings can be corrected without delay.[11] Problems do, however, arise in the written mode. This is especially true of the translation of texts from other languages into Finnish, where various, sometimes cumbersome, ways of circumlocution are required; as a reaction to this, there have been occasional half-serious calls for the introduction of gender-specific orthographic forms.

While such linguistic engineering is unlikely to take place, the point serves to illustrate the difference between the situation in Finnish, on the one hand, and the situation facing a Southwest Midland scribe in the late ME period, on the other. For the latter, unlike for present-day writers of Finnish, gender-specific written forms were readily available for selection. By the fourteenth century, the written form *she* would have been familiar to any scribe or reader of northern or eastern texts, and would naturally come to be copied by scribes translating out of these dialects, as part of their passive repertoire. During the late ME period, with a steady increase in the production and circulation of texts, the use of *she* in writing could hardly have failed to spread.

Even in areas where a *she* type pronoun did not belong to the spoken system, its use in writing would have had two important communicative advantages: *she* was already widely used and understood, as well as being gender-specific. Accordingly, it may be assumed that written forms of the *she* type eventually came to spread in active use as well as in passive repertoires, becoming part of the 'loose change' effect that preceded standardisation (cf. 2.2 above).

Before this took place, however, the most obvious solution was simply to retain the traditional spelling *heo*. Throughout the ME period, the Southwest Midland area seems to show a comparatively large-scale production of texts and, as a result, a continuous and somewhat conservative spelling tradition. The loss of distinction between /e:/ and /ø:/ during the thirteenth and fourteenth centuries had, first of all, lead to the confusion of the graphs <e> and <eo> and, eventually, to the latter being discarded in most contexts. However, as the distinction between *he* and *heo* served a very useful function, retaining the traditional spelling for the feminine pronoun made very good sense. As <e> and <eo> had simply become variant spellings for a single sound, *heo* (quite apart from being the traditional form) would have had the advantage over *she* in its closer correspondence to the spoken form. Eventually, however, the wide geographical currency of *she* had to weigh in the favour of the latter.

Accordingly, in the fifteenth century, the spelling <heo> appears to have been purely a feature of the written language, and there is thus no need to postulate a prolonged survival of the phoneme /ø:/ in a single word. This explanation also makes sense of the correspondence between the medieval and modern patterns in Figure 5. The modern form *er* is not directly derived from late ME *heo;* instead, the latter is simply an orthographic form used in areas where the spoken system was already characterized by the lack of formal gender distinction still typical of the traditional Herefordshire dialect.

4. *Conclusions*

The example of the forms of 'she' in the late ME materials illustrates at least three important points about the relationship between the written and spoken media in ME. The first concerns the degree of correspondence between the two systems. As shown above, the medieval Herefordshire material contains three main types of forms for 'she', as well as various minor forms. It appears that only one type may have been current in the spoken mode, and that this one corresponds to the least common of the written types. Accordingly, in this particular case the written and spoken media

function as fairly independent systems.

Secondly, the number of distinct types appears to be higher in writing than in speech. This relates to the model shown in Figure 3: both the written and spoken modes are variable, and the influence of the former on the latter is not necessarily always conservative. A ME scribe would simply have one repertoire of spoken variants and another of written ones; while the two are connected, the actual choices will reflect different factors, such as different communicative needs.

The third point has been made numerous times before, but may still be worth stating. Even if we had, in the study of ME, the luxury of unlimited materials, the complexity of the material means that the specific character of each individual text cannot be ignored. Unlike modern dialectologists, we cannot choose our informants, nor can we go back to ask them additional questions when the surviving data is insufficient. Every text has, therefore, to be analysed carefully before its status as evidence can be evaluated: simply looking at overall figures means that we are likely to miss the significant patterns. Accordingly, the methods of, for example, corpus linguistics cannot be directly transferred onto the ME material with the expectation that the results will make immediate sense. What we can do, instead, is to take the informants on their own terms, and with some patience and collaboration we might be able to make the dead speak.

NOTES

[1] This work is planned to form the first stage of the recently launched Middle English Grammar Project, the eventual aim of which is to produce comprehensive accounts of Middle English on all levels of language. A survey of ME orthography is now under way, the principal co-workers being Drs Jeremy Smith and Simon Horobin (University of Glasgow) and the present writer (Stavanger College).

[2] Cf also Smith (1996: 17).

[3] These texts formed the basic material in Black (1996), and include all scribal texts localized in Herefordshire in LALME, with the exception of four texts (LALME Linguistic Profiles 7290, 7363, parts of 7400 and 7480), which were deemed unsuitable as evidence for Herefordshire usage.

[4] See LALME, I: 27. For example, dialect vocabulary formed a major part of the collected SED data, while lexical data are a very minor element in LALME.

[5] According to Jordan (1968: 63, 65, 99), the unrounding began in the north and east

already in the OE period, while the rounded vowel remained in the Southwest Midland area until the fourteenth, or even fifteenth century; such a late dating does not, however, agree with the findings of more recent studies (Sundby 1963; Kristensson 1987). Jordan's dating seems to be based on the assumption that *eo* spellings always imply a rounded vowel, and his fifteenth-century dating may in part simply reflect the frequent occurrence of the written form *heo* 'she'.

[6] The original development of the *she* type falls outside the present argument; for a useful discussion of the traditional theories, see Duncan (1972).

[7] A fourth type, *s(c)heo,* is relatively minor in the material, and probably represents simply a 'derived variant' or orthographic permutation of the *she* type, the graphs <e> and <eo> having become functionally equivalent, at least in most contexts (for the concept of derived variants, see Benskin and Laing 1981: 77). It is significant that the two texts in which it occurs frequently (LPs 7330 and 7520) are ones that show widespread, presumably hyperadaptive, use of <eo> for expected *e*; for example, the regular form of the definite article in LP 7520 is *þeo.* The forms *a, ho, hy, hi(j)* occur as very minor variants only, and are also not included in the discussion; of these, the *ho* and *hy/hi(j)* types are relict forms that reflect a (north)western and southeastern textual background respectively.

[8] The only texts in the Herefordshire material in which <eo> appears more than sporadically in words other than *heo* are BL MSS Add. 46919 and Harley 2253, both from the first half of the fourteenth century, and (with less frequent <eo>) Bodleian MS Digby 171 and Longleat, Marquess of Bath's MS 5, both from the late fourteenth century. Moreover, with the exception of Harley 2253, all these texts contain numerous examples of confusion between <eo> (or the equivalent digraph) and <e>, suggesting that the distinction was no longer a living feature in the dialect of the scribe. Harley 2253 seems to have no back spellings of <eo> (or equivalent) for <e>, but contains a large proportion of <e> for historical *eo,* including the frequent occurrence of *he* 'she'.

[9] See Smith (1985: 91 ff.) and Okumura (1991, *passim*). The text seems to be a composite one, of which the first part is copied from a Southwest Midland exemplar; apart from occasional Gowerian relicts, this part of the text shows a strongly coloured northeastern Herefordshire usage. Gower himself seems to have used the *she* type, with a possible variant <scheo> (Smith 1985: 83, note 21).

[10] Samuels (1985 [1988]: 80) holds that 'Langland was simply availing himself of two of the variants that existed in his own dialect, and not, as Chambers thought, combining his own provincial *heo* with a *she*-form that he learned in London'. There is, however, no particular need to assume that the *she* type formed part of Langland's spoken usage: it is demanded by the alliteration less frequently than *heo* (or *he*), and his command of a wide range of dialectal forms, used for the purposes of alliteration, is well known (see

Kane 1981; cf also Black 1998b and forthcoming).

[11] It may be noted that casual spoken styles take the syncretism still further, with the single form *se* – used for animates and inanimates alike - covering the full range (apart from impersonal use) of the English pronouns *he, she* and *it*.

REFERENCES

Black, M. 1997. Studies in the dialect materials of medieval Herefordshire. Diss. Ph.D., University of Glasgow.

Black, M. 1998a. Lollardy, language contact and the Great Vowel Shift: spellings in the defence papers of William Swinderby, *Neuphilologische Mitteilungen* 99: 53-69.

Black, M. 1998b. A scribal translation of *Piers Plowman, Medium Ævum* 67 (2): 257-90.

Black, M. Forthcoming. Putting words in their place: an approach to Middle English word geography.

Benskin, M. 1991. The 'fit'-technique explained. In F. Riddy (ed) *Regionalism in Late Medieval Manuscripts and Text,* pp. 9-26. Cambridge: D. S. Brewer.

Benskin, M. and M. Laing. 1981. Translations and *Mischsprachen* in Middle English manuscripts. In M. Benskin and M.L. Samuels (eds) *So meny people longages and tonges: Philological essays in Scots and mediaeval English presented to Angus McIntosh*, pp. 55-106. Edinburgh: Middle English Dialect Project.

Duncan, P. 1972. Forms of the feminine pronoun in Modern English dialects. In M. Wakelin (ed) *Patterns in the folk speech of the British Isles*, pp. 182-200. London: The Athlone Press.

Ihalainen, O. 1994. The dialects of England since 1776. In R. Burchfield (ed) *The Cambridge History of the English Language.* Vol 5, pp. 197-274. Cambridge: University Press.

Jordan, R. 1968. *Handbuch der mittelenglischen Grammatik: Lautlehre.* 3. Auflage, Heidelberg: Carl Winter / Universitätsverlag.

Kane, G. 1981. Music 'neither unpleasant nor monotonous'. In P.L. Heyworth (ed) *Medieval Studies for J. A.W. Bennett*, pp. 43-63. Oxford: Clarendon Press.

Kristensson, G. 1977. A case of the role of functional load. *Anglia* 95: 450-53.

Kristensson, G. 1987. *A Survey of Middle English Dialects 1290-1350: the West Midland counties*. Lund: University Press.

LALME = *A Linguistic Atlas of Late Mediaeval English.* See McIntosh et al. (1986).

Leeds, W. 1985. *Herefordshire Speech.* Revised edition. Leominster: ARCH.

McIntosh, A. 1963. A new approach to Middle English dialectology. *English Studies* 44: 1-11. Reprinted in M. Laing (ed) (1989) *Middle English Dialectology: Essays on Some*

Principles and Problems, pp. 22-31. Aberdeen: University Press.

McIntosh, A. 1974. Towards an inventory of Middle English scribes. *Neuphilologische Mitteilungen* 75: 602-24. Reprinted in Laing (1989), pp. 46-63.

McIntosh, A. 1975. Scribal profiles from Middle English texts. *Neuphilologische Mitteilungen* 76: 218-35. Reprinted in Laing (1989), pp. 32-45.

Mcintosh, A., M.L. Samuels and M. Benskin, with M. Laing and K. Williamson. 1986. *A Linguistic Atlas of Late Mediaeval English.* 4 vols. Aberdeen: University Press.

Okumura, Y. 1991. A study of the language of Cambridge, St John's College, MS B 12 (34)'. Diss. M.Phil., University of Glasgow.

Orton, H., S. Sanderson and J. Widdowson. 1978. *The Linguistic Atlas of England.* London: Croom Helm.

Samuels, M.L. 1972. *Linguistic Evolution, with Special Reference to English.* Cambridge: University Press.

Samuels, M.L. 1981. Spelling and dialect in the late and post-Middle English periods. In M. Benskin and M.L. Samuels (eds) *So Meny People Longages and Tonges: Philological Essays in Scots and Mediaeval English Presented to Angus McIntosh*, pp. 43-54. Edinburgh: Middle English Dialect Project. Reprinted in J. J. Smith (ed). 1988. *The English of Chaucer and his Contemporaries*, pp. 86-95. Aberdeen: University Press.

Samuels, M.L. 1985. Langland's dialect. *Medium Ævum* 54: 232-47. Reprinted in J.J. Smith (ed). 1988. *The English of Chaucer and his Contemporaries*, pp. 70-85. Aberdeen: University Press.

SED = The Survey of English Dialects. See Orton et al. (1978), Upton et al. (1994) and Upton and Widdowson (1996).

Smith, J.J. 1985. Studies in the language of some manuscripts of Gower's Confessio Amantis. Diss. Ph.D., University of Glasgow.

Smith, J.J. 1995. The Great Vowel Shift in the North of England, and some spellings in manuscripts of Chaucer's *Reeve's Tale. Neuphilologische Mitteilungen* 95: 433-37.

Smith, J.J. 1996. *An Historical Study of English: Function, Form and Change.* London: Routledge.

Sundby, B. 1963. *Studies in the Middle English Dialect Material of Worcestershire Records.* Bergen: Norwegian Universities Press.

Tolkien, J.R.R. 1929. *Ancrene Wisse* and *Hali Meiðhad. Essays and Studies* 14: 104-26.

Upton, C., D. Parry and J.D.A. Widdowson. 1994. *Survey of English Dialects: The Dictionary and Grammar.* London: Routledge.

Upton, C. and J.D.A. Widdowson. 1996. *An Atlas of English Dialects.* Oxford: University Press.

Wakelin, M.F. 1982. Evidence for spoken regional English in the sixteenth century. *Revista Canaria de Estudios Ingleses* 5: 1-25.

Address
Merja Black
School of Humanities
Stavanger College
PB 2557 Ullandhaug
4004 Stavanger
Norway
<Merja.Black@hum.his.no>

Explanations of Sound Change: Contradictions between Dialect Data and Theories of Chain Shifting

Robert Stockwell and Donka Minkova

Abstract

We argue that certain putatively explanatory principles claimed to govern chain shifting in vowel systems, in particular the 14-15C. English vowel shift, cannot be taken as explanatory because they are readily falsified by facts documented in the Orton data for several dialects outside Southern British English. We assume the strongest possible form of the uniformitarian hypothesis, namely that chain shifts of the past had the same activating principles that can be inferred from the study of living chain shifts, to the extent that such inferences can be made at all. From this point of view we review the New York City shift, the Northern American Cities Shift, the Popular London and Cockney shifts, and the Southern States shift. Finally we examine the facts of the North Midlands shift and show that virtually all the principles which have been evoked to account for the SBE shift are falsified there. This in turn suggests that the SBE shift started from a vowel system which was less 'pure' (pure long/short) than has been believed: possibly full of diphthongs of the types that are demonstrated in modern chains to be likely to engage in shifting.

1. *Terminological preliminaries*

We refer frequently to the work of Labov,[1] especially his 1994 comprehensive survey of his many years of research on chain shifts. However, in two respects we use basic terms and concepts in a significantly different way. The first is the term *peripheral*, and the second is the phonetic values assigned to symbols.

1.1. *Tense/lax vs. peripherality*

We use *peripheral*, a term first introduced into the (historical) phonological literature in Stockwell 1973, as Lindau (1978) later recommended for phonetic purposes generally, namely as a phonetic feature representing what has often been referred to as 'tense/lax'. Tense/lax, she found, have no measurable phonetic correlates, unlike peripherality, which has clear articulatory correlates and is therefore to be preferred. Labov uses [+/- periph] not as a distinctive feature but rather to refer to two tracks, a track on the periphery of the vowel space, and a track removed toward the center from the periphery, but not in the center. These uses are very nearly equivalent, and certainly for 'pure' (non-gliding, non-diphthongal) vowels of the IPA type, we would not differ with Labov: [i e a u o ɒ) are peripheral and [ɪ ɛ æ ʊ ʌ ɔ) are non-peripheral. Presumably [ɨ ə ɑ] should be classified as non-peripheral, but there can be no contrast with respect to peripherality among central vowels.

1.2. *Diphthongs and peripherality*

The difference between Labov and us depends on the question of how diphthongs are to be regarded with respect to peripherality. We take it that all out-gliding front and back diphthongs – Vy and Vw – are peripheral. Labov takes these all to be non-peripheral (1994: 234).

The diphthongs starting in the central area, [ɨy ɨw əy əw ɑw ɑy), are arguably also peripheral because the glide moves to the periphery. In some theories of the English vowel shift, including Dobson's (1957/1968) and our own (1988), these are the intermediate stages in the development from [i:] to [ay] and of [u:] to [aw] (and are replicated in the London shift discussed below). Labov, whose theory of the directionality of shifting will be tested against the Orton data in this paper, does not take a position on the peripherality of these diphthongs.

There are two arguments for taking front and back Vy and Vw as [+periph]. First, at least the high and mid diphthongs alternate both idiolectally and dialectally with long pure vowels of the same height: {[i:] - [iy] - [ɪy)], {[e:] - [ey] - [ɛy]}, {[u:] - [uw] - [ʊw)], [(o:) - (ow) - (ʌw)]. Second, the glides represented by [-y, -w] are potentially at the peripheral extremes and one must assume a strong assimilatory attraction in that direction; indeed, it is hard to imagine that a complex nucleus beginning with any non-central vowel and ending with either of these glides could be viewed, taken as a whole, as non-peripheral. We insist, therefore, that the out-gliding diphthongs which have peripheral (tense) long monophthongal congeners are [+periph].

By the inverse of this logic, we argue that centering diphthongs like [ɪə ɛə æə ʊə

ɔə ɑə] are [-periph]: first, they glide toward the non-periphery (the center of the vowel space), and second – even more compellingly – they begin with [-periph] elements, namely [ɪ ɛ æ ʊ ʌ ɑ).

Labov, however, assumes that there are two concentric circles, an outer ring (peripheral) and an inner ring (non-peripheral) (Labov 1994: 234). Long pure vowels rise up the outer ring, and the -Vy and -Vw diphthongs fall along the inner (but not central) ring. To anticipate our conclusions, we believe that long pure vowels are stable, and that the ones which rose in the English vowel shift must have been in-gliding (centering) diphthongs like [ɪə ɛə æə ʊə ɔə ɑə]. These are the ones that can clearly be documented to rise in living vowel shifts. The ones that fall are the out-gliding diphthongs. Since, as argued above, out-gliding diphthongs are peripheral, it cannot be peripherality which governs chain shifts. Rather it is the diphthongal **type** which determines direction of change. And there is a reason for that, namely dissimilation between the components of the diphthong, for greater perceptual clarity. We return to this point, below.

1.3 *Do peripheral vowels rise?*

The arguments above force full re-evaluation of the well-known Labovian dictum that peripheral vowels rise in chain shifts, i.e. the view that they go up the outside track while something else is free to move down the non-peripheral track. The true peripheral vowels, the ones that everyone would agree are on the periphery, namely [i: e: u: o:), do not participate in observable on-going chain shifts. Most of the ones that do participate in observable on-going chain shifts, the ones documented by Labov, are [ɪə ɛə æə ʊə ɔə ɑə]. So some principle distinct from peripherality must be invoked. (The situation reported by Trudgill in Norwich (Trudgill 1974) involves fronting of high back vowels as well as shortening of many diphthongs; it does not seem to fit any coherent picture of peripheral vowel chain shifting.)

1.4. *Phonetic symbols*

Though trivial in principle, one's choice of phonetic symbols can lead to massive misunderstandings, and diagrams can look much better, or much worse, than they ought to because, for example, Labov does not follow international standards of vowel representation, and his version of Trager and Smith's system from the '50's is misleading because he has tried to simplify it: for example writing [o], referred to as 'short o', for the American vowel[2] of POT, HOT, COT, which in fact is [ɑ]; and writing [oh] for [ɔə], which thereby distorts the chain shift of back vowels, making it appear that the vowel of POT, HOT, COT is contained in the 'long' complex nucleus

[oh] when in fact it is not: that vowel would be [ɑh], the vowel of FATHER.

Another notational problem anticipates our discussion of Orton. We reject the presumption of a regular distinction between [ey] and [ɛy], and between [eə] and [ɛə]. Orton, in the 'Introduction' to the great *Atlas* (1978), levels out the first of these distinctions recorded by the fieldworkers, but not the second one. The result is the presumption of phonemic contrasts which do not exist anywhere, so far as we have been able to determine. We have bitten the bullet and leveled them both in a way which appears perhaps to have been Orton's intent also, namely [ey] vs. [ɛə]. We take [e] and [ɛ] to be allophones of the same phoneme in these contexts, with the higher (peripheral) vowel occurring before the fronting glide, and the lower (non-peripheral) vowel occurring before the centering glide. Where we are drawing diagrams of chain shifts, however, we have consistently stayed with a single symbol at each level of vowel height (i.e., we have written consistently [ɛy] rather than [ey], to avoid the appearance of graphing a change from [ɛy] to [ey] where none has occurred except allophonically).

2. *The Uniformitarian Hypothesis*

As characterized by Lass (1997: 26), this hypothesis, familiar from the 19th century, states that 'Nothing that is now impossible *in principle* was ever the case in the past'. But to be useful in the discussion of chain shifts, the hypothesis has to be restated in probabilistic terms, as Lass has done (1997: 26): 'The general distribution of likelihood in a given domain was always the same in the past as it is now.'

How does this apply to our data? As adumbrated above, if we look only at on-going chain shifts, Labov's first principle, that peripheral[3] vowels rise in chain shifts, does not appear to be likely. Only on-going shifts count as basic to understanding the mechanisms. This rules out all 'after-the-fact' shifts, since they are what we seek to explain. It turns out that only a few on-going shifts have been observed in sufficient detail and with sufficient reliability to count as evidence. The ones that we feel sure should carry evidential weight about the nature of vowel shifting in English are the New York City shifts [Labov 1966: (oh) and (eh) variables, *passim*], the Northern American Cities shift (Labov 1994, 188-91 *et passim*), and the Popular London and Cockney shifts (Wells 1982:177); the Norwich shift (Trudgill 1974, 1988) does not 'chain' in a way parallel to these others, since it involves complete fronting of high back vowels. The Southern American shift, which we will discuss below, we argue is not a shift but something else, though highly relevant. All of these can be reasonably

said to have been 'caught in the act', since they are on-going.

3. *The evidence of well-documented modern shifts*

We have discussed a number of issues related to these shifts and their interpretation in Stockwell-Minkova (1997); what follows is an extension of that discussion especially with reference to the dialect data presented in Orton's *Survey of English Dialects.*

3.1 *New York City and Northern American Cities*

The striking fact to note about these well-known shifts is that in-gliding diphthongs of which the first segment is non-peripheral are the ones which rise up the periphery. The NYC shift in CAN'T, HALVE, BATH, and COFFEE, DOG, FOG, CAUSE can be represented as in Figure 1:

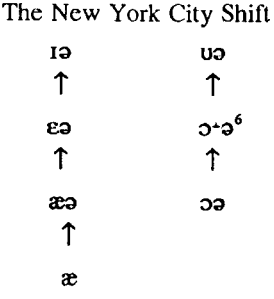


Figure 1

Labov's diagram of the Northern Cities Shift (1994: 191) is shown in Figure 2.

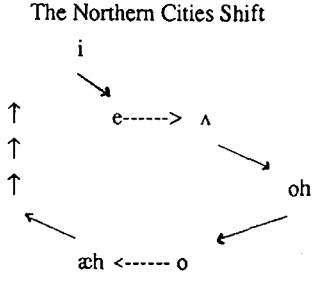


Figure 2

These are confusing pictures, because the link of [oə] to [o], the first stage of the Northern Cities Shift, and the first step in the NYC front vowel shift, [æ] > [æə], – are not shifting but simple shortening ([o]) and lengthening ([æə]), and as such are not proper parts of chain-shifting at all. In the latter case, once lengthened, then the New York chain is activated. In the former case, the [o] appears to be unrounded, fronted, then lengthened again to enter into the front chain. But this appearance is misleading: Labov's [oh] represents a variable over the set [uə oə ɔə], and [o] is a variable over [a ɑɒ]. There is no simple link of the type represented in the diagram. What is clear is that the Northern Cities shift is replicating NYC front vowel raising; and it is introducing a new shift, lowering and centralizing of short front vowels. In both localities, NYC and the Northern Cities, the **raising** is agreed by everyone to occur with in-gliding – i.e., non-peripheral – vowels, and the peripheral vowels of the BAIT BEET BOAT BOOT words are stable.

3.2. *London*

In all of these cases the diphthongs which lower and centralize their first element start out as homorganic out-gliding diphthongs – the things Labov writes VY and VW. In the first stage these are peripheral. In the second and third stages they start as central and glide to the periphery. What defines them as a natural class is not peripherality but glide directionality, diphthongal type as defined in 1.2 above. The last two columns, starting with RP [ay] and [ɔy], are non-homorganic out-gliding diphthongs, and the directionality of the change in them is raising and rounding of the nucleus, which leads to further end-point differentiation. Wells's diagram of the London shift (Wells 1982: 308), with trivial changes made for consistency with our phonetic notation, is shown as Figure 3:

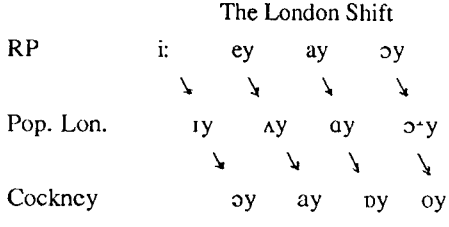


Figure 3

Even more explicitly, then, this shift can be defined as spreading the diphthong, distancing the nucleus from the glide. When the glide is homorganic, the distancing

takes the form of lowering and centralizing of the nucleus. When the glide is non-homorganic and the nucleus is round, the distancing consists primarily of greater rounding of the nuclear vowel: RP /ɔy/ –> Pop.Lon. /ɔɹy/, Pop.Lon. /ɔɹy/ –> Cockney /oy/.

3.3. *Southern States*

The salient properties which distinguish virtually all Southern States accents, from the Gulf coast up northward through the mountains to the Potomac River in the east and the Ohio and Mississippi rivers in the midlands, and westward to the Rio Grande, are these:

(1) the starting point of the nucleus in -Y and -W high and mid diphthongs is non-peripheral, significantly more-so than in other modern American accents; thus the nucleus of BEET, BEAD is [ɪy], that of BAIT, BAYED is [ɛy], that of BOOT, BOOED is (ɯw] or even [ɨw], and that of COAT, CODE generally has no rounding at the beginning, thus [ʌ w].

(2) the out-gliding diphthongs which in other accents have nuclear low vowels, thus words like MINE, TIDE, DOWN, LOUD – all with variation depending on whether the coda is voiced – tend strongly toward monophthongization: [ma:n], [ta:d]; or toward resyllabification: [dæyɨn, (læyɨd].

(3) most notably, the syllabic nucleus of LAW, DOG, BOUGHT is out-gliding [-w]. It is not the case that long open o is changing positions. Rather, the direction of its off-glide has changed – from what Labov would write with -H to what he would write with -W, from in-gliding (centering) to out-gliding [ɔə] > [ɔw]

We argue that these changes are not parts of shifts, though (1) suggests a possible first stage for the historical long [i:] shifting to [ay], and (2) suggests what must have happened to Old English -w diphthongs spelled *eo* and *ea*. It is possible that (3) gives us a clue about how Old English [a:] might have become Middle English [o:].

3.4. *Other shifts*

There are other (partial) shifts that are sufficiently clear and on-going to be noted and used as evidence to support or deny a theory of chain shifting. In Philadelphia, the centralizing of [iy] to [ɨy] in ME, SEE, FEE words resembles or even replicates what

many scholars assume was the first stage of the historical English vowel shift – a stage which takes place only after the historical [i:] had become [iy], the diphthongal type most subject to dissimilation. In Southern British English, the centralizing of [ow] to [ʌw] or even [ɛw] in KNOW, GO, HOME words exemplifies one of the possible directions of dissimilation in out-gliding diphthongs, with the nuclear vowel moving from back to central to front. In Australia the backing and rounding of [ɑy] to [ɒy] or even [ɔy] in MINE, FIGHT, LIED exemplifies further dissimilation in this diphthongal type. In Australia the lowering of [ey] to [æy] in MATE, LAID, SHAME illustrates the same principle.

4. *Modern dialect data with resemblances to the on-going shifts*

4.1. *The North Midlands*

Orton (1952) and Orton et al. (1978) set forth a number of facts about rural dialects spoken in the North Midlands by speakers over 60 years of age in the 1950's who must have acquired their accents by approximately the turn of the century. These dialects, unlike southern British English, appear to have developed through chains similar to those that Labov has found in NYC and in the Northern Cities.

We drew heavily on Orton's description (Stockwell and Minkova, 1988) in one of our efforts to show that the traditional view of the mechanisms by which the southern vowel shift took place were much over-simplified.[4] The diagram which we drew (Stockwell and Minkova 1988: 371) to represent what may be called 'The North Midlands Vowel Shift' is rather unclear. Figure 4 provides a clearer representation:

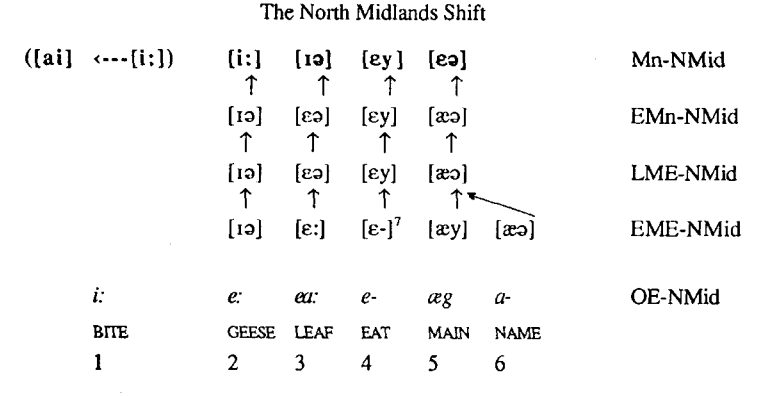


Figure 4

There are certain properties of this vowel shift which we call especially to your attention:

(1) It is not the case that it can be accounted for by the general statements that peripheral vowels rise and non-peripheral vowels fall (Labov 1994: 234). Note in particular that the vowel of column 5 becomes non-peripheral; and the vowels of columns 3, 5, and 6 are clear instances of non-peripheral vowels rising. On the other hand, the vowel of column 4, which was the focus of Orton's attention in this paper, is an instance of a peripheral vowel which has been stable, not rising, since ME.

(2) It is apparently possible for [V:] to contrast at any date with [Və], since they do in modern English (North Midlands) at both high and mid levels. The interpretation of Stockwell (1978) and of Stockwell and Minkova (1988) takes the position that the OE long vowels were of only three types – Vy, Vw, Və. This was evidently not rich enough (given the North Midland facts), contrary to criticism, which has claimed that it was too rich, that only V: and V are needed. It is clear that all three plus V: are needed, contrastively at the most basic phonological as well as phonetic levels.

Given that we greatly admire Harold Orton and his work, and remembering that this is written for a celebratory centenary occasion, we bring the following point forward with a due and proper measure of respect. Orton, to our surprise, appears to have placed theory ahead of dialects. To follow our summary, it will be useful to refer to our appendix which reproduces the OE sources[5] of the (1) EAT, (2) LEAF, (3) CLEAN, (4) GEESE, (5) TREE, (6) STREET, (7) NEED, (8) NAME, and (9) NAIL words. If the OE sources immediately leap to mind, the appendix is otiose.

4.1.1. Pattern 1

We begin with Orton's own summary (1952: 124-25) of what he calls 'four different patterns of development': that is to say, four – quite different – vowel shifting patterns within a small area of England. In Figures 5-10 our symbols are directly translatable into a kind of approximate IPA. Our normalizations are reasonable: We substitute [-y] for the palatal glide in rising diphthongs where Orton uses the IPA iota [ɪ], thus [ɛy] is to be read as [ɛɪ], though [e:] and [i:] are read as standard IPA. Modern

English endpoints are shown in boldface.

Pattern 1, exemplified in Lancashire and South Yorkshire:

EAT	> [ɛy]
LEAF, CLEAN	> [iə]
GEESE, TREE, STREET, NEED	> [i:], [iy]
NAME, NAIL	> [e:], [ɛə]

The form of the Pattern 1 shift is displayed in Figure 5, where it is apparent that only the development of [e:] corresponds with SBE.

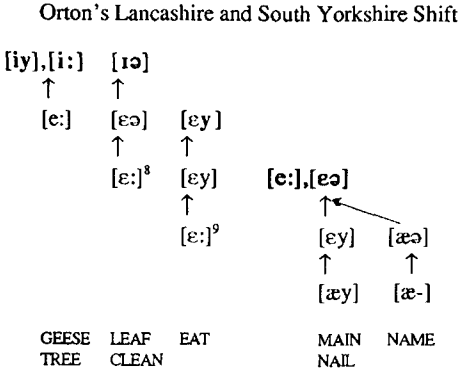


Figure 5

4.1.2. Pattern 2, exemplified by Oldham:

EAT, NAME, NAIL	> [e:], [ɛy]
LEAF, CLEAN	> [ɪə]
GEESE, TREE, STREET, NEED	> [i:]

The form of the Pattern 2 shift is displayed in Figure 6. Note especially the raising of the clearly non-peripheral [tɛə), and the stability, the failure to rise, of the peripheral [ɛy] (=[ey]). Note also the merger of [tɛə] with [ɛy] and subsequent stability.

4.1.3. Pattern 3, exemplified by Lincolnshire:

EAT, LEAF, CLEAN	> [ɪə]

GEESE, TREE, STREET, NEED	> [i:]
NAME, TAIL	> [ɛə]

The form of the Pattern 3 shift is displayed in Figure 7. Note the merger of [ɛy] with [ɪə] and the loss of the former. The main point to note is that after the LEAF and EAT words develop in-glides, they rise.

Orton's Oldham Shift

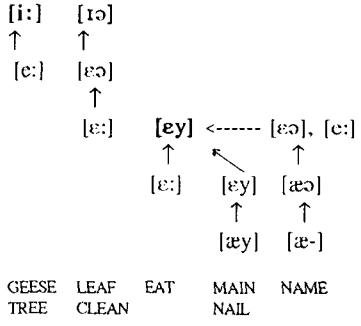


Figure 6

Orton's Lincolnshire Shift

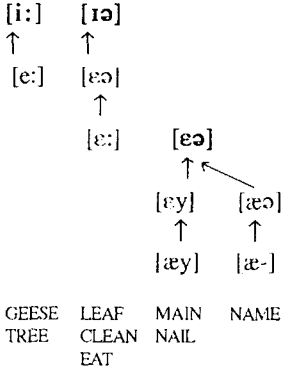


Figure 7

4.1.4. Pattern 4, exemplified by Biddulph Moor, Staffordshire:

EAT, LEAF, CLEAN, NAME, HAIL	> [i:] (note mergers)
GEESE, TREE, STREET, NEED	> [ɛy] (note – different – mergers)

The form of the Pattern 4 shift is displayed in Figure 8. The rather startling fact to note is that while there are many more mergers than in Patterns 1-3, the mergers of the second group are counter-indicated by the principle that peripheral vowels rise in chain shifts.

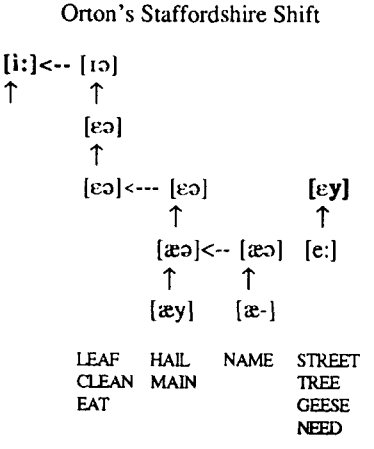


Figure 8

4.1.5. Pattern 5 is Southern British English, virtually the inverse of Pattern 4, with everything rising to the top except for the NAME and HAIL words. The ones that rise are in-gliding, non-peripheral. The ones that remain stable and do not rise are out-gliding, peripheral.

EAT, LEAF, CLEAN, GEESE, TREE, STREET, NEED	[i:]
NAME, HAIL	[ɛy]

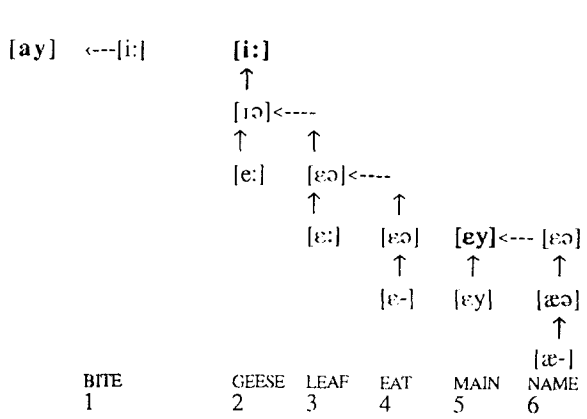


Figure 9

4.1.6. Pattern 6 is Standard American English, differing only in the loss of distinctive length (no figure is needed to represent this additional change).

EAT, LEAF, CLEAN, GEESE, TREE, STREET, NEED	[iy]
NAME, HAIL	**[ɛy]**

Orton offered a more detailed account of intermediate stages only for Pattern 1 (1952: 128). He found it 'incredible that this supposed intermediate sound could for any appreciable length of time preserve its separate identity without colliding and subsequently being levelled with either' e-1 or e-2 (1952: 127). He proposes instead that the EAT words developed from a vowel which was more open than e-2. What is fascinating is that he cites no dialect evidence to support this view. His entire argument is based on merger-avoidance. We reproduce the content of his chart of the Pattern 1 intermediate stages (using our symbols) below, and Orton asserts clearly that the in-gliding stage is purely theoretical: 'On purely theoretical grounds, it seems to me quite possible that so far as concerns Lancashire and South Yorkshire (viz. Pattern 1), the ME sounds we have been considering developed as' [shown below in Figure 10]:

Merger-Avoidance in Lancashire and South Yorkshire (Orton)

	ME	LME	ENE	NE
4, 5, 6, 7	> [e:]	> [i:]	> [i:]	> [i:]
2,3	> [ɛ:]	> [e:]	> [ɛɔ]	> [ɪə]
1	> [ɛ:] (lowered)	> [ɛ:]	> [ɛy]	> [ɐy]
8	> [a:]	> [æ:]	> [æɔ]	> [ɛə]
9	> [ay]/(?[æy])	> [æ:]	> [æɔ]	> [ɛə]

Figure 10

4.2. *Implications for vowel shift principles*

Let us review the data: in Lancashire and South Yorkshire the GEESE, TREE, STREET, NEED words went through a vowel shift which was like the one we find in the South, namely they were raised from [e:] to [i:].

But the LEAF (2) and CLEAN (3) words not only did not merge with them, as in the South, but they developed into a nucleus which is not found at all in the South, namely [ɪə]. Furthermore, Orton (1952: 128) speculated that the intermediate stage between the 'long open <e>' of ME and the present-day [ɪə] was [eə], which we regularize to [ɛə]. Orton provided no independent argument to support this speculation, but in fact there was a rather strong 'systemic' argument available, namely the parallel development of the later stages of the NAME and NAIL words to [ɛə], one slot lower in the system. Stockwell (1962: 668, 1978: 341) speculated that even in the South where centering glides do not occur in PDE at all – except as replacements for [-r], a completely independent issue – such glides played a role in the vowel shift. In Orton's Lancashire and Yorkshire data these glides occur with great regularity in the reflexes of the LEAF, CLEAN, NAME, and NAIL words. This fact is itself sufficient basis to support Orton's speculation that [ɛə] was the ancestor of Lancashire and Yorkshire [ɪə]. The centering glides are clearly the norm in all of these categories, and the rare [iy] or [ey] that shows up in the data must be taken as interdialectal borrowing from the prestige accent (Orton's suggestion, and, we believe, correct).

And most surprising, the EAT (1) words developed an out-gliding nucleus [ɛy] with utter disrespect for the rest of the system. The rest of the system is strongly in-gliding, even in the nuclei that were out-gliding in OE, ON, and OFr (set 9, the NAIL

words). It is hard to emphasize sufficiently the contrariness of these developments. No current theory of chain shifting predicts this; and in fact the principal existing theory of vowel shifting, namely Labov's, predicts that exactly the opposite should have occurred in all but the top row, which follows the prediction that peripheral vowels rise.

5. *Principles of chain shifting*

In order for an analysis of a historical English shift to be supported by the evidence of modern English chain shifts, it appears, from the above Modern English dialect evidence and the on-going shifts, that :

a. nuclei which move upward have centering glides, and
b. nuclei which move downward have homorganic out-glides, front with front vowels and back with back vowels.

There is no evidence in living dialects – except Norwich, to which we return below – that any other kind of shifting in complex nuclei is possible: there are no cases of pure long vowels rising, for example. (Remember, the GVS is not counter-evidence to this claim: no one has proved that the shifting nuclei were long pure vowels, and indeed the very fact of their shifting suggests that they were not.)

Our main point about the use of dialect evidence in reconstruction, then, is this: if there is an obvious dominant result – in this case, in Pattern 1 as in Lancashire and South Yorkshire, with massive development of in-gliding (centering) nuclei – our reconstruction should say, these are not innovative but inherited.

But if we say the centering complex nuclei were not innovative but inherited, the question becomes, inherited from how far back in time? What is the time-depth of the centering nuclei? Orton reconstructed, at the EMnE stage, in-gliding nuclei for the LEAF, CLEAN, NAME, and NAIL words. But the *only*, repeat *only* evidence for this reconstruction is the modern dialect evidence. On the other hand, the only evidence *against* this reconstruction is flimsy: the orthography of Anglo-Norman scribes. These scribes were not phoneticians. They had Latin spelling traditions, and had never heard of in-glides and out-glides, to say nothing of schwa. We have no comprehensive descriptions of English pronunciation before the latter half of the 16th century (John Hart is really the first such). Virtually all philologists and linguists have taken the spellings at face value: but should we? We know they didn't mark length: why should

we assume they marked vowel quality reliably? If we don't assume that, then 'some sort of long <e>' could just as well have been [ɛə] as [e:], all the way back, in fact, to Anglo-Saxon and even West Germanic times. And *mutatis mutandis* throughout the system.

6. *Centering glides today*

We have seen that centering glides are rampant in the north. How about the south? We know they don't exist in London and SBE generally (except as reflexes of [-r], which are irrelevant). Is there evidence that centering glides once existed widely in the south? Well, yes, there is, but it's not very strong. It consists of a small number of rising diphthongs (actually a large number if we include place names) which must have resulted from *Akzentumsprung*. *Akzentumsprung* can *only* operate on complex nuclei – two non-identical vocalic elements, V1+V2, with the further stipulation, at least for English, that V2 must be lower than V1 (and indeed we can probably stipulate that the first element must be not only higher than V2 but must have become, if it was not already, [i] or [u]). If there were no other evidence (there is considerable spelling of strange-looking ME diphthongs in the South, but it is not certain that these complex vowel spellings must be interpreted as in-gliding diphthongs), this should make one suspect that in-gliding diphthongization must have existed as much in the south as in the midlands and north.

One may push a step further and suggest, on the basis of this evidence supporting substantial amounts of in-gliding diphthongization both north and south of the home counties, perhaps the vowel shift even in the home counties was based on the same (and, we believe, if not only, at least primary) phonetic motivation we find evident all over the NW Midlands, namely 'distancing' between the two elements of the diphthong for perceptual optimality. (Labov has called this 'nucleus-glide differentiation'.)

Assuming that the ultimate target of a centering diphthong is a point maximally distanced from the out-glide end-points, i.e. the -y and -w of the peripheral diphthongs, namely some kind of low central [a] or [ɑ], we can argue that the reason that in-gliding diphthongs raise the first element is perceptual optimization: [æə] is worse than [ɛə] which is worse than [ɪə]. In the back, [ɒə] is worse than [ɔə] which is worse than [ɔɨə] which is worse than [ʊə]. Put another way, Labov has the motivations for chain-shifting in English (and indeed throughout Germanic) backwards: it is not that peripheral vowels rise, because there is no phonetic motivation for that claim to be

true; rather, it is that the elements of in-gliding diphthongs distance themselves from each other for optimal perception, which raises the first element. In fact, true peripheral vowels, rather than rising, are quite stable. [ey] and [ow] have been stable since late Anglo-Saxon times (*they, grow*). Whenever [ɪə] has become [i:] it has remained stable (but [ɪy] races down the middle to [ay] by the same distancing principle we saw above, which forces the nucleus of out-gliding diphthongs to fall, as it has done in London, Philadelphia, and of course most famously Australia and New Zealand). Whenever [ʊə] has become [u:] it has remained stable unless, as in the upper Rocky Mountain states, *all* the back vowels have lost their rounding (but [ʊw] historically becomes [aw] by the same principle just noted for [ay]). Needless to say, this entails that by the time of the vowel shift, the nuclei which participated in it had all become diphthongs. They were not long vowels, as traditionally conceived. The ones that rose were in-gliding. The ones that fell were out-gliding. The ones that underwent *Akzentumsprung* were certainly in-gliding.

But how, then, do we explain the lack of in-glides (discounting those which derive from post-vocalic -R) among the modern SBE vowels? This is our ultimate wild speculation. Standardization is a kind of purifying and stabilizing process. It throws out excessive variety, and it regularizes spacing and style of articulation. Because of [-r] vocalization, a large number of new in-gliding vowels were created. There was no hope of retaining older ones.

NOTES

[1] We accept the *facts* documented by Labov and colleagues, but not the *interpretation* of these facts. Indeed, we believe the interpretation is to a considerable extent the reverse of what is happening and has happened. We will present arguments concerning the correctness of this reversal, below.

[2] The examples which follow here and throughout the text written with small capital letters are **types**, not tokens. Wherever they appear, one can read 'words like X, Y, Z'. In this we are following the excellent example of Wells (1982), though we have not selected the same exemplars of the types unless by accident.

[3] Sometimes referred to as 'long' or as 'tense' vowels by Labov; but in diagrams of shifts, like the one on p. 234 of his 1994 book, it is clear that these are the ones he understands to be peripheral.

[4] The passage being referred to is found on pp. 370-71. In reviewing these arguments recently, we noticed that there is a minor error which must have made every careful reader of it wonder what we were saying about the vowel shift in Lancashire and south Yorkshire. On p. 371 where we should have written, 'The Modern English set of contrasts for the above developments is [ai, i:, ɪə, ɛi, ɛə],' we wrote 'The Middle English set of contrasts ... *et seq'*, for which we apologize.

[5] Taken from Orton (1952: 99-101).

[6] The vowel represented by open o(ɔ) plus the raising/closing sign (ˬ) here is the rounded congener of [ʌ]. It is sometimes referred to by some such phrase as the 'coastal New England short "o"'.

[7] From MEOSL (OE short *e*-).

[8] From OE long *æ* and long *ea*.

[9] From MEOSL (OE short *e*-).

REFERENCES

Dobson, Eric. 1957/1968. *English Pronunciation 1500-1700.* Oxford: Clarendon Press.

Labov, William. 1966. *The Social Stratification of English in New York City.* Washington, D.C.: Center for Applied Linguistics.

Labov, William. 1994. *Principles of Linguistic Change: Internal Factors.* Cambridge, Mass.: Blackwell.

Lass, Roger. 1997. *Historical Linguistics and Language Change.* Cambridge: Cambridge University Press.

Lindau, Mona. 1978. Vowel features, *Language* 54, 541-63.

Orton, Harold. 1933. *The Phonology of a South Durham Dialect.* London: Kegan Paul, Trench, Trubner and Co., Ltd.

Orton, Harold. 1952. The isolative treatment in living North-Midland dialects of OE *ĕ* lengthened in open syllables in Middle English, *Leeds Studies in English and Kindred Languages*, Numbers 7 and 8, 97-129.

Orton, Harold, Stewart Sanderson, and John Widdowson (eds). 1978. *The Linguistic Atlas of England.* Atlantic Highlands, New Jersey: Humanities Press.

Stockwell, Robert P. 1962. On the utility of an overall pattern in historical English phonology. In Horace G. Lunt (ed) *Proceedings of the Ninth International Congress of Linguists, Cambridge, Mass., 1962*, pp. 663-71. The Hague: Mouton.

Stockwell, Robert P. 1973. Problems in the interpretation of the great English vowel

shift. In M. Estellie Smith (ed) *Studies in Linguistics: Papers in Honor of George L. Trager*, pp. 344-62. The Hague: Mouton.

Stockwell, Robert P. 1978. Perseverance in the English vowel shift. In Jacek Fisiak (ed) *Recent Developments in Historical Phonology*, pp. 337-48. The Hague: Mouton.

Stockwell, Robert P. and Donka Minkova. 1988. The English vowel shift: Problems of coherence and explanation. In Dieter Kastovsky and Gero Bauer (eds) *Luick Revisited* pp. 335-417. Günter Narr Verlag: Tübingen.

Stockwell, Robert P. and Donka Minkova. 1997. On drifts and shifts, *Studia Anglica Posnaniensia* 31: 283-303.

Trudgill, Peter. 1974. *The Social Differentiation of English in Norwich.* Cambridge: Cambridge University Press.

Trudgill, Peter. 1988. Norwich revisited: Recent linguistic changes in an English urban dialect, *English World-Wide* 9: 33-49.

Wells, John C. 1982. *Accents of English 2: The British Isles*. Cambridge: Cambridge University Press.

Address

Robert Stockwell and Donka Minkova
Department of Linguistics
University of California, Los Angeles
405 Hilgard Avenue
Los Angeles
CA 90095
USA
<Stockwel@humnet.ucla.edu>
<Minkova@humnet.ucla.edu>

APPENDIX

1. The EAT words: OE short <e> in an open syllable (*e-*) *–steal* ('handle'), *eat, meat, kneading, speak, steal, tread, weaner* ('piglet'), *meal, break, besom, fever,* Scand *leak*
2. The LEAF words: OE long <ea> – *bean, beam, beat, dead, deaf, death, great, heap, eastward, clee* ('claw of pig's foot'), *lead, leaf, leap* ('basket'), *seam, sheaf, steamer, threap* ('argue'), *team, cheap, head, cleat* ('metal plate'), *reamy, flay*
   [Numbers 1-2 are so-called 'long e-2']
3. The CLEAN words: OE long <æ> (*i*-mut of long <*a*>) – *clean, deal, heat, heathen, each, lead, lean, leavings, mean, wreath, sheath, sea, spread, sweat, tease, wheat*
4. The GEESE words: OE long <e> (*i*-mut of long <*o*>) – *breed, breeches, feed, feel, feet, geese, gleed* ('glowing'), *green, heed, heel, keep, meet, seech* ('seek'), *teeth*
5. The TREE words: OE long <eo> – *bee, fleece, fleetings* ('milk curds'), *lief, -kneed, reest* ('ploughshare') *see, tree, weeds, wheel, three*
6. The STREET words: Angl OE long <e> from Germ long <æ> – *ate* (pt pl), *greet* ('weep'), *let, needle, read, seeds, street, cheek, breathe, dread, thread*
7. The NEED words: Angl OE long <e> (i-mut of long <ea>) – *need, reech* ('reek'), *sheet, sleeve*
   [Numbers 4-7 are so-called 'long e-1']
8. The NAME words:
   (a) OE short <æ> or <a> in an open syllable – *ache, blade, father, lading, lame, late, name, rake, rather, shape, shave, slade* ('slope'), *snake, spade, stake, swath, tale wade, water*;
   (b) Scand short <a> in an open syllable – *cake, gape, gate*;
   (c) OFr short <a> in an open syllable – *bacon, braces, case, dateless* ('foolish'), *face, favor, lace, pale, scales, space, place, spane* ('wean')
9. The NAIL words:
   (a) OE <æg> – *day-, fain, maiden, main, nail, snail, tail;*
   (b) OE <eg> – *ail, braid, -lay, played, sail, way;*
   (c) Angl long <eg> – *grey*;
   (d) Scand <ei> – *baitings, gradely, grain* ('prong'), *lake* ('play'), *lait* ('search'), *raik* ('wander'), *nay, weakly;*
   (e) Angl long <æg> (i-mut of long <a> + g) – *either, key*
   (f) Ofr <ai> – *bailiff, gay, quay, pay, ray, train, complaint, chain*
   (g) Ofr <ei> – *pray. Rail, paint, sprain, pain, reinings* ('reins')

The Phonology of Definite Article Reduction

Mark J. Jones

Abstract

This article aims to investigate the geographical and phonological distribution of reduced definite articles in northern English dialects. Previous research into the phenomenon is surveyed and a new analysis of distributional variation is presented, using existing data sources. Although the existing data does not allow a survey of the depth originally desired, a number of points of interest come to light which suggest areas for future research. Distributional maps for the whole area are presented, and the sensitivity of phonetic forms to non-segmental phonology is noted, as is the existence of differing phonological distributions across areas which possess identical phonetic forms. The article concludes that much remains to be done in this field, especially with regard to the actual phonetic realisations in different segmental environments.

1. *Introduction*

Definite Article Reduction (DAR) occurs in dialects spoken across northern England and is the conventional way of indicating definiteness within these dialects. The area affected by DAR covers all northern English counties with the exception of Northumberland and some parts of Durham. The counties included are the historic counties of Yorkshire and Lancashire and all counties as far south as Cheshire, northern Staffordshire and northern Nottinghamshire.

The term *Definite Article Reduction* is a historical one, and should not be interpreted as suggesting that realisations are somehow derived synchronically from forms approximating the Standard English (and general dialectal) *the* (cf. Lodge, 1984: 38 ff.). It should be noted that all who use DAR seem also to use the standard English *the*. Fluctuation between reduced and non-reduced forms does not appear to be

phonologically motivated.

No precise phonetic description exists of DAR, but the following phonetic variants occur according to transcriptions in the SED:

[t] [d] [d̥] [tʰ] [ð] [θ] [t͡ð] [t͡θ] [ʔ] [ʔʰ] [t͡ʔ]
[ʔθ] [ʔ͡θ] [ʔʰ] [t͡ʔt] [t͡ʔ͡θ]

As can be seen from the above, the reduced forms typically involve simple and glottalised plosives, simple and glottalised fricatives, affricates and glottal stops. Though the fricative realisations are frequently voiceless, confusion could possibly arise between the commonly occurring elision of the vocalic portion of standard English *the* leaving [ð], and voiced fricative realisations of DAR where these are known to occur in a particular variety. This syncopated article is found across a wider area than DAR (see Barry, 1972: 168) and always before vowels.

Much remains unclear about the impressionistically transcribed realisations above. The phonetic character of the glottal stops frequently involves no stop at all, but glottal stricture (Shorrocks, 1991: 174). In light of this, and of the diffuse phenomena referred to in phonetic literature as *glottalised* (Henton et al., 1992: 73), the realisations typically transcribed using the IPA glottal stop symbol [ʔ] will be referred to as *laryngealised* or *laryngeal* forms. The symbol [ʔ] will be retained, though its IPA value is only one of the possibilities considered here. The terms *laryngeal* or *laryngealised* are also preferred as less confusion arises when discussing forms such as [t͡ʔ], involving a combination of oral and glottal occlusion. These forms will be referred to as *glottalised*, though in the new research presented here, *all* plosive realisations, whether glottalised or not, are transcribed broadly as [t], leaving their exact quality to be determined. Square brackets are used rather than slash brackets to indicate the *realisational* significance of the forms within those brackets, as it is not clear what underlying form can be posited for DAR as a whole. Lodge (1984: 38 ff. and 134 ff.) derives DAR forms in Stockport from underlying /ðə/, resembling standard English *the*. The reduced article may be realised as a fricative in Stockport, but for other areas where no fricative forms occur it is not clear that an underlying form should possess a fricative. Similarly the presence of an underlying vowel merely serves as a link to the standard English form. One might as well propose an underlying form /water/ for both *water* [wɔːtə] and *Wasser* [vasɐʁ] produced by 'bi-dialectal' German speakers who know English. Both words undoubtedly have the same origin, but different historical processes have applied to the reflexes of that Proto-Germanic word which are indicative of the development and divergence of the two

varieties. One would not, *prima facie*, assume that use of both by one speaker constituted synchronic derivation from a common underlying form. It is preferable not to regard dialectal forms as secondary to (and derived from) apparently innate standard forms in dialect speakers' linguistic competence.

The phonetic variation reflected in the above transcriptional possibilities is a factor of geographical location and most dialects seem to have more than one reduced form, i.e. some dialects possess the plosive and laryngeal forms, some only the laryngeal form, and others possess fricative, laryngeal and plosive forms. The main aim of this paper is to investigate the geographical and phonological distribution of plosive and laryngealised forms. As such fricative and affricate realisations will not be discussed.

DAR is represented in literature by *t'* for the plosive and laryngeal forms, and *th'* for the fricative forms, e.g. the speech of Joseph, Heathcliff's servant from Brontë's *Wuthering Heights* (1847: 24): 'T' maister's down i' t' fowld. Go round by th' end ot' laith, if ye want to spake to him.' The dialect spelling practice of representing both [t] and [ʔ] with one symbol seems to have influenced previous studies of DAR, which have not distinguished between the plosive and laryngeal forms.

2. *Research Review*

Three large scale surveys have been conducted since the latter end of the 19th century, all of which mention the phonetic variation in forms across the area.

1) **Ellis** (1889) investigated localities across the British Isles using word lists, specimen texts and conversations for translation into the local dialect using a form of phonetic transcription (dialect palaeotype). Informants were educated natives and non-natives. This formed the basis for Joseph Wright's account of DAR in his *English Dialect Grammar* (1905).

2) **Jones** (1950) surveyed the realisations of DAR for 60 localities in the county of Yorkshire with the intention of refining distribution maps of the phonetic realisations identified by Ellis.

3) ***The Survey of English Dialects (SED).*** The forms of the definite article were not specifically examined by any one question in the Survey questionnaire. Responses to questions and the so-called *Incidental Material*, notes made by the

fieldworker in addition to the required information, inevitably contain many examples. Some of this information was examined by Barry (1972).

Summary of Surveys (pre-1990)

Wright (1905: 237 f.) identifies the following phonetic types occurring in the areas listed.

1. t in mid-east Northumberland; Cumbria; Westmorland; north, east, north-mid, south-west and south Yorkshire; north-west Lancashire; north Lincolnshire.
2. θ in mid and south-east Lancashire; west-mid Staffordshire.
3. t/θ in south-mid and west Yorkshire; north, east-mid, south-west and south Lancashire; Cheshire; north Staffordshire; Derbyshire; Nottinghamshire.
4. d/t in west Durham; north-east Yorkshire.
5. d/t/θ in north-west and east Yorkshire.

As Wright's account indicates, more than one realisation occurs in most localities, and this variation in realisations within one variety was taken to be dictated explicitly by the following segment. Ellis attempts to resolve some issues concerning the phonological distribution of DAR forms (Ellis, 1889: 295, 517, 619), but wavers between citing the preceding and following segments as being determining factors. Since Wright, the role of the following segment has been considered primary, specifically whether that segment was a consonant or vowel (Barry, 1972: 166f.; Jones, 1952: 86 ff.; Wright, 1905: 237 ff.). Presumably, this was assumed because the article functions with the following noun/adjective phrase as a syntactic unit.

Both Ellis and Wright refer to *suspended* or *modified* [t]'s. Neither of them refer explicitly to laryngeal stricture. Ellis's suspended [t] requires some comment. He himself takes some pains to explain its articulation (Ellis, 1889: 317) (Ellis's palaeotype has been replaced in the following by his glosses where necessary):

> 'The suspension of consonants is quite different from the **suspended (t')** for the definite article... The mode in which [the article] makes its presence felt is peculiar. When it is possible it hangs by a glide to the previous vowel or consonant, as in **in t' cart**... but in **t' cart's comin'**... this is impossible. It then modifies the position for the organs for (k), so that the glide on to (aa) in **t' car** is quite different from that in simple **car**. Before (t,

> d) as **t' tongue, t' dog** it intensifies the (t, d) in a remarkable manner. It never properly runs on to the following vowel, **t' old chap** and **told t' chap** have different effects as well as meanings... In no case must voice or flatus [aspiration] intervene...'

In the notes to his fieldworkers, Ellis (1889: 10*) has the following to say about suspended [t]'s :

> 'Note also particularly whether **the** does not always become a suspended **t**^{l} when it is possible, as when it follows another word, as **from t' school**, or, when this is not possible, whether it becomes just perceptible by a dull kind of minute thud, due to trying to speak without moving the tongue from the palate, as **t' man, t' ass** (not **tass**)...'

Ellis indicates that the suspended form of the definite article is *not* a long (or geminate) consonant ('... the suspension of consonants... is quite different...' etc. above), but that it seems to consist of laryngeal stricture ('...modifies the position of the organs for (k) so that the glide onto (aa)...' etc.). This 'modification of the organs' does not seem to involve the lingual articulation. Ellis has his own symbols for palatalisation etc. which are not used here. There is also the reference to the 'dull thud' caused by a non-lingual articulation, which is suggestive of laryngeal activity. He speaks of sounds becoming 'intensified'. The affected consonant 'never properly runs on to the following vowel' but that voice and aspiration do not intervene. Elsewhere, however, Ellis considers suspension to arise through assimilation to neighbouring alveolar plosives (or fricatives): 'note whether the (th) or (th^{l}) is not assimilated to (d) or (t), causing a **suspension** of the (t) or (d), by the tongue remaining a sensible time against the palate...' (Ellis 1889: 10). Here duration *is* clearly the most important feature, and long (or geminated) /t/ and /d/ are the result. Ellis uses the same terms and symbols to refer to what appear to be at least two separate articulations, the plosive [t] and laryngeal [ʔ] realisations of this paper, and possibly also the glottalised forms.

Wright has the following to say about the suspended realisation (Wright, 1905: 238): 'It is to be observed that in those dialects where the definite article has the form t, should the following word begin with a dental, the only trace of the article is the suspension of the dental.' Note here that he implicitly considers the *suspended* form of

the article to occur in dental environments only, [t] occurring elsewhere, and the quote suggests that Wright felt the articulation in this environment to consist of a long or geminate [t] (or [d] as no mention of voicing is made) only, without glottal constriction. In so doing Wright differs from Ellis, who apparently uses the term *suspended* [t] to refer to either lengthened or laryngeal articulations and does not restrict the distribution of such realisations to dental environments only. Wright considers the conditioning environments of DAR realisations to be following word-initial vowels, dentals and other consonants, with a place of articulation distinction drawn in the consonantal environments. This tripartite phonological distinction is adopted by Barry (1972).

The term *suspended* is also defined by Jones (1952: 87) as meaning an 'audible suspension of breath'. The term is phonetically vague, as Jones notes. There is no way of knowing when this suspension occurs precisely. Applied to an alveolar plosive, for example, it may occur before the closure, which would suggest [ʔ ˷t]; as an extended closure period, suggestive of a long /t/ and presumably the result of assimilation; or after the closure release, perhaps as increased voice onset time (VOT) and suggestive of changes in laryngeal stricture. Jones uses the term himself, though reluctantly, and transcribes forms involving *suspension* as[t̚]. Glottal closure is taken to be involved and to occur simultaneously (or almost simultaneously) with a closure at the alveolar ridge. His *suspended* forms are the *glottalised* forms ([t̚ʔ]) of this paper, but not, it seems, always the same as Ellis's.

Jones distinguishes between laryngealised, plosive and glottalised plosive forms in his phonetic transcription, but he makes no attempt at discovering the conditioning environments for these different forms at the localities he considered. His aim is to refine the boundaries between the three areas in the historic county Yorkshire defined by Ellis on the basis of the phonetic forms used (Jones, 1952: 81, ff.):

Type I: [t] *'or some modification of it'* before a consonant or a vowel.

Type II: [t] or *'modified [t]'* before consonants,
[θ] before vowels.

Type III: Ø realisation.

The surveys of Ellis, Jones and the SED note an area of Ø realisations in the Holderness area of eastern Yorkshire (type III above). Wright makes no mention of

this. The possibility of Ø realisations will be considered further in the conclusion.

The SED questionnaire contained twelve questions designed to elicit responses including a form of the definite article. Barry seems to have taken three questions as the basis of his investigation, reflecting the conditioning environments proposed by Wright; vowels, dentals and other consonants. The questions used were the following:

V.6.6	Where do you bake the bread? *In the oven.*
V.8.12	When you put things on the table ready for a meal, what do you say you do? *To lay the table.*
IX.2.3	In summer, you don't water your garden in the middle of the day; you wait [gesture] ... *till the sun goes down.*

Barry mapped laryngealised forms as these were transcribed by SED fieldworkers, but did not consider the conditioning environments for plosive and laryngeal forms. Jones and the SED fieldworkers attempted to distinguish impressionistically between phonetic [t], [ʔ] and [t̚ʔ] in *phonetically ambiguous* environments. This point will be discussed further below.

The following list summarises research on DAR prior to 1990:

- Realisation varies phonetically across the region.
- The exact nature of suspended forms is not clear.
- Jones does distinguish between [ʔ] and [t] phonetically, but not phonologically.
- There is alternation within one variety due to *following* segment. Wright identifies three environments, before vowels, dentals, or other consonants.
- Some attempts are made to distinguish between [ʔ] and [t̚ʔ] *before* and *after* lexical /t/ and in other phonetically ambiguous environments.

Muldowney (1990) investigated DAR realisations in the dialect of the Vale of York using tape recordings of free conversation made for the Tape Recorded Survey of Yorkshire Speech. His study focuses on two villages (Riccall and North Duffield). Neither of these villages was chosen as a locality for the SED and had not been covered by the surveys of Ellis or Jones. The nearest SED localities are York (Y19) to the north, Y24 (Cawood) to the west (the nearest SED locality), Y25 (Newbald) to the east, and Y27 (Carleton) to the south-west. Muldowney's study is based on impressionistic phonetic transcriptions of free-speech recorded between 1986 and 1988 and consists of data from 12 informants, 7 men and 5 women.

Muldowney distinguishes between plosive ([ʔ]) and laryngeal ([ʔ]) realisations

and attempts to determine what causes the alternation. The symbol [ʔ] indicates that Muldowney, like Jones (1950, 1952), considers there to be (near) simultaneous alveolar and glottal closure.

Muldowney realised that the form the article takes cannot be due entirely to the following segment, as the examples below demonstrate (Muldowney, 1990: table 6):

do t' housework [duːʔ aʊswɔːk]	**laryngeal realisation**
in t' house [ɪn ʔ aʊs]	**plosive realisation**

Muldowney (1990: 3) states that: '... the glottal stop [ʔ] and its varying relationship with the sounds on either side [of DAR] is vital to our understanding of the phenomenon.'.

This study breaks with previous studies in investigating the entire segmental context as a conditioning environment and in *separately* considering the conditioning environments for plosive and laryngealised forms of DAR. In his study of the SED material, Barry does question whether the realisation of the preposition *in* as *i'* throughout the northern counties would affect the form of the article (Barry, 1972: 167), but does no more to investigate the entire segmental context of realisations. Muldowney concludes that the determining factor in the selection of [ʔ] or [ʔ] is the presence of a preceding or following alveolar consonant. One problem associated with this analysis concerns the ability to positively distinguish impressionistically between [ʔ] and [ʔ] in ambiguous environments: we have no way of satisfactorily determining impressionistically which realisation actually occurs if the preceding or following environment is an alveolar plosive.

Like the previous studies an attempt is made to distinguish between [ʔ], [t] and [t͡ʔ] realisations in *phonetically ambiguous* environments. These environments involve final and initial alveolar plosives ([t] and [d]), and initial affricates ([tʃ]and [dʒ]), as in the following example (Muldowney, 1990: 26):

[fɛtʃt ʔ ɔːs] *fetched t' horse* (laryngeal realisation)

could equally well be interpreted as:

[fɛtʃt t͡ʔ ɔːs] (glottalised plosive realisation)

Reliable impressionistic transcription of the different forms is very difficult in these environments. Any differences which do occur are likely to be not so much qualitative

as quantitative (as in Wright's interpretation of long initial dentals above), and quantitative differences are less amenable to reliable transcription, especially if long and short instances are not available for direct comparison. As for final /d/, Shorrocks comments on devoicing of [d] before the laryngealised forms in his study of the definite article in Farnworth (1992: 174), making distinctions between voiced and voiceless alveolars difficult to draw consistently. These are best excluded from impressionistic analysis as well. The inability to determine which form has occurred discounts one set of alveolar consonants from Muldowney's conclusion. Muldowney himself excludes alveolar /r/'s, which do not condition a plosive realisation in the Vale of York data. In view of these points, Muldowney's conclusion that DAR forms are determined by *place* of articulation seems premature. An attempt to present a phonological analysis of DAR at Stockport in Lancashire (nearest SED localities Derbyshire 1 and Cheshire 2) by Lodge (1984: 38 ff., and 134 ff.) also suggests that place of articulation dictates the presence of the [t] form in this dialect., but similar exceptions to the conclusion can also be found here, e.g. [en ʔ dosben] and [pas ʔ sɑlt (Lodge 1984: 39, 135). We must conclude that place is not the determinant of the reduced article realisation in these varieties.

Another case in point involves the responses to SED question VII.2.14, intended to elicit the standard English response *we two*. Many responses in the north (Cu2, We1, We2, La3, La6, La10, Y2, Y3, Y4, Y6, Y9, Y10, Y14, Y15, Y17, Y20, Y23, Y24, Y28) and most in the west Midlands have a response glossed orthographically as '(the) two on us', with an optional article. Although the presence of the article in this environment is not required by standard English, there is no reason to equate standard and dialectal usages, and dialect use of the article is suggested by standard forms at (He7, Gl6, O4). The presence of a reduced definite article is indicated at localities with fricative realisations (La7). It is clear that transcription of DAR in these ambiguous environments is difficult to carry out reliably. In her Ph.D. on West Yorkshire dialects, Melchers (1972: 49) has the following to say about the problems of DAR transcription: 'There were often difficulties in identifying reduced forms and especially in establishing whether the definite article, the indefinite article, or no article at all had been implied.' Only a detailed instrumental analysis and comparison of DAR and non-DAR contexts may reliably tell us which is the correct analysis in these phonetically ambiguous environments.

3. *A New Survey of DAR*

A review of earlier work suggests the need for phonetically and phonologically more sophisticated approach in which:

- Plosive and laryngeal variants are distinguished.
- Entire segmental environment are considered.
- Phonetically ambiguous environments are excluded.

There are two existing data sources which have not been fully utilised for an analysis of DAR:

1) All examples of DAR from the Basic Material (BM) of the SED.
2) The recordings made as part of the SED.

The present study examined both of these, but this paper is based on an analysis of the recordings. These were made as part of the SED survey (see Klemola and Jones, this volume, for details on the recordings). Some DAR localities had no recording available (Du 6; La 3, 6, 7, 13, 14; Y 10, 12, 25; Ch 2; Db 2, 3; Nt 2). Seventy two possible DAR localities had extant recordings. Of these, 71 were found to exhibit DAR, and 63 were used for this study. The remaining 8 had very few examples of DAR. The reasons for the lack of DAR are varied, and generally seem to involve informants who were very familiar with the standard and accommodated to the fieldworkers, who in some cases were non-native English speakers.

In total around 15 hours of recordings were analysed, and over 2300 examples of DAR collected, including ambiguous environments. Once these had been excluded, around 1800 examples remained to be transcribed impressionistically.

4. *Method of Analysis*

Like other surveys to date, an impressionistic analysis of broadly defined phonetic variants was undertaken. Care was taken to distinguish between laryngeal and plosive articulations, but glottalised realisations such as [t ˭?] were subsumed under plosives. Even when phonetically ambiguous environments were excluded, unclear articulations were not analysed. In other alveolar environments, such as *vowel_/n/*, W.E. Jones frequently transcribed the DAR form as [t], stating that the 'approach to [t]

closure is clearly heard' (Jones, 1950: 8). In the analysis presented here [t] has been transcribed only when transitions to an alveolar place of articulation are heard in non-alveolar environments, or if a plosive burst is heard in any environment. Thus in an environment such as *vowel_/b/* the transcription [t] would occur if either of the above cues were judged to occur. In an environment containing a non-plosive alveolar (/s/, /z/, /r/, /l/, /n/), such as the above, the only *non-ambiguous* cue is the plosive burst, as the tongue tip has already attained the alveolar place of articulation. A transcription of [t] in such environments is indicative of that burst. If no burst is heard the form is transcribed as [ʔ]. In an environment involving a preceding alveolar and a following plosive, such as */n/_/p/*, alveolar transitions occur because of the preceding alveolar [n] and, it might be argued, a plosive burst indicating the [t] form is unlikely to occur due to the following stop. Instrumental analysis of such environments might demonstrate an additional gesture for the [t] realisation of the article but until such an analysis is carried out the situation is not entirely clear.

5. *Results*

Limitations of the Data

Like all other surveys of DAR with the exception of Jones's, the data for analysis had not been collected specifically with the intention of investigating DAR and the data for analysis were taken from free-speech recordings. In this the present study is identical to all previous studies. This has set enormous limitations on the results. If all the possible initial and final segments are considered, including syllabic /n/ and /l/ and clusters counted separately, over 500 possible DAR environments occur. The recordings include a total of 255 environments, i.e. almost half are missing. The maximum number of environments present in the recordings for one locality is 66 (Y13), but most have far fewer represented in the recordings. As a result, the localities contain few comparable environments and a region-wide survey is thus not possible to the extent desired at the outset of this study. Occasionally variation is found at one locality in identical segmental conditions. This variation may be the result of fast speech processes or speech errors (usually too few examples exist to rule the latter out), or non-segmental effects on DAR, such as stress or syntactic boundaries. Not enough data exists within the present study to speculate further, and examples are in any event seldom. Pre-pausally DAR forms seem to occur as [t] throughout the area.

The above points have made analysis of the data difficult, and have stressed the

need for systematically collected data.

An analysis has however led to the following results:

1) *Refinement of geographical distribution.*

Figure 1 shows the localities at which DAR forms occur, and shows which forms alternate in which areas. This now includes [ʔ] forms, and location of varieties having alternation between phonetically distinct variants. Wright's account of geographical variation is included above and the two generally correspond, though Wright includes more phonetic detail on stop voicing. The recordings do contain voiced allomorphs of the reduced definite article in Cumberland and Durham, which have been counted as plosives here. The area of fricative-only realisations identified by Wright in central Lancashire is not indicated in the recordings, though coverage of this area is poor. The BM data from the SED includes transcriptions of plosive, laryngeal and fricative realisations from this area. The area of Staffordshire containing fricative realisations is supported by the recordings, though here too laryngeal forms also occur. It should also be noted that fricative forms are not limited to vocalic onsets as suggested by previous studies, occurring at Cheshire 3 in the environment */n/_/k/*, as in the phrase *in th' corner* [ɪn θ kɔɹnəɹ]. No locality is without laryngealised forms, and the largest area is occupied by varieties possessing [t]~[ʔ] variation.

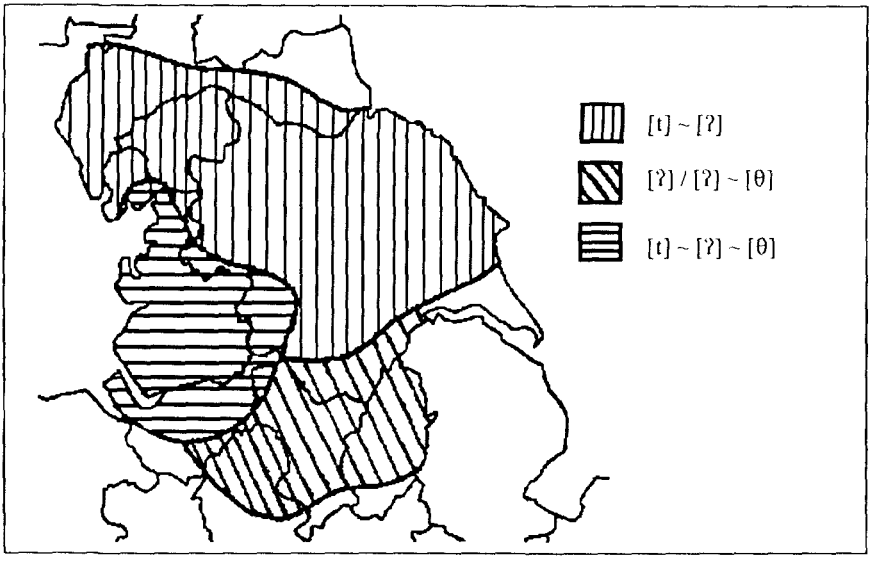


Figure 1: The Geographical Distribution of Phonetically Distinct Realisations of DAR.

2) *Clusters regarded as separate phonological units.*

Initial clusters occur infrequently within the data sample, but the following observation can be made: DAR realisations before clusters may differ from those before simple segments which occur initially in recorded clusters. Some localities having [t]~[ʔ] alternation have [t] before simple segments and [ʔ] before clusters involving those segments, e.g. Cu5, Cu6 and La4 have [t] in the environment *vowel-_/s/*, but the laryngealised realisation [ʔ] in the environment *vowel_/st/*. Viewed as a linear sequence of independent units, /s/ occurs initially in both and hence the same DAR form would be expected. The fact that this is not actually found at the localities above in the data analysed suggests that cluster initial /s/ and syllable initial /s/ are, phonologically speaking, different creatures.

3) *Phonological Variation in [t]~[ʔ] area.*

Though a large area of the DAR region possesses the same broad phonetic possibilities for DAR realisations, plosive [t] and laryngeal [ʔ], these forms do not have the same phonological distribution across this area, i.e. the same conditioning environment triggers different realisations in different localities. In the environment */n/_/r/*, We4 has a [t] realisation but Y23 has the [ʔ] form. Although comparable data is limited, this is not the only environment affected. This phonological difference affects the following preceding environments across the data sample:

Ø (utterance initial), [l], [n] and the vowels,

And the following initial segments:

[b], [k], [f], [v], [s], [ʃ], [l], [r], [w], [j], and the vowels.

The geographical distribution of these forms in the environments *vowel_/f/* and *vowel_/s/* differs (Figure 2). This demonstrates that DAR realisations do not only vary phonetically, but that different phonological types exist as well, a point missed by previous studies which did not consider the entire segmental context or laryngeal forms separately from plosives.

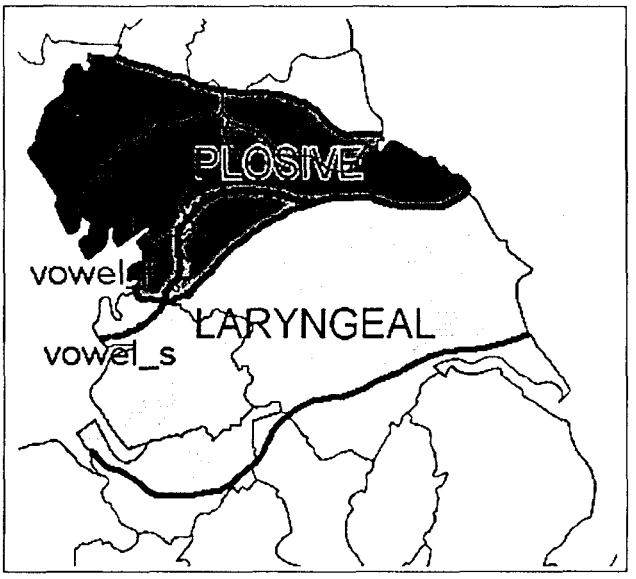


Figure 2: Phonological Variation in the Distribution of Plosive and Laryngeal Forms for the Environments vowel _f and vowel _s

6. *Discussion*

This study has attempted to apply a more sophisticated phonological and phonetic analysis to the phenomenon of definite article reduction based on recordings made at localities across the north of England. The recordings were not made specifically to obtain data on DAR, which is reflected in the poor number of possible environments represented overall. Any future studies must involve systematically collected data for phonetic and phonological analysis. A controlled sample of environments needs to be analysed to elucidate the phonological processes operating across the DAR area as many segment types do not occur at all within the data for this study. Final and initial clusters, syllabic nasals and liquids and final /r/ need particular consideration. Non-segmental aspects of DAR require further investigation. Despite these shortcomings, a number of points of interest have come to light.

It has been possible to demonstrate that the accepted geographical distribution of the phonetic variants of DAR was correct, but that the phonological treatment of DAR varies geographically for localities within the area showing [t]~[ʔ] alternation.

Northern and western localities frequently have plosive realisations in environments where other localities have [ʔ]. The environments affected have been identified, though the list should not be considered exhaustive or representative of any one variety at this stage. This difference in conditioning environment may reflect the operation of two different phonological processes in the development of DAR (such as lenition and assimilation, see below), but this remains speculation without more environments for analysis. Equally, this pattern could reflect a northward spread of environments in which [ʔ] is the only possible realisation. As can be seen in figure 2, the southernmost part of the DAR area does contain varieties in which only [ʔ] occurs in normal speech environments. The non-congruence of phonological variance depicted in figure 2 is suggestive of the spreading sound change hypothesis. Further research is needed to understand how such a change might proceed and to ascertain which hypothesis is in fact correct. It has also been shown that some localities treat clusters as distinct phonological units from simple segments. The above example showed that DAR realisations are not always identical in the environment *vowel_/s/* and *vowel_/st/*. This indicates that realisations can be dictated by non-linear phonological considerations such as branching onsets. Data for all clusters is scarce and needs further attention.

Future studies of DAR must involve a thorough phonetic analysis to ascertain the precise nature of the phonetically different realisations in different environments. The exact character of the laryngealised realisation needs to be thoroughly investigated prior to a complete phonological study of all environments (including initial/final alveolars). The realisation before initial alveolars when studied may indicate whether Wright and others were correct in differentiating between these and other plosive articulations. The geographical and/or phonological distribution of glottalised versus simple plosive realisations, a phonetic distinction ignored in this paper, needs to be examined, as it may prove crucial to the identification of differing phonological processes, lenition and assimilation, within the DAR area. This is not the place to conduct a lengthy discussion of similarities and differences between these two phonological processes. A few comments, however, will make the relevance to DAR articulations clear.

Lenition operates to increase the degree of stricture present in a certain articulation, e.g. [p] > [f] > [h] > Ø. Lenition of syllable final [t] to [ʔ] is common in British English dialects (Harris, 1990: 284 ff.), and DAR alternations between plosive and laryngeal realisations may represent the outcome of lenition. Equally, however, if a variety possesses a glottalised reduced article [t̰ʔ], laryngeal forms could result from assimilation to neighbouring segments of the supraglottal gesture in the glottalised

plosive form. This would leave the glottal gesture and, presumably, a lengthened 'assimilee' consonant. Though he never explicitly mentions laryngeal stricture, Ellis considers assimilation to contiguous alveolar plosives and /s/ to operate at several localities (Ellis, 1889: 10, 295, 448), producing suspended forms at the place of articulation. These suspended forms can only be interpreted as long consonants in this context. It is not inconceivable that durational differences alone may cue the DAR/non-DAR distinction, without any laryngeal stricture. Assimilatory processes affecting stops before other stops are discussed in Ohala (1990: 258 ff.), and examples of Latin developments into Italian are given there and in Maiden (1995: 71). Examples of regressive assimilation of stops to following fricatives and nasals also occur in the latter, e.g. Latin DIXIT > Italian ['disse] and NEC+MINUS > [nem'me:no], and for Korean /t/ to [n] before /n/ in Kim (1987: 888). Old Norse also provides such an example, e.g. [tn] as in *vatn* 'water' has become [vann] in Norwegian. This discussion does not include examples of *progressive* assimilation, which seems to be rarer in any case (Ohala, 1990: 258 ff., 271 n. 1).

It is not always clear how lenition and assimilation can be distinguished in terms of their operating environments; VC_1C_2V sequences can be affected by both, as Latin DICTUM > Italian [detto] (Maiden, 1995: 71) and the realisation of 'pitbull' in my variety of English as [pɪʔbʊl] demonstrate. The development of Sanskrit consonant clusters (VC_1C_2V) into New Indo-Aryan single segments (VC_2V) would suggest complete lenition of the first cluster member were it not for textual evidence of lengthening (i.e. assimilatory processes) in intermediate languages such as Pali (though Murray 1982 suggests a different interpretation). Intervocalic position clearly favours lenition, as Kenstowicz (1994: 35) notes. An analysis of intervocalic realisations of DAR from the SED recordings (including initial vowels resulting from 'h' dropping) does tentatively indicate a geographical difference, though the paucity of the data must be emphasised. A more thorough survey of intervocalic realisations might thus favour the hypothesis that assimilation has been at work in the north of the DAR area and lenition in the south.

Mention must be made of the fact that most preceding unstressed vowel environments consist of prepositions, such as *o'* for 'of' and *wi'* for 'with' and *i'* for 'in'. The phonology of function words frequently needs special consideration, e.g. the almost exclusive presence of initial /ð/ in standard English pronouns, adverbs and demonstratives. Consequently, it may be wise to consider the possibility that reduced articles after prepositions do not represent the outcome of any productive process, but are fossilised function words comparable to Italian *col* < *con il*, *colla* < *con la* (with the same meaning as *wi't'*) or German *zum* < *zu dem* 'to the'. Instances of other words

ending in unstressed vowels + a DAR noun phrase are needed to set the matter straight. Once again the data raises more questions than it answers and indicates what factors an in-depth survey would have to take into account.

It remains to be seen, in fact, how the phonological distributions of the laryngealised, plosive and glottalised forms differ and which varieties possess all three. Instrumental studies of durational changes accompanying laryngealised forms may indicate whether assimilation or lenition has occurred. It must also be remarked that the glottalised form [t˗ʔ] is hard to place in a lenition series. Lodge (1984: 143) notes that a change from [t] to [t ˗ʔ] (normalised transcription) hardly constitutes lenition if the latter is defined as a reduction in stricture. The glottalised form involves an *increase* in stricture, with additional closure at the glottis. Realisations involving fricatives and affricates also need to be analysed.

The question of Ø realisations in Holderness also needs to be addressed. Whilst Ø is the expected final outcome of a lenition process, acoustic analysis of data may show up some unexpected results. It is not beyond the bounds of possibility that durational factors, amplitudinal changes, fundamental frequency perturbations of surrounding vowels or other non-segmental cues, difficult to transcribe impressionistically, are used to signal definiteness. In view of this, Ø should perhaps at best be considered a possible realisation pending an instrumental investigation.

This study has demonstrated the limits of 'second-hand' data in analysing DAR and shown that it is a much more complex phenomenon than has previously been realised. This must be reflected in the methods and techniques of analysis applied to it in order to fully investigate the phenomenon.

REFERENCES

Barry, Michael. 1972. The morphemic distribution of the definite article in contemporary regional English. In Martyn Wakelin (ed) *Patterns in the Folk Speech of the British Isles*, pp. 164-81. London: Athlone Press.

Brontë, Emily. 1847. *Wuthering Heights*. Penguin Classics Edition. 1994.

Ellis, A.J. 1889. *On Early English Pronunciation, vol. V*. Early English Text Society. London: Trübner & Co.

Harris, John. 1990. Segmental complexity and phonological government, *Phonology* 7: 255-300.

Henton, Caroline, Peter Ladefoged and Ian Maddieson. 1992. Stops in the world's

languages, *Phonetica* 49: 65-101.

Jones, W.E. 1950. *An Investigation into the Phonetics of the Definite Article in Yorkshire.* Unpublished MA thesis, University of Leeds

Jones, W.E. 1952. The definite article in living Yorkshire dialect, *Leeds Studies in English 7-8*: 81-91.

Kenstowicz, Michael. 1994. *Phonology in Generative Grammar.* Oxford: Blackwell.

Kim, N-K. 1987. Korean. In B. Comrie (ed) *The World's Major Languages.* London.: Croom Helm.

Lodge, Ken R. 1984. *Studies in the Phonology of Colloquial English.* London: Croom Helm.

Maiden, Martin. 1995. *A Linguistic History of Italian.* London: Longman.

Melchers, Gunnell. 1972. *Studies in Yorkshire Dialects.* Unpublished Ph.D. thesis, University of Stockholm.

Muldowney, J.E. 1990. *The Realisation of the Definite Article in the Dialect of the Vale of York.* Unpublished MA thesis, University of Leeds

Murray, Robert W. 1982. Consonant cluster developments in Pali, *Folia Linguistica Historica 3* (2): 163-84.

Ohala, John J. 1990. The phonetics and phonology of aspects of assimilation. In J. Kingston and M. Beckman (eds) *Papers in Laboratory Phonology 1*: 258-75. Cambridge: Cambridge University Press.

Orton, Harold and M. Barry (eds). 1969-71. *Survey of English Dialects. The Basic Material,* vol. 2 (3 parts): *The West Midland Counties.* Leeds: E.J. Arnold.

Orton, Harold and W. Halliday (eds). 1962-3. *Survey of English Dialects. The Basic Material,* vol. 1 (3 parts): *The Six Northern Counties and the Isle of Man.* Leeds: E.J. Arnold.

Orton, Harold and P.M. Tilling (eds). 1969-71. *Survey of English Dialects. The Basic Material,* vol. 3 (3 parts): *The East Midland Counties and East Anglia.* Leeds: E.J. Arnold.

Shorrocks, Graham. 1985-1987 [1991]. The definite article in the dialect of Farnworth and district (Greater Manchester county, formerly Lancashire), *Orbis XXXVI*: 173-86.

Wells, J.C. 1982. *Accents of English 2, The British Isles.* Cambridge: Cambridge University Press.

Wright, Joseph. 1905. *English Dialect Grammar.* Oxford: Oxford University Press.

This material is also presented, in a slightly different form, on the DAR web-site (http://members.tripod.com/~definite_article).

Address
Mark J. Jones
51 Lydgate Hall Crescent
Crosspool
Sheffield
S10 5NE
England
<markjjones@hotmail.com>

The East-West New England Dialect Boundary: Another Look at the Evidence

Lawrence M. Davis, Charles L. Houck and Veronika K. Horváth

Abstract

This paper analyses twenty-one list manuscripts from the Connecticut portion of the Linguistic Atlas of New England (1939-43) (LANE) to examine further the concept of dialect boundaries. We chose Connecticut because of Kurath's (1939) claim that the Connecticut River forms a clear east/west boundary for New England. A second reason for choosing Connecticut is that, further north, almost all of New England falls in LANE's eastern dialect area.

So far as we know, no one before us has questioned Kurath's (1939) claim in the Handbook of the Linguistic Geography of New England that 'New England has two major dialect areas, an Eastern and a Western', and that 'the "seam" between these two settlement areas runs straight north from the mouth of the Connecticut River (between [subjects] 30 and 31) through Connecticut . . .' (p. 8). Given our findings, we think that the time has come to take another look at that analysis and the assumptions behind it.

The analysis involves a multidimensional scaling (MDS) procedure to analyze the lexical and phonological responses in the LANE records of the sixty-seven Connecticut subjects. We believe that the results from this analysis will contribute further to the continuing discussions regarding what exactly we mean when we use the term dialect boundary.

This article analyses twenty-one list manuscripts from the Connecticut portion of the *Linguistic Atlas of New England* (1939-43) (LANE) to examine further the concept of dialect boundaries. We chose Connecticut because of Kurath's (1939) claim that the Connecticut River forms a clear east/west boundary for New England. A

second reason for choosing Connecticut is that, further north, almost all of New England falls in LANE's eastern dialect area. Figure 1 shows a map of the sites sampled by LANE in Connecticut:

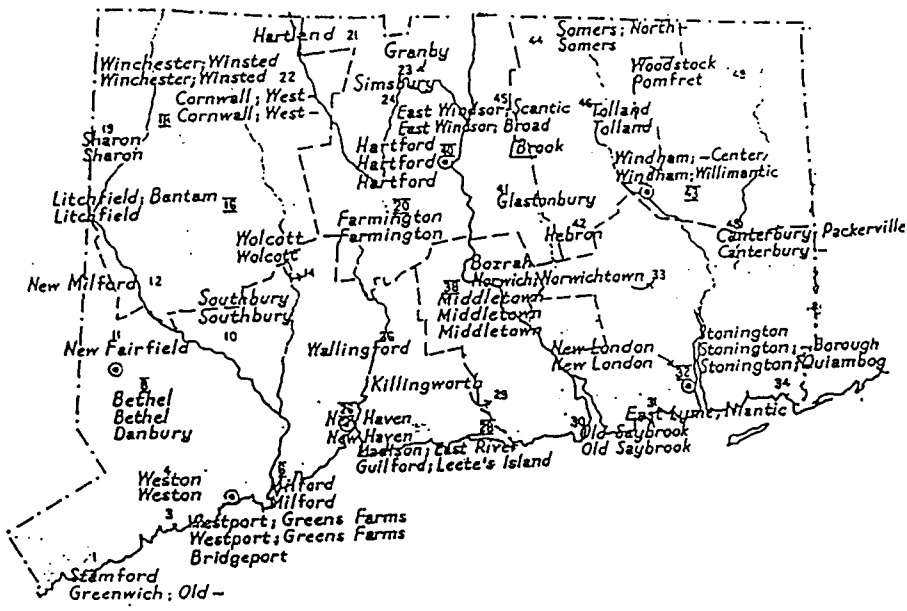


Figure 1

So far as we know, no one before us has questioned Kurath's (1939) claim in the *Handbook of the Linguistic Geography of New England* that 'New England has two major dialect areas, an Eastern and a Western', and that 'the "seam" between these two settlement areas runs straight north from the mouth of the Connecticut River (between [subjects] 30 and 31) through Connecticut . . .' (p. 8). Given our findings, we think that the time has come to take another look at that analysis and the assumptions behind it.

This article will present the results of a multidimensional scaling (MDS)[1] procedure to analyze the lexical and phonological responses in the LANE records of the sixty-seven Connecticut subjects. We believe that the results from this analysis will contribute further to the continuing discussions regarding what exactly we mean when we use the term *dialect boundary*.

Method

First we had to choose a method of analysis that would provide a descriptive statistical tool that helps to discover underlying structures, relationships, or affinities in data which contain considerable variation and error. MDS is such a tool. We view the major advantage of MDS over other statistical techniques used on dialect data to be that one can use MDS to analyze simultaneously the similarities and differences between the individual subjects' lexical and phonological inventories on the one hand, and similarities and differences between the behavior of the lexical and phonological items themselves on the other.

MDS analyses give two sets of scores: the object scores and the category quantifications. The object scores are assigned to the subjects so that those who show similar patterns of presence and absence of the selected lexical and phonological items will receive similar scores. For example, two subjects who show exactly the same pattern of presence or absence of certain items will receive exactly the same object score. A subject whose responses differ only by the presence of one lexical or phonological item will receive a score closer to the first two subjects than one who differs by the presence of two different lexical or phonological items, etc.

The category quantifications provide information on the actual dialect terms in the subjects' speech. These scores characterize the presence or absence of lexical items. For example, if two lexical items, like *bucket* and *gunny sack* or *gutters* and *tassel* are present in the inventory of the same subjects, the category 'presence' for the two items will receive similar quantifications, i.e. similar values. Both the object scores and the category scores can be plotted, and can be further analyzed using correlation coefficients. MDS also calculates a Variable Dimension Score for each variable. This score indicates the relative importance of every variable in the overall solution, and thus provides a way to quantify the results of the item-based results we get from traditional dialectology.

All this means that multidimensional scaling, like other forms of multivariate statistics, allows dialectologists to group similar subjects together rather than responses. This capability is not insignificant since, at least in the real world, we tend to think of dialects as groups of people speaking similarly rather than groups of unrelated responses. For example, the methods of traditional linguistic geography can tell us whether a Survey of English Dialects subject said [θundə] or [θʌndə], but, if the former pronunciation was recorded, it cannot tell us whether he or she also said [butə], or whether he or she uses other northern forms such as *beasts* vs. *cattle*. While we are not the first dialectologists to use multivariate statistics to group subjects

(Linn 1981, Linn and Regal, 1985,1988, Wachal 1986, Cichocki, Péronnet and Babich 1988), we would argue that, given modern statistical models, dialectologists should undertake more efforts in that direction.

Of course, this is not the method employed in traditional American dialectology. That method involved making list manuscripts of each item on the questionnaire (i.e., the different terms for the dragon fly, the pronunciation of *four*, and so on). The dialectologist would then map these terms, and, where patterns seemed to exist, he or she would draw an isogloss between them, much as Orton and his colleagues did in *A Word Geography of England* (1974) and *The Linguistic Atlas of England* (1978). At that point, the American and English methods diverged: the English, of course, mapped the isoglosses only, while the Americans went further and also tried to find major 'bundles' of isoglosses in order to establish major and minor dialect boundaries. Figure 2 shows one such bundle of isoglosses for Connecticut (Kurath et al. 1939: 30):

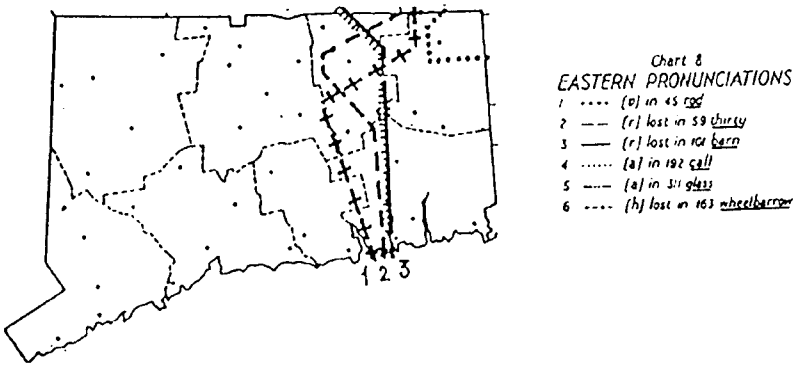


Figure 2

For our study, we first had to select and code the data. Table 1 lists the ten phonological items used by Kurath to show the presence or absence of constriction of postvocalic /-r/:

barn
father
afternoon
girl
water
beard
morning
corncrib
this year
thirty

Table 1

Table 2 lists the remaining selected four items with their respective contrasts which, according to Kurath, also evidence distinct isoglosses:

rod [ɑ] versus [ʊ]

calf [æ:] versus [ə]

glass [æ:] versus [ə]

tassel [æ:] versus [ə]

Table 2

Table 3 lists the lexical items which Kurath used:

1. sour milk/lapp(bb)ered milk/curdled milk/curdled klabber/bonny kapper (milk)/ klabber/klabber(ed) milk/thick milk
2. wheat bread/bread/loaf bread/raised bread/light bread/raised wheat bread
3. quilt/bed quilt/comfortable/tak comfortable/comforter, comfort quilt/ comfort/batwork comfort/batswork quilt/feather/down comfortable/patchwork quilt/puff
4. stone boat/stone board/stone drag/stone drug, drog/drag/drag board/drog, drug/sled
5. tassel, tossel/corn tassel, tossel/topgallant/top out/top/tip/pole /corn top/spindle/blow/tassel/tossel out/hound's ears
6. stovepipe(bibe)/smokepipe/stove funnel/funnel/pipe/
7. griddle cake/pancake/flap over/flapjack/slapjack/flannel cake/griddles
8. gutter(s)/eave(s) trough(s)/gutter pipe/trough(s)/conductor(s)/eave(s) pipe(s)/spout

Table 3

The presence or absence of each variant given in response to a lexical or phonological question was coded as a separate variable. Since MDS works best when all the categories contain at least ten percent of the possible answers, we excluded items such as *bread, loaf bread, raised bread, light bread,* and *raised wheat bread,* but were able to include the more ubiquitous *wheat bread.* More specifically, since the maximum possible frequency for any one form was sixty-seven – the number of LANE subjects in Connecticut – we included in the analysis only the items which had at least six occurrences. As a result, we were able to use twenty-three lexical items and fourteen phonological ones. In fact, the frequency range for the lexical responses was from six to sixty-one, meaning that as few as six subjects used the least frequent form and as many as sixty-one used the most frequent one.

Results

Our first task was to examine the object scores to ascertain whether Kurath was justified in proclaiming such a definitive east-west boundary along the Connecticut River. Looking at the scores for the constriction of postvocalic /r/, we did find it to be every bit the east-west New England marker that Kurath claimed it to be. Indeed, with one lone exception east of the river, Connecticut subjects were r-less in at least some of their responses, and, generally, the further east that they reside, the more r-less they became. The situation west of the river is more complex than indicated by Kurath's claim. First of all, Hartford, the state capital, is r-less, and lies just west of the river. Middletown and Old Saybrook, also just west of the river, are mixed, but the former is only weakly so. New Haven, the site of Yale University, is also mixed, and lies some thirty-five miles west of the river. Wallingford, some twenty miles west of the river, and Milford, over forty-five miles west, both evidence clear r-less responses. In addition, when one looks at the Variable Dimension Scores, there is no doubt that this phonological feature has discriminating power. The scores range from a high of .875 to .545, with *beard* the highest, and *corncrib* the lowest. One can say as Kurath did, however, that in general speakers west of the river are r-full.

In order to run a regression analysis on the data, we arbitrarily chose eleven sites in central Connecticut[2] and used their object scores for the analysis.[3] Where there were two object scores, we took the average, but these scores were close enough so as not to skew the results. The Spearman Rank Correlation analysis on the relationship between the subjects' geographical location and their pronunciation of constricted postvocalic /r/ revealed a very high correlation between the two (0.932) where the

probability of getting such results by chance are less than one in 100 ($p < .01$). Figure 3 presents a scattergram of these findings:

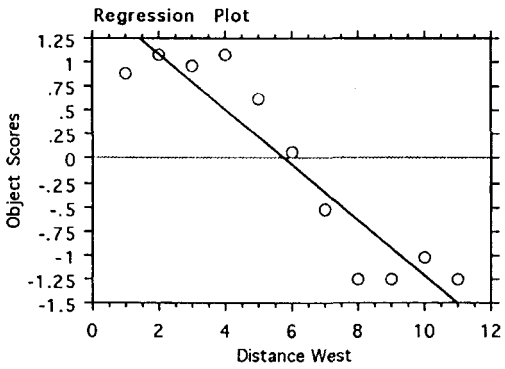


Figure 3

This inverse correlation between the absence of postvocalic /r/ constriction and distance west, coupled with the high Variable Dimension Scores (.875 to .545) leaves little doubt as to the importance of postvocalic /r/ in the LANE. It does suggest, however, that there is no clear, sharp boundary between eastern and western Connecticut; rather, there is basically a gradual lessening of r-lessness as one moves westward.

The other phonological items that Kurath used to establish his east-west boundary are not very discriminating at all, in spite of maps such as Figure 2 noted earlier. The [æ:/a] contrast in *glass*, *calf*, and *tassel* does not show any central distributional tendency. Both [æ:] and [a] occur quite frequently on both sides of the Connecticut River. Ninety-one percent of all the subjects have [glæs], seventy-seven percent have [kæf], and fifty-two percent have [tæsəl], and in no case is there an east-west distribution of these terms. The latter item, *t[æ]ssel*, is probably nondiscriminating since that distinction was carried well west of the Mississippi, let alone west of the Alleghenies. The [rɑd / rɒd] distinction is almost as mixed and nondiscriminating as *tassel*. Sixty-seven percent have [rɒd] and twenty-three percent have [rɑd]. Later dialectologists have in general considered this more of a northern/southern distinction than east/west. These high percentages of incidence for *glass*, *calf*, and *tassel* among Connecticut speakers indicate that [æ] is just likely to occur east of the river as is [a] and cannot be used in any definitive way in establishing an east-west dialect boundary, despite the fact that eleven occurrences of [a] are in fact located east of the river. Four

occur west of the river as well.

The presence or absence of postvocalic /-r/ appears to be the only east-west marker, and the Connecticut River seems to be the boundary, only because it is in the middle of the state. Our analysis indicates a clear linear relationship for postvocalic /r/ rather than a clear dialect boundary. Too many r-less speakers exist in the LANE records west of the Connecticut River. R-lessness, at least in 1931-33, was probably enough to geographically mark Connecticut speakers, even though what seemed to be other related markers were probably more perceptual than actual. They certainly were not as discrete as the Kurath (1939) statements make them out to be.

The lexical object scores, which measure how similar were the subjects' responses, reflect the east-west boundary to some extent, but overall the scores are much lower. For example seventeen of the lexical scores are below .50, while only four subjects evidenced r-object scores that low. This suggests, of course, that the lexical data lack much similarity or agreement.

The category quantifications of the lexical items themselves are even less revealing. The values are very low, with only two showing any discriminating power at all. These two lexical items are *lapp(bb)ered* milk with a value of .474 and *stone boat* with a value of .331. The first can be compared with two other members of this lexical set: *curdled milk* with a very low value of .002 and *bonny klapper* with a value of .235. The value of *stone boat*, .331, can be compared to that of *stone drag* and *drag* with their values of .219 and .048. The only other somewhat high value was the ubiquitous *pancake* with .224, hardly a discriminating term. And it is only relatively high when matched against *griddle cake* (.020), *flapjack* (.032), and *slapjack* (.004).

The category quantifications are all relative. The **highest** category quantification score for the lexical items was .474, while the **lowest** score for postvocalic /-r/ constriction was .575. Since the highest score for postvocalic /-r/ constriction is .875, it seems clear enough that /-r/ is a much better indication of regional difference than any of the lexical items. Furthermore, a look at the actual LANE maps confirms these results. As Harold Orton liked to say, there it is – right on the ground. No lexical item we examined came even close to being representative of regional distribution as was *corn crib*, which had the lowest category quantification score of all the examples of postvocalic /-r/ in our study.

Conclusion

Our conclusion for the LANE Connecticut results is that there is no set of

lexical and/or phonological features which, seen together, can justify Kurath's positing the Connecticut River as the east-west boundary. What does appear, however, is the undeniably strong east-west correlation for postvocalic /r/ constriction. The other phonological items that Kurath used to establish the east-west boundary proved to be less than helpful, showing little or no distributional patterning. The distribution of the lexical object scores also fails to support Kurath's east-west division. Furthermore, the lexical results were not isomorphic with the scores for postvocalic /r/; that is, a subject with a high r-less object score may only have a lexical object score half that of the r-less object score. In addition, the weak category quantification scores for the lexical items do not allow us to place great stock in their discriminating power. We suspect that where Kurath found dialect boundaries, there are only occasional isoglosses. In no instance did we find the necessary bundling of isoglosses that would indicate a major or even a minor dialect boundary. Kurath's Connecticut isoglosses in Figure 2 here and in the *Handbook of the Linguistic Geography of New England* do not provide a categorical division between lexical items or pronunciations any more than his *wheelbarrow* isogloss did so further north (see Davis and Houck 1995: 380-81). Many r-less speakers, for example, are found west of the Connecticut River as well as east of it.

These findings should be considered in the context of earlier studies of ours. Our studies of the upper midwest (Davis and Houck 1995) and Iowa (Horvath and Houck 1996) as well as the paper we presented at Methods IX on the dialect situation here in England (Davis, Houck and Upton 1997) give us real reason to question the traditional American notion of dialect boundaries. In all three of these works, we tried to show that certain so-called major dialect boundaries are a function of which forms the dialectologist chooses to combine and to map. In the Methods IX paper we noted earlier (Davis, Houck and Upton 1997), we quoted from Gaston Paris' (1888) 'Les parlers de France,' who in turn quoted from Paul Meyer. The following is probably truer today for the United States and England than it is for much of western Europe, including the area that Paris was discussing:

> . . . dans une masse linguistique de même origine que la nôtre, il n'y a réellement pas de dialectes; il n'y a que des traits linguistiques qui entrent respectivement dans des combinaisons diverses. . . (163)

Given the findings of our previous work and this one as well, we must conclude that dialect boundaries in both England and the United States have both a psychological and a perceptual reality that typically are very difficult to verify

objectively. It is difficult to say at what point someone becomes dialectally different from another – what composite of phonological and lexical features permits us to recognize another as being dialectally different from ourselves. We all make these distinctions, but how we do so is not obvious from our data or from any other English or American data we have seen.[4]

The Connecticut isogloss for postvocalic /-r/ is not at all easy to draw, and it may not be possible to draw it at all. It is conceivable that one might have more success if one were to use a statistical model such as that suggested by Kretzschmar (1992), a method which uses statistical techniques very different from multi-dimensional scaling to establish isoglosses. It would be most interesting to use Kretzschmar's (1992) method on the Connecticut data to see if his results would replicate ours. We suspect that they would, given the actual numbers of occurrences for both postvocalic /-r/ constriction and the lexical items examined here. To use Harold Orton's expression negatively, the Connecticut boundary is not at all apparent on the ground.

In *American Tongues*, a film produced some years ago and still used at a number of American universities, Roger Shuy tells the story that, as one passes from western to eastern New England, one goes from /park jə kar/ to /pak jə ka/. Our evidence suggests that things are far from that simple. To complicate matters further, we found that r-lessness occurred more frequently, but not statistically so, before nasals and in open position, so *your* and *car* would evidence more r-lessness than with *park*.

Michael Linn has observed (personal communication) that sometimes just one perceptual isogloss, if it is important enough, might be considered a dialect boundary. Charles Houck tells the story of riding southward in the Midlands with Stanley Ellis when Ellis asked, 'Did you feel that bump?' After Houck replied that he had felt nothing, Ellis responded with, 'Well, we just crossed the [bugə/[bʌgə] line!' Unfortunately, this particular distinction was not published in Orton's *Linguistic Atlas of England*, but we can be fairly certain that it would not match completely with the lines marking variant pronunciations of *thunder* and *butter*.

Bloomfield (1933: 328-29) recognized this problem when he noted that the [hus] / [hys] boundary in Dutch was different from the [mus] / [mys] boundary. He, like Gilliéron before him, concluded that 'every word has its own history'. The multidimensional scaling analysis of the Connecticut data certainly lends additional support to this hypothesis. It is altogether fitting to end this paper by pointing out that we Americans probably should have long ago taken our clue from English dialectologists, and should have just observed, recorded, and mapped the regional differences. Our research indicates that, whatever discoveries future dialectologists

might make about English in England and America, they will not involve finding bundles of isoglosses justifying major and minor dialect boundaries. Harold Orton and Eugen Dieth, on this question, surely knew what they were about.

NOTES

We would like to thank William A. Kretzschmar, Jr. and Michael D. Linn, who read earlier versions of this article and made some very candid and very helpful suggestions for its improvement.

[1] The MDS algorithm used in this paper is the Homogeneity Analysis via Alternating Least Squares or HOMALS, an SPSS program developed by the Department of Data Theory, University of Leiden.

[2] The towns sampled were, from east to west, Canterbury, Windham, Norwichtown, Hebron, Glastonbury, Middletown, Farmington, Wolcott, Southbury, New Milford, and New Fairfield. We ran similar correlations through both the northern and southern parts of the state. Since our findings for all three of these correlations were essentially the same, we decided to present the results for the central part of the state here.

[3] We decided to use Spearman's Rank Correlation Coefficient, a non-parametric correlation coefficient, because the data are ordinal in nature.

[4] While isoglosses may not bundle into clearly-defined dialect boundaries in England and the United States, Guillaume Schiltz has noted (personal communication) that a very different situation obtains in much of western Europe. He notes, for example, that the *Schwarzwaldschranke* 'runs over the northern and middle Black Forest and divides the Low-Alemanic dialects in the West from the Swabian in the East.' In addition, Klausmann 1997 demonstrates that the Allemanic-Franconian dialect boundary, a function of the Medieval boundary between Franconia and Swabia, is still alive and well, evidenced by a number of isogloss bundles.

REFERENCES

Bloomfield, Leonard. 1933. *Language*. New York: Holt, Rinehart, and Winston.

Cichocki, Wladyslaw. 1988. The use of dual scaling in social dialectology: multidimensional analysis of vowel variation. In Alan Thomas (ed) *Methods in Dialectology*, pp. 187-99. Clevedon and Philadelphia: Multilingual Matters.

Cichocki, Wladyslaw, Louise Péronnet and Rose Mary Babich. 1988. Atlas linguistique du vocabulaire maritime acadien: a final progress report. In Alan Thomas (ed) *Methods in Dialectology*, pp 109-19. Clevedon and Philadelphia: Multilingual Matters.

Davis, Lawrence M., and Charles L. Houck. 1995. What determines a dialect area?: Evidence from the Linguistic Atlas of the Upper Midwest. *American Speech* 70: 371-86.

Davis, Lawrence M., Charles L. Houck, and Clive Upton. 1997. The question of dialect boundaries: the SED and the American atlases. In Alan Thomas (ed) *Issues and Methods in Dialectology*, pp. 271-83. Bangor: University of Wales.

Horváth, Veronika K. and Charles L. Houck. 1996. The Northern-Midland Boundary in Iowa Again?: A Global Approach to Dialectology. Midwest American Dialect Society. Chicago, November.

Houck, Charles L. 1967. A computerized statistical methodology for linguistic geography: a pilot study, *Folia Linguistica* 1: 80-95.

Houck, Charles L. 1969. A Statistical and Computerized Methodology for Analyzing Dialect Materials. Diss. University of Iowa.

Houck, Charles L. 1985. The Iowa Northern-Midland Boundary Revisited: A Computerized and Statistical Analysis. Midwest American Dialect Society. St. Louis, November.

Houck, Charles L. 1986. Multidimensional Scaling as a Statistical Analytical Procedure in a Traditional Dialect Study. Midwest American Dialect Society. Chicago, November.

Klausmann, Hubert. 1997. Der Ellwanger Sprachraum – ein Ostschwäbisches Randgebiet. In A. Ruoff and Peter Löffelad (eds) *Syntax und Stilistik der Alltagssprache: Beiträge der 12. Arbeitstagung zur alemannischen Dialektologie*, pp. 65-83. Tübingen: Max Niemeyer Verlag.

Kretzschmar, William A., Jr. 1992. Isoglosses and predictive modeling, *American Speech* 67: 227-49.

Kurath, Hans. 1939. *Handbook of the Linguistic Atlas of New England.* Providence, RI: Brown University for the American Council of Learned Societies.

Kurath, Hans. 1939-43. *Linguistic Atlas of New England.* 3 vols. in 6 pts. Providence, RI: Brown University for the American Council of Learned Societies.

Kurath, Hans. 1949. *A Word Geography of the Eastern United States.* Ann Arbor: University of Michigan Press.

Linn, Michael D. 1981. A statistical model for classifying dialect speakers. In H. Warkentyne (ed) *Papers from the Fourth International Conference on Methods in Dialectology*, pp. 244-53. Victoria, B.C.: University of Victoria Press.

Linn, Michael D. and Ronald Regal. 1985. Numerical taxonomy as a tool in dialect research. In H. Warkentyne (ed) *Papers from the Fourth International Conference on Methods in Dialectology*, pp. 245-62. Victoria, B.C.: University of Victoria Press.

Linn, Michael D. 1988. Verb analysis of the Linguistic Atlas of the North Central States: A case study in preliminary analysis of large data sets. In Alan Thomas (ed) *Methods in Dialectology*, pp. 138-55. Clevedon and Philadelphia: Multilingual Matters.

Paris, Gaston. 1888. Les parlers de France, *Revue des patois gallo-romans*. 2: 161-75.

Wachal, Robert S. 1986. An Investigation into the Use of Inverse Factor Analysis for Forming Clusters of Informants Based on Linguistic Variables. Midwest American Dialect Society. Chicago, November.

Addresses

Lawrence M. Davis
Department of English
Wichita State University
Wichita
KS 67207
USA
<davis@wsuhub.uc.twsu,edu>

Veronika K. Horváth
Department of Computer Science
Southern Illinois University
Carbondale
IL 62901
USA

Charles L. Houck
Department of English
Ball State University
Muncie
IN 47306
USA

Attitudes Towards British English Dialects in the 19th Century

Manfred Görlach

Abstract[1]

The 19th century saw the emergence of ModE, the standard language now becoming accessible to wide ranges of the population, as a consequence of literacy spreading through various types of schools before 1870/72 and compulsory education thereafter, urbanization and a degree of geographical and social mobility unheard of in the much stabler society of the previous period, and the availability of cheap reading material: newspapers, novels and teach-yourself manuals. Concurrently, the functional range of broad dialect receded; village dialects gave way to regional dialects and later to modified standard spoken with whatever heavy accent, and attitudes towards dialect changed from widespread negligence, to amusement and comic attraction to fiercely negative and, finally, nostalgic evaluation.

Although the topic has been treated in a few publications (Bailey, 1996; Görlach, 1995; Honey, 1988; Ihalainen, 1994; Phillipps, 1984 and Wakelin, 1977) the use of the available sources has been too selective, and the analysis frequently too one-sided, to permit a comprehensive and balanced overview.

1. *Introduction*

The English language in 19th-century England was largely determined by the heritage of the 18th. This is true for concepts of good styles and correctness in grammar and pronunciation (where the pattern set by Lowth, Murray and Walker provided the norms well into the 19th century); the statement also applies to attitudes towards social and geographical varieties of English. An analysis of late 18th-century views on non-standard English is therefore important; it has to take

into account the following authors and text types:

1) The two dictionaries by Francis Grose, *A Classical Dictionary of the Vulgar Tongue* (1785) and *A Provincial Glossary* (1787), attempted to draw together the available evidence on slang and dialect lexis. Neither was based on original research, but the collection of non-standard lexis from various sources was apparently seen as a necessary and worthwhile complement to the efforts of Johnson who had intentionally excluded both types of vocabulary, in his attempt to codify the respectable standard lexis attested by the best authors.
2) The dialect collections of the lexis of five regions by W. Marshall (1787-96), who included a wealth of regional folkloristic material besides dialect proper, but gave due recognition to the apparent interest in the provinces.
3) Dialect poetry, among which *An Exmoor Scolding* and *Exmoor Courtship* (Anon., 1727?, 1746, ^{7}1771/^{9}1782) and John Collier's (Tim Bobbin's) *View of the Lancashire Dialect* were possibly best known. Published simultaneously[2] in 1746 in opposite parts of England they were felt to be similar enough to be sometimes bound together (as in the Bodley copy I used), and complemented by discussions of the value of dialect in later editions. I here quote a relevant passage from the ninth edition of the *Scolding* 'on the propriety and decency of the Exmoor language':[3]

> **Q1** It may be proper to advertise such of our Readers as may be Strangers to the *Devonshire* Dialects, that the following is a genuine Specimen thereof, as spoken in those Parts of the County where the Scene is laid; (the Phraseology being also agreeable thereto, and the Similes, &c. properly adapted to the Characters of the Speakers;) and not an arbitrary Collection of ill-connected clownish Words, like those introduced into the Journals of some late Sentimental Travellers as well as in the Productions of some Dramatic Writers, whose Clowns no more speak in their own proper Dialects, than a dull School-boy makes elegant and classical *Latin*; their suppos'd *Language* being such as would be no less unintelligible to the Rusticks themselves, than to those polite Pretenders to Criticism who thereby mean to make them ridiculous. (...)
>
> (v) And even near *Exmoor*, none but the very lowest Class of

> People generally speak the Language here exemplified; but were it more commonly spoken by their Betters, perhaps it might not be so much to their Discredit as some may imagine; most of the antiquated Words being so expressive as not to be despised, tho' now grown obsolete, and no longer used by the politer Devonians, who in general speak as good modern English as those of any other County. (...) Hence every County has its peculiar Dialect, at least in respect to the vulgar *Language* of their Rusticks, insomuch that those of different Counties can't easily understand each other. Among Persons engag'd in Commerce indeed, or who have had a liberal Education, we may better distinguish their several Counties by their Accent, than by any Impropriety in their *Language*: But we are here speaking only of the lower Class of People in each County; and that these have in several Parts of England a more uncouth and barbarous Jargon than the worst among the Devonians, might be easily shewn (...).
>
> (Preface, Exeter, *January*, 1782)[4]

A negative evaluation of dialect was common at the beginning of the 19th century. Crombie in 1802 flatly denied the practicability of remedial education for the vulgar in the countryside – and what is even more significant is that these remarks were reprinted without changes in later editions of his book (the latest I have checked is the ninth of 1865!):

> **Q2** (...) to define the proper province of the grammarian, I proceed to observe, that this usage, which gives law to language, in order to establish its authority, or to entitle its suffrage to our assent, must be, in the first place, *reputable*.
>
> The vulgar in this, as in every other country, are, from their want of education, necessarily illiterate. Their native language is known to them no farther, than is requisite for the most common purposes of life. Their ideas are few, and consequently their stock of words, poor and scanty. Nay, their poverty, in this respect, is not their only evil. Their narrow competence they abuse, and pervert. Some words they misapply, others they corrupt; while many are employed by them, which

> have no sanction, but provincial, or local authority. Hence the language of the vulgar, in one province, is sometimes hardly intelligible in another. Add to this, that debarred by their occupations from study, or generally averse to literary pursuits, they are necessarily strangers to the scientific improvements of a cultivated mind; and are therefore entirely unacquainted with that diction, which concerns the higher attainments of life. Ignorant of any general principles respecting language, to which they may appeal; unable to discriminate between right and wrong; every one therefore prone to adopt whatever usage casual circumstances may present; it is no wonder, if the language of the vulgar be a mixture of incongruity and error, neither perfectly consistent with itself, nor universally intelligible even to them. Their usage, therefore, is not the standard, to which we must appeal for decisive authority; a usage so discordant and various, that we may justly apply to it the words of a celebrated critic, *Bellua multorum est capitum; nam quid sequar aut quem*?

John Poole, who employed the Bell and Lancaster system in Enmore School in Somerset, advised very drastic measures against the students' use of dialect features:[5]

> **Q3** Even a coarse or provincial way of pronouncing a word, though sanctioned by the general practice of the district, is immediately noticed by the teacher; and exposes the child, who uses it, as much to the correction of those below him, and consequently to the loss of his place, as any other impropriety in reading would do.
> (J. Poole, *The Village School Improved*, 1813: 40-41, quoted from Mugglestone, 1995: 293)

Confidence that teachers would prove competent in dealing with the problem increased during the 19th century, but in 1860 Foster & Foster were still quite diffident, pointing to the long tradition of village-school teaching which had achieved few of the intended results (1860/1995, I: 334).

In spite of these critical voices, there was a great deal of interest in, and sympathy for, regional dialects at the beginning of the 19th century. We might

wish to argue that a more positive attitude towards dialect is normal to develop when it is no longer felt as a threat to the standard language – and this may happen long before the loss of dialect is seen as imminent. However, such general assumptions are impossible to verify, at least for England around 1800. It is difficult to say whether concepts like Wordsworth's, who stressed the importance of plain language as a poetical medium in the Preface to the Lyrical Ballads, has anything to do with our topic – after all, he did not use dialect in his poetry, which would have appeared too uncultivated. There is, in England, apparently no parallel to the re-evaluation of dialect that took place in Germany at the same time, and where Herder's interest in folk traditions provided the intellectual background which made J.P. Hebel's poetry (and that of his followers) possible. If we look at dialect in English literature, it is clear that the stimulus came from outside, mainly from Maria Edgworth in Ireland and particularly Sir Walter Scott (cf. Blake, 1981). I will return to dialect literature later on.

However, the main interest in dialect was because it was seen to preserve valuable linguistic documentation of 'pure' stages of English. Thus, a new self-assertive attitude towards dialects made itself felt from the early 19th century onwards – even though it was, of course, mostly voiced by non-dialect speakers, namely gentlemen and schoolmasters. The statement by 'a native of Craven' who stresses the ancient status of the dialect and the threat of corruption by the outside world is a good example of this type of evaluation:

> **Q4** I have attempted to make the second edition of the *Craven Glossary* more worthy of the readers attention, by a large addition of words, and by numerous authorities, collected from ancient writers. Though this has been the most laborious part of my work, it has, at the same time, been the source of the greatest pleasure; for whenever I found a Craven word thus sanctioned by antiquity, I was more and more convinced, that my native language is not the contemptible slang and *patois*, which the refined inhabitants of the Southern part of the kingdom are apt to account it; but that it is the language of crowned heads, of the court, and of the most eminent English historians, divines, and poets, of former ages. (...)
>
> Pent up in their native mountains, and principally engaged in agricultural pursuits, the inhabitants of this district had no opportunity of corrupting the purity of their language by the adoption of foreign idioms. But it has become a subject of much

> regret that, since the introduction of commerce, and, in consquence of that, a greater intercourse, the simplicity of the language has, of late years, been much corrupted. Anxious, therefore, to hand it down to posterity unadulterated, the author has attempted to express, in a familiar dialogue, the chaste and nervous language of its unlettered natives. (1828)

The similarity to the statement made by Forby on the dialects of East Anglia will be obvious:

> **Q5** From a writer who offers to the public a volume on a *Provincial Dialect*, and ventures to announce his intention of confirming, by *authority* and *etymology*, the strange words and phrases he is about to produce, some introductory explanation of his design may reasonably be required. The very mention of such an undertaking is likely to be received with ridicule, contempt, or even disgust; as if little or nothing more could be expected, than from analysing the rude jargon of some semi-barbarous tribe; as if, being merely oral, and existing only among the unlettered rustic vulgar of a particular district, *Provincial Language* were of little concern to general readers, of still less to persons of refined education, and much below the notice of philologists.
>
> However justly this censure may be pronounced on a fabricated farrago of cant, slang, or what has more recently been denominated *flash language*, spoken by vagabonds, mendicants, and outcasts; by sharpers, swindlers, and felons; for the better concealment of their illegal practices, and for their more effectual separation from the 'good men and true' of regular and decent society; it certainly is by no means applicable to any form whatsoever of a *National Language*, constituting the vernacular tongue of any province of that nation. Such forms, be they as many and as various as they may, are all, in substance, remnants and derivatives of the language of past ages, which were, at some time or other, in common use, though in long process of time they have become only locally used and understood.
>
> (Forby, 1830: 1-2)

Note the value Forby ascribes to dialects for their preservation of old words – and his scathing criticism of cant and slang. Forty years later the attitude remained much the same, though the desire to preserve the dialect had become more urgent; however, the greatest threat to local dialects was, of course, still to come with the introduction of compulsory education in 1870. Huntley's very detailed description of the sociolinguistic change in the Cotswolds deserves to be quoted at length:

> **Q6** Another reason, which at this present time renders dialects more worthy of remembrance, is the universal presence of the village schoolmaster. This personage usually considers that he places himself on the right point of elevation above his pupils, in proportion as he distinguishes his speech by classical or semi-classical expressions; while the pastor of the parish, trained in the schools still more deeply, is very commonly unable to speak in a language fully 'understanded of the people,' and is a stranger to the vernacular tongue of those over whom he is set; so that he is daily giving an example which may bring in a latinized slip-slop. In addition to this, our commercial pursuits are continually introducing American solecisms and vulgarisms. Each of these sources of change threaten deterioration. Many homely but powerful and manly words in our mother tongue appear to totter on the verge of oblivion. As long, however, as we can keep sacred our inestimable translation of the Word of God, to which let us add also our Prayer-book, together with that most wonderful production of the mind of man, the works of Shakespeare, we may hope that we possess sheet-anchors, which will keep us from drifting very far into insignificance or vulgarity, and may trust that the strength of the British tongue may not be lost among the nations.
>
> It has, moreover, been well observed that a knowledge of dialects is very necessary to the formation of an exact dictionary of our language. Many words are in common use only among our labouring classes, and accounted therefore vulgar, which are in fact nothing less than ancient terms, usually possessing much roundness, pathos, or power; and, what is more, found in frequent use with our best writers of the Elizabethan period. The works of Shakepeare abound in examples of the Cotswold

> dialect, which indeed is to be expected, as his connexions and early life are to be found in the districts where it is entirely spoken; and if, as has been thought, he spent some part of his younger days in concealment in the neighbourhood of Dursley, he could not have been better placed to mature, in all its richness, any early knowledge which he might have gained of our words and expressions. This, however, is certain that the terms and phrases in common use in the Cotswold dialect are very constantly found in his dialogue; they add much strength and feeling to it; and his obscurities, in many cases, have been only satisfactorily elucidated by the commentators who have been best acquainted with the dialect in question.
>
> The Cotswold dialect is remarkable for a change of letters in many words; for the addition or omission of letters; for frequent and usually harsh contractions and unusual idioms, with a copious use of pure Saxon words now obsolete, or nearly so. If these words were merely vulgar introductions, like the pert and ever-changing slang of the London population, we should look upon them as undeserving of notice; but as they are still almost all to be drawn from undoubted and legitimate roots, as they are found in use in the works of ancient and eminent authors, and as they are in themselves so numerous as to render the dialect hard to be understood by those not acquainted with them, they become worthy of explanation; and then they bring proof of the strength and manliness of the ancient English tongue, and they will generally compel us to acknowledge, that while our modern speech may possibly have gained in elegance and exactness from the Latin or Greek, it has lost, on the other hand, impressiveness and power. (...)

The contrary opinion was held by the Hon. Samuel Best, who, obviously expressing a commonly accepted view, stated in the fourth edition of his *Elementary Grammar for the use of village schools* of 1857:

> **Q7** The classically-educated man cannot, if it were desirable, so ignore his education as to address a congregation in the jargon and patois of the village. [...] We may and ought to raise them to

> our standard; we cannot, without profaneness in sacred things, descend to theirs.
>
> (Best 1857, quoted from Michael, 1987: 351)

Demographic evidence explains why even in the major industrial centres (at least in the north) dialect remained stable because the vast majority of migrants came from the immediate neighbourhood. This meant that extreme forms of village dialects were given up in the new melting pots, but regional dialects in a somewhat levelled form were strengthened.[6]

There were obvious regional differences in the amount of deviation from the London-based standard, and in attitudes to rural speech. Halliwell (1847) found Derbyshire dialect 'broad', but Buckingham close to standard, and Northern, Southwestern, East Anglian and London varieties best known (cf. Ihalainen's summary, 1994: 212). The degree to which regional dialects were accepted by their speakers as badges of identity is partly reflected by the number of publications the English Dialect Society were able to use as sources. Ihalainen's count of the pages devoted to each county in the 1877 bibliography (1994: 273) gives the following rough proportions:

Lancashire	32%
Yorkshire	18%
Cumberland	14%
Cornwall	7%
Northumberland	5%
Devon	4%
Westmorland	3%
Others	16%

2. *The erosion of dialect*

Grievous concerns about the imminent loss of dialects began around 1870. There is W. A. Wright's classic appeal of 1870 for the founding of a dialect society:

> **Q8** It has long been my conviction that some more systematic effort ought to be made for the collection and preservation of our

> provincial words. In a few years it will be too late. Railroads and certificated teachers are doing their work.
>
> (*Notes & Queries* 1870, quoted from Petyt, 1980: 76)

When the English Dialect Society was started in 1873, two motives were prevalent for the move:

1) The data collected from rural forms of speech were intended to broaden the data basis for linguistic history and reconstruction, an objective which was motivated by the comparative method in linguistics, then also dominant in Britain, though with some delay.
2) Dialect lexis in particular was seen as rapidly disappearing in view of increasing mobility and the end of the relative isolation of many villages and entire regions. Hardy saw in retrospect what happened as a consequence of social stigmatization:

> **Q9** ... education ... has gone on with its silent and inevitable effacements, reducing the speech of this country to uniformity, and obliterating every year a fine old local word. The process is always the same: the word is ridiculed by the newly taught; it gets into disgrace; it is heard in holes and corners only; it dies, and worst of all, it leaves no synonym.
>
> (Hardy, 1908: iii, quoted from Jacobs, 1952: 10)

Elworthy's account in 1888 of how the Board Schools modified the speech of the rural working classes is very relevant in this context:

> **Q10** The children have all learnt to read, and have been taught the 'correct' form of all the verbs they use. The girl would come home, and her mother would say, 'Lize! you didn ought to a-wear'd your best shoes to school.' Eliza would say, 'Well, mother, I wore my tothers all last year, and they be a-wore out.' In this way parents become familiar with the strong forms of literary verbs, but they have no notion of dropping the past inflection to which they have always been accustomed, while at the same time they wish to profit by their children's 'schoolin'. Consequently the next time the occasion arrives, Eliza is told she should have *a-*

> *wor'd* her tother hat, &c., and thus *wor'd* and *a-wor'd* soon become household words with the parents; and the same or a like process is repeated by them with respect to other words all through their vocabulary. All children naturally copy their parents' accent, tone and sayings ... Consequently the school-teaching sets the model for written language, and the home influences that for everyday talk. The result is that at the present moment our people are learning two distinct tongues – distinct in pronunciation, in grammar and in syntax. A child, who in class or even at home can read correctly, giving accent, aspirates (painfully), intonation, and all the rest of it, according to rule, will at home, and amongst his fellows, go back to his vernacular, and never even deviate into the right path he has been taught at school.
>
> (Elworthy, 1888; quoted from Phillipps, 1984: 87)

The expansion of the standard language, then, levelled out the major regional differences. The process is symbolized (although this was not the author's intention!) in Hardy's account (1883, in Golby, 1996: 300-1) of the urbanization of Devon rural society. He noted the disappearance of the old rural costume (~ dialect) which was being replaced by second-hand and ill-advised metropolitan dress, often in bad taste (~ modified standard) – the mirror-image in the dress:language symbol in Shakespeare's Edgar is obvious. As a result, bidialectalism became more common than it had ever been in England (and lasted for two generations at least until broad dialect was lost). Hardy, again, remarked on the fact in 1891 for the same region that Elworthy had commented on three years before him; he described the sociolinguistic situation in a much-quoted passage from *Tess*:

> **Q11** Mrs Durbeyfield habitually spoke the dialect; her daughter, who had passed the Sixth Standard in the National School under a London-trained mistress, spoke two languages; the dialect at home, more or less; ordinary English abroad and to persons of quality.
>
> (Ch.3, also quoted in Phillipps, 1984: 88)

What remained of regional characteristics was, at least for educated speakers, an

unmistakable accent, tolerated until the mid-century even among leading politicians and other members of the upper crust (cf. Honey, 1988, 1989). For the West Country accent that was left in the speech of even one of Her Majesty's Inspectors compare F.H. Spencer's self-characterization:

> **Q12** In speech I acquired the accent and the intonation of the common people. That is sometimes a trifling inconvenience; and though the population was of so diverse an origin, so strong is the local speech of the countryside that we all spoke the mid-Wessex speech, the speech of Gloucester, Berks and Wilts, which thickens into Somerset as you go West ... Fifty years of intercourse with people of all kinds and much travel have seemed not entirely to dissipate all traces of that speech ... The grammatical peculiarities of the Wiltshire yokel ... were not, however, ours. We did not say 'Her be gwaain whoam' for 'she is going home'. Nor did we use *thic* or *thuc* or a dozen other locutions still in common use in North Wiltshire. But most people who know the vowel sounds of the English provinces, and can recognize the Wessex *r*, would place me today as a native of mid-Wessex.
> (On the 1880s, in Spencer, 1938, quoted from Phillipps, 1984: 88)

Dialect must have been a veritable problem of English lessons in the schools. One of the earliest reflexes of this is certainly Thomas Batchelor's *Grammar* of 1809 which was intended for 'provincial schools' and in which he provided an analysis of 'the minute varieties which constitute a depraved or provincial pronunciation' (cf. Honey, 1997: 80). However, such explicit concern is very rare among writers of school grammars and we must assume that not even all teachers, however much they admired the standard language and never dared to question its relevance and prestige, were capable of speaking it fluently without any interferences.[7] It therefore comes as a surprise that there are so few remarks relating to these problems in the grammar books of the 19th century: we are led to assume that since school English was taught in a diglossic situation in much of the country, teachers and grammarians may have thought that the less mention of parallels in the two systems was made, the less likely was the possibility of transfers.[8] Contrastive aspects of language teaching, then, made their way only

very occasionally into 19th-century grammars.[9] One of the few I found is Pearson (1865) whose *The Self-help Grammar of the English Language; Intended for Reading, Dictation, Parsing, Composition and Home-Work in the Second and Third Classes in an Elementary School* gave some attention to the specific learners' problems of Yorkshire children. Amongst other things, he warned against the use of *thou/thee*, since it was obsolete in St E (and had been virtually given up even by the Quakers). His warning reminds us of a similar statement made by Marshall (1788, quoted by Ihalainen, 1994: 229) on *thou* as a Yorkshire provincialism.

Complaints by Her Majesty's Inspectors on the poor quality of regional English, in particular pronunciation, became frequent only towards the end of the century (Honey, 1988: 219-21), but they do not add up to a picture of how dominant dialect still was among the pupils and how it affected their correct acquisition of the spoken and written standard. Compare the quotations above and an early statement by HMI F.C. Cook in 1845:

> **Q13** I could easily make out a long list of the gross mistakes, omissions and mispronunciations of the principal words, and perversions of the sense, which are almost universally made by the young children, and which are in many cases unobserved, or uncorrected, by the monitors.
>
> (quoted from Honey; 1988: 219)

By contrast, positive attitudes towards dialect were certainly rare exceptions among HMIs. Honey points out one of these:

> **Q 14** In Rochdale, an interesting preservationist inspector of schools (HMI Mr Wylie) caused controversy in the local press in 1890 by his attempts to foster the use of local dialect in school. The response of some parents is illuminating: 'Keep the old Lancashire dialect out of the schools, Mr Wylie, for I want my children to talk smart when they're grown up.'
>
> (1997: 100-01)

3. *Dialect literature*

Even for people who were in favour of dialect in written compositions, its use was restricted to certain text types. Thus, it occurs much less in prose than in verse, and prose writing is found mainly in humorous stories; wider ranges, as exemplified in the Yorkshire almanacs, were obviously found only in the north.

We must not make too much of the absence of dialect in biblical translation. Foster & Foster in 1860 saw the great advantage for parishioners in

> using the vernacular in the way of familiar exposition and exhortation, than to attempt a style of composition which they cannot manage without blundering.

However, even they firmly stated that

> There can be no question about the propriety of Ministers using provincial dialects in their stated ministrations.

This attitude does not come as a surprise - that the Bible had to be in the accepted standard language remained true even for 19th-century Scotland – and the situation is not much different today after Lorimer's *New Testament in Scots* has become available. Also, there is still no bible in Jamaican Creole or, further afield, for Occitan or Luxemburgish, and bibles in Low German are little used. Ironically, then, it is Lucien Bonaparte's philological interest that gave us at least some biblical passages in 19th-century English dialects. This statement is true even though Shorrocks (fc.) rightly points to 18th-century clergymen's attempts 'to learn local dialects in order to understand their parishioners better'. One such effort was made by the Rev. William Hutton whose *A Bran New Wark* was 'written in a mixture of literary English and Westmorland dialect'; the author 'explained in his introductory remarks that his parishioners were more likely to heed his homilies if these were written in the local dialect' (Shorrocks' summary).

The sociolinguistic change sketched above obviously affected the frequency and functions of dialect writing. In Hollingworth's (1977) somewhat controversial opinion, Lancashire dialect poetry developed in three stages:

1) A phase of predominantly oral poetry (of which a few famous specimens like 'Th' Owdham weyver' came to be written down after 1840).

2) The 'golden age' of Lancashire dialect poetry, 1856-1870, represented by Waugh, Laycock and Ramsbottom: 'the amount of poetry produced, and its quality [...] considering what came before and what came afterwards, are truly amazing' (1977: 2). Hollingworth attributes much of the flowering of this tradition to the belated influence of Burns and to 'the rapid and transient movement of dialect poetry at this time from an oral tradition in which it was already well established, though poorly recorded, into a written form where it became more permanent but quickly lost vitality' (1977: 3). 'Waugh in particular was often referred to as the Lancashire Burns, and he clearly relished the title. In his commonplace book he carefully preserved a letter from Spencer J. Hall of Burnley written in 1874. "You and your confreres [have] done for Lancashire what Burns and Hogg [have] done for the Lowlands of Scotland – you [have] immortalised a dialect and made it classical"' (1977: 3).[10]
3) A final phase, in which the tradition 'moved away from a living expression of the "songs of the people" to an antiquarian and rather nostalgic attempt to conserve a dying culture and language' (1977: 5), an interpretation used to explain why the three poets (who lived on until 1891, 1893 and 1901 respectively) had 'burnt themselves out' (1977: 5).

There are various difficulties with Hollingworth's persuasive hypothesis, especially if we wish to look at his evidence from the viewpoint of the historical sociolinguist:

1) Since dialect prose is much more difficult to read than (short) poems, how can we explain that it lived on in the Yorkshire almanacs well into the 1920s?
2) How can we believe that the impact of general education in 1870 was as immediate as Hollingworth suggests? Potential readers of these poems who had gone through the new school system would not have read them as children, and therefore not before 1890 – when the tradition had been dead for some time.

Although the situation in Scotland was slightly different, especially with the much firmer hold Scottish dialects had on the countryside, it is interesting to see the parallels not just in the impact that Burns and Scott had on English writers, but also in what Scottish Kailyard poetry and newspaper prose (Donaldson, 1986) had

in common with Lancashire dialect poetry and the Yorkshire almanacs, namely the fact that they relied on dialect being dominant in everyday life.

Even 19th-century English novelists, who only made restricted use of dialect in their dialogues, depended on this vitality of spoken dialect when they wished to be moderately 'realistic' (cf. Dickens[11] and Thackeray for London and East Anglia, Mrs. Gaskell for Lancashire,[12] George Eliot for North Warwickshire and Hardy for 'Wessex'/Dorset[13]). Emily Brontë's representation of Haworth dialect in the speech of Joseph is a much quoted example:

> **Q14** 'Nelly', he said, 'we's hae a Crahnr's 'quest enah, at ahr folks. One on 'em's a'most getten his finger cut off wi' hauding t'other froo' sticking hisseln loike a cawlf. That's maister, yah knaw, ut's soa up uh going tuh t'grand 'sizes. He's noan feared uh t'Bench uh judges, norther Paul, nur Peter, nur John, nor Mathew, nor noan on 'em, nut he! He fair likes he langs tuh set his brazened face agean 'em! And yon bonny lad Heathcliff, yah mind, he's a rare un! He can girn a laugh as weel's onybody at a raight divil's jest. Does he niver say nowt of his fine living amang us, when he goas tuh t'Grange (...)?
> (*Wuthering Heights*, ch.10, quoted from Petyt, 1970: 47-48; also in Blake, 1981: 149)

Note that, largely true to the Scott tradition (and likely to reflect the sociolinguistic realities in the Haworth parsonage), broad dialect is reserved to the semi-literate Joseph (cf. for a close dialectal analysis, Petyt, 1970). Moreover, the passage was dedialectized by Charlotte for the second edition in 1850[14] to make it intelligible to readers outside Yorkshire (and not so much because of more general concerns about propriety).

Most authors saw the limitations on the use of dialect quite clearly, as, for instance, Hardy did:

> **Q15** An author may be said to fairly convey the spirit of intelligent peasant talk if he retains the idiom, compass, and characteristic expressions, although he may not encumber the page with obsolete pronunciations of the purely English words, and with mispronunciations of those derived from Latin and Greek. [...] If a writer attempts to exhibit on paper the precise

> accents of a rustic speaker, he disturbs the proper balance of a true representation by unduly insisting upon the grotesque element.
>
> (Hardy 1878, quoted from Blake 1981:166)

This statement is apparently not quite true of the north – a consequence of the type of urban in-migration mentioned above. One of the most convincing uses of 19th-century literary dialect is found in Burnett's *That Lass o'Lowries* (1877), a story with a Lancashire industrial setting; the fact that this successful novel was part of a wider fashion for dialect prose is indicated by *Punch*'s immediate reaction (20 Oct.-17 Nov. 1877) which made fun of such use of dialect.

> **Q16a** 'Let's hear,' cried a third member of the company.
>
> 'Gi' us th' tale owt an' owt, owd lad. Tha'rt th' one to do it graidely.'
>
> Sammy applied a lucifer to the fragrant weed, and sucked at his pipe deliberately.
>
> 'It's noan so much of a tale,' he said, with an air of disparagement and indifference. 'Yo' chaps mak' so much out o' nowt. Th' parson's well enow i' his way, but,' in naïve self-satisfaction, 'I mun say he's a foo', and th' biggest foo' fur his size I ivver had th' pleasure o' seein'.'
>
> They knew the right chord was touched. A laugh went round, but there was no other interruption and Sammy proceeded.
>
> 'Whatten yo' lads think as th' first thing he says to me wur?' puffing vigorously. 'Why, he coos in an' sets hissen down, an' he swells hissen out loike a frog i' trouble, an' ses he, "My friend, I hope you cling to th' rock o' ages." An' ses I, "No I dunnot nowt o' th' soart, an' be dom'd to yo'. It wur na hos*pit*ible," with a momentary touch of deprecation, – 'An' I dunnot say as it wur hospitible, but I wur na i' th' mood to be hospitible just at th' toime. It tuk him back too, but he gettin round after a bit, an' he tacklet me again, an' we had it back'ard and for'ard betwixt us for a good half hour. He said it wur Providence, an' I said, happen it wur, an' happen it wurn't. (...)
>
> (Burnett 1877:91-93)

Q16b OUR NEW NOVEL. **THAT LASS 'O TOWERY'S.**

By the Authors of Several other Things, &c. &c.

(...) It was NEGUS BARCROW, The Young Engineer.

'Yo domm'd dummer-tailed bolthead,' roared DAN BEERIE, fiercely. 'Yo yung poopy-cur snig-snagged boler! oil jewdy thee putty tupped naws, an giv yo siccan shuv i' th' oi, as yo'll reccomember fur ivvur, domm'd av oi doon't, th' oi gows t' gallus fur't!'

And he sprang on him with a ferocity that would have sent into the second week of the middle of the year after next (counting from the date of his receiving the blow) any man less physically powerful than NEGUS BARCROW.

EMMY raised a loud cry.

'A fyt! a fyt! T' owd feether's fytin an' millin Yung Ing'neer! Coom an, av yo be coomin! tiz ar reel beet o' jam, tiz!'

And she clapped her hands together, crying, 'Gow't, Yung Ing'neer. Gow't t' owd feether!' (...)

* At present we will offer no opinion as to the *quality* of the dialect. We have sent a Special Commissioner to the North, who, being a gentleman of considerable imitative power, will give us, on his return, some idea of what the dialect may be. We don't recollect anything exactly like it, but perhaps Mr. TAYLOR, who plays the part of *'Owd Sammy'* in *Liz*, will step in, and give us some explanation. In the meantime we will be cautious. – ED.

* Our Special Commissioner with a dictionary has not yet returned from the North, nor has he sent us either a line or a telegram. He was sent there to inquire into the dialect and the character of the people as represented in this story. Perhaps Miss ROSE LECLERQ, who is now performing most admirably in *Liz*, would kindly look in one morning and give us her opinion on the subject. Need we say we should be only too delighted to profit by her experience. – ED.

Extract from Letter of the three Co-Authors of the New Provincial Novel Company Limited, to the Editor. – 'We say!

> Isn't it going on capitally? Here's your fine fresh dialect, eh? Post the tin, sagacious *Redacteur, et croyez en nous à jamais*, as we say in Old Gaul. Never was such local colouring, eh? Worth all the money! And then the Curate! *that* fetches the Sunday readers.- No harm where there?s a Curate. 'Yours ever, CO-AUTHORS.' (...)
>
> (*Punch*, 1877)

That Lass is, however, significant also for another aspect which may come as a surprise to modern readers – the willingness of earlier readers to enjoy dialects not their own. How else could we explain the success of the book in America where it was first published, or the impressive overseas sales of Bell's *Wee MacGreegor* (1901), in which the very frequent dialogues are all written in dense Glaswegian? (cf. Görlach, 1992).

Dialect poems are shorter and easier to read. Tennyson's Lincolnshire and Barnes' Dorset poems still have a claim to be authentic dialect. However, Barnes saw the danger of an artificial 'revivalist' type of poetry and compared writing in a dying dialect to writing in snow on a spring day (quoted in Görlach, 1992), but he also saw the genuine, straightforward and honest character of local dialect. This attitude becomes apparent in his facetious 'translation' of stilted passages from the Queen's speech into plain Dorset dialect.

> **Q17** In somewhat of a merry mood, I was one day minded to see how far our homely Dorset speech could give the meaning of the seemingly ministerial wording of the so-called Queen's speech on the opening of Parliament in 1884. Her Majesty's speech as written and read in Her Majesty's name. Here are samples of a few clauses -
>
> 'My Lords and Gentlemen. – The satisfaction with which I ordinarily release you from discharging the duties of the Session is on the present occasion qualified by a sincere regret that an important part of your labours should have failed to result in a legislative enactment.'
>
> *(1) The lightheartedness I do mwostly veel when I do let ye off vrom the business upon your hands in the Sessions, is theäse time a little bit damped, owen to a ranklen in my mind, that a goodish*

lot o' your work vell short o' comen into anything lik laws.

'The most friendly intercourse continues to subsist between myself and all foreign Powers.'

(2) The very best o' veelens be still a-kept up, in deälens between myzelf an' all o' the outlandish powers.

(Barnes 1886)

There is, then, a distinction to be made in dialect writing according to the audience aimed at: it can raise respect for the living dialect and give it a kind of prestige as long as it is an exclusively spoken, but common, form; however, it tends to become artificial and nostalgic when the basis of everyday use is gone – a development that has increasingly affected writing in English dialect in the 20th century and is becoming a danger to Scots (for the complete context see my discussion in Görlach, 1992).

4. The question of Cockney

The development which came to consider rural dialect (as represented by the 'best' speakers, usually NORMS = 'non-mobile old rural males') as 'pure' was a 19th-century innovation. The nostalgic reverence of dialect as a phenomenon of the lost golden age found its counterpart in the depreciation of urban lower-class speech – in Britain represented by London, since the other urban centres were still on their way to develop urban norms. Cockney received two quite different evaluations, both, however, making it inacceptable as 'the regional dialect of London' (as in Mayhew's account).

1) Seen from the perspective of a traditional dialectologist, the speech was unbearably mixed and discredited by its connection with poverty and crime, as Forby had found as early as 1830. Later on, Halliwell (1847/1881) was even more outspoken:

Q18 The metropolitan county presents little in its dialect worthy of remark, being for the most part merely a coarse pronunciation of London slang and vulgarity.

(Halliwell, 1881: xxiv, quoted from Ihalainen, 1994: 212)

2) From the 18th century onwards, Cockney speech had been used for comic characters on the stage (Matthews, 1938) and later on in dialogue in narrative. The music-hall and Dickens are probably the most typical representatives of this tradition – which became quickly stereotyped and fossilized. When Shaw looked back on the tradition in the notes to *Captain Brassbounds Conversion* (1900), he found it was already a matter of the past.

19th-century Cockney can be regarded as a blend of regional dialect (being confined to London, and characterized by pronunciation, syntax and lexis) and sociolect (being restricted to informal uses of the lower classes), which is further stigmatized by its conventional association with comic characters of low intelligence and less education.

5. *Dialect research*

The methods of 19th-century dialectology in England are not my major concern here (cf. Petyt's summary, 1980: 68-81, and Shorrocks, fc.).[15] However, it is important to interpret the motivations that led to a noteworthy peak in scholarly dialectology between 1870 and 1898 as a reflexion of attitudes in the academia and educated circles outside the universities: whereas the organization of the great research projects and the analysis and publication of the findings was in the hands of Ellis and Wright and other members of the EDS, the data collection could not have been undertaken without the willing and unpaid assistance of thousands of helpers who served as correspondents, informants and collectors of data. This reflected a widespread positive interest in dialects in the English society of the time – evidenced by an impressive return rate of the questionnaires, which it would be difficult to repeat in our times. It may be good to distinguish between a more popular attitude and the linguistic interest and point out a few characteristics of the main researchers:

1) Many of the major scholars had been broad dialect speakers in their youth – such as Joseph. Wright, who described his idiolect in his epoch-making monograph on the dialect of Windhill (1892), or the Scotsman James Murray.
2) Most scholars were philologists devoted to the study of Old and Middle

English texts (most notably Ellis, Wright and Skeat), were active as editors for the EETS and all were committed to the comparative-historical method, which some had studied in Germany (such as Skeat, Sweet and Wright, who even published some of their research in German).

The background was, then, the conviction that comparative philology in combination with the rapidly developing discipline of phonetics would raise linguistics to a science at long last, capable to compete with the natural sciences. Dialect research was important since the data preserved many features lost from the standard languages but necessary for a proper reconstruction of the history of English. This unique combination of scientific rigidity (which culminated in the neogrammarian tenet that sound changes permit no exceptions) and the 19th-century fascination with history gave dialectology a special place in linguistic investigation, made more urgent by the widely accepted belief that the researchers were members of the last generation to find a bidialectal speech community with living dialects. Note that both comparative philology and phonetics led to an increased interest in dialect pronunciations, whereas earlier collectors had been mainly concerned with vocabulary.

Is a scholarly description of 19th-century dialects possible on the basis of the surviving data and interpretations? The prevailing attitudes, the topic of my reflexions, tend to distort the evidence – whether they are negative (as in HMIs' reports or in statements about urban speech) or positive (as in the nostalgic idea of a dialect-using golden age then being destroyed by railways and modern communication).

The elicitation of dialect data and their analysis has also been repeatedly criticized: Ellis' material was regarded by Dieth as 'a tragedy ... a huge store of information which every dialectologist consults, but, more often than not, rejects as inaccurate and wrong' (quoted from Petyt, 1980:76).[16] Not all regions were covered with the same degree of thoroughness by either Ellis or Wright, and Wright's exclusion of spoken data (in the tradition of the *OED*) is very difficult to understand and impossible to justify. If dialect was specific of certain social classes, mainly restricted to spoken uses and prevalent (or even obligatory) in informal registers in individual speech communities, then a full description should have included data from these ranges, and the researcher should have noted social and stylistic restrictions. The time was obviously not ripe for a dialectology systematically investigating the social aspects of dialect use (cf. Mattheier, 1980). Nor was anyone in the 19th century keen on making a survey of

attitudes to dialect and perceptual dialectology (cf. for the U.S. Preston, 1989). But since there has not been any modern investigation of nonlinguists' views of regional differences in BrE and their evaluation, either, the absence in the 19th century is less surprising. However, scholars investigating 19th-century attitudes have not even a modern description available for comparison if they wish to undertake a properly diachronic study of the topic, and to raise the discussion from the impressionistic and anecdotal level.

Apart from such obvious limitations, the specific motivation of the researchers in the 19th century was even more restrictive: the treatment of dialect in the poetry of Barnes, Tennyson or Hardy was as much a middle-class affair as that of Ellis and Wright, however different their motives and linguistic qualifications. It is obvious that their research did not raise the prestige of the dialects among the native speakers and did not convince them that dialects were worth handing on to the next generation (as publications of the Yorkshire almanacs may have done at least for a time).

Such a conclusion is of course not specific for the 19th century – but the period was possibly the beginning of a development which has reduced the frequency of the use of dialects as well as minority languages all over 20th-century Europe and in consequence impoverished them in functions, expressiveness, modernity and structural complexity, a development which results in a vicious circle, since the dialects tend to be used even less – unless they are supported by a very strong connection between language and local identity.

NOTES

[1] For helpful comments on draft versions I want to thank Clive Upton and Katie Wales. For a survey stressing different aspects of 19th-century English dialects and dialectology see in particular Ihalainen (1994). A much shorter version of this paper will also be included as a chapter in Görlach (1999).

[2] The Exmoor pieces are likely to be reprints of texts originally composed in 1727 (cf. Shorrocks, fc.)

[3] The seventh edition of the two Exmoor pieces was reprinted for the EDS in 1879; F.T. Elworthy accompanied the text by the phonetic transcription he developed for rendering English dialects and a joint glossary (1879: 24-176). The *Scolding* is also unique being published in a 'bilingual' edition, with accompanying translation in StE (Exeter, 1792).

[4] Note that the loss of dialect was felt to be an ongoing process in another stronghold at the same time. John Collier alias Tim Bobbin stated in the 1775 introduction to his *View on the Lancashire Dialect*:

> But as Trade in a general Way has now flourish'd for near a Century, the Inhabitants not only Travel, but encourage all Sorts of useful Learning; so that among Hills, and Places formerly frequented by Strangers, the People begin within the few Years of the Authors Observations to speak much better English. If it can properly be called so.
>
> (1775: iv)

[5] Cf. similar statements quoted by Mugglestone (1995: 306 and 308): In a report of 1860, Mr Grant claimed 'teachers were lying in wait for provincialisms', and, as they occur, making constant interruptions in order to align 'faulty' utterances with the normative paradigms of 'proper' speech. *Chambers' Educational Course* (n.d.) demanded that '*Clearness of articulation* should be most carefully inculcated, as indistinctness acquired in childhood can hardly ever be removed. By these means, vulgar and provincial dialects will be gradually extirpated, and purity of speech introduced.'

[6] In-migration from the neighbouring rural districts appears to have been a special feature of the newly-developing industrial centres of the Midlands and the North, which created large speech communities of speakers of regional dialects (though no longer village varieties). There is an obvious parallel in 20th-century urban German dialects; at least before 1945 newcomers to Cologne, Mannheim, Stuttgart and Munich, to name only a few centres, migrated from the surrounding countryside, which has made the regional urban dialects very stable to the present day (cf. Mattheier, 1980).

[7] There is indirect evidence of the ineffectiveness of the teaching of St E in emigrants' letters; the need to communicate in written form made many people write letters who would not have done so in England, and the proportion of semi-literates among them is particularly high. However, the value for these documents is slighter than claimed is by Garcia-Bermejo Giner & Montgomery (1997).

[8] Modern attempts in a framework of 'enlightened' linguistics may be contrasted (all, however, apparently discontinued): there were schoolbook texts in Black English developed by Wolfram and Fasold (1969) intended to smooth the students' transition to St E and an entire series, was published in Germany (Besch, 1976-78) which provided linguistic analyses contrasting one regional dialect with *Schriftdeutsch* and offered teaching aids based on predictable difficulties.

[9] This is also true for grammars published for use outside England. No special method was developed for the acquisition of English in, say, Ireland or India, and even grammar books published in Scotland usually contain lists of Scoticisms — at best.

[10] For Yorkshire, compare William Wright, who styled himself 'The Yorkshire Burns',

and wore a plaid cap to support his claim (Arnold Kellett, p.c.).

[11] For the most recent summary of dialect in Victorian fiction see Chapman (1994: 50-66). Dickens' attempts at Lancashire dialect (e.g. Stephen Blackpool in *Hard Times*) are not very convincing — but it is likely that his readers did not care whether its representation was realistic or not.

[12] There is a link with the poetry discussed above in that Mrs Gaskell's *Mary Barton* contains a (toned-down) version of 'Th'Owdham weyver' (Hollingworth, 1977: 128), apart from a good deal of dialect in the dialogues.

[13] For an insightful comparison of Eliot's and Hardy's uses of dialect cf. Ingham (1970).

[14] She wrote to the publisher:

> "It seems to me advisable to modify the orthography of the old servant Joseph's speech; for though as it stands it exactly renders the Yorkshire dialect to a Yorkshire ear, yet I am sure Southerners must find it unintelligible; and thus one of the most graphic characters is lost on them. (Quoted from Petyt, 1970: 2)

Petyt comments: 'I think there can be little doubt that Charlotte did more harm than good (...) the resulting dialect speech is mixed, artificial and inconsistent' (1970: 3).

[15] Graham Shorrocks' paper read at the Harold Orton conference somewhat overlaps with mine and should be compared for this section; it will be published elsewhere.

[16] However, compare Shorrocks' fair statements about the achievement, with all necessary restrictions, of the *EDD* and Wright's contribution to it.

REFERENCES

Bailey, Richard W. 1996. *Nineteenth-century English.* Ann Arbor: University of Michigan Press.

Barnes, William. 1878. *An Outline of English Speech-Craft.* London: Kegan Paul.

Batchelor, Thomas. 1809. *An Orthoepical Analysis of the English Language.* London: for Didier and Tebiett.

Besch, Werner (ed). 1976-78. *Dialekt/Hochsprache — Kontrastiv.* Sprachhefte für den Deutschunterricht. Düsseldorf: Bagel.

Blake, N.F. 1981. *Non-standard Language in English Literature.* London: Deutsch.

Burnett, Frances Hodgson. 1877. *That Lass o' Lowrie's.* (Repr. Woodbridge: Boydell Press, 1985).

Chapman, Raymond. 1994. *Forms of Speech in Victorian Fiction.* London: Longman.

[Clarke, Hewson]. 1811. *Lexicon Balatronicum: A Dictionary of the Vulgar Tongue.* A

dictionary of buckish slang, university wit, and pickpocket eloquence. London: for C. Chappel. (Repr. Northfield, III.: Digest Books, 1971).

Cobbett, William. 1818. *A Grammar of the English Language, in a Series of Letters.* New York/London: Clayton & Kingsland, for the author, sold by T. Dolby. Revised ed. 1823, London. New ed. 1866 with an additional chapter by J.M. Cobbett. (1823 ed. repr., with an introduction by Robert Burchfield, Oxford: University Press, 1984).

[Collier, John], 1775. *The Miscellaneous Works of Tim Bobbin, Esq.; Containing his View on the Lancashire Dialect; with large Additions and Improvement.* Manchester: for the Author. ([1]1746).

Donaldson, William. 1986. *Popular Literature in Victorian Scotland: Language, Fiction and the Press.* Aberdeen: Aberdeen University Press.

Ellis, Alexander J. 1869-74. *On Early English Pronunciation.* London: Asher.

Elworthy, Frederick. 1888. *The West Somerset Work-Book.* London.

Forby, Robert. 1830. *The Vocabulary of East Anglia*, vol I. London: J.B. Nichols & Son. Repr. 1970, Newton Abbot: David & Charles.

Forster, John. [3]1977. *Class Struggle and the Industrial Revolution: Early Industrial Capitalism in Three English Towns.* London.

Foster, A.F. and Foster, M.E., 1860. English literary and Vernacular, repr. in Harris 1995 (I): 312-38

Garcia-Bermejo Giner, Maria F. and Michael Montgomery. 1997. British regional English in the 19th century: the evidence from emigrant letters. In Alan R. Thomas (ed) *Issues and Methods in Dialectology* 167-83. Bangor: University of Wales.

Golby, J.M. ed. 1996. *Culture and Society in Britain 1850-1890.* Oxford: Oxford University Press.

Görlach, Manfred. 1992. Sociolinguistic determinants for literature in dialects and minority languages. In J.J. Simon and Alain Sinner (eds) *Proceedings of the Third Conference on the Literature of Region and Nation* pp. 130-60. Luxembourg,.

Görlach, Manfred. 1995a. *New Studies in the History of English.* (AF232). Heidelberg: Winter.

Görlach, Manfred. 1995b. Dialect lexis in Early Modern English dictionaries. In Görlach 1995a: 82-127.

Görlach, Manfred. 1999. *English in 19th-century England.* Cambridge: University Press.

Grose, Francis. 1785. *A Classical Dictionary of the Vulgar Tongue.* London; repr. Menston: Scolar (EL 80).

Grose, Francis. 1787. *A Provincial Glossary, with a collection of local proverbs and popular superstitions.* London: for S. Hooper; repr. Menston: Scolar (EL 63).

Halliwell, J.O. 1847. *A Dictionary of Archaic and Provincial Words* 2 vols. London: John Russell Smith. ([10]1881).

Harris, Roy (ed). 1995. *English Language and Language-Teaching 1800-1865.* London:Routledge/Thoemmes Press.

Hollingworth, Brian (ed). 1977. *Songs of the People: Lancashire Dialect Poetry of the Industrial Revolution.* Manchester: University Press.

Honey, John. 1988. 'Talking proper': schooling and the establishment of English 'Received Pronunciation'. In Graham Nixon and John Honey (eds) *An Historic Tongue: Studies in English Linguistics in Memory of Barbara Strang* pp. 209-27. London & New York: Routledge.

Honey, John. 1989. *Does Accent Matter? The Pygmalion Factor.* London & Boston: Faber & Faber.

Honey, John. 1997. *Language is Power: The Story of Standard English and its Enemies.* London: Faber & Faber.

Huntley, Revd R.W. 1868. *A Glossary of the Cotswold (Gloucestershire) Dialect.* London: John Russell.

Ihalainen, Ossi. 1994. The dialects of England since 1776. In R.W. Burchfield (ed) *English in Britain and Overseas: Origins and Development* pp. 197-274. Cambridge: University Press,

Ingham, Patricia. 1970. Dialect in the novels of Hardy and George Eliot. In George Watson (ed) *Literary English Since Shakespeare* pp. 347-63. Oxford: University Press,.

Jacobs, Willis D. 1952. *William Barnes Linguist.* Albuquerque: University of New Mexico Press.

Johnson, Samuel. 1755. *A Dictionary of the English Language.* London: for J. & P. Knapton, etc. (Repr.: London: Times Books, 1983; revised ed. by Todd, London; for Longman etc., 1818).

Kellett, Arnold and Ian Dewhurst (eds). 1997. *A Century of Yorkshire Dialect.* Otley: Smith Settle.

Lloyd, J.R. 1899. *Northern English.* Leipzig: Teubner. (^{2}1908).

Lorimer, William Laughton. 1983. *The New Testament in Scots.* Edinburgh: Southside.

Marshall, W. 1787-96. *Provincialisms of East Norfolk/East Yorkshire/Vale of Glocester/Midland Counties/West Devonshire*; repr. English Dialect Society 1 (1873).

Mattheier, Klaus. 1980. *Pragmatik und Soziologie der Dialekte.* (UTB 994). Heidelberg: Quelle & Meyer.

Matthews, W. 1938. *Cockney Past and Present: A Short History of the Dialect of London.* London: Routledge & Kegan Paul.

Michael, Ian. 1987. *The Teaching of English: From the Sixteenth Century to 1870.* Cambridge: Cambridge University Press.

Mugglesetone, Lynda. 1995. *'Talking Paper': The Rise of Accent as Social Symbol.* Oxford: Clarendon Press.

Pearson, William. 1965. *The Self-help Grammar of the English Language; Intended for Reading, Dictation, Parsing, Composition and Home-Work in the Second and Third Classes in an Elementary School.* London, etc.: Hamilton, etc.

Petyt, K.M. 1970. *Emily Brontë and the Haworth Dialect.* Menston: Yorkshire Dialect Society.

Petyt, K.M. 1980. *The Study of Dialect: An Introduction to Dialectology.* London: Deutsch.

Phillipps, K.C. 1984. *Language & Class in Victorian England.* Oxford: Blackwell.

Poole, John. 1813. *The Village School Improved.* London.

Preston, Dennis R. 1989. *Perceptual Dialectology: Nonlinguists' Views of Areal Linguistics.* Dordrecht: Foris.

Viereck, Wolfgang. 1970. The English Dialect Society and its dictionary. *Transactions of the Yorkshire Dialect Society*, 70. (Reihe? Zeitschrift?)

Wakelin, Martyn F.[2]1977. *English Dialects: An Introduction.* London: Athlone.

Walker, John. 1791. *A Critical Pronouncing Dictionary and Expositor of the English Language.* London; repr. Menston: Scolar (EL 117).

Wolfram, Walter A. and Ralph W. Fasold. 1969. Toward reading materials for speakers of Black English. In Joan C. Baratz and Roger W. Shuy (eds) *Teaching Black Children to Read* pp. 138-55. Washington, D.C.: Center for Applied Linguistics,.

Wright, Joseph. 1892. *A Grammar of the District of Windhill, in the West Riding of Yorkshire.* London: Kegan Paul & Trübner.

Wright, Joseph. (ed) 1898-1905. *English Dialect Dictionary* (*EDD*). 6 vols. London: Frowde.

Wright, Thomas. 1857. *Dictionary of Obsolete and Provincial English.* London.

Address

Prof. Dr Manfred Görlach
Universität zu Köln
Englisches Seminar
Albertus Magnus-Platz
D-50923 Köln
Germany
<manfred.goerlach@uni-koeln.de>

Scots Language Attitudes and Language Maintenance

Caroline Macafee and Briege McGarrity

Abstract

This paper attempts to correlate language attitudes in Aberdeen with a range of social variables, and with a simple measure of language maintenance.

A random sample with non-locals eliminated was taken from three districts of Aberdeen. 75 individuals filled in a postal questionnaire on their attitudes towards the traditional dialect of the North-East of Scotland, and of these 62 agreed to an interview. This allowed their attitudes to be explored in more depth, and they also responded to a lexical questionnaire which tested their knowledge and asked about their use of 96 dialect terms to do with weather.

Attitudes were on the whole very positive, and question by question there were no significant differences (at the 0.05 level) by gender, and very few by age or class. The lexical scores, however, showed, as expected, a very steep age gradient, although knowledge was not high even in the oldest age group (41 out of 96 items). Regression showed an average loss from frequent use of one word (i.e. almost 1% of this small specimen of lexis) for every 6 years of age.

In order to test for a relationship between attitudes and language maintenance, factor analysis was used to extract three attitudinal factors from the questionnaire responses, and these were correlated with the lexical scores. The factors were interpreted as Defensiveness, Positiveness and Participation. Participation and Positiveness did correlate with some measures of claimed use of the vocabulary, but there was no correlation with measures of knowledge.

The main aim of this paper is to explore the possibility of measuring and quantifying the relationship between language attitudes and language maintenance. It also has a secondary aim of exploring differences in attitudes between the Central Belt of Scotland and the North-East.

Anyone acquainted with the literature on the present state of the Scots language will be familiar with the low status that it has in Glasgow and Edinburgh, and its strong links there with social class stratification. Scots has been absorbed into a continuum with English in the Central Belt, to the extent that it is possible to conduct Labovian sociolinguistic studies, treating the Scots element as merely variation within English. A number of recent studies (Menzies, 1991; Macafee, 1994; Máté, 1996) indicate that speakers of Scots dialects in the Central Belt (and the South-West) often perceive themselves as speaking a mixed language, or even a slang form of Scots or English.

As we move north, however, we escape the anglicising influence of Edinburgh and the proletarian influence of Glasgow and, by the time we reach the North-Eastern corner of Scotland, we find a self-sufficient, locally rooted, proudly Scottish bourgeoisie who set a quite different tone with regard to local and Scottish culture, including the Scots language. North-Eastern Scots (sometimes known by the journalistic term 'the Doric') is held in high esteem, and is universally understood and spoken at least on some occasions by individuals all the way up the social scale. More than in Central Scotland, Scots in the North-East is perceived to be a linguistic entity distinct from English, and is focused by code-switching, rather than the code-mixing characteristic of the Central Belt.

Two recent studies which asked members of the public to say whether they spoke Scots both showed considerable regional differences. (The local government regions have now been replaced by a country structure.) The larger study (Máté, 1996; Macafee, in progress) used three separate market research studies with approximately one thousand respondents in each. A selection of the results is shown in Table 1.

Table 1: Proportion of Scots speakers in three regions of Scotland, based on Máté, 1996: Table 5. Figures from Survey 2 have been omitted as invalid (Macafee, in progress)

Region	First Survey Scots speakers			Second Survey Scots speakers		
		n	%		*n*	%
Grampian	66	110	*60*	63	108	*58*
Lothian	41	178	*23*	46	136	*34*
Strathclyde	129	455	*28*	137	487	*28*

The earlier – and ground-breaking – study by Murdoch (1995) has higher figures overall (probably in part because the methodology involved face-to-face interviews rather than cold questioning as part of a routine market research survey, and also in part because of a high proportion of 'non-native speakers', apparently people with passive knowledge of the language). In both studies, however, the same disproportion between the Central Belt and the North-East is evident. Murdoch has separate figures by the region in which people were interviewed and by the region in which they received their schooling. The figures for schooling show a very high proportion of the Grampian sample identifying themselves as Scots speakers (97%). See Table 2.

Table 2: Scots spickers resident an Scots spickers skweelit be selectit regions (based on Murdoch and Gordon)

Place of schooling /residence	Scots speakers schooled			Resident Scots speakers		
		n	%		*n*	%
Grampian	23	24	*96*	30	30	*100*
Aberdeen	33	34	*97*	24	30	*80*
Total Grampian	**56**	**58**	***97***	**54**	**60**	***90***
Lothian	12	21	*57*	-	-	-
Edinburgh	8	13	*62*	-	-	-
Total Lothian	**20**	**34**	***59***	**16**	**30**	***53***
Strathclyde	19	38	*50*	9	30	*30*
Glasgow	20	48	*42*	16	30	*53*
Total Strathclyde	**39**	**86**	***45***	**25**	**60**	***42***
All	**255**	**402***	***63***	**258**	**450**	***57***

*The nummers dinna add up tae 450 as fowk educatit ootside Scotlan hae been omittit

Given these background differences, we should not be surprised to find that language attitudes in the North-East are rather different from those described in the sociolinguistic literature for Glasgow and Edinburgh, having more in common with the language loyalty reported from Shetland (Melchers, 1985), but combined, as in

Shetland, with the internalised ideologies of 'correctness' and 'clarity' attributed to the super-imposed Standard English (cf. Graham, 1983 and Roberts, 1993).

The research reported here was carried out in Aberdeen by Briege McGarrity under the supervision of Caroline Macafee. It was an attempt to build on the 1994 Glasgow research of Macafee, in which qualitative methods were used to explore language attitudes, and the attitudes expressed by the people interviewed did go a long way towards explaining the low status of Glasgow dialect, and its apparent decline in terms of lexical richness. Macafee's lexical questionnaire of about one hundred words was not designed as a measure of decline (on the contrary, it included children's vocabulary and slang as well as traditional Scots words), but as a stimulus for conversation with the interviewees. The Glasgow research consequently lacked quantitative measures of vocabulary maintenance and language attitudes, and there was therefore no way of systematically relating the two. In the Aberdeen research, an attempt was made to put this relationship on a quantitative footing.

The 1991 Census statistics for the Aberdeen wards were used to choose three areas of the city that it was hoped would produce a representative sample by social class. In practice, differential response rates produced a sample rather skewed towards the middle classes. The voters' roll was used to take a random sample of the adult population. A profile of the sample is given in Tables 3a-3d.

Tables 3a-3d: Profile of samples randomly selected in 3 wards of Aberdeen (non-natives dropped). 75 individuals replied to a postal questionnaire on attitudes, of whom 62 were interviewed and completed a lexical proficiency test.

Table 3a: Respondents by age group

Age group	No of respondents: lexis	No of respondents: attitudes
20-33yrs	23	25
34-45yrs	21	25
46-59yrs	11	13
60+ yrs	7	11
All	62	74*

* missing data for 1 individual

Table 3b: Respondents by educational level.

Educational level	No of respondents: lexis	No of respondents: attitudes
Minimum	22	25
Secondary	9	13
Tertiary	31	37
All	62	75

Table 3c: Respondents by gender.

Gender	No of respondents: lexis	No of respondents: attitudes
Males	26	34
Females	36	41
All	62	75

Table 3d: Respondents by manual versus non-manual occupation.

Occupation	No of respondents: lexis	No of respondents: attitudes
Non-manual	39	49
Manual	23	26
All	62	75

McGarrity (1998) gives a breakdown into five occupational classes, but unfortunately only two unskilled manual respondents were obtained, and many of the analyses by occupational class in the research become statistically significant only when the occupational groups are collapsed into the two broad categories given here.

The initial contact was made by post. An attitude questionnaire and a questionnaire asking for basic personal data were sent out. The people contacted were also asked if they would agree to an interview. Non-locals were eliminated when the postal questionnaires were returned, leaving 75 individuals, of whom 62 agreed to be interviewed. Macafee's experience in Glasgow had led us to attach a high value to qualitative interviews as a way of countering the reductive tendency of attitude questionnaires, and of getting to the complexity and internally contradictory nature of real people's responses to linguistic pressures. We do not have space to explore this aspect in the present paper, but see McGarrity (1998).

The interview also included the administration of the lexical questionnaire. This was, in effect, a very crude proficiency test. What we did was to select an area of vocabulary and establish what we hoped was a full word list for the North-Eastern Scots dialect. The area chosen was the weather, which then had to be considerably whittled down to precipitation only, in order to get a small enough word list. We considered using parts of the body, and in retrospect this would probably have been a better measure, but we were afraid that this was too intimate a topic. Initially, we thought that weather was something that was not likely to be affected by external factors (in the way that much dialect vocabulary is moribund because the referent is out of date), but of course, the vocabulary of the weather is affected by changes in lifestyle and working conditions – people simply do not have to pay so much attention to the vagaries of the climate, living as we do a more protected and indoors life than previous generations. The word list was compiled from *The Concise Scots Dictionary* and recently-published local word lists (Buchan and Toulmin, 1989: Wilson, 1993) to provide a baseline of vocabulary traditional to the area. The questionnaire was administered in two parts, first of all supplying the English word and asking for Doric equivalents, and then presenting the Doric words and asking if the person knew and used them. 'Use' was graded on a three-point scale from 'rarely' through 'occasionally' to 'frequently'. There were 96 weather terms, including words distinctive to Scots (e.g. onding ëa downpourí) or to the Doric (e.g. ondag ëa downpourí, goor ëslush in running waterí, drabblichy ëdrizzlyí) and distinctive forms of words shared with Standard English (e.g. weetie ëwetí). The scoring system was as follows:

> 0 = no knowledge, 1 = passive knowledge, 2 = rare use, 3 = occasional use, 4 = frequent use.

This gives a number of possible ways of scoring the lexical results. Taking the

results as they stand, the possible score for any word is 0-4. It is also possible to produce a Knowledge score by keeping 0 as 0 and taking any other response (i.e. knowledge or any other degree of use) as 1. Similarly, a Use score can be constructed by taking knowledge without use as 0 and any use as 1. Scores were also calculated for Occasional/Frequent Use and for Frequent Use. In brief, the composite lexical scores were calculated as follows:

Overall = average score
Knowledge = recode 0 = 0, 1/2/3/4 = 1
Use = recode 0/1 = 0, 2/3/4 = 1
Occasional/Frequent use = recode 0/1/2 = 0, 3/4 = 1
Frequent use = recode 0/1/2/3 = 0, 4 = 1.

A summary of the scores by age, gender, education and occupation is presented in Table 4.

Table 4: Group scores, lexical questionnaire

Social Groups	Overall possible score = 384	Knowledge possible score = 96	Use possible score = 96	Occasional/ Frequent Use possible score = 96	Frequent Use possible score = 96
20-33	47	24	10	8	5
34-45	60	30	13	11	6
46-59	**82****	**38***	**17***	**15***	**11***
60+	**93****	**41***	**23****	**18****	**11***
Male	63	31	14	12	8
Female	62	30	13	12	7
Minimum	**74***	33	**17***	**15***	**10***
Secondary	57	26	13	12	7
Tertiary	55	30	11	10	5
Non- manual	56	29	12	10	5
Manual	**73***	36	**16***	**14***	**10***

The asterisks denote statistical significance at the 0.05 (**) or the 0.1 (*) level.

It will be seen that there is a relationship of both Knowledge and Use (in various degrees) – and therefore also the overall score – with education and with occupation. (Although not statistically significant, the difference between non-manual and manual occupations for Knowledge is in the expected direction.) There is no relationship with gender.

One of the difficulties in measuring language maintenance is the establishment of a realistic baseline. The main reference sources – the Scots dictionaries and *The Linguistic Atlas of Scotland* – are all at least partly antiquarian in intention and probably do not represent the typical Scots vocabulary of the average member of the public at the time the materials were collected. Various studies of vocabulary maintenance have found a very steep decline relative to these sources (for a summary of such studies, see Macafee, 1997; and add Hendry, 1997). In the present study, we found that even the oldest age group in the sample knew less than half of the vocabulary (see the column 'Knowledge' in Table 4). The number of words in Frequent Use was very small. For other age groups, the loss of vocabulary was considerable: a regression on age showed an average loss of one word (i.e. about 1% of this small specimen word list) from Frequent Use for every six years of age. However, this is not so steep as the decline in Knowledge. The figures for all four categories of Knowledge and Use are as follows:

> Knowledge: coefficient = 0.38, i.e. approximately 1 word/ word form lost for each 2.5 years of age
> Use: coefficient = 0.26, i.e. approximately 1 word/ word form lost for each 4 years of age
> Occasional/Frequent use: coefficient = 0.20, i.e. approximately 1 word/ word form lost for each 5 years of age
> Frequent use: coefficient = 0.16, i.e. approximately 1 word/ word form lost for each 6 years of age.

In general, the decreasing rate of loss through these categories suggests that Frequent Use may be bottoming out at the end of an S-curve, leaving a residuum of vocabulary with a good chance of indefinite survival.

The attitude questionnaire is given in Table 5 (see Appendix). It includes questions of belief (e.g. 'Women know and use more Doric words and phrases than men') as well as value judgements (e.g. 'The Doric is old-fashioned') and readiness for action (e.g. 'Schools in the North-East should encourage their pupils to become better acquainted with the Doric and Scots language in general'). There are also factual

questions about language (e.g. 'I can understand the Doric in its written form'). The respondents were presented with a five-point scale from 'strongly agree' to 'agree' to 'neutral' to 'disagree' to 'strongly disagree'. The presentation of the questions was varied so that respondents did not fall into a pattern of expecting always to agree or disagree. Most of the questions can be regarded as having positive or negative implications for the language, and the scores are adjusted in Table 5 so that a high score (above 3) is always a sign of positive rather than negative attitudes (if applicable). A quick glance at the mean scores in Table 5 will show that the attitudes expressed were generally very positive, despite this being a mostly middle class, urban sample. There tended to be agreement with some pessimistic statements (7, 26), giving low scores; and statements concerning participation (4, 15, 27) also produced some scores neutral or below. Statements 20 and 21, to which we will return, are also neutral or low.

There is no space here to discuss the scores for the individual attitude questions (see McGarrity, 1998). The most notable findings were that the elderly were more likely to agree that the Doric was old-fashioned (statement 17), but never agreed that they used Doric words only in jest (statement 28). The professional and managerial class mostly disagreed that they spoke the Doric (statement 4). (Surprisingly, the two unskilled manual respondents were also divided between disagreement and neutrality.) Skilled non-manual workers were particularly likely to disagree that lack of contact between grandparents and grandchildren was killing the Doric (statement 26). They were more neutral than other occupational classes on whether the Doric forms part of the North-East identity (statement 29) and on whether North-Easterners should make efforts to preserve Doric words and phrases (statement 30) (compare Labov's classic findings concerning the linguistic aspirations and insecurity of the second-highest class). Dividing the sample into two occupational groups, as here, non-manual respondents were more likely to use Doric words in jest (statement 28). Those with tertiary education were less likely to say that they spoke the Doric, but were more likely to speak the Doric only in jest. Unlike those with minimum or tertiary education, those with secondary education showed no agreement at all with the statement that the Doric is mostly spoken by uneducated people (statement 11). Again, there were no significant differences by gender.

In order to test for a correlation between attitude scores and lexical proficiency (or language maintenance), we attempted to summarise the 30 questions in terms of broad, underlying attitudinal dispositions. These composite attitude scores were constructed using a Principal Components Analysis. This is a way of comparing the results for the 30 attitude questions in order to group together those questions that are,

in some sense, measuring the same thing. In this analysis, one statement (5) was dropped, as the interviews had revealed the responses to be unreliable. Four others (2, 10, 20 and 27) were dropped for low sampling adequacy as revealed by an anti-image correlation matrix (they did not differentiate amongst individuals – in other words there was a large measure of agreement in the responses). The analysis produced 8 factors with an Eigen value of 1.00 or more, cumulatively explaining 72.1% of the variance. The first 3 factors were selected for discussion, as showing meaningful groupings of the statements, and these 3 factors cumulatively explain 47.8% of the variance. The statements that make up each factor are noted in the first column of Table 5 (in abbreviated form). Notice that a particular statement can contribute to more than one factor. The first factor was termed Defensiveness. Its Eigen value was 5.53, and it explained 22.1% of the variance. The strength of this factor, however, may be due to the similarity amongst some of the constituent questions. The second factor was termed Positiveness, with an Eigen value of 4.48, explaining 18.0% of the variance. The third factor was termed Participation, with an Eigen value of 1.92, explaining 7.7% of the variance. There was only one significant relationship between these scores and social variables, namely Participation with occupation. Recall that we also found only a few significant relationships between the individual attitude questions and the pre-determined social factors. This suggests that there are unidentified extra-linguistic factors at work. The three components together explain about half of the variance in the data, but they are measuring some unknown aspects of life experience that do not coincide neatly with age, education or occupation.

One of the main goals of the research was to relate attitudes to language maintenance. This was done by giving each individual a score on each of the three attitudinal factors and correlating these scores with scores for lexical knowledge and use. There were only three significant correlations, all of them positive. The factor Positiveness correlated with one of the measures of lexical use. The factor Participation correlated with two of the measures of use, as one would hope, since it includes statements about language use. The figures were as follows:

Use correlates positively with Positiveness (+0.333, *)
Use correlates positively with Participation (+0.509, **)
Occasional/Frequent Use correlates positively with Participation (+0.440, **)
(* = significance, ** = strong significance)

In general, however, the usual social variables do not go a long way towards

explaining language attitudes in this community, and language attitudes do not go very far in explaining lexical decline. The attitudes are largely positive and this is shared by the middle as well as the working classes, and yet the traditional vocabulary seems to be rapidly disappearing from use (and thereafter, of course, from knowledge).

We conclude that attitudes, as we have been able to measure them, may to some extent guide the use of the traditional dialect, particularly a readiness to participate and the assignment of a high value to the dialect (elements of the factor we have named Participation); but that the continuing decline in the corpus of vocabulary remains to be explained by other factors, such as change in material culture (in this case less exposure to and dependence on the vagaries of the elements). There were, nevertheless, clearly internal conflicts in speaker attitudes. It should also be noted that the Aberdeen dialect is perceived as degraded relative to the traditional rural dialect, and that in stating their attitudes towards 'the Doric', respondents may have had the idealised rural dialect in mind. Two of the few questions on which speakers were not on average positive (see Table 5) were statement 20, 'Standard English speakers have more success in the modern world than dialect speakers' (i.e. there tended to be agreement with this), and statement 21, 'Standard English should be spoken to children' (neutral). The face-to-face interviews also brought out this old, familiar story of a tension between loyalty to the traditional dialect and a belief that it is a material handicap. To some extent, previous studies of attitudes (as suggested in Macafee, 1994) may have unintentionally focused the respondents' attention on issues of 'inferiority' relative to Standard English, but the regional figures for Scots speaking, cited at the beginning of this paper, suggest that language loyalty may indeed be particularly strong in the North-East of Scotland.

ACKNOWLEDGEMENTS

We would like to thank Tony Glendinning and Alan Anderson for their invaluable help and advice with the statistical aspects of this study.

REFERENCES

Buchan, Peter and David Toulmin. 1989. *Buchan Claik. The Saut an the Glaur o't. A Compendium of Words and Phrases from the North-East of Scotland.* Edinburgh: Gordon Wright.

The Concise Scots Dictionary. 1985. Ed.-in-chief Mairi Robinson. Edinburgh: Chambers.

Graham, John. 1983. Social changes during the Quinquennium. In Donald Withrington (ed) *Shetland and the Outside World 1469-1969*, pp 216-33. Oxford: Oxford University Press for the University of Aberdeen.

Hendry, Ian. 1997. *Doric – An Investigation into its Use amongst Primary School Children in the North East of Scotland.* Unpublished University of Aberdeen MLitt thesis.

The Linguistic Atlas of Scotland vols. I and II. 1975 and 1977. Eds J.Y. Mather and H.H. Speitel. London: Croom Helm.

Macafee, Caroline. 1994. *Traditional Dialect in the Modern World: A Glasgow Case Study.* Frankfurt am Main: Peter Lang.

Macafee, Caroline. 1997. Ongoing change in Modern Scots: the social dimension. In Charles Jones (ed) *The Edinburgh History of the Scots Language*, pp. 514-50. Edinburgh: Edinburgh University Press.

Macafee, Caroline. in progress. The 1996 Question Testing: Demographic Aspects.

McGarrity, Briege. 1998. *A Sociolinguistic Study of Attitudes towards and Proficiency in the Doric Dialect in Aberdeen.* Unpublished University of Aberdeen MPhil thesis.

[Máté, Ian]. 1996. *Scots Language. A Report on the Scots Language Research carried out by the General Register Office for Scotland in 1996.* Edinburgh: General Register Office.

Melchers, Gunnel. 1985. 'Knappin', 'Proper English', 'Modified Scottish': some language attitudes in the Shetland Isles. In Manfred Görlach (ed) *Focus on Scotland*, pp. 87-100. Amsterdam: Benjamins.

Menzies, Janet. 1991. An investigation of attitudes to Scots and Glasgow dialect among secondary school pupils, *Scottish Language* 10, pp. 30-46.

Murdoch, Steve. 1995. *Language Politics in Scotland.* Aberdeen: Aberdeen University Scots Leid Quorum.

Murdoch, Steve and R. Gordon. 1995. The Language in Politics Database. Aberdeen: privately published.

Roberts, Lesley. 1993. Ambiguity in the classroom, *Scotland's Languages* launch issue: pp. 7-8.

Wilson, William Morrice. 1993. *Speak o' the North-East.* n.p.: NES Publications.

Addresses
Caroline Macafee
Department of English
University of Aberdeen
Aberdeen
AB24 2UB
Scotland
<c.macafee@abdn.ac.uk>

Briege Mcgarrity
53 Dunamore Road
Cookstwn
Co. Tyrone
BT80 9NT
N. Ireland

APPENDIX

Table 5: Average scores for 30 attitudinal statements

		Score					
	Statements 1-30	1	2	3	4	5	Mean Score
1 Pa	The Doric is a dialect of Scots	1.3	5.3	6.7	52	34.7	4.1
2	The Doric is peculiar to the North-East of Scotland	4.0	9.3	5.3	46.7	34.7	3.9
3	The Doric still in current use in the North-East of Scotland	0	2.7	10.7	62.7	24	4.0
4 Pa	I speak the Doric	13.3	20	26.7	26.7	13.3	3.0
5	I understand the Doric but I do not speak it	12	24	13.3	42.7	8	3.1
6 Pa	I understand the Doric in its written form	2.7	14.7	20	52	10.7	3.5
7 D	The Doric is being successfully passed down to the younger generation	10.7	50.7	21.3	16	1.3	2.5
8 Po	* The Doric is sub-standard English	40	32	21.3	5.3	1.3	4.0
9 D	The influx of incomers from outside of the North-East has adversely affected the survival of the Doric	2.7	20	20	45.3	12	3.4
10	The Doric is more commonly spoken in rural areas	1.3	8.0	2.7	64	24	4.0
11	* The Doric is mostly spoken by uneducated people	33.3	40	16	8	2.7	3.9
12	Women know and use more Doric words and phrases than men	12.0	33.3	49.3	5.3	0	3.9
13 Pa	There should be more entertaining programmes on television and radio in the Doric	0	24	36	34.7	5.3	3.2
14 Pa	There should be more serious, high quality programmes on television and radio in the Doric	1.3	18.7	44	32	4	3.1
15	I participate in activities involving the Doric, e.g. singing, poetry reciting,	29.3	42.7	18.7	5.3	4.0	2.1

	storytelling, party pieces, festivals, attending Doric classes						
16 Po	Schools in the North-East should encourage their pupils to become better acquainted with the Doric and Scots language in general	2.7	9.3	17.3	49.3	21.3	3.7
17 Po	The Doric is old-fashioned	22.7	41.3	18.7	16	1.3	3.7
18	* The Doric is mostly spoken by working-class people	17.3	29.3	17.3	34.7	1.3	3.3
19 D	City people do not speak the traditional Doric	4.0	25.3	20	45.3	5.3	3.2
20	* Speakers of Standard English are more successful in the modern world than dialect speakers	6.7	18.7	17.3	48	9.3	2.7
21 Po	* You should speak Standard English to children	8.0	24	29.3	37.3	1.3	3.0
22 D	Television is killing the Doric	1.3	20	30.7	32.0	16	3.4
23 D	Newspapers are killing the Doric	1.3	32	26.7	37.3	2.7	3.0
24 D	Radio is killing the Doric	1.3	16	20	52	10.7	3.5
25 D	Formal education is killing the Doric	1.3	18.7	30.7	42.7	6.7	3.3
26 D	Lack of contact between grandparents and grandchildren is killing the Doric	12	53.3	14.7	16	4	2.5
27	* I confine the Doric to informal situations, i.e. speaking to family, friends and neighbours	4.0	16	14.7	53.3	12	2.5
28	* I only use Doric words in jest	14.7	41.3	25.3	17.3	1.3	3.5
29 Po Pa	The Doric forms an important part of our North-East identity	1.3	1.3	14.7	37.3	45.3	4.2
30 Po Pa	The people of the North-East should make every effort to preserve the wealth of Doric words and phrases	1.3	4.0	13.3	40	41.3	4.2

Scores run from 1-5 (3 = neutral, poles are (dis)agreement and strong (dis)agreement), adjusted so that the higher the score, the more positive the response. Items marked with an asterisk have been reversed in the scoring system.
D= Defensiveness, Po = Positiveness, Pa = Participation (Principal Components)

Glottals Past and Present: a Study of T-glottalling in Glaswegian

Jane Stuart-Smith

Abstract

Since the turn of the century the realisation of /t/ with a glottal stop (T-glottalling; Wells, 1982: 261) has been noted as a highly-stigmatised pronunciation of Glaswegian vernacular (e.g. Macafee, 1983). To date one sociolinguistic study has examined this linguistic variable in Glaswegian (Macaulay and Trevelyan, 1973; Macaulay, 1977) and found systematic sociolinguistic variation in T-glottalling particularly in certain phonetic contexts. Since Macaulay's Glasgow study, sociolinguistic patterns of T-glottalling (and glottalization) have been studied in increasing detail, particularly in British urban accents.

This paper presents the results of the first study on T-glottalling in Glaswegian since 1973. Informal conversations and read word-lists were digitally recorded from self-selected pairs of same-sex informants, from two distinct age groups, and from two areas of the city, broadly reflecting middle and working class differences. This methodology resulted in high quality recordings of relaxed conversational speech suitable for both auditory and fine-grained acoustic analysis of phonetic variants. Here quantitative and qualitative analysis of an initial auditory transcription of the sociolinguistic patterning of T-glottalling in the two speech styles is given, which reveals clear sociolinguistic stratification and sharp stylistic variation, as might be expected. The present work offers an opportunity for a real-time observation of change in T-glottalling in Glaswegian since 1973, although with the reservation that the two studies show necessary methodological differences. Cautious comparison with Macaulay's findings suggests that there may be evidence for a sound change in progress.

1. *Introduction*

Recent studies have revealed the rampant progress of the use of the glottal stop for /t/ ('T-glottalling'; Wells, 1982: 261) throughout accents of British English. However, while research has considered glottalling in a number British urban accents, it is now almost 25 years since glottalling in Glasgow – possibly the original home of the glottal stop – has been investigated. In this paper I report the first quantitative sociolinguistic study of T-glottalling in Glaswegian since that of Macaulay and Trevelyan in 1973.

I consider first T-glottalling in general and in Glaswegian speech in particular. I then outline the linguistic situation in urban Scotland, which is relevant for any sociolinguistic study of Glaswegian. I give the methodology of the study – data collection and analysis – and then present the results. First the broad quantitative results are given, and then a qualitative analysis, in terms of the phonetic patterning of glottalling in different phonetic environments across speakers. Both quantitative and qualitative differences in T-glottalling are apparent between working class and middle class speakers, and between younger and older speakers. The present data allows an opportunity for a real-time study of T-glottalling between 1973 and 1997. I conclude the paper with a cautious comparison of the results which suggests that there may be real-time evidence for sound change in progress.

1.1. *T-glottalling in accents of English*

Wells (1982: 261) adopts the term 'T-glottalling' to describe the linguistic phenomenon of the replacement of /t/ with a glottal stop. T-glottalling is found in several different phonetic environments. Relevant for this study, note:

> *prepausal*, _#: at the end of a word before a pause, either utterance or turn-final, e.g. ... *and that*
>
> *prevocalic*, _#V: at the end of a word, and before a following word which begins with a vowel, e.g. *a lot of*
>
> *intervocalic*, V_V: between vowels in the same word, e.g. *butter*
>
> *preconsonantal*, _#C: at the end of a word, and before a following word which begins with a consonant, e.g. *but that*

A well-known feature of Cockney English, T-glottalling is becoming increasingly common in RP, and even more so in other social and regional accents of English, particularly urban. Wells' statements have been confirmed for urban accents in a large number of studies including: Norwich – Trudgill (1988) (following 1974); Cardiff – Mees (1987); Milton Keynes – e.g. Kerswill and Williams (1994), and Hull – e.g. Kerswill and Williams (1997); Newcastle – e.g. Docherty *et al* (1997). T-glottalling tends to be associated with male working class speech, although not always. Note the preference for glottalling in females in Cardiff, Tyneside and Hull, all areas where T-glottalling is not a feature of the local vernacular.

T-glottalling is also very common throughout Scotland (Johnston, 1997: 500-01). Macaulay and Trevelyan's Labovian study of Glaswegian in 1973, referred to hereafter in Macaulay's revised and published version of 1977, included glottal stops for /t/ as the only consonantal variable. This was quickly followed by research into the speech of Edinburgh schoolchildren (Romaine, 1975; Reid, 1978). This work, summarised in Macaulay (1991: 31f.), together with other studies (for a review, see Macafee, 1997) has confirmed T-glottalling as particularly prevalent in working-class speech.

1.2. *T-glottalling in Glaswegian*

Glasgow has a special place in any discussion of T-glottalling in English. It is reputed to be the original source of the glottal stop in urban British English (e.g. Macafee, 1997: 528). Whether or not this derivation is correct, it is certainly the case that T-glottalling has been particularly noted as a highly-stigmatised feature of Glasgow vernacular speech since the nineteenth century. A letter of 1892 notes: 'Strangers hurl at us a sort of shibboleth such sentences as 'pass the wa'er bo"le, Mr Pa'erson' (in Macafee, 1994: 27, n. 20). The social stigma attached to T-glottalling is well illustrated in McAllister's tirade against 'this degenerate tendency in modern speech'; McAllister (1963: 71). Such attitudes were still rife when Macaulay carried out his work.

To date there is only one study of T-glottalling in Glaswegian. Macaulay (1977) considered the speech of 48 speakers, 16 10-year old children, 16 15-year old children, and 16 adults (parents). These informants were stratified into four social classes according to occupational category: I, IIa, IIb, III. T-glottalling was examined in two speech styles,[1] a relatively formal interview, and reading some sentences aloud. Sociolinguistic differences were clearest 'before a pause' (= prepausal) and 'before a vowel' (= prevocalic and intervocalic) (Macaulay, 1977: 45).

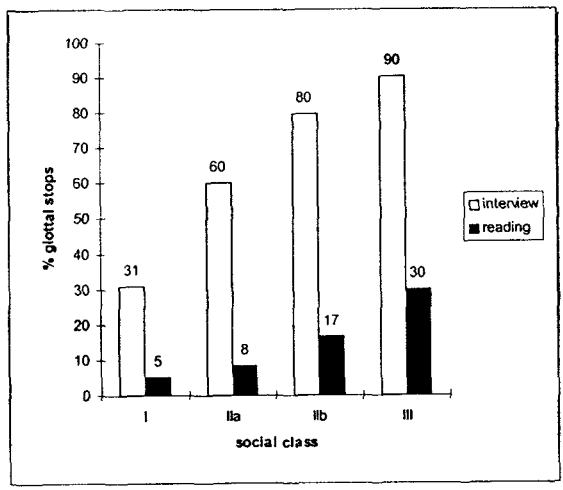


Figure 1

The results revealed clear sociolinguistic and stylistic variation; see Figure 1. Analysis of tokens 'before a vowel' into 'word-medial' (e.g. *water, started*) and 'word-final' (e.g. *that is*) showed that glottal stops were categorically absent in word-medial position for Class I adults. Younger speakers tended to use more glottals than older speakers, at least for Classes I, IIa and IIb. Differences of sex were most obvious in Class IIa and IIb informants, with males producing more glottal stops than females. The stigmatised nature of T-glottalling was confirmed by the results for reading aloud, which generally showed a shift to the prestige variant [t].

1.3. *The linguistic situation in Glasgow*

Macaulay's study, following Labov's early work, assumed that for each linguistic variable there was a continuous dimension of variation; the position of variation along the continuum would correlate with independent variables, such as social class, age, and sex. Sociolinguistic patterns would be apparent in terms of a linear increase (or decrease) in variation. Macaulay's quantitative results for T-glottalling could be used to argue for the existence of a sociolinguistic continuum for this variable in Glaswegian: higher social class informants used less glottal stops and lower class more. However another dimension may also be important when considering sociolinguistic variation in Glaswegian, the linguistic variety spoken by speakers of different class backgrounds in the city.

The linguistic situation in Glasgow is complex; for discussion, see Macafee

(1983, 1994, 1997). Nevertheless, for the purposes of description, it is possible to recognise the existence of two historically distinct language varieties: Scottish Standard English and Glasgow dialect. Like Scotland in general, Glasgow represents a 'traditional' dialect area, where a 'coherent alternative language variety' (Wells, 1982: 4) exists alongside the standard. Within Wells' framework for describing English accents, an accent in a traditional dialect area can be expected to show differences from the standard most clearly in lexical incidence, but also in terms of phonemic system, phonotactic structure, and allophonic realisation.

Scottish Standard English (SSE) – a variety of English similar to standard English English spoken with a Scottish accent (e.g. Aitken, 1979; Abercrombie, 1979) – is spoken by many middle class speakers. Glasgow dialect (or vernacular) is spoken by many working class speakers. Historically, this variety continues a variety of west Central Scots, with influence from Irish English, its own distinctive slang, and the results of continual dialect levelling towards SSE (Macafee, 1994: 26f.). There are generally very negative attitudes towards Glasgow dialect (e.g. Menzies, 1991), which has become identified with a large industrial city in decline and its associated social problems. The relationship between these varieties can be modelled as a linguistic continuum (after Aitken, e.g. 1979). In Glasgow and much of the urban Central Belt, working class speakers tend to style-drift up and down the Glasgow dialect (Scots)/SSE continuum according to sociolinguistic context.

Johnston (1983: 1), after Knowles (1978), points out that the range of variation in standard and vernacular varieties is not always organised along a linear continuum. Historically, the varieties of the middle and working classes in Glasgow are derived from two distinct, yet related sources. Macaulay's finding of variation in T-glottalling according to social background is entirely expected. Working class speech continues urban Scots, which has shown T-glottalling for at least a century. Middle class speech, typically SSE, has no recorded history of T-glottalling beyond what is assumed for other standard varieties of English. However, given that Glasgow is a traditional dialect area, continuing two once distinct linguistic systems, it is not impossible that while appearing quantitatively continuous, T-glottalling may in fact be qualitatively discrete for speakers of working and middle-class backgrounds respectively.

1.4. *Research questions for this study*

Macaulay's work was on Glaswegian speech collected over 25 years ago. This paper considers the results of T-glottalling in speech data recorded from a sample of Glaswegian speakers in 1997. There are three main research questions:

1. What is the sociolinguistic nature of T-glottalling in Glasgow now? Is there still sociolinguistic patterning in terms of class, age and gender, and how is this reflected quantitatively and qualitatively?
2. Are all glottals the same? Is there any evidence to suggest systemic differences in T-glottalling which correlate with social class and the Glasgow dialect/SSE distribution?

3. Can we identify differences in T-glottalling since 1973? Studies across Britain suggest an apparent increase in T-glottalling – do these real-time results support this for Glaswegian?

2. *Methodology*

This study of T-glottalling is variationist, carried out in the Labovian paradigm of quantitative sociolinguistics. While there have been a number of quantitative studies of Scottish English (see Macafee, 1997), there has also been discussion about the relevance of such methodology in the Scottish context (e.g. Macafee, 1994; 1997), particularly with concern about the focus on phonological variables, which are not always the most appropriate to capture sociolinguistic differences in Scottish speech (e.g. Macaulay, 1991). Johnston (1983) examined the validity of the Labovian approach in the Scottish context, with particular reference to stylistic variation. His data from Edinburgh showed some irregular patterns of stylistic variation, which he explained in terms of the salience of the variable for the speaker as of 'high', 'medium' or 'low' consciousness with respect to the dialect. Johnston's conclusion was not to abandon the Labovian approach, but rather to modify the predictions of the model so that data from traditional dialect areas could be accounted for. Note that in Johnston's data T-glottalling emerged as a 'high-consciousness' variable, in other words, a stereotypical feature of the dialect, overtly commented on by informants, which showed regular patterns of variation.

Variationist methodology is used because the primary object of study is a phonetic/phonological variable. Such a methodology is entirely appropriate provided that (a) the linguistic context of Glaswegian, and its potential effect on patterns of variation, is remembered; hence (b) quantitative analysis is used in conjunction with qualitative analysis; and (c) the claims made on the basis of the findings are recognised for what there are, namely the speech behaviour of a very small number of individuals, which can only be indicative, but not necessarily representative, of the city as a

whole.

2.1. *Data collection*

The data from which the results are drawn were collected as part of a wider study into Glaswegian, with particular emphasis on the analysis of phonetic and phonological variation and change in Glaswegian (see also Stuart-Smith, forthcoming). Following the interesting results of the ESRC-funded Phonological Variation and Change project on Newcastle and Derby English, our data were collected with a view to possible acoustic analysis, and a similar methodology was used.

The speech data for the study were collected in summer of 1997. They comprise a set of high-quality digital (DAT) recordings. 32 speakers were recorded in same-sex dyadic conversations of up to 45 minutes. They then read out a word-list. The conversational speech is for the most part relaxed and casual (the least casual speech is to be found amongst the adult middle-class speakers). The social profile of the speakers is given in Table 1.

Speakers	**adolescent (13-14)**	**adult (40 +)**
working class (male)	4	4
working class (female)	4	4
middle class (male)	4	4
middle class (female)	4	4
Criteria		
working class (inner city)	born/raised and resident in area	born/raised and resident in area
middle class (suburban)	born/raised and resident in area	resident in area at least 20 years

Table 1

The speakers were drawn from two areas of the Glaswegian conurbation, representing broadly urban working-class and suburban middle-class areas. These two areas were primarily selected by choosing two representative schools, using external educational and social criteria (percentage of exam passes; percentage of school leavers going on to further education; percentage of children receiving clothing grants). ('middle class' is roughly equivalent to Macaulay's Class I; 'working class' to his Class IIb and III; R. Macaulay, personal communication.) The adults were approached via existing social networks, such as the university, a local women's centre, a teacher training college, and a local church. Both adolescents and adults chose their conversational partner.

2.2. *Data analysis*

The discussion of T-glottalling in this paper is based on the auditory transcription of all potential sites for glottal stops in the conversations and wordlists. The transcription, which was cross-checked for reliability, was carried out by the author, and by two final-year students experienced in practical phonetics. The number of instances of /t/ transcribed in the conversations varied according to speaker (average number: 75). Variants of /t/ were transcribed in three phonetic environments: prepausal (e.g. *but #*); prevocalic (e.g. *a lot of*); and intervocalic (e.g. *water*). The wordlist offered examples for all three positions.

Following Macaulay (1991: 33), preconsonantal tokens were not transcribed, both because of the difficulty of perceiving a clear glottal stop before a following consonant (particularly alveolar) and because this environment was least informative in terms of social variation (Macaulay, 1977: 45). The wordlist recordings allowed relatively narrow transcription of variant pronunciations, the conversations less so, given the fast speech rate of some speakers.

Three main phonetic variants were identified auditorily:

– a released stop, often dental, transcribed [t]. Preglottalization of the type discussed in Wells (1982: 260-61) was observed in the wordlist data, but was more difficult to identify in the conversations. No distinction is made for the purposes of this paper.

– a glottal stop, transcribed [ʔ]. [ʔ] is assumed here to be the complete replacement of the oral articulation by glottal stop, although it is acknowledged that it is very difficult to ascertain with certainty complete lack of articulatory closure in all cases. Preliminary acoustic analysis of the wordlist tokens show that differing acoustic

patterns of glottal stops are present in this data (cf. Docherty and Foulkes, 1995).

– an alveolar tap, transcribed [ɾ]. A tap was found sporadically in a few speakers. As might be expected, it occurred regularly in imitations of American speech, e.g. the phrase *alrighty then* in two working class boys.

Two other variants occurred less often. One, which was difficult to distinguish from the tap, is described here as a 'voiced [t]': [t̬] (possibly [d]). This only occurred in prevocalic position, in the speech of two of the middle class men (in e.g. *got it, about a*) and in one working class girl (in *got it, about it*). Wells (1982: 326) reports a similar variant for London Cockney. The second variant was found in the speech of only one speaker, a middle class girl, and there only in prepausal position (e.g. ... *but#, .. and that#*). She showed complete deletion of the glottal stop, sometimes with breathy-voiced aspiration. This variant may be an idiosyncrasy, but it is noted here given the observation of a similar sounding [h] variant for /t/ in Irish English by Kallen (1998).

3. *T-glottalling in Glaswegian: quantitative results*

The quantitative results for T-glottalling in wordlists and conversations are presented in Table 2. In this and all figures descriptive statistics are presented which show only numbers of glottal stop variants, given as percentages for ease of comparison. The unequal numbers of variants in the wordlist and conversational data mean that these data are less suitable for parametric statistical tests. The non-parametric Mann Whitney U-Test (independent samples) was used to test for statistical significance of differences of data groupings ($p < 0.05$). The results are presented and compared briefly with previous findings; the implications of the results in terms of real-time changes since Macaulay's 1973 study are discussed in section 5 below.

3.1. *Social class and T-glottalling*

Working class speakers showed considerably more glottal stop variants than middle class speakers ($p = 0$) across both speech styles. This result was also found within each style (wordlists $p = 0.001$; conversations $p = 0$). The sharp divide in T-glottalling across social class is entirely in line with Macaulay and Trevelyan's results (1973: 61f.; cf. Figure 1), and also those of Edinburgh schoolchildren in Reid (1978: 166f.), as well as more generally in accents of English across Britain.

	Conversation			WordList		
	MC	**WC**	**Total**	**MC**	**WC**	**Total**
	% (n/N)	% (n/N)	% (n/N)	% (n/N)	% (n/N)	% (n/N)
adult female	45 (35/77)	92 (74/81)	69 (55/79)	0 (0/36)	8 (3/36)	5(2/36)
adult male	43 (34/81)	90 (67/74)	67 (51/78)	3 (1/36)	8 (3/36)	5 (2/36)
all adults	**44 (35/79)**	**91 (71/78)**	**68 (53/79)**	**3 (1/36)**	**8 (3/36)**	**5 (2/36)**
young female	81 (58/72)	99 (80/81)	90 (69/77)	5 (2/36)	84 (30/36)	45 (16/36)
young male	74 (57/78)	89 (45/51)	82 (51/65)	5 (2/36)	62 (22/36)	33 (12/36)
all young	**78 (58/75)**	**94 (62/66)**	**86 (60/70)**	**5 (2/36)**	**73 (26/36)**	**39 (14/36)**
Total	**61 (47/77)**	**93 (67/72)**	**76 (57/75)**	**5 (2/36)**	**41 (16/36)**	**23 (9/36)**

Table 2

3.2. *Gender and T-glottalling*

Perhaps surprisingly no quantitative gender differences were found, either overall, within the two speech styles, or within the social classes (see e.g. the results for the conversations, Table 2). At first sight this might seem at odds with Macaulay's results, where male speakers generally used more glottal stops than female speakers in both speech styles (Macaulay, 1977: 45, 47). However, a closer look at the earlier findings for interview style reveals that while gender differences were pronounced in Class IIa speakers, they were virtually non-existent for Class I and III speakers. Otherwise in 1973, marked gender differences for working class speakers were only found tentatively in the reading style, where male speakers (Class III men, and 15 year old boys from Class III and IIb) used more glottal stops (Macaulay, 1977: 52).

3.3. *Age and T-glottalling*

Overall, younger speakers tended to produce more glottal stops than older speakers ($p = 0.015$), although the pattern for the two speech styles was different for working class and middle class speakers (see Table 2). When reading the wordlists, the middle class children produced about the same, very low, percentage of glottal stops, as the middle class adults. The working class children, however, produced far more glottals than the working class adults ($p = 0.002$). The situation is reversed in the conversations. Here, the working class speakers of both age groups produce a very high percentage of glottal stop variants, but the middle class speakers vary, with younger middle class speakers producing far more glottal stops than older speakers ($p = 0.016$).

Macaulay also found that T-glottalling was more likely in younger speakers than older speakers. In his interviews age differences tended to be more pronounced in higher class speakers than lower class speakers; Class III speakers showed almost no age differences, with all speakers using very high percentages of glottal stops, as here. Reading aloud his results were slightly different. 15 year-olds from Classes III and IIb did show considerably more glottal stops, but this was restricted to boys; Class III men also showed a similarly high degree of T-glottalling.

3.4. *Style and T-glottalling*

A comparison of T-glottalling across all speakers shows a marked difference between casual conversation and reading aloud the wordlist ($p = 0$), with far more glottal stops produced in the conversations. This difference persists within both working and middle class speakers ($p = 0$), although working class speakers show a greater tendency for glottalling when reading the wordlist than middle class speakers ($p = 0.001$). The reason for this lies in the high degree of glottalling by younger working class speakers in the wordlists. The results are generally similar to the tentative findings for style shifting in Macaulay (1977: 52f.); see also Reid (1978), and elsewhere in Britain (e.g. Trudgill, 1988).

4. *T-glottalling in Glaswegian: qualitative analysis*

The quantitative analysis of T-glottalling in the 1997 data revealed sharp differences according to social class and age. These results are in general in line with Macaulay's earlier results, reflecting a sociolinguistic continuum, with increase in social class correlating with decreasing numbers of glottal stops (and vice versa).

However these numbers did not reflect my intuitive feelings about the patterning of T-glottalling. It seemed to me that there was a definite systematicity in the distribution of T-glottalling according to phonetic environment which correlated with age and social class.

In order to investigate this, a qualitative analysis of the patterning of glottal stops according to phonetic environment was carried out for the conversational data. Most studies of T-glottalling in accents of English have assumed that phonetic environment is potentially important, and have considered the relative frequency of glottalling in different environments, sometimes with reference to social class.

The relative frequency of glottal stops according to phonetic environment across all speakers in the 1997 conversational data for Glasgow was compared with results from four other Scottish studies (Romaine, 1975; Reid, 1978; Macaulay, 1977; summarised in Macaulay, 1991: 31f.); see Table 3.

Researcher	**Macaulay**	**Romaine**	**Reid**	**Macaulay**	**Stuart-Smith**
Location	**Glasgow**	**Edinburgh**	**Edinburgh**	**Ayr**	**Glasgow**
most frequent	(#C)	#V	(#C)		
↑	#V	##	##	##	##
↓	##	(#C)	#V	#V	#V
least frequent	V__V	V__V	V__V	V__V	V__V

Table 3

The new Glasgow data agree in showing glottal stops to be least likely in intervocalic position. The ranking of the other positions is most like that found in Edinburgh schoolchildren by Reid, and Ayr adults by Macaulay. (Romaine's results seem unusual; Macaulay, 1991: 32.) The ranking does not, however, agree with that of the earlier Glasgow study. Macaulay's finding that glottals are less likely in prepausal than prevocalic position may perhaps be explained by the formality of the

discourse and by the relative position of the tokens in the overall discourse structure. The spontaneous speech in the current study was generally casual. Macaulay's data were from interviews which elicited 'careful, rather formal speech' (p. 21). Moreover, in Tyneside speech it emerged that glottalling is very rare turn-finally (Docherty et al, 1997). Macaulay's 'prepausal' position includes utterance-final and turn-final tokens. It is just possible that a similar process was operating in Macaulay's sample of Glaswegian, perhaps provoked by the format and formality of the interview.

4.1. *Phonetic patterning: older speakers*

The most illuminating analysis of the 1997 data was found when the phonetic patterning of T-glottalling was examined with respect to social class and age. The average number of glottal stops in each of the three phonetic environments is shown for older working class and middle class speakers in Figure 2a.

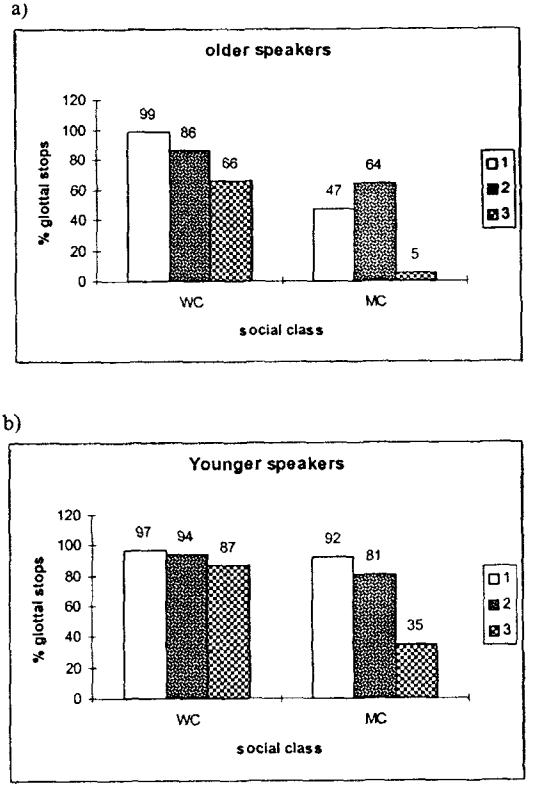


Figure 2

The bar chart clearly shows a difference in patterning according to phonetic environment in working class and middle class speakers. When all instances of exceptions to T-glottalling were examined, a set of descriptive rules to describe T-glottalling emerged for each social class; see Table 4.

Social Class	**Phonetic environment**		
	(1) Prepausal (_#)	**(2) prevocalic (_#V)**	**(3) intervocalic (V_V)**
working class	obligatory	usually ([t] e.g. *put it; at all*)	usually ([t] style-drifting)
middle class	optional	optional	rarely

Table 4

T-glottalling is very usual for working class speakers. Prepausally there was only one exception, which might be described as emphatic. The woman concerned was listening to her conversational partner, but began to sound slightly impatient; eventually she broke in with *'Righ[t]'* with a released stop. Prevocalically the majority of exceptions to glottal stops fell into two categories. One was when /t/ was first in a sequence of two, in e.g. *put it, get it*. Glottalling seems to be avoided in this environment in many accents of English (see e.g. Trudgill, 1974: 174-75), not only in Scottish English (e.g. Macaulay, 1991: 35-36). The stop may also be preserved by a particularly Scottish tendency to resyllabify the sequence into one where the first /t/ begins the next syllable, e.g. /gɛ'tɪt/. This prosodic tendency to make a final consonant begin the following syllable, if this begins in a vowel, accounts for the second group of exceptions, such as *at all* /a'tal/ (cf. Abercrombie, 1979: 68; Macaulay, 1991: 35).

The use of glottal stops was also common intervocalically. In some speakers at

least, [t] in this position was clearly linked to style-shifting. The usual pattern was for speakers to begin using [t], perhaps because of a conscious attempt to use 'better' speech. One woman begins with *be[t]er, la[t]er* but about halfway through the conversation switches to forms with glottals. Instances of [t] are found mostly in past participles in *-ed* to stems ending in /t/, e.g. *shouted*, or in what Reid (1978: 162) calls more 'learned' words, e.g. *hospital, university*. The most striking instance of this style-shifting was in the speech of one of the working class men. He used [t] intervocalically categorically during his conversation until he began to talk about the rats that used to come out at nightfall where he lived as a boy. At this point he switched to glottals – and only used [t] subsequently for expected exceptions (*voted, nominated*).

For middle class speakers [t] is the norm in intervocalic position, even in quite casual speech. While some of the exceptions to glottalling in prevocalic position matched those of working class speakers, in e.g. sequences such as *put it, but on*, overall the use of [t] was much more difficult to explain. It would seem that glottal stops could be used prepausally and prevocalically, but [t] was also an equal possibility. Note that for these speakers alone prepausal glottal stops were less frequent than prevocalic ones, possibly due to a higher incidence of released stops turn-finally (though this is yet to be investigated).

There is an important point to note about style-shifting and T-glottalling with respect to the older speakers. When a working class speaker shifted to a more formal style, [t] was only inserted intervocalically, but not prepausally or prevocalically. In particular, the glottal stop was always maintained in prepausal position. A view of T-glottalling for working class and middle class speakers as a linguistic variable on the same sociolinguistic continuum might predict that working class speakers would 'improve' their speech by adding [t] across the board. However, this does not occur. [t] is only added intervocalically – in the most stigmatized position. This means that working class speakers show a different patterning of glottal stops even when style-shifting up, since the resulting pattern does not reflect that of a middle class speaker.

4.2. *Phonetic patterning: Younger speakers*

A similar analysis was carried out for younger speakers, and the results are presented in Figure 2b, and schematically in Table 5.

The pattern of T-glottalling in working class children seems to reflect an extreme form of the pattern found in working class adults. Here T-glottalling can only be described as obligatory in all positions. It is possible to explain every exception to glottalling (i.e. instance of [t]) in much the same way as for adult speech: [t] occurs as

the result of emphasis, prosody, and style-shifting.

Social Class	Phonetic environment		
	(1) Prepausal (_#)	**(2) prevocalic (_#V)**	**(3) intervocalic (V_V)**
working class	obligatory	obligatory ([t] emphatic)	obligatory ([t] style-drifting)
middle class	very likely	likely	rarely

Table 5

In prepausal position there were only two instances of [t], both emphatic. The first occurred when one of the boys wanted to change topic:

> *R: Shut up you wee shite. Just cos you're right into Mrs X's [= teacher] gear.*
> *L: shh. (pause) Righ[t] Who do you think is the nicest teacher in this school? ...*

The second was in the second of the boys' interviews, when one of the pair broke the clip-on microphone from its clip, with an emphatic *'Shi[t]'*. (This was followed by his friend reassuring him: *'Just say it fell off and broke'*!) Prevocalically, exceptions to glottalling were rare, but again explicable, either occurring (often emphatically) in 'two /t/' sequences, e.g. *at it, doubt it*, or in emphatic sequences where resyllabification might be expected, e.g. *Shu[t] up, cut i[t] off!* (horrified suspicion of what another girl had done to her pet gerbil's tail – in fact she had put it in antiseptic). Use of [t] intervocalically was limited to either *-ed* past participles, e.g. *shouted, related*, or to obvious instances of style-shifting. A bottle of spring water was left in

the room as refreshment, and reading of the label resulted in *Strathmore Carbona[t]ed Spring wa[t]er.* A polite admonishment from one boy to another was *You'd be[t]er behave yourself.* Once the microphone had been separated from its clip, the boys became a pair of stage commentators for the listener, e.g. *and now we're going to fuck abou[t] with the compu[t]er.* This speech style involved an extended intonation and more frequent use of [t] intervocalically. This is reminiscent of the mimicry reported in Edinburgh boys by Reid (1978: 165), when the number of glottal stops reduced dramatically during 'commentator' speech.

Middle class children also showed a similar pattern to that found in middle class adults. Even in casual conversation intervocalic glottalling was rare. Prevocalically and prepausally glottals were much more common, but released stops were also possible, and could not be explained away as easily as for the working class children. It is difficult to assume that T-glottalling is obligatory for middle class children in the same way as for working class children. The potential difference between T-glottalling in the two groups is highlighted by a subgroup of three middle class children (two boys and one girl) who showed high percentages of glottals, equalling working class scores. They also seemed to show a similar patterning to the working class glottalling, although with the difference that the exceptional instances of [t] could not all be explained. This finding is somewhat similar to that found in the style-shifting of working class adults. When young middle class speakers use a lot of glottal stops, perhaps to 'lower' their speech, they do not just add glottals across the board, they seem to try to move towards the working class pattern.

4.3. *Phonetic patterning: Summary*

This qualitative analysis of T-glottalling suggests that, in answer to the second research question posed in 1.4, not all glottals are the same. Linguistic attempts to shift socially up and down do not involve simply increasing or decreasing the number of glottal stops used, as might be the impression from former studies of T-glottalling in Scottish English. Rather, the results here seem to suggest that the allophonic patterning of T-glottalling for working class and middle class speakers is systematically different for the two groups. Movement up or down sociolinguistically seems to involve a systematic shift: middle class children moving 'down' approximate the working class pattern, but are not entirely successful, continuing traces of their middle class pattern in the unexceptional use of [t]; working class adults trying to move 'up' approximate their middle class peers in intervocalic position, but again retain their working class pattern in the categorical use of glottals prepausally.

Thus even the humble glottal stop may continue a systematic patterning which

reflects the Scots or English heritage of the speaker. Working class Glasgow dialect (Scots) speakers show a systematic allophonic pattern in T-glottalling, which is being maintained by younger speakers. The SSE-based system of middle class speakers shows a different patterning, which is stable for adults, but which is moving towards the working class, Scots-based, system for some younger speakers. The effects of increased glottalling in non-standard English English varieties on Scottish English is difficult to ascertain, although not to be completely discounted, given the recent findings of e.g. /f/ for /θ/ in these same (working class) children (Stuart-Smith, forthcoming). It would be interesting to see whether further research on English English T-glottalling also revealed qualitative as well as quantitative difference in glottals.

These findings seem to me to confirm the Labovian approach in the Scottish context. If quantitative analysis is informed by qualitative discussion, even phonetic and phonological variables can be useful in characterising salient, if subtle, sociolinguistic differences. Perhaps it is not which variables are chosen as the subject of analysis, but how they are analysed which may be important.

5. *Glottals past and present: A real-time change in Glaswegian?*

The existence of the two studies into T-glottalling from 1973 and 1997 would seem to present a good opportunity for the observation of any potential real-time changes in this linguistic feature. However, such a comparison must be made with care. The difficulty of real-time studies is discussed, among others, by Trudgill (1988); see also Labov (1994: 72f.). Obviously the clearest obstacle in comparing the Glaswegian data directly is the different methodology adopted for each project. These are different in a number of respects including (in no particular order): background and gender of fieldworker; numbers of participants; regional location of participants (R. Macaulay, personal communication); assignment of social class; reading stimuli; spontaneous speech situation (interview v. conversation alone); number of variants transcribed. The 1997 data collection was not designed as a restudy of 1973, but in order to obtain high quality speech data appropriate for the analysis (also acoustic) of phonetic and phonological variation.

The differences between these two studies make direct quantitative comparison inadvisable. However, cautious indirect comparison is possible. In terms of social class, Macaulay's Class I can be compared with 'middle class', and Class IIb and III with 'working class'. The adolescents of this study were aged 13-14 years, and so are

roughly comparable with Macaulay's 15 year-old group. Accordingly the figures from both studies were calculated and compared visually; see Figure 3.

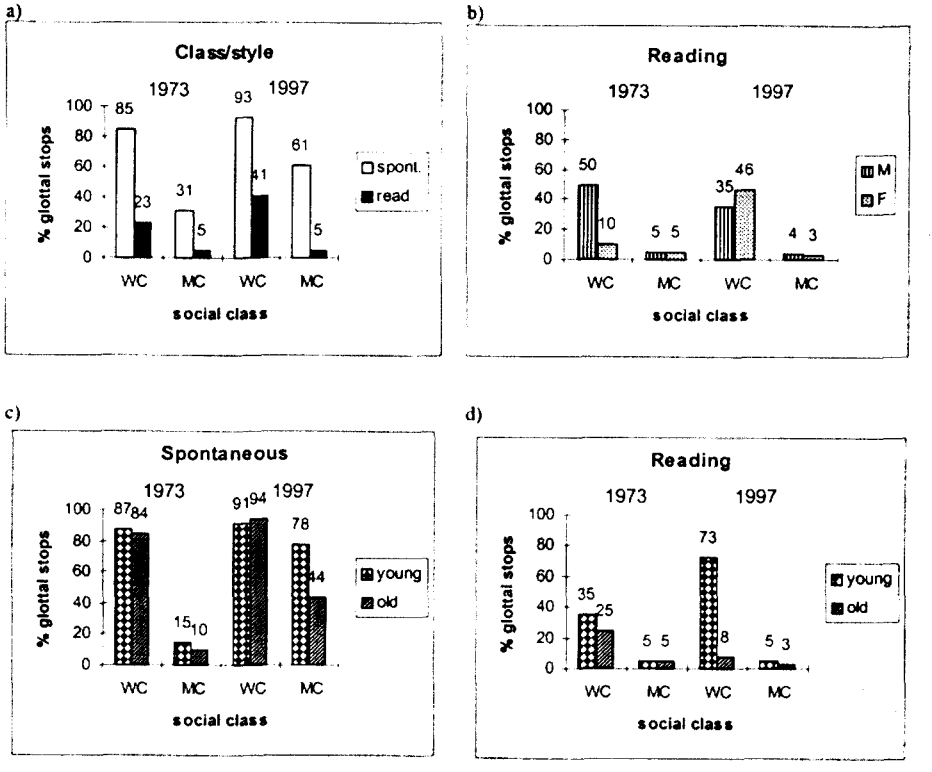


Figure 3

In both 1973 and 1997 far more glottal stops are produced in spontaneous speech than read speech across both social classes (Figure 3a). The only difference is for middle class speakers, who show more glottals in spontaneous speech in 1997 than in 1973. This may be an artefact of the difference between the type of spontaneous speech recorded - interviews in 1973, and unobserved conservations in 1997; note that the working class too show slightly more glottal stops in 1997. Alternatively this may reflect a true change of middle class speakers using more glottals. Such an interpretation would be consistent with Wells' observations for RP (Wells, 1982: 261).

Gender differences overall were not found in the 1997 data, nor in the

spontaneous speech in 1973 (for the comparable classes). There is however a difference in the read speech; see Figure 3b. In 1973 working class male speakers used more glottal stops than female speakers. In 1997 there is no longer a difference (statistically) between male and female working class speech. This is partly caused by working class men in 1997 using far fewer glottals than they did in 1973, but also by working class girls using far more (old WC males 1973: 40% ~ 1997: 8%; young WC females 1973: 10% ~ 1997: 84%). It seems unlikely that working class men have increased their use of [t] since 1973. Perhaps an aspect of the data collection is to blame, such as the gender of the fieldworker, male in 1973 but female in 1997. A young female fieldworker might also account for the increased glottalling in young working class girls. Alternatively, this may reflect an actual change in progress. If so, it would show a female preference for T-glottalling in an area where glottals are a well-established feature of the vernacular.

Perhaps the most striking differences between the 1973 and 1997 data occur when age and class are considered; see Figures 3c and d. In spontaneous speech in 1973 both age groups show a similar number of glottal stops across the social classes. In 1997 there seems to be a marked increase in the number of glottal stops in young middle class speakers. While this increase could be caused by methodological differences in the two studies, it is more difficult to isolate a particular reason. This may reflect a real increase in T-glottalling in young middle class speakers, but only in spontaneous speech. When reading aloud, young middle class speakers show similarly low numbers of glottals to adults, suggesting that for these speakers T-glottalling continues to be a strongly stigmatized.

However, this may no longer be the case for young working class speakers, who show a marked increase in T-glottalling in read speech in 1997; Figure 3d. However, the majority of the children did shift with respect to another stereotypical Glasgow dialect feature, the vowel of *house*, typically /u/ in the vernacular but /ʌu/ in SSE. It is possible that vernacular /u/ could only be elicited through an overt spelling such as *hoose*. Alternatively, the presence of /ʌu/ might suggest that reading the wordlist was exerting some form of pressure to shift linguistically, but in such a way that T-glottalling is no longer excluded as stigmatized. This finding is similar to Trudgill's results for Norwich, where younger speakers in 1983 showed a dramatic increase in glottalling in formal styles (Trudgill, 1988: 44-45).

6. *Concluding Remarks*

T-glottalling continues to be a vigorous feature of Glaswegian speech. The quantitative analysis of the socially-stratified speech data collected in 1997 reveals differences in T-glottalling according to social class and age, but not gender, with working class and younger speakers using more glottal stops. Qualitative analysis suggests that numerical differences in glottalling between social classes are not simply incremental along a continuum, but reflect two distinct types of allophonic patterning, possibly reflecting the Scots/Scottish Standard English linguistic heritage of working and middle class speakers respectively. The apparent time-change results suggest an increase in glottalling in younger speakers. Real-time change can be glimpsed by a necessarily cautious comparison of the 1997 results with those of Macaulay and Trevelyan's 1973 study. It appears that certain changes may be in progress, all in terms of an increase in T-glottalling: middle class speakers in general seem to be using more glottal stops; young working class girls now equal glottalling of their male peers. Perhaps the apparent time-change does reflect actual change: the real-time comparison shows younger speakers using more glottals, although differently according to class. Middle class children are T-glottalling more, but only in spontaneous speech: read speech still inhibits glottals. Not so for young working class speakers, who show a high degree of glottal stops even in read speech, perhaps reflecting that for these speakers at least the glottal stop is no longer the object of stigma it once used to be.

NOTES

[1] I use the term 'style' with reservation here, as indeed does Macaulay himself (R. Macaulay: personal communication; forthcoming). While I use 'style' in this paper to refer to differences in conversational and read speech (following e.g. Trudgill 1974), I acknowledge that these differences do not simply result from 'formality', but may also be due to other factors, such as differences in linguistic activity or the influence of the written standard.

ACKNOWLEDGEMENTS

I am very grateful to Claire Timmins who acted as the fieldworker for the data collection, and to Barbara McGinley and Eleanor Lawson who helped transcribe the wordlist data. I would also like to thank Catherine Macafee, Ronald Macaulay, and an anonymous reviewer for helpful and constructive comments; I of course am responsible for the views expressed here. I am grateful to the Faculty of Arts Research Committee, University of Glasgow, for financial assistance towards this research.

REFERENCES

Abercrombie, D. 1979. The accents of standard English in Scotland. In A.J. Aitken and T. McArthur (eds) *The Languages of Scotland.* pp 54-70. Edinburgh: Chambers.

Aitken, A.J. 1979. Scottish speech: A historical view with special reference to the Standard English of Scotland. In A.J. Aitken and T. McArthur (eds) *The Languages of Scotland.* pp. 85-118. Edinburgh: Chambers.

Docherty, G. and Foulkes, P. 1995. Acoustic profiling of glottal and glottalised variants of English stops. In *Proceedings of ICPhS 95 Stockholm.* pp. 350-53.

Docherty, G., Foulkes, P., Milroy, J., Milroy, L., and Walshaw, D. 1997. Descriptive adequacy in phonology: a variationist perspective. *Journal of Linguistics.* 33 (2): 275-310.

Johnston, P. 1983. Irregular style variation patterns in Edinburgh speech. *Scottish Language.* 2: 1-19.

Johnston, P. 1997. Regional Variation. In Jones (ed) (1997). pp. 433-513.

Jones, C. (ed). 1997. *The Edinburgh History of the Scots Language.* Edinburgh: Edinburgh University Press.

Kallen, J. 1998. Lenition of /t/ in Irish English: What is converging with what? Paper presented at conference on the convergence and divergence of dialects in a changing Europe, Reading, September.

Kerswill, P. and Williams, A. 1994. *A New Dialect in a New City: Children's and Adults' Speech in Milton Keynes*, Final Report to the ESRC on grant no. R000232376.

Kerswill, P. and Williams, A. 1997. Investigating social and linguistic identity in three British schools. In U.B. Kotsinas, A.B. Stenstrom, and A.M. Karlsson (eds) *Ungdomssprak i Norden.* pp 159-76. Stockholm: University of Stockholm Press.

Knowles, G. 1978. The nature of phonological variables in Scouse. In P. Trudgill (ed)

Sociolinguistic Patterns in British English, pp. 80-90. London: Edward Arnold.

Labov, W. 1994. *Principles of Linguistic Change. Vol. 1: Internal Factors.* Oxford: Blackwell.

McAllister, A.H. 1963. *A Year's Course in Speech Training.* 9th edn. London: University of London Press.

Macafee, C. 1983. *Varieties of English around the World: Glasgow.* Amsterdam: Benjamin.

Macafee, C. 1994. *Traditional Dialect in the Modern World: A Glasgow Case Study.* Frankfurt: Peter Lang.

Macafee, C. 1997. Ongoing change in modern Scots: the social dimension. In Jones (ed.). pp. 514-48.

Macaulay, R.K.S. and Trevelyan, G. D. 1973. *Language, Education and Employment in Glasgow.* Final report to the SSRC.

Macaulay, R.K.S. 1977. *Language, Social Class and Education: A Glasgow Study.* Edinburgh: Edinburgh University Press.

Macaulay, R.K.S. 1991. *Locating Dialect in Discourse: The Language of Honest Men and Bonnie Lassies in Ayr.* New York: Oxford University Press.

Macaulay, R. K. S. forthcoming. Is sociolinguistics lacking in style?

Mees, I. 1987. Glottal stop as a prestigious feature in Cardiff English. *English Worldwide.* 8: 25-39.

Menzies, J. 1991. An investigation of attitudes to Scots and Glasgow dialect among secondary school pupils. *Scottish Language.* 10: 30-46

Reid, E. 1978. Social and stylistic variation in the speech of children: some evidence from Edinburgh. In Trudgill, P. (ed) *Sociolinguistic Patterns in British English*,. pp. 158-71. London: Edward Arnold.

Romaine, S. 1975. Linguistic variability in the speech of some Edinburgh schoolchildren. Unpublished MLitt. Thesis. University of Edinburgh.

Stuart-Smith, J. 1999. Glasgow: accent and voice quality. In P. Foulkes and G. Docherty (eds) *Urban Voices: variation and change in British accents* 203-22. London: Edward Arnold.

Trudgill, P. 1974. *The Social Differentiation of English in Norwich.* Cambridge: Cambridge University Press.

Trudgill, P. 1988. Norwich revisited: recent linguistic changes in an English urban dialect. *English Worldwide.* 9 (1): 33-49.

Wells, J.C. 1982. *Accents of English.* 3 vols, Cambridge: Cambridge University Press.

Address
Jane Stuart-Smith
Department of English Language
University of Glasgow
Glasgow G12 8QQ
Scotland
<j.stuart-smith@englang.arts.gla.ac.uk>

Dialect Recognition and Speech Community Focusing in New and Old Towns in England: the Effects of Dialect Levelling, Demography and Social Networks[1]

Paul Kerswill and Ann Williams

Abstract

In Britain, the past few decades have seen the erosion of regional dialects and the spread of levelled, non-standard varieties centring on larger conurbations. This process of dialect levelling has been attributed to increased social and geographical mobility in post-war Britain and has been shown to occur in areas where there is a high level of dialect contact. The study we report here aimed, first, to investigate whether levelling is more advanced in highly mobile populations such as new towns, where the speech community is new and diffuse, than in stable, focused speech communities; and second, by testing participants' ability to recognise their own varieties, to account for the social psychological mechanisms behind dialect levelling. In this article, we discuss the relationship between the dialect perception data and the linguistic results. Three British towns of similar size, but with different demographic profiles were chosen: Milton Keynes (a new town) and Reading in the south, and Hull in the north. The linguistic analysis shows that the accents of Milton Keynes and Reading are converging by a process of levelling: older regional variants are rejected and either standard or new variants are being adopted – changes which reflect abrupt social discontinuity in Milton Keynes and rapid, but less dramatic social change in Reading. In Hull, where we find the kind of close-knit networks that inhibit linguistic change, the adolescents retained traditional features. The dialect recognition results parallel the linguistic data in that overall recognition rates were low for the two levelled, southern accents, but high for Hull. This leads us to claim that accurate dialect recognition is an integral part of focusing in a stable speech community. Conversely, where there is rap[id linguistic change, giving rise to greater diffuseness, dialect recognition is less predictable. Specifically, in both Reading and Milton Keynes, we find that the young

judges do not recognise elderly local speakers, but identify their age peers more accurately. This suggests a discontinuity in the speech communities across three generations, a find which is expected in the new town of Milton Keynes, but is more surprising in Reading, where there is a higher degree of social continuity. We conclude that dialect recognition might be considered a measure of the rapidity of change within a speech community.

1. *Introduction*

1.1 *Folk linguistics and language variation and change*

In a recent article, Preston has made a plea for the systematic study of non-linguists' opinions about language varieties to complement professional linguists' insights about 'scientifically discovered aspects of language structure and use' (1996a: 72). His concern in that article is partly with implications for public policy, though, as we shall argue, such folklinguistic opinions also bear strongly on issues of language variation and change. Preston's research on Americans' perceptions of United States English dialect areas has enabled him to present both quantitative and ethnographic evidence of a range of social psychological and linguistic factors which influence folklinguistic awareness. If we assume that people's awareness of a particular linguistic feature is related to its sociolinguistic patterning in a speech community, then Preston's approach is of obvious value to sociolinguists investigating language variation. Indeed, the 'modes of awareness' he hypothesises (Availability, Accuracy, Detail and Control) seem to us to have the potential significantly to extend and deepen Trudgill's notion of 'salience' (Trudgill 1986: 11).

This article centres on just one means of collecting folklinguistic data: non-linguists' identification of regional and/or sociolectal varieties presented on a test tape. Surprisingly, this procedure has been relatively neglected despite Preston's plea for it (1989: 3), even though, as we hope to show, the evidence it provides is directly relevant to the understanding of language variation.

1.2. *Dialect recognition and the attribution of speakers to one's own community*

A brief review of three dialect recognition studies will illustrate the range of possible links with language variation.[2] Preston (1996b: 320-29) considered the extent to which non-linguist respondents can differentiate between voices with regional

phonology, but no lexical or grammatical cues as to their origin. Preston asked non-linguist Americans to allocate nine Eastern United States varieties to their correct positions on a north-south scale. The overall result was that, with the judgements pooled, listeners placed virtually all the voices correctly on this scale – even though in many cases the difference in the judges' average ranking of adjacent locations was small. One particular pattern emerged which sheds light on the social psychology of variation. There were clear differences between a Michigan (northern) and an Indiana (central) group of judges: surprisingly, the Michigan judges did not differentiate the northern voices as clearly as did the Indiana judges, a finding which, Preston suggests, reflects the 'unity of that territory' as displayed in the hand-drawn dialect area maps produced by subjects from this region (1996b: 324). However, Preston does not explicitly consider the cause of the Michiganers' apparently inferior discrimination abilities. It seems to us that the result may actually be a consequence of the Michiganers' enthusiastic identification with a broad northern area, differentiated clearly from the South, such that the placement task is somehow downgraded when individual voices are perceived as belonging to in-group members, and thus deemed socially attractive.

That this is a possibility is suggested by findings from the second study we consider: this is Williams, Garrett and Coupland's (1999) exploration of Welsh teenagers' recognition of and attitudes to regional accents of Welsh English as spoken by teenagers of the same age as themselves. Two voices from each of six locations were played to judges from the same six locations. A not unsurprising result was that the teenagers were generally more successful at recognising voices from their own location than from elsewhere, and that overall recognition scores for individual voices were fairly low (ranging from 21% to 42%). However, individual voices varied greatly in terms of whether judges from the same locations as the voices could recognise them: the highest rate was 100%, the lowest 13.8%. Equally surprising was the fact that there were often considerable differences in recognition rates between two voices from the same location. Williams et al. find that, in general, it is the voices whose owners are perceived as 'likeable' and 'a good laugh', possibly due to the content of the narratives, which tend to be 'claimed' as belonging to the judge's own regional group regardless of the actual provenance of the voice. They point to this as indicating the complexity of the dialect identification task; it is an example of a social psychological factor mediating between the ostensibly stimulus-based task (here, based on segmental and suprasegmental accent features) and the response. This type of explanation seems to throw additional light on Preston's American findings, as we saw above – though we suggested that perceived in-group membership *in itself* had the power to make a

voice 'attractive', perhaps as an effect separate from other possible (paralinguistic) features of the voice which, if Williams et al.'s conclusions are correct, might be relevant.

The dialect recognition task we report in this article likewise presents two young voices from each research site, again with adolescent listeners. Additionally, we included older voices from the same locations. As we shall see, we too found differences in the recognition rates for voices from same location. However, we adopt a different, but complementary angle in the interpretation of these differences: we examine the particular accent features of the voices, and those voices' relation to the processes of dialect levelling and speech community focusing.

The outcomes of the American and Welsh studies also indicate that the dialect identification technique can help in the interpretation of variation and change in speech communities. The key notion, already alluded to above, is *focusing* (Le Page 1978): a speech community is said to be focused if there is relatively little variation and if the variation that remains is clearly patterned. Such communities are socially stable, and linguistic change is likely go be slow. 'Diffuse' communities, on the other hand, do not have such clear norms, reasons for this usually lying in a more volatile social structure. (See Kerswill 1993 for an example of a diffuse in-migrant speech community interacting with a focused urban speech community.) The link between focusing and perceptual dialectology is this: in a focused community, one would expect members to be more successful at recognising other members' language varieties than the case would be in diffuse communities.

It was as an attempt to test this hypothesis that the third study was conceived. Kerswill's investigation of dialect perception in the Bergen region of Norway aimed to investigate the focusing of the Bergen speech community by testing native Bergeners' sensitivity both to very small and to somewhat greater deviations from canonical Bergen speech (Kerswill 1985, briefly reported in Kerswill 1993). The study used a test tape containing the voices of rural migrants from the immediate hinterland who had accommodated in varying degrees to the Bergen urban dialect, along with a native Bergen speaker. The results showed that not only could the Bergen judges tell the difference (to a statistically significant degree) between the most 'accommodated' rural speaker and the genuine Bergen voice, but they could also rank the remaining speakers in terms of their degree of 'ruralness', the ranking being identical to that established by applying a dialect index based on a range of morpholexical features. It was suggested that the Bergen speech community is exceptionally focused in that the phonetic criteria for 'membership' are extremely subtle and yet salient, though they could not be picked up by a careful phonetic comparison of the Bergener and the apparently fully

'accommodated' rural speaker on the tape.

The Bergen study did not provide any comparative data to evaluate either the method or the conclusions reached. The study we will be reporting partially provides this comparison, as well as taking account of the findings of the American and Welsh studies.

1.3 *Mediating factors affecting dialect recognition*

On the face of it, a dialect recognition task is simply a test of sensitivity to linguistic (usually restricted to phonetic) differences; this is true whether or not the task involves 'own-community' or 'other-community' identification (as with the three studies we have just reported, ours tests both of these). Yet, as we have seen, the recognition process will be mediated by a number of other factors, including:

1. The life experience of the judges (relating especially to whether their social networks are close-knit or open, and to whether they as individuals have been socially and geographically mobile).
2. The absolute linguistic differences between the varieties being offered for recognition, and the differences between these and other varieties known to the judges. This factor will itself be affected by the *salience* of the features differentiating them or, to use Preston's (1996a) terminology, their *availability* for perception and comment, and the *accuracy* and *detail* with which they are perceived.
3. The sociolinguistic maturity of the judges (relating mainly but not exclusively to age – see Kerswill 1996, Williams et al. 1999: 370-71).
4. The subjectively perceived social attractiveness of the speaker due to paralinguistic factors (voice quality, tempo, pitch range, content) which one might presume to be unrelated to the identification of the varieties.

Ideally, a study of dialect recognition should either test or control for these factors. The present study tests the first (the life experience of the judges) by systematically varying judges by social class/social network and town. It also tests the second (the effect of linguistic differences), but in a qualitative way. It explicitly controls for the third (sociolinguistic maturity) by using judges of the same age. The fourth (the perceived social attractiveness of the voices due to paralinguistic features) can be be approached by the use of a questionnaire, as it was in the Welsh study, though this does not answer the question of *which* features actually influence the perception.

Alternatively, the Matched Guise technique (Lambert, Hodgson & Fillenbaum 1960) may be used. In the present study, the latter was rejected because we felt it essential to preserve the naturalness of the material presented. Instead, we focused on the phonetic features contained in the authentic extracts which the subjects heard.

2. *Dialect recognition in urban England*

2.1 *The context of the dialect recognition task*

2.1.1 The Dialect Levelling project

In Britain, as in other European countries, there has been a steady trend towards the loss of regional dialects, resulting in new, compromise varieties combining some of the original dialect/accent features, some new forms, and some forms adopted from a relevant standard. These varieties have a geographical spread that is greater than the old regional dialects, and in a few cases they function as regional standards rather than completely ousting the old dialects. In all cases, they are the result of *dialect levelling* – the reduction in regionally marked forms and the adoption of regionally more widespread features.[3]

It was in order to explore the linguistic and social mechanisms behind dialect levelling that the project *The role of adolescents in dialect levelling*[4] was set up. An important facet of the project was the exploration of subjective factors affecting people whose dialect is involved in levelling, part of this investigation being the dialect recognition task reported here.

Before discussing the task, we place it in the context of the wider study. The project had the following premises:

1. In areas of high population movement, there may be rapid changes in dialect and accent features, including levelling. The speech community is *diffuse*.
2. Membership of a close-knit, stable social network with strong local ties leads to linguistic conformity (i.e. not 'stepping out of line'). This inhibits change, including that manifesting as levelling. The speech community is *focused*.
3. The distance of a town from a national metropolis (in this case London) is inversely proportional to the degree to which the town adopts linguistic features from that metropolis (the gravity model: see Trudgill 1983).
4. Language change is most visible through the comparison of teenage language with older adults' speech and with the speech of younger children.

Additionally, the project built on Milroy & Milroy's (1992) contention that, in urban societies, there are clear differences in the social networks contracted by people of different social classes. The more privileged middle classes tend to have ties outside their immediate neighbourhoods and families: they are often geographically mobile, and are likely to have been socially mobile, too. Working-class people, especially in times of adverse economic circumstances, tend to have closer ties with family, neighbours and work colleagues. The difference is reflected in speech, in that middle-class people use less localised and more standardised varieties than do working-class people. Conversely, working-class speech is more likely than middle-class speech to symbolise a local identity.

Despite these differences, both middle-class and working-class speech undergoes change, including levelling, and the project aimed to document this. We also tested the hypothesis that geographical mobility and open networks (both held to be middle-class traits: Milroy & Milroy 1992) affect the speed and direction of change (towards forms which are both more standard and less localised) *independently* of social class. We did this by investigating (1) both middle-class and working-class teenagers, and (2) towns which differ greatly in terms of the *overall* degree of mobility of their populations.

Two of the towns chosen, Reading and Milton Keynes, are about the same distance from London (c. 70 kms) and have similar population figures, economies and commuting patterns. (See Figure 1 for a location map.) Both are prosperous, with low unemployment. Crucially, they differ in their recent demographic histories. Parts of Reading have large, stable populations with strong local ties, while the town as a whole has seen considerable in-migration. By contrast, Milton Keynes was designated only in 1967, and since then has seen a massive, and continuing, population increase due to in-migration, mainly from the south-east of England. The third town, Hull (official name: Kingston-upon-Hull), contrasts with Reading and Milton Keynes in its distance from London (340 kms.), in its geographical isolation on the north-east coast in East Yorkshire, and in its declining industries reflected both in high unemployment and falling population figures. Even more than in Reading, a large proportion of its inhabitants have strong local ties. (See Table 1 for a demographic summary of the three towns.) We expect, then, that dialect levelling (based on the rise of London and general south-eastern features) will be further advanced in Milton Keynes than in Reading, that changes in Hull will be less rapid and will follow a relatively independent course, and that the use of levelled and standard features will be greater among the middle-class teenagers in all three towns. Some of the results of the project are reported in Kerswill & Williams (1997, 1999) and Williams & Kerswill (1999).

Figure 1: Map showing location of places mentioned in this article

	New Town?	Close to London?	Population 1991	Population change 1981-91	% skilled manual+un-skilled*	Unemploy-ment*
HULL	no	no (340 kms)	254,000	-8.7%	63.0%	12.02%
READING	no	yes (60 kms)	129,000 (not counting Wokingham)	-5.1% (increase with Wokingham added)	42.8%	4.25%
MILTON KEYNES	yes, founded 1967 (pop. 44,000)	yes (70 kms)	176,000	+39.2%	44.9%	4.75%

*1991 Census, taken from *BBC Constituency Guide 1997*

Table 1: Summary of demographic characteristics of Reading, Milton Keynes and Hull

2.1.2 Dialect recognition and dialect levelling: hypotheses

As suggested earlier, it is likely that members of focused speech communities will recognise each other on the basis of voice samples more easily than people whose communities are diffuse. This expectation can be extended to dialect levelling: speakers whose communities are undergoing rapid levelling will find this kind of 'own-community recognition' more problematic than speakers from communities not subject to levelling. However, as we have seen, recognition is mediated by several other factors, both social psychological and linguistic (see 2.1). Thus, we arrive at the following hypotheses:

Hypothesis 1: Own-community recognition will be better among people with strong local ties (working-class judges in Reading and Hull will be more successful than middle class groups in the same towns, but working-class Milton Keynes judges will not have the same advantage).

Hypothesis 2: Own-community recognition will be better in towns with relatively little mobility (Hull > Reading > Milton Keynes).

Hypothesis 3: Own-community recognition of an accent with strongly localised phonetic features will be better than that of accents without such distinctive features. In the present study, this potentially confounds, or at least interacts with Hypothesis 2.

Hypothesis 4: Recognition of an accent from outside the judge's own community depends on how familiar that accent is to the listener (familiarity being a function of a number of disparate factors, especially personal contacts and the broadcast media). We refer to this as the *familiarity hypothesis.*

Hypothesis 5: To judge from the experience of Williams et al. (1999), different voices from the same town (even if there is no age difference between the speakers) will not be recognised at the same rate by members of that speech community. Reasons for this are likely to be complex: in this article, we focus mainly on the linguistic features of voices.

Hypothesis 6: We extend Hypothesis 5 by further hypothesising that recognition rates will be influenced by the perceived age of the speakers: own-community speakers close to the teenage judges' age will be more successfully identified than speakers who are significantly older. This arises mainly from the assumption that a judge is more likely to recognise an accent similar to his or her own than an accent that is different. Linguistic differences within a community can be a function of age, resulting from rapid language change, and these can lead to older voices not being recognised by younger judges. However, other things (such as phonetic

features) being equal, an accent is likely to be 'claimed' if it is perceived that the speaker is of a similar age to the listener, and hence potentially socially attractive.

2.1.3 Subjects, materials, methods

The dialect recognition task was conducted as one component of a language-related discussion which formed part of the fieldwork for the project. The subjects attended one of six schools, two in each of Hull, Milton Keynes and Reading. The selection of the schools was made according to the criteria for the main project: since we wanted to investigate dialect levelling among people with either locally-based, close-knit networks or more open, less local networks, we selected schools whose pupils could be expected to conform to one of these two broad categories. In Hull and Reading, this meant targeting schools in mainly low-income districts with high continuity of population and schools in middle-income districts with a high proportion of incomers. In Milton Keynes, there are no districts with high continuity of population: this gave us the opportunity to study levelling among high mobility, low-income groups (see Kerswill & Williams 1997).

An important corollary of this procedure is that the two groups are likely to be either broadly working class, using local accents, or middle class with less localised forms of speech. For ease of reference, we refer to the two groups as 'WC' and 'MC', respectively.

24 group interviews were conducted, 22 by AW and two by PK, following an agreed format. The groups were composed of four (very occasionally five or six) 14-15 year olds, each of whom had previously taken part in an individual sociolinguistic interview with the same fieldworker. A total of 96 adolescents took part in the project, a figure which gives 32 in each town and 16 in each school. The subjects participated in a number of activities designed to tap their language awareness. These began with the dialect recognition task reported here, followed by a questionnaire inviting discussion of regional grammatical features, and a general linguistic discussion covering issues such as 'good' and 'bad' speech and correction by parents or teachers.

For the dialect recognition task, subjects in each town were presented with taped samples of ten speakers, chosen so as to be both locally relevant to the judges while still allowing us to compare identifications of some of the same speakers across the three towns. Thus, three different, but substantially overlapping tapes were prepared (six voices being shared), the extracts being taken from interviews we had conducted previously or which had been conducted for us, or which had been recorded off-air. We ensured the samples contained phonetic features characteristic of their regional origins.

In each case, the sample was an extract from a personal narrative.

Tape presented to judges in:	Voices —>									
HULL	1 Hull F83	2 Milton Keynes F13	3 Durham M55	4 Middles-brough F17	5 Reading F50	6 Hull M9	7 Public school M14	8 Yorks. East Riding M80	9 London M13	10 Hull M15
READING	1 Reading M82	2 Hull M15	3 London F35	4 Reading M15	5 Durham M55	6 London M13	7 Public school **M14**	8 Reading F50	9 Milton Keynes **F13**	10 Reading F18
MILTON KEYNES	1 Milton Keynes F82	2 Hull M15	3 London F35	4 Reading M15	5 Durham M55	6 London M13	7 Public school M14	8 Reading F50	9 Milton Keynes F13	10 Milton Keynes M9

Table 2: Voices presented to judges in Hull, Reading and Milton Keynes

The tapes were composed of the voices given in Table 2, coded by sex and age as shown. For each town, there are one elderly speaker and two young speakers. The 'public school' voice was that of a pupil at a prestigious fee-paying school in the south of England. The subjects were given a form on which they were asked to fill in answers to three questions while the tape was being played: 'Where do you think this person comes from?', 'About how old do you think this person is?', and 'Do you think this person lives in a town or in the country?'. (The last two questions were mainly included in order to give all the subjects a chance of getting at least some answers right, while making the task more interesting.) Afterwards, the fieldworker led a discussion about any features of the voices that might have influenced the subjects in their judgements. For each voice for each town, there is a maximum of 32 judgements; in practice, some subjects failed to make an entry for every voice: the average number of judgements is therefore closer to 30.

2.2 *Recognising voices from one's own speech community: overview of results*

We return to one of the main issues of this article: linguistic focusing. As in the Bergen study mentioned above, we can investigate this indirectly by considering people's recognition of voices from their own town. However, by systematically varying both the judges and the 'native' voices (those from the judges' own town), we

are in a position to answer much more specific questions about the nature of dialect recognition and its relationship to focusing.

Figures 2a–2c (Appendix) show the recognition of the two young voices from the home towns of the judges: scores for the WC and MC groups are given separately. Three results stand out. First, both the Hull groups are much more successful than any of the other groups – in conformity with Hypothesis 2 ('judges from towns with little mobility are well attuned to local speech') – though Hypothesis 3 ('highly distinctive dialects are likely to be more easily recognised than less distinctive dialects') may be a confounding factor. Second, within Hull, the WC group is the more successful – this time in conformity with Hypothesis 1 ('people with local ties are attuned to local speech'). The third is perhaps more surprising: this is the fact that the Reading subjects are even less successful at the task than their Milton Keynes counterparts – contrary to both Hypotheses 1 and 2. We turn first to the Hull data.

2.3 *Focusing in Hull*

2.3.1 Local networks and localised dialect as factors favouring dialect recognition

Figures 3a–3f (see Appendix) show the Hull results in more detail. In addition to showing the 'Hull' identifications (dark shading), they show the number of times the voices were heard as being from Yorkshire, the county in which Hull is situated (light shading). The recognition rates for an elderly Hull speaker, F83, are also given. Note the overall greater success of the WC judges (their range being 86.7% to 94.1%, as opposed to the MCs' 53.3% to 80.0%), a result which is in line with Hypothesis 1. However, we cannot confirm this interpretation until we have shown that the WC group actually has stronger local networks. We did this by asking the judges where their parents were born, on the assumption that local parents are a reflection of locally-based networks. Table 3 shows that, of the working-class parents, 94% of the mothers and 87% of the fathers were born in Hull – the vast majority of them born on the estate where they currently reside – while the figure for Hull-born middle-class mothers and fathers is much lower at 53%.

On the face of it, we have evidence that the strongly local networks of the working-class judges facilitate their recognition of Hull voices. As we shall see later, this interpretation is confirmed by a comparison with the recognition patterns found in Reading and Milton Keynes: to anticipate, the Reading WC group is more successful at recognising own-community voices than the corresponding MC group, while the equivalent pattern is not present in Milton Keynes, where few of the families have local origins.

	Working class			Middle class		
	Born	Mother's birthplace	Father's birthplace	Born	Mother's birthplace	Father's birthplace
Girls						
1	**Hull***	Withernsea		**Hull**	**Hull**	**Hull**
2	**Hull***	**Hull***	**Hull***			
3	**Hull***	**Hull**	**Hull**	**Hull**	East Riding	Barnsley
4	**Hull***	**Hull***	**Hull***	**Hull**	**Hull**	**Hull**
5	**Hull***	**Hull***	**Hull***	**Hull**	East Riding	East Riding
6	**Hull***	**Hull***	**Hull***	**Hull**	**Hull**	Canada
7	**Hull***	**Hull**	**Hull**	Nuneaton	**Hull**	Nuneaton
8	**Hull**	**Hull**	Spain	**Hull**	**Hull**	**Hull**
Boys						
1	**Hull***	**Hull***	**Hull***	**Hull**	**Hull**	**Hull**
2	**Hull**	**Hull**	**Hull**	Birmingham	Kurdistan	Wales
3	**Hull***	**Hull***	**Hull***	**Hull**	Manchester	**Hull**
4	**Hull**	**Hull**	**Hull**	Manchester	Manchester	Wales
5	Chester	**Hull***	Lincs	**Hull**	East Riding	East Riding
6	**Hull***	**Hull***	**Hull***	**Hull**	**Hull**	**Hull**
7	**Hull***	**Hull***	**Hull**	**Hull**	**Hull**	**Hull**
8	**Hull***	**Hull***	**Hull**	**Hull**	Manchester	**Hull**
% born in Hull	93.7	93.7	86.7	80.0	53.3	53.3

* indicates individuals born on the estate where they currently live
Note: for ease of identification, 'Hull' is printed in bold type.

Table 3: Birthplace of Hull subjects and their parents

While we have demonstrated a clear working-class advantage in recognition rates, we have not explained why the overall rate for Hull is so much higher than that for the southern towns. As we have already mentioned, the accent may contain features which distinguish it sharply from neighbouring varieties and which act as positive identification markers (Hypothesis 3). One segmental feature appears to be unique to Hull and its immediate hinterland. This is the vowel in the lexical set of PRICE (Wells 1982), which, in Hull, has two very distinct allophones: a diphthong [aɪ] before voiceless consonants, as in *price* itself as well as in *bright*, *bike* and *knife*, and a monophthong [a:] before voiced consonants, as in *bride*, *five* and *pint*. A typical Hull pronunciation of *night time*, then, is [naɪʔ ta:m]. Table 4 shows the use of the two variants in voiceless and voiced environments among WC and MC adolescents and WC elderly people. The most striking result is the virtual absence of the distinction among MC speakers: an RP-like diphthong [aɪ] is used fairly consistently in both environments. The picture is very different for the WC subjects: even in the reading list from which the adolescent data is taken, the distinction is categorically

maintained by all except two of our speakers (a girl and a boy). Overall, the WC dialect appears to be maintaining this old, complex, localised feature, which was described in detail as early as 1877 (Ross, Stead & Holderness 1877: 9).

(a) with following voiceless consonant, e.g. *bright*

	% [aɪ ~ aˑɪ]	% [aː]
WC elderly (N=4)	100	0
WC boys (N=8)	100	0
WC girls (N=8)	100	0
MC boys (N=8)	100	0
MC girls (N=8)	100	0

(b) with following voiced consonant, e.g. *bride*

	% [aɪ ~ aˑɪ]	% [aː]
WC elderly (N=4)	0	100
WC boys (N=8)	17.5	82.5
WC girls (N=8)	25.7	74.2
MC boys (N=8)	95.0	5.0
MC girls (N=8)	100	0

Note: Each adolescent read the following words: *bright, knife, lighter, bike, whiter; bride, five, pint, smile, wider.* Scores for the elderly are derived from the interview data: 20 tokens per speaker were transcribed

Table 4: The PRICE vowel with following voiceless and voiced consonants, Hull speakers (adapted from Williams & Kerswill 1999, Table 7)

We now face the question of whether the Hull listeners consciously or unconsciously attended to this particular feature when carrying out the listening task: the two environments (voiceless and voiced) were indeed present in the recorded extracts for two of the speakers, M9 and M15, with the distinction clearly made. However, because the decision had been taken to use only sections of spontaneous discourse as stimuli, and to use a variety of stimulus voices, we could not easily test for any direct effect, using, say, the technique of Labov's Subjective Reaction test (Labov 1972: 146-9). However, assuming that a dialect is normally perceived as a whole, rather than by listening for individual features, we can conclude that the high recognition scores are the result of a dialect *gestalt* which is linguistically well demarcated from other dialects. If this is so, it implies that investigating responses to the PRICE vowel may not be relevant in this context: the distinctiveness of this vowel may be indicative of the distinctiveness of the accent as a whole.

If we allow the definition of focusing to encompass at least some shared phonetic features, then we can conclude that, like Bergen, Hull appears to be a focused speech community. Yet we have also uncovered differences within the community: it is the WC judges who appear to be more 'focused' than the MC, first, in having a more localised accent and, second, in recognising the voices better. We have already argued that this greater ability is at least *partly* due to these judges' greater familiarity with the local variety owing to their more strongly local networks; this factor should, we argue, be seen as combining with the fact that the WC judges are being asked to recognise accents that are actually similar to theirs (Hypothesis 6).

At this stage of the argument, we cannot of course make any statement about the degree of focusing in relation to other communities: we approach this issue below in our discussion of Milton Keynes and Reading. But first, we address a surprising difference in the identification patterns of the WC and the MC judges.

2.3.2 Hull or East Riding? The mediating effect of (socio)linguistic exposure

The discussion of focusing does not, however, explain why the two groups of judges identify the elderly Hull speaker in such different ways, with high identifications as 'Hull' by the WC, and no such identifications by the MC, who instead favour a 'Yorkshire' identification. Figures 4 and 5 show the non-generic 'Northern' identifications (that is, excluding not only identifications as 'southern', 'Midlands', etc., but also generic 'Northern') for the three Hull voices and the two geographically closest voices, East Riding of Yorkshire (corresponding to the rural hinterland of Hull) and Middlesbrough (a large town some 100 kms. to the north).

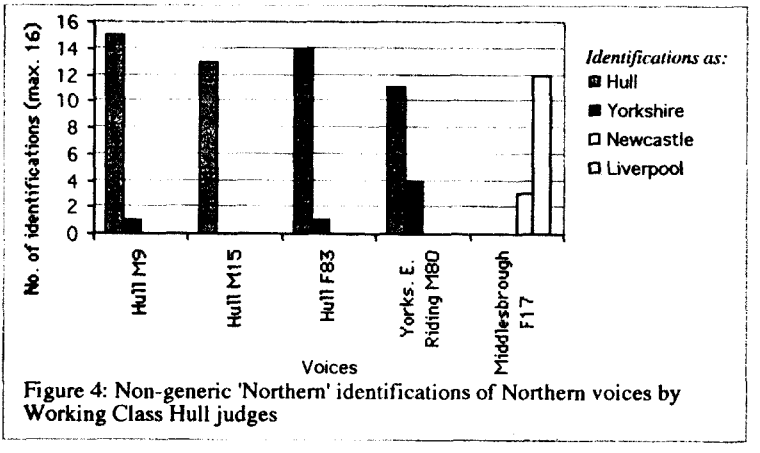


Figure 4: Non-generic 'Northern' identifications of Northern voices by Working Class Hull judges

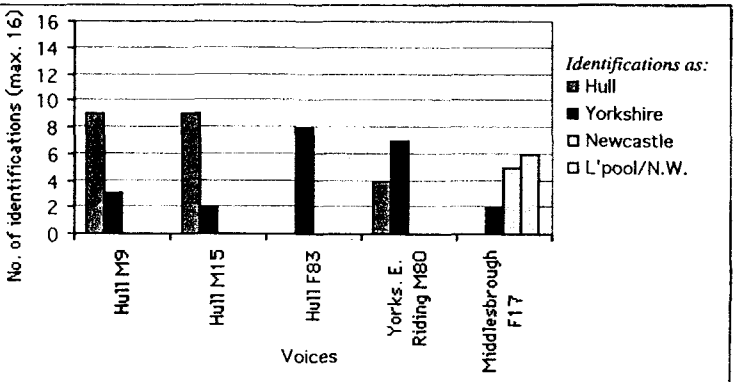


Figure 5: Non-generic 'Northern' identifications of Northern voices by Middle Class Hull judges

Middlesbrough is incorrectly identified by all the judges, something which suggests a lack of familiarity with the accent, due, probably, to a minimal amount of contact between that town and Hull. This would support Hypothesis 4 (the 'familiarity hypothesis'). By far the largest number opted for Liverpool (in the north-west of England), though Newcastle, which lies some 60 kms. to the north of Middlesbrough, was a popular choice – both accents having been made familiar in recent years in popular television soaps and children's programmes. Given the unfamiliarity of the Middlesbrough accent, it is likely that three phonetic factors contribute to these very specific, but erroneous identifications of Middlesbrough: first, intonation is subjectively similar to that of both Liverpool and Newcastle, where a rise-plateau pattern is associated with finality (Cruttenden 1995; Local 1986). Second, the taped extract contains Newcastle-type glottally reinforced pre-vocalic /t/ (Milroy, Milroy & Hartley 1994). The third factor involves the vowel system: while some realisations on the tape, especially those of FACE and GOAT (half-close monophthongs), are similar to those of Newcastle, others are similar to Liverpool, particularly [ɛ:] for NURSE (though it is not merged with SQUARE as it may be in Liverpool); indeed, Llamas, in discussing the NURSE vowel, comments that migration from Ireland and Wales 'may explain the similarity of [Middlesbrough] to Scouse [Liverpool] with regard to this and other variants' (1998: 109).

On the other hand, the scores for the elderly East Riding voice are similar to those for Hull F83, suggesting both the local 'relevance' of the accent and a considerable phonetic similarity to the Hull accent. As with F83, the WC judges place him in Hull, while the MC prefer a 'Yorkshire' identification. A possible explanation for this is that many of the MC judges do not live in the city, but in the dormitory

villages just outside the city boundary. They are therefore more likely to identify elderly speakers as rural because it is in a more rural context that they encounter them. The WC group, all of whom live on the council estate in the north of the city, encounter elderly people mainly in the city.

This argumentation could be extended to explain why the MC judges are nevertheless able to identify the young Hull voices: visits to the city are likely to bring them into contact with young Hull people. However, this possibility is complicated by the presence of a local dialect levelling process, by which features of Hull speech are spreading to the rural hinterland (Middleton 1999). This means that the young voices on the tape to a significant extent resemble younger WC speakers in the villages, where many MC judges live. On the face of it, this ought to lead to an identification of the young voices as 'Yorkshire' rather than 'Hull'. The fact that this is not the case may perhaps be due to a blurring of the city/country distinction for younger people, with dormitory villages increasingly seen as suburbs of the city. This interpretation must, however, remain suggestive.

In this study, the young judges are in general linguistically more similar to the younger voices they are being asked to judge than they are to the older voices. This will in itself lead to higher recognition rates (Hypothesis 6), as will become even clearer when we come to consider Reading and Milton Keynes. As we have already suggested, this has the corollary that voices perceived to belong to people similar to the judges are likely to be socially attractive, and hence 'claimed' by the judges. In conclusion: the relationship between dialect recognition and focusing is not direct, but mediated, affected as it demonstrably is by the judges' own social networks, their exposure to different varieties, and their linguistic similarity to the voices they are judging.

2.4 *Dialect levelling in Reading*

2.4.1 Familiarity and local networks as factors in the recognition of Reading accents

We hypothesised that dialect recognition in Reading would be a little less consistent than in Hull, but considerably more consistent than in Milton Keynes. This turns out not to be so: recognition rates in Reading are much lower than in Hull, and lower even than in Milton Keynes. We had two grounds for the hypothesis: the greater distinctiveness of the Reading accent as compared to Milton Keynes, and the strongly local working-class networks in the town (see Kerswill & Williams 1999 for a discussion of these points). Figures 6a–6h show the identifications of four Reading voices, including the two young voices already shown in Figure 2b. No-one

recognised the elderly speaker as a Reading voice; instead, approximately half the judges placed him in the Southwest, with responses such as 'Devon', 'Cornwall', 'Somerset' and 'Bristol', as well as 'West Country', which is the widely accepted generic term for this region of England. Most of the remainder favoured 'Northern' identifications, indicating a complete failure to associate the voice with the south of the country at all. We will return to the reason for this lack of recognition after we have considered speaker F50.

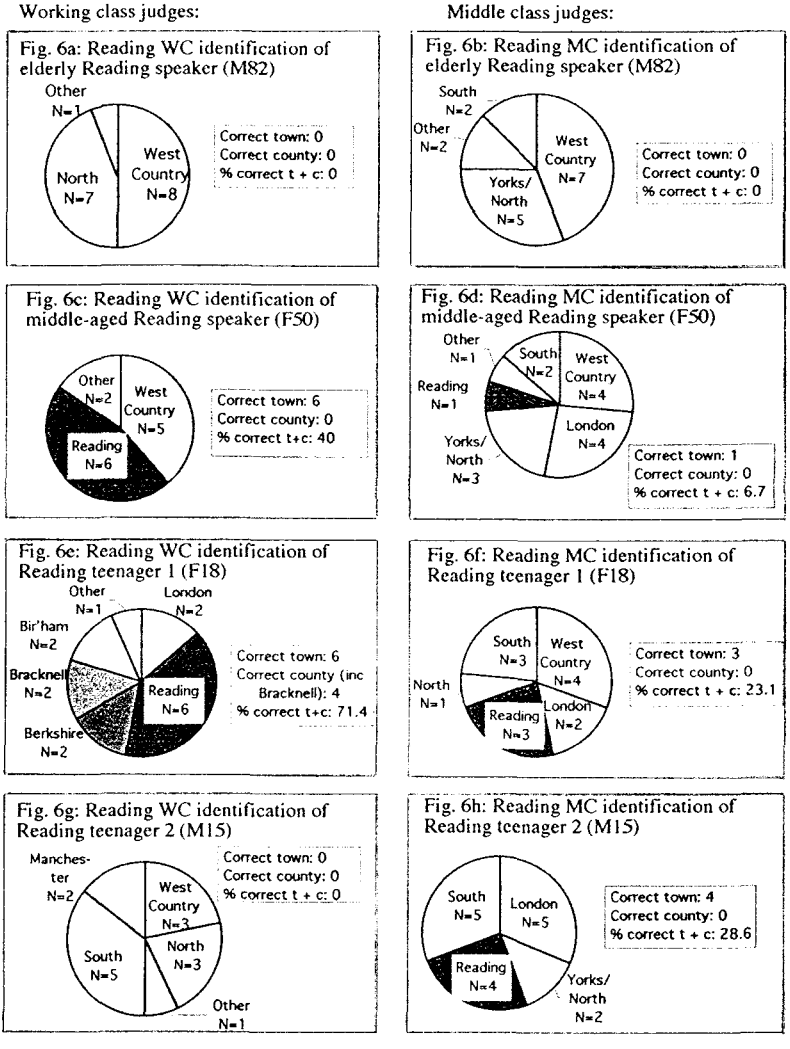


Figure 6: Reading identifications of speakers from Reading

The picture is more optimistic for the second oldest speaker, F50: 40% of the WC and one of the MC judges correctly identified her – though 'West Country' remains, overall, the most popular option. The recognition pattern for F50 is not surprising, since she represents the generation of the WC judges' (older) parents or (younger) grandparents, and so is a familiar 'voice' in the community. This also explains the MC judges' failure to identify her: almost none of the judges' families originate from Reading, with the result that this is a much less familiar voice for them, being encountered only outside the family. Table 5 shows the differences between the birthplaces of the parents clearly: almost all the WC parents are Reading-born, while only a very small proportion of the MC parents are born there. The WC–MC divide in terms of birthplace is even stronger than it is in Hull (Table 3), and this appears to be reflected in the recognition patterns.

	Working class			Middle class		
	Born	Mother's birthplace	Father's birthplace	Born	Mother's birthplace	Father's birthplace
Girls						
1	**Reading**	**Reading**	**Reading**	**Reading**	Barbados	Barbados
2	**Reading**	**Reading**	**Reading**	Warrington	Yorkshire	Yorkshire
3	**Reading**	Guyana	Guyana	**Reading**	Essex	Essex
4	**Reading**	**Reading**	**Reading**	**Reading**		
5	Germany	India	**Reading**	I. of Wight	**Reading**	I. of Wight
6	**Reading**	Cambridge	**Reading**	Ascot	London	Portsmouth
7	**Reading**	**Reading**	**Reading**	**Reading**	**Reading**	Tadley
8	**Reading**	**Reading**	**Reading**	**Reading**	Watford	Yorkshire
Boys						
1	**Reading**	**Reading**	**Reading**	**Reading**	**Reading**	**Reading**
2	**Reading**	**Reading**	**Reading**	Slough	**Reading**	Somerset
3	**Reading**	**Reading**	**Reading**	**Reading**	Wolverhampton	London
4	**Reading**	**Reading**	**Reading**	**Reading**	Sussex	Hastings
5	**Reading**	**Reading**	**Reading**	Hillingdon	Hastings	**Reading**
6	**Reading**	**Reading**	**Reading**	**Reading**	Newcastle	Newcastle
7	**Reading**	**Reading**	London	London	London	London
8	**Reading**	**Reading**	Ireland	**Reading**	Germany	Devon
% born in Reading	93.7	81.2	81.2	62.5	26.7	11.8

Note: For ease of identification, 'Reading' is printed in bold type.

Table 5: Birthplace of Reading judges and their parents

We still have to answer the question of why the oldest speaker, M82, is not recognised at all. We suggest that this voice belongs to the generation of the great-

grandparents of the judges, and is thus relatively unfamiliar. Such an explanation would be in line with the 'familiarity hypothesis' we have just discussed. Phonetically the voice is very different from that of the judges: he is fully rhotic (that is, he pronounces non-prevocalic /r/), and uses a strongly retroflex articulation of /r/ – neither of which is true of the judges. He does not use [f] for /θ/ or [v] for intervocalic /ð/, as do most of the WC judges.

The two youngest speakers attract strikingly different recognition patterns. F18 is recognised as 'Reading', or at least 'Berkshire' (the county in which Reading is situated), by 71.4% of the WC speakers. Given that this is a WC voice, this is not surprising, and the lower success rate of the MC speakers (23.1%) is in line with both Hypothesis 1 ('people with local ties are attuned to local speech') and Hypothesis 4 (the 'familiarity hypothesis'). M15, however, is recognised as a Reading speaker only by four MC judges, no WC judges correctly identifying him. Most of the remainder opted for an undifferentiated 'South', suggesting a measure of recognition coupled with uncertainty. Listening to the extract, however, gives a clue to the reason for this, and (as we shall see in the next section) suggests the direction in which the Reading accent is changing: although the voice can be heard to use a range of non-standard phonetic features, such as h-dropping in the items *happening* and stressed *he*, categorical glottal replacement of intervocalic /t/, vocalisation of non-initial /l/, and a broad diphthong [ɛɪ] in FACE, he does not use any marked Reading features. F18, in addition to using all the features mentioned, uses a central vowel [ə] in the items *funny* and *stuff*, (cf. M15: [ɐ]) and a diphthong [ɔɪ] in *inside* (cf. M15: [ɒɪ]) – both of these being features not widely found in London-influenced south-eastern accents, and the latter being specially mentioned as a Reading feature by some judges in the discussion following the identification task. M15 uses a more levelled variety than F18, in the sense discussed in Williams & Kerswill (1999): he does not use strong Reading features, but nor does he use marked London features. Instead, he uses the set of south-eastern features which are spreading throughout the region and further afield, including those which are in evidence in the taped extract. This, in turn, makes his accent more similar to that of the MC judges than is F18's: this is why, we suggest, four of the MC judges accepted him as 'Reading'. Next, we consider whether this boy's speech represents a stage in the change in the Reading accent.

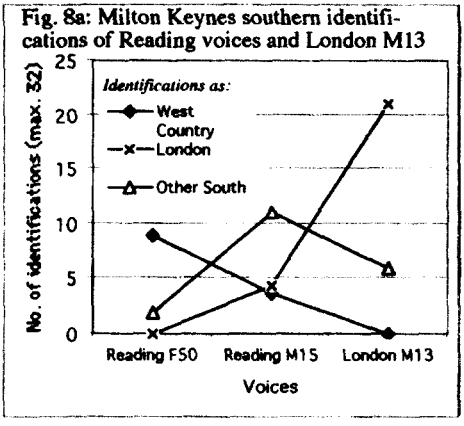
2.4.2 The de-focusing of Reading

Our data shows that the identification of the Reading accent as 'West Country' diminishes with the decreasing age of the speaker. This apparently simple fact masks a complex issue: that of the effect of the time dimension. Would this result have been

obtained for an older panel of judges, or if the survey had been done 25 or 50 years previously? Our supposition is that older listeners, or judges in an earlier period, would have been less likely to adjudge M82 a 'West Country' speaker than today's adolescents were. This is because there has demonstrably been change in Reading, most of which can be considered part of regional dialect levelling (Williams & Kerswill 1999). Thus, from a contemporary adolescent's vantage point, the speech of elderly speakers can seem very remote not only in time, but in place. We now explore the apparent 'de-Westcountrification' of the accent, and consider the direction in which it is heading: in particular, is it becoming 'Cockneyfied' (that is, 'Londonised')? Figure 7 (see Appendix) shows the 'West Country', 'Reading' and 'London' identifications for the four Reading voices and those for the two London speakers. Other identifications, including generic 'South', have been omitted. Figure 7a shows the 'Reading' identifications: the impression given by this graph, which ranks the four Reading voices by descending age, is that the accent is becoming less south-western, with only three judges deeming M15 to be 'West Country', which puts him nearly into line with the two Londoners. Figure 7b shows a gradual 'improvement' in 'Reading' identifications, though this stops with F18, with even London M13 being heard as more 'Reading' than M15.

So far, we could be tempted to use the analogy of the accent 'travelling' rapidly in an easterly direction towards the capital. But Figure 7c destroys the analogy. The two Londoners, F35 and M13, are overwhelmingly identified as such (with scores of 78.1% and 69.7%, respectively), while for M15, who is the Reading speaker who receives the highest 'London' identification and whose speech contains the fewest Reading features, the figure is only 16.7%. The picture emerging is that the Reading accent, for all the levelling it has been subject to, remains distinct. Moreover, inner-London speech, even that of the youngest age group, is still easily identifiable by outsiders. Indeed, the extract of London M13's speech contains a number of London features, including [ɛ:] for the vowel of MOUTH, a relatively front vowel, [a̲], for STRUT, and the vowel /eɪ/ (as in FACE) in the auxiliary *ain't*, an item in which Reading speakers tend to use /e/ (as in DRESS).

Interestingly, these judgements are not specific to Reading judges, since the Milton Keynes and, more surprisingly, the Hull judges gave similar identifications. Figure 8 shows the 'West Country', 'London' and other southern identifications of the southern voices which were presented to those judges: clearly, there are stable phonetic features in a London accent and, apparently, in a Reading accent which are nationally salient and available (in Preston's 1996a sense) and which lead to 'London' and 'West Country' identifications, respectively.



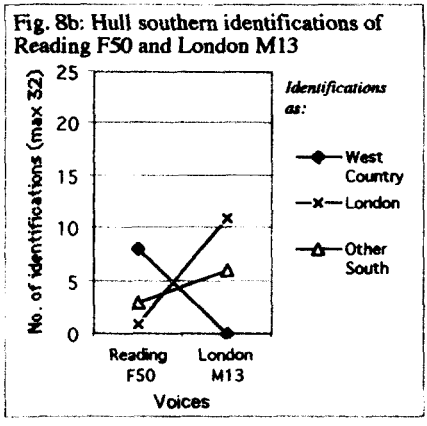


Figure 8: Milton Keynes and Hull southern identifications of Reading and London voices

This is evidence of the continued presence of a degree of focusing in Reading, yet there are also clear signs of the 'de-focusing' of the speech of the town: change has been sufficiently rapid for the oldest generations to be no longer identified as natives of the town. This is clearly not true of Hull for either elderly or young speakers, who are recognised at a very high rate. Our research shows that the continued focusing of Hull is mirrored by a slower rate of change than in Reading.

A particular consequence for Reading of the reduction in focusing (assuming it was greater in earlier decades) is that its residents, particularly the younger ones, seem to associate its accent with the West Country. This 'perceptual dislocation' of the accent reflects, we believe, the rapid social changes in the town over the past 50 years. In 1950, it was a market town dominated by agriculture (its university was founded as an agricultural college) and industries related to horticulture, food manufacture and brewing. Today, it is one of the principal national centres for high-technology computer-based industries, financial services and retailing. In this environment, the link with agriculture has been lost, and it is not surprising that the oldest speakers and their accents have been marginalised.

One particular phonetic feature is a specific cue to the perception of the older accent as south-western: the non-prevocalic /r/, which was mentioned by a number of judges in the discussion sessions as a feature they attended to in arriving at a 'West Country' identification of F50. Anecdotally, we can mention that Reading young people regularly report being accused of talking 'country' when they visit London –

even though they do not use the non-prevocalic /r/; and a middle-aged speaker reported being surprised at how 'country' she sounded the first time she heard her voice on tape. All this is tied in with a strong negative stereotype of south-western speech as being that of unsophisticated farmers, the word 'farmers' itself being the vehicle through which the stereotype is often expressed, with both *r*'s being realised in mocking imitation of West Country speech. Reading's geographical and dialectal position near the boundary between the stereotypically rural South and Southwest and the stereotypically urban Southeast, coupled with the rapid economic changes noted above, makes it particularly vulnerable to the 'farmer' stereotype.

The example of Reading shows that de-focusing goes hand in hand with dialect levelling and a rapid rate of change. Levelling potentially robs people of the possibility of using strongly local speech to mark allegiance to groups based on territory, class or ethnicity (see Kerswill & Williams 1997 and 1999 for discussions of language used as an identity marker). With the perceptual dislocation of traditional Reading speech to another region, and the lack of a distinctive replacement, Reading speakers seem to be losing this possibility. Likewise, the rate of change there is sufficiently fast for there to be a disjunction between the oldest and the youngest speakers, at least in terms of young people's recognition of old people as part of the speech community. It may be realistic to talk of a move away from strong local identifications towards identities based on other groupings, including class, age, gender and ethnicity, with regional identities subsumed into a sense of being 'from the south-east'.

In the next section, we turn to the New Town of Milton Keynes, where there is by definition a sharp break in continuity between the oldest and the youngest speakers (Kerswill & Williams 2000 forthcoming).

2.5 *Milton Keynes: an incipient focused, but levelled speech community?*

2.5.1 Non-local networks and the recognition of Milton Keynes voices

Figure 9 shows the recognition patterns for Milton Keynes, as before with the two class groups' identifications shown separately. As we noted earlier (Section 2.2), the Milton Keynes judges are more successful at their task than are the Reading judges, a finding which goes against our hypotheses. However, closer examination shows that the results pattern quite differently from those of Reading, in a way consistent with Milton Keynes's status as a new community whose younger families have no time-depth in the town.

The failure of any of the WC judges to recognise the elderly speaker, F82, comes

as no surprise: in addition to the factors we have already adduced for the parallel finding in Reading, a reason must also be the fact that very few of these judges have any family connections with older people in the town. Elsewhere, we have argued that this lack of continuity is reflected in the linguistic production data (Williams & Kerswill 1999; Kerswill & Williams forthcoming); what we are dealing with here is the effect a lack of continuity has on dialect recognition. Table 6 shows the judges' place of birth and that of their parents. There is a striking difference between this table and the equivalent tables for Hull and Reading: in Milton Keynes, there are only slightly more locally-born parents among the WC group than among the MC group (around 13%, as opposed to 3% for the MC), whereas the percentage of locally-born WC parents in the other towns was extremely high (80-90%).

	Working class			Middle class		
	Born	Mother's birthplace	Father's birthplace	Born	Mother's birthplace	Father's birthplace
Girls						
1	Scotland	Scotland	Scotland	**M. Keynes**	Newbury	St. Helena
2	**M. Keynes**	Halifax	London	**M. Keynes**	London	Leeds
3	Luton	Portsmouth	Watford	Oxford	Oxford	Oxford
4	London	London	London	**M. Keynes**	Lowestoft	**Bletchley**
5	**M. Keynes**	**Bletchley**	**Bletchley**	Cranfield	Leicester	Bucks.
6	Lancashire	Lancashire	Liverpool			
7	Blackpool	London		Glasgow	Inverness	Inverness
8	**Bletchley**	Stevenage	Ireland	**M. Keynes**	Kenya	Kenya
Boys						
1	**M. Keynes**	**Bletchley**	**Bletchley**	Birkenhead	Birkenhead	Birkenhead
2	London	Essex	London	London	Luton	Luton
3	**M. Keynes**	London	London	Kent	Manchester	Dorset
4	**M. Keynes**	Gt. Yarmouth	Ireland	Aylesbury	Poland	Manchester
5	Newbury	Newbury	Tadley	Northampton	Newport Pagnell	Newport Pagnell
6	Ireland	Halifax	Ireland	Bristol	Bristol	Manchester
7	**M. Keynes**	London	London	Northampton	Newcastle	'North'
8	**M. Keynes**	London	Jamaica	Brighton	Northants.	Leicester
% born in M. Keynes	50.0	12.5	13.3	26.7	0	6.7

Note: For ease of identification, 'Milton Keynes' and 'Bletchley' are printed in bold type (Bletchley lies within the borough of Milton Keynes).

Table 6: Birthplace of Milton Keynes judges and their parents

This does not explain why the positive identification of F82 as 'Milton Keynes' or 'Buckinghamshire' (the county in which Milton Keynes lies) is relatively high for the MC judges (38.5%). The same argumentation could perhaps be used as for the

Hull MC judges: many of the Milton Keynes MC judges lived in villages near the town. Elderly speakers like F82 would be encountered in the villages, and her 'voice' would be familiar. At present, however, this interpretation is somewhat speculative.

The fact that few of the Milton Keynes judges have locally-born parents means that 'localness of network' ceases to be a possible factor in the explanation of differences in judgements, as it was in both Hull and Reading, where there was a marked tendency for the WC judges to recognise own-community voices better than MC judges. The Milton Keynes results, *when taken together with the results for Hull and Reading*, in fact strongly support the relevance of networks as an independent factor: Figures 9c–9f (see Appendix) show that there is practically no difference in the recognition of the younger voices between the two classes (20% vs. 25% for M9, 67% vs. 64% for F13). This leads us to the conclusion that it is network, and not class that is the decisive factor in own-community dialect recognition. This conclusion constitutes powerful support for Hypothesis 1 – though we argue elsewhere that class has a decisive effect in other areas: those of language and identity (Kerswill & Williams 1997), and the patterning of linguistic variables (Kerswill & Williams 1999).

2.5.2 Milton Keynes and Reading: converging accents following different paths

We have previously noted the Reading judges' lack of success in recognising Reading accents, and we ascribed this to dialect levelling and rapid change. These factors should apply even more in Milton Keynes, though the slightly better own-community identifications seem to refute this. This means that we may be witnessing an incipient 'focused' speech community, which is developing out of the diffuse melting pot of the incomers' generation. Our research shows that both towns are subject to the same dialect levelling, leading to a number of shared features. For dialect perception, the equivalent of linguistic levelling is increasing similarity in patterns of recognition – and here we find that the overall frequencies are indeed similar, though there are detailed differences which we can relate to differences in the localness of the judges' networks.

However, when we consider the phonetic features of the young Milton Keynes and Reading voices, a striking difference emerges between the two towns. We saw above how the more strongly localised Reading voice, F18, was perceived as 'Reading' much more frequently than the more levelled voice, M15, whose provenance listeners were unwilling to commit themselves to. The Milton Keynes voices, on the other hand, show precisely the reverse pattern. F13 does not use any marked regional features; thus, she does not use the older Buckinghamshire [ʌɪ] for PRICE, but instead

uses [ɑɪ], and she uses [aʊ] for MOUTH. This makes her accent subjectively quite similar to Reading M15. On the other hand, Milton Keynes M9, who is identified as a Londoner by 71.8% of the judges, has a rather different accent: in particular, he uses London [ɛ:] for the MOUTH vowel – acquired, no doubt, from his parents, who are from London.

The pattern is that, whereas in Reading it is the *less levelled* accent that is the better identified, in Milton Keynes it is the *more levelled* accent. The difference can be related directly to the history of the two towns and their dialects over the past 30 years. Reading's dialect has long contained localised features, and these survive sufficiently (albeit weakly) for them to be markers of the Reading origin of a young speaker. By contrast, there are few if any young linguistic inheritors of the older North Buckinghamshire dialect of the area now occupied by modern Milton Keynes: younger members of local families are, presumably, now linguistically absorbed into the new, in-migrant mainstream. Of the two young Milton Keynes speakers, the one with the more localised pronunciation traits (M9) in fact derives his accent from elsewhere, in this case London. The fact that 35% of the in-migrants came from the capital means that M9's accent will be widely heard in Milton Keynes – more so, probably, than a young North Buckinghamshire-derived accent. Nevertheless, it is the levelled accent represented by F13 that is probably numerically in the ascendancy, and to which young speakers accommodate as they reach their teens: indeed, our previous research (Kerswill & Williams forthcoming) suggests that speakers like M9 tend to modify their accent towards that represented by F13 as they reach their teens. It is this adolescent age group, we argue, that is establishing the 'new' accent of Milton Keynes. In consequence, F13's accent is the one perceived as characteristic of the younger speakers, and this is reflected in the relatively high recognition scores for her voice.

2.5.3 Does own-community perception co-vary with linguistic features?

As we have shown elsewhere (Williams & Kerswill 1999), the accents of Milton Keynes and Reading are converging by a process of levelling, though they are taking different routes. The dialect perception data adds to the linguistic performance data by giving more detail to those routes. First, it accurately reflects the linguistic discontinuity between older and younger generations in Milton Keynes; however, contrary to expectations, it shows that the same discontinuity applies in Reading, though only in terms of perception, linguistic features showing considerable continuity despite the rapidity of change. Second, it shows that perception patterns correlate with the strength of the listener's local networks, and that these networks are

in turn reflected in the degree to which the listener's own speech is localised, and hence localisable. However, in an exceptionally fluid community, such as that in a new town, it seems that this relationship does not apply: it may be the more levelled speakers, with fewer localised features, who are perceived as local. Clearly, the relationship between dialect perception and dialect production is not straightforward, affected as it is by a number of social factors. Careful examination, however, yields insights into dialect levelling which are not available from the linguistic data alone.

3. *Discussion: Dialect perception and focusing*

3.1 *Social structures, linguistic distinctiveness and familiarity as factors in focusing*

Before we return to the main theme of this article, we will summarise our findings in relation to the hypotheses.

There was ample support for **Hypothesis 1**: 'Own-community recognition will be better among people with strong local ties': in both Hull and Reading, the WC groups showed better own-community recognition. In Milton Keynes, there was no WC advantage. However, while in Hull the WC showed better recognition rates for all the Hull voices, in Reading it was the MC who recognised the more 'levelled' speaker (M15) the better. This suggests that his greater similarity to the MC judges' own accents might have played a part. It may well be that the hypothesis only holds for the recognition of voices with strongly localised accents.

There turned out to be a close relation between **Hypothesis 2**: 'Judges from towns with little mobility are well attuned to local speech' and **Hypothesis 3**: 'Highly distinctive dialects are likely to be more easily recognised than less distinctive dialects'. In the context of the present study, they must be interpreted together. They form part and parcel of an emerging, multifaceted picture of focusing in which large-scale social patterns, especially mobility and social networks, interact with language use: a focused speech community is one in which highly distinctive dialect features coupled with a slow rate of language change co-occur with strongly local networks and low geographical mobility. A corollary of all four factors in tandem (distinctive dialect, slow rate of change, local networks and low mobility) is the high recognition rates noted for Hull. On an individual level, there will of course be differences, particularly those dealt with by Hypothesis 1 and by Hypothesis 4, to which we turn next.

It is almost a tautology to say that accents which are familiar to the judge will

be better recognised than those which are not (cf. **Hypothesis 4**). However, the range of factors contributing to familiarity is wide. The most important distinction may be between those factors which promote the recognition of an own-community accent and those which facilitate the identification of accents from elsewhere. Local networks and family ties influence own-community recognition, as this research has shown; however, for the recognition of other accents, three factors in particular may be important: (1) the degree of contact between one's own community and the community represented by the voice, (2) whether a voice sounds like someone the judge happens to know, and (3) the influence of the broadcast media. In the contemporary world, the broadcast media are a crucial means by which familiarity with varieties is spread, and this becomes very clear from the discussion sessions following the dialect recognition task. Of the six voices heard by subjects in all three towns, Durham M55 has the most consistent identification: 63.7% identified the voice as 'Newcastle', which we accepted as correct since Durham lies just 25 kms. south of Newcastle and has an accent sharing many features with that of Newcastle. In the discussion sessions, many judges said that they knew the accent from *Byker Grove*, a popular and long-running children's soap set in Newcastle and using local child and teenage actors. Additionally, judges in one school cited the fact that one of their teachers had a Newcastle accent. The second best identified voice was that of London M13, who was recognised by 60.4% of the judges; in this case, the popular soap *EastEnders* would have been a factor making London accents familiar, though London voices are heard over a wide sector of radio and television broadcasting. In contrast, Hull M15 was identified as coming from Hull by no-one outside that city – even though Hull has a population figure that is 65% of that of Newcastle (254,000 vs. Newcastle's 389,000). However, he was correctly located in Yorkshire by 24.6% of Reading and Milton Keynes judges, this being the single most common identification ('North', 'Liverpool' and 'Manchester' being popular, but less common choices). There are no television series set in Hull, nor are there any icons of popular culture from there. A comparison of the results for Hull M15 and Durham M55 strongly supports the hypothesis that familiarity through media exposure is a decisive factor.

We have found ample support for **Hypothesis 5**: 'Different voices from the same town (even if there is no age difference between the speakers) will not be recognised at the same rate by members of that speech community'. However, the reasons we adduced for this finding can be related to factors other than those discussed by Williams et al. (1999): though of course we do not deny that 'social attractiveness' due to paralinguistic and content factors plays a part, we were able to relate the differences to the degree of focusing of the speech community and the amount of

dialect levelling.

In our study, **Hypothesis 6**: 'Own-community voices close to the age of the judges will be relatively easily recognised' forms an extension of Hypothesis 5. We found that a lack of recognition of elderly speakers only occurs in the two towns with rapid language change: Reading and Milton Keynes. This means that we can add 'rapidity of change' to loss of focusing and the presence of dialect levelling as a factor impeding dialect recognition. However, the intervening variable which reflects the factors directly affecting a judge's success, is the combination of a distinctive dialect and the judge's familiarity with that dialect.

3.2 *Degrees of focusing*

Finally, we return to the main theme of this article, the relationship between dialect perception and speech community focusing. It is clear that there is no direct correlation between the two: measuring focusing by means of dialect perception leads to a complex picture, and the results must be interpreted against the background of a number of mediating social and social psychological variables. Changes in dialect perception over 'apparent time', that is, comparing the recognition of older and younger voices, shows that both rapid linguistic change and a break in contact across the youngest and oldest speakers lead to an apparent discontinuity in a speech community and, we may assume, a reduction in focusing. This reduction can be reversed, as we can see in Milton Keynes where some degree of dialect recognition is beginning to appear and less diffuse social networks are developing. Overall, we can say that Hull is the most focused of the three towns: recognition is mainly high, and there is little loss of recognition across three or even four generations. Although the Bergen study did not investigate age differences of this sort, it did show that there are unequivocal phonetic cues to speech community membership there, and that they are exploited by native judges. Despite a different methodology, the high success rates suggest that the same is true for Hull.

Reading and Milton Keynes must be regarded as occupying the same, much lower position on a putative subjective focusing scale: both show a loss of recognition across generations and relatively low levels of recognition within the same generation. But this simple picture for Reading and Milton Keynes belies greater complexity: our comparison of the social networks of the judges, and our discussion of the degree of levelling of the voices presented on the tape, suggests differences in the sociolinguistic structure of these towns, due, we argue, to their very different demographic histories. Most striking of all was the very frequent identification of

older Reading speakers as 'West Country' and therefore rural, suggesting what we called a 'perceptual dislocation' of the town's accent. We argued that this was a consequence of rapid social changes in the town. Reading contrasts with Milton Keynes, which was founded on an in-migrant base and consequently started from a state of diffusion, which is gradually being replaced by a degree of focusing.

The relationship between this 'perceptual focusing' (that is, degree of own-community dialect recognition) and linguistic behaviour is one we have touched upon in this article. Perceptual focusing is closely linked to Hypothesis 3 ('highly distinctive dialects are likely to be more easily recognised than less distinctive dialects'), and yet the relationship is not straightforward. As we have seen, there is a clear effect of the judge's familiarity with the accent being presented, including that of his or her own town. This familiarity is in turn linked to the judge's social network characteristics. As we argued in the discussion of Hypotheses 2 and 3, dialect recognition forms part of a broader view of focusing, in which social structures, especially local networks and low mobility, combine with distinctiveness of dialect, clear sociolinguistic patterns and slow linguistic change to form a focused speech community.

All this, of course, confirms dialect recognition as an aspect of human sociolinguistic behaviour that is mediated by, and interacts with, a range of highly disparate factors. As a result, it has a complex, but nonetheless investigable, relationship with other sociolinguistic processes, including dialect levelling and other forms of language change.

NOTES

[1] A version of this paper also appears in Daniel Long and D. Preston (ed) (forthcoming) *A Handbook of Perceptual Dialectology* Vol. 2. Amsterdam: Benjamins.

[2] There is a problem of terminology here. The reader will note that we will be using the terms *accent* and *dialect* somewhat loosely. In British linguistic tradition, 'accent' refers to pronunciation features (e.g., Hughes & Trudgill 1996: 3). This covers subphonemic variation, but also variations in phonological inventory and the predictable difference in phonemic incidence this leads to (e.g., Southern English and Scots /kʌp/ *cup* corresponds to Northern English /kʊp/, because of the absence of /ʌ/ in Northern English varieties). 'Accent' also covers phonologically predictable differences in incidence, such as the Southern English use of /ɑ:/ as against Northern /æ/ before voiceless fricatives in items such as *bath*. 'Dialect', on the other hand, refers to grammatical and lexical features, as well

as non-predictable differences in phonological incidence, such as Durham English /θaʊt/ in *thought*, for which Received Pronunciation has /θɔːt/. In this article, our use of 'accent' and 'dialect' on the whole reflects this division. However, following usual practice we use 'dialect' as a modifier referring to all aspects of regional and social variation in the terms *dialect recognition* and *dialect levelling* '.

[3] For discussions of levelling in Europe, see Cheshire, Edwards & Whittle 1993; Thelander 1982; Hinskens 1996; Trumper & Maddalon 1988; Sandøy 1998; Kerswill 1996b; and papers in Vol. 10 of *Sociolinguistica*.

[4] Funded by the Economic and Social Research Council of Great Britain, 1995-8, ref. R000236180. Award holders: Ann Williams, Paul Kerswill and Jenny Cheshire. Research Fellows: Ann Williams and Ann Gillett. See Kerswill & Williams (1997), Williams & Kerswill (1999).

REFERENCES

Cheshire, Jenny, Edwards, Viv & Whittle, Pam. 1993. Urban British dialect grammar: the question of dialect levelling. In James Milroy & Lesley Milroy (eds) *Real English: The Grammar of English Dialects in the British Isles*, pp. 54-96. London: Longman,.

Cruttenden, Alan. 1995. Rises in English. In Jack Windsor Lewis (ed.) *Studies in General and English Phonetics*, pp. 155-74. London: Routledge.

Hinskens, Frans. 1996. *Dialect Levelling in Limburg. Structural and Sociolinguistic Aspects*. Tübingen: Niemeyer.

Hughes, Arthur & Trudgill, Peter. 1996. *English Accents and Dialects* (3rd edition). London: Arnold.

Kerswill, Paul. 1985. Native dialect and dialect mixing in Bergen: a perception experiment. *Cambridge Papers in Phonetics and Experimental Linguistics* 4.

Kerswill, Paul. 1993. Rural dialect speakers in an urban speech community: the role of dialect contact in defining a sociolinguistic concept. *International Journal of Applied Linguistics* 3(1): 33-56.

Kerswill, Paul. 1996a. Children, adolescents and language change. *Language Variation and Change* 8(2): 177-202.

Kerswill, Paul. 1996b. Divergence and convergence of sociolinguistic structures in Norway and England. *Sociolinguistica* 10: 90-104.

Kerswill, Paul & Williams, Ann. 1997. Investigating social and linguistic identity in three British schools. In U.-B. Kotsinas, A.-B. Stenström & A.-M. Malin (eds)

Ungdomsspråk i Norden. Föredrag från ett Forskarsymposium [Youth language in the Nordic countries. Papers from a Research Symposium]. Series: MINS, No. 43. 159-76. Stockholm: University of Stockholm, Department of Nordic Languages and Literature,

Kerswill, Paul & Williams, Ann. 1999. Mobility versus social class in dialect levelling: evidence from new and old towns in England. *Cuadernos de Filología Inglesa* 8: .

Kerswill, Paul & Williams, Ann. forthcoming. Creating a new town koine: children and language change in Milton Keynes. *Language in Society* 29(1).

Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

Lambert, Wallace. E., Hodgson, R.C. & Fillenbaum, S. 1960. Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology* 60: 44-51.

Le Page, R.B. 1978. Projection, focusing, diffusion, or, steps towards a sociolinguistic theory of language, illustrated from the Sociolinguistic Survey of Multilingual Communities. Trinidad: University of the West Indies, St. Augustine. *Occasional Paper No. 9 of the Society for Caribbean Linguistics*. Also in *York Papers in Linguistics* 9, 1980, 9-32.

Local, John. 1986. Patterns and problems in a study of Tyneside intonation. In Catherine Johns-Lewis (ed) *Intonation in Discourse*, pp. 181-97. London: Croom Helm.

Llamas, Carmen. 1998. Language variation and innovation in Middlesbrough: a pilot study. In Paul Foulkes (ed) *Leeds Working Papers in Linguistics and Phonetics* 6: 97-114.

Middleton, Lucy. 1999. The levelling of the traditional East Riding accent towards the mainstream Hull accent. Unpublished undergraduate dissertation, Department of Linguistic Science, University of Reading.

Milroy, James, Milroy, Lesley & Hartley, Sue. 1994. Local and supra-local change in British English: the case of glottalisation. *English World Wide* 15: 1-33.

Milroy, Lesley & Milroy, James. 1992. Social network and social class: toward an integrated sociolinguistic model. *Language in Society* 21: 1-26.

Preston, Dennis R. 1989. *Perceptual Dialectology*. Dordrecht: Foris.

Preston, Dennis R. 1996a. Whaddayaknow? The modes of folk linguistic awareness. *Language Awareness* 5: 40-73.

Preston, Dennis R. 1996b. Where the worst English is spoken. In Edgar Schneider (ed) *Focus on the USA*, pp. 297-360. Amsterdam: Benjamins.

Ross, F., Stead, R. & Holderness, T. 1877. *A Glossary of Words Used in Holderness in the East Riding of Yorkshire*. London: Trubner & Co.

Sandøy, Helge. 1998. The diffusion of a new morphology in Norwegian dialects. *Folia Linguistica* 32/1-2:83-100.

Thelander, Mats. 1982. A qualitative approach to the quantitative data of speech variation. In Suzanne Romaine (ed.) *Sociolinguistic Variation in Speech Communities*, pp. 65-83. London: Arnold,

Trudgill, Peter. 1983. *On Dialect.* Oxford: Blackwell.

Trudgill, Peter. 1986. *Dialects in Contact.* Oxford: Blackwell.

Trumper, John & Maddalon, M. 1988. Converging divergence and diverging convergence: the dialect-language conflict and contrasting evolutionary trends in Modern Italy. In Peter Auer & Aldo di Luzio (eds) *Variation and Convergence. Studies in Social Dialectology*, pp. 216-58. Berlin: de Gruyter.

Wells, John C. 1982. *Accents of English.* Cambridge: Cambridge University Press.

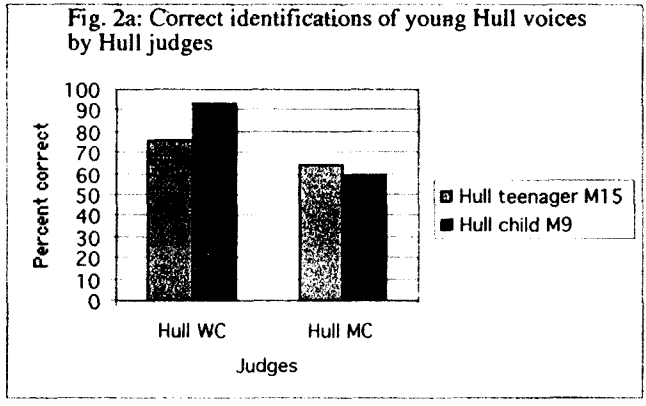
Williams, Angie, Garrett, Peter & Coupland, Nikolas. 1999. Dialect recognition. In Dennis Preston (ed) *Handbook of Perceptual Dialectology*, pp. 369-83. Amsterdam: Benjamins.

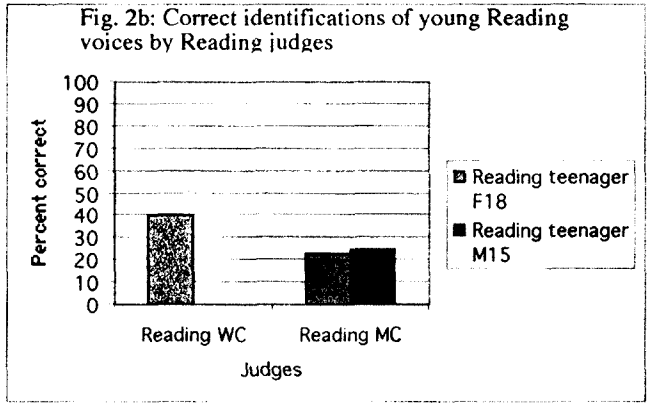
Williams, Ann & Kerswill, Paul. 1999. Dialect levelling: continuity vs. change in Milton Keynes, Reading and Hull. In Paul Foulkes & Gerry Docherty (eds) *Urban voices.* London: Arnold.

Address

Paul Kerswill and Ann Williams
Department of Linguistic Science
University of Reading
Whiteknights
PO Box 218
Reading RG6 6AA
England
<p.e.kerswill@reading.ac.uk>

APPENDIX





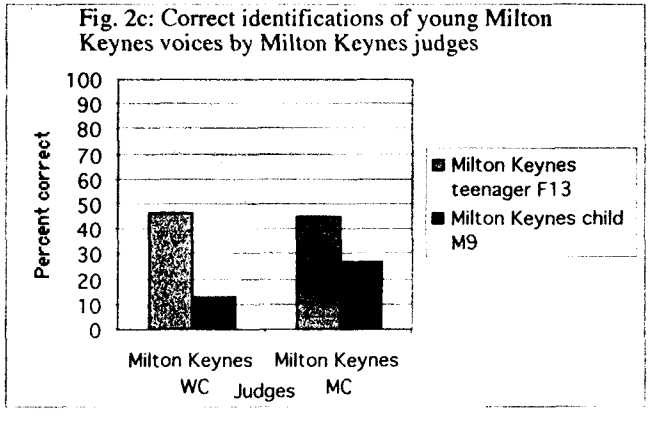
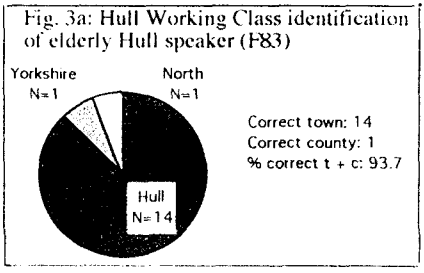
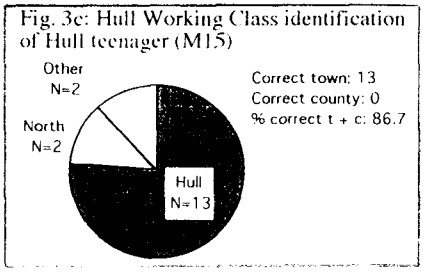
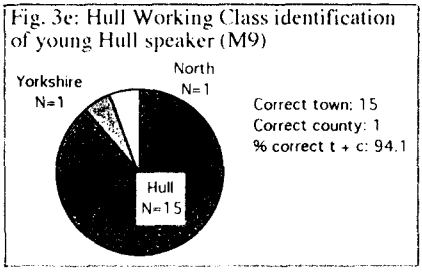


Figure 2: Correct identifications of young voices from judges' home towns

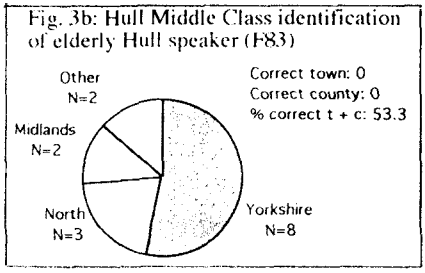
Working class judges:

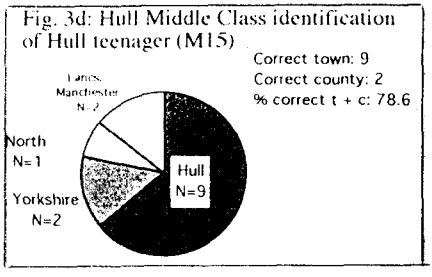






Middle class judges:





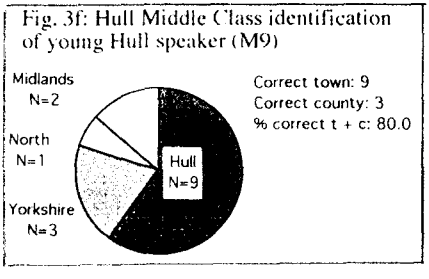


Figure 3: Hull identifications of speakers from Hull

English Intonation in the British Isles

Kimberley Farrar, Esther Grabe and Francis Nolan

Abstract

Large-scale research projects such as the Survey of English Dialects, together with studies of individual dialects, have yielded a great deal of information on how the dialects of the British Isles differ from one another on the segmental phonetic level. Although it is recognised that intonation is also important in characterising dialectal diversity, there have been no parallel systematic cross-dialect studies. One reason for this is the lack of a consensus on how to represent intonation patterns. No system which is comparable to the IPA in segmental phonetics is available for transcription. This paper discusses a new project funded by the ESRC, which will examine the intonational diversity in the British Isles. The main aim of the project is to develop a database of recordings from seven dialects of British English. This database will be available on CD-ROM, and will contain speech data, as well as files containing orthographic transcripts. Intonational transcriptions will also be provided, using a new transcription system for intonation, specifically developed with comparative intonation analysis in mind.

The findings of the project will therefore provide the first comparative description of intonational diversity in the British Isles and provide other researchers with a new methodology for further investigation. At the time of writing, work is in its early stages. This paper therefore concentrates on the methodology developed for the project.

1. *Introduction*

Large-scale research projects such as the Survey of English Dialects have yielded a great deal of information on how the dialects of the British Isles differ from one another on the segmental phonetic level. Although it is recognised that intonation is

also important in characterising dialectal diversity, there have been no parallel systematic **cross**-dialect studies.

For example, one of the most well-known studies of dialectal diversity in English, John Wells' *Accents of English*, concentrates primarily on segmental features of different dialects. Wells acknowledges that intonation is important in distinguishing dialects, but notes that 'our ignorance in this area is still very great' (1982: 91). It is the lack of any large body of work on intonation differences which precludes Wells from making any more detailed comments in his subsequent descriptions of individual dialects.

This is not to say that no work has been done on the intonation of different dialects of British English. Work by Pellowe and Jones (1978) and by Local (1986) has given some detail on intonation in Tyneside English for example. Cruttenden and Coward (1997) have looked at intonation patterns in Mancunian English, and recent work by Lowry (1997) focuses on the use of different patterns in Belfast English.

However, such studies have concentrated on particular dialects, and any comparisons that are drawn therein tend to be comparisons between those dialects and RP. A systematic cross-dialect study concentrating on intonation is still lacking.[1] Nor would it be easy to make comparisons between say Tyneside English and Mancunian, on the basis of work already done, since the studies I have mentioned used completely different approaches, employing different transcription methods for their data, and different methods and materials to elicit that data in the first place. Therefore, although their findings are interesting, and perhaps indicate areas of interest for us to look at, they do not facilitate comparative work, and we are still no closer to being able to describe the prosodic differences between dialects of British English.

In this paper I discuss a project which is designed to help fill this gap. I intend to concentrate here on the methodology that has been designed for the project, since it is clearly differences in methodology that go at least some of the way to explaining why existing studies do not allow interdialectal comparison. I hope to show how approaches to the study of intonation often demand a different methodological approach to those with which dialectologists may be more familiar, and how methodology and materials have been designed with these concerns in mind.

2. *Methodology*

The project is called *English Intonation in the British Isles*. It is ESRC funded for three years and began in October 1997.[2] Data will be collected from seven different

dialects of British English, and made available to other researchers on CD-ROM. The dialects in the corpus are General Southern British,[3] Leeds, Belfast, Dublin, Newcastle, London Jamaican English, and London Gujerati English.[4] On the CD-ROM, data will be labelled for intonation patterns, using a transcription system developed for the project (see Section 4).

Much of the work on the project so far has involved the development of a suitable methodology for the recordings. As I have already stressed, the aim of the project is to provide comparable data from the different dialects in order to facilitate the kind of comparative work that has so far largely been lacking in this field. In order to do this, the approach and materials have to be consistent across the different dialects. It is also the case that the focus on intonation patterns in particular presents certain difficulties that must be taken into account in the planning of the data collection.

Firstly the choice of which dialects to focus on has been made according to several different criteria. All of the recordings will be made in more or less urban areas. Since early diachronic dialectology necessarily focused on rural areas and traditional dialects, such study tended to ignore varieties spoken by the majority of the population who live in towns and cities. This is particularly true of a heavily urbanised country like England, for example, where some 90% of the population are urban dwellers.

Some of the dialects chosen have been selected because other work on them has indicated that there are interesting features of intonation there, which seem to contrast with patterns found elsewhere, even if no direct comparision has been possible thus far to test this. There are many comments in the literature about the tendency in northern urban areas to use rises in simple statements, for example, where in General Southern British (henceforth GSB) a fall would be expected. Cruttenden, for example, cites these rises as a feature of Urban Northern British, and prevalent therefore in cities such as Belfast, Tyneside, Glasgow, Birmingham and Liverpool (Cruttenden 1997: 133). Several different types of rise are found. In Glasgow, for example, there is a simple rising glide on the final pitch accent, as shown in Figure 1 for the sentence *Peter is here* (the capital letters show the location of the pitch accent).

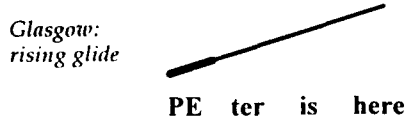


Figure 1

Data from Belfast English shows a different pattern however. Here there is a rise on the following unaccented syllable, and the maintenance of this pitch level on succeeding unaccented syllables (see, for example, Lowry 1997). This pattern is shown in Figure 2. The pattern is therefore usually described as a rise-plateau. Interestingly it is the mirror image of the default pattern on statements in RP, as can be seen in Figure 3. Cruttenden reports that the rise-plateau pattern is also characteristic of Tyneside speech (1997: 133).

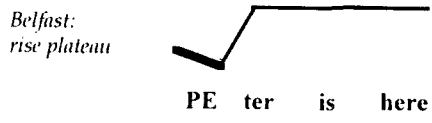


Figure 2

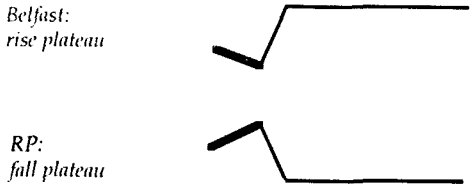


Figure 3

Belfast and Newcastle have therefore been chosen to form part of the new corpus. Using the data recorded, the patterns in these dialects will be described, and a close comparison of the rise patterns used in different dialects will be possible. One aim of the analysis will be to see whether the rise patterns found in the different dialects do indeed represent the same phenomenon.

3. *Materials for Data Collection*

A number of considerations have also been necessary in the development of materials. Firstly, a large number of possible intonation patterns needs to be represented in the data. There is no sense in having data that consists of statement patterns alone for example.

The recordings take the form of a structured interview in five stages. The first three are individual tasks.

3.1 *Sentences*

The subject is first required to read a series of sentences. Included are various types of questions for example, such as Wh-questions, questions marked by inversion, and questions without morphosyntactic markers. Also included of course are simple statements:

EXAMPLES:	*Wh-Question*	Where is the manual?
	Inversion Question	May I leave the meal early?
	Unmarked Question	You remembered the lilies?
	Simple Statement	We live in Ealing.

Additionally, we have a set of sentences which tests for phonetic differences in the realisation of particular patterns. Data from Swedish has shown that different dialects of Swedish differ in what happens when a particular pitch movement is produced on a word which is very short. On very short words, some Swedish dialects compress pitch movements – when there is little time for a particular movement, the movement is speeded up so it can be finished anyway. This is known as *compression* and is shown in Figure 4.

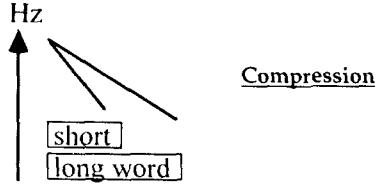


Figure 4

In other Swedish dialects, we find a different solution to this problem – when there is little time, the movement is simply not completed. This is referred to as *truncation*, as shown in Figure 5.

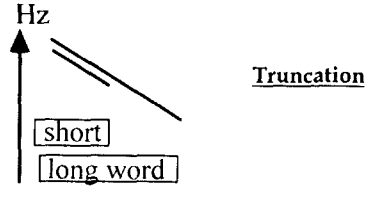


Figure 5

So far, nobody has tested whether similar effects are found in different dialects of English. Sections such as the following have therefore been included in the materials to test for this:

> Anna and Peter were having dinner. Peter said: "Who do you think I met in the market today, Anna?
> "Mr. Sheafer!" she replied.
> "Mr. Sheaf!" she replied.
> "Mr. Shift! " she replied.

Subjects read the first paragraph setting the scene, and then each of the possible replies in turn. A falling contour is expected on the surname, and the amount of material on which this fall can occur is varied. In the first example there are two syllables, in the second only one syllable with a long vowel, and in the third one syllable with a short vowel.

3.2 *Reading Passage*

The second individual task is to read a story. This is a version of the Cinderella story, an extract of which is given below:

> Once upon a time there was a girl called Cinderella. But everyone called her Cinders. Cinders lived with her mother and two stepsisters called Lily and Rosa. Lily and Rosa were very unfriendly and they were lazy girls. They spent all their time buying new clothes and going to parties. Poor Cinders had to wear all their old hand-me-downs! And she had to do the cleaning!

A reading passage is a good starting point for intonation analysis, because the boundaries of intonation phrases can be determined with a higher degree of certainty than in spontaneous speech for example. Using a passage which is normally read to children is useful because such speech is often produced relatively slowly, and with an extended pitch range. This helps when it comes to identifying the inventory of intonational patterns in a particular dialect.

In the design of the passage and other materials, consideration has also been given to the type of patterns we **expect** to find in the different dialects, and the

materials have been designed to give scope for these. One such case concerns the possible deaccenting of old or given material. In GSB when a repeated item functions as old or given information, it does not generally receive a pitch accent when mentioned for the second time. This can be seen in the following example:

General Southern British:

I went to the shop to buy MARS BARS, but they'd totally run OUT of Mars Bars.

In GSB, the second use of *Mars Bars* does not receive a pitch accent, and the final pitch accent occurs earlier in the intonational phrase, here on *out* (shown by the use of capitals). Deaccenting is obligatory in GSB in this context, and appears to be so for most dialects of British English (Cruttenden, forthcoming). However, some varieties do not deaccent old information. One example is Singapore English, as recent work by Low (1998) has shown. Deaccenting in this context is also reported to be absent in Caribbean varieties of English. Therefore for speakers of Singapore English or Caribbean English, in the same example, the second mention of *Mars Bars* would receive a pitch accent:

Caribbean English:

I went to the shop to buy MARS BARS, but they'd totally run out of MARS BARS.

London West Indian varieties of English are heavily influenced by Caribbean English, and therefore we will be looking to see whether such failure to deaccent is evident in the data from London Jamaican English. In order to test this, sentences such as the following are included in the passage:

> They were in a bad mood. They'd wanted to buy some new gowns, but their mother said that they had enough gowns.

3.3 *Retold Story*

The third individual task is the retold story: subjects have to retell the Cinderella story using only a series of pictures as a reminder. Since speakers are no longer simply reading aloud, a more spontaneous form of speech is to be expected here. The

need to concentrate on the task in hand also means that they are likely to be paying less attention to their speech.

3.4 *Paired Subject Tasks*

3.4.1 Map Task

The first task to be completed by subjects in pairs is a map task. In this, subjects are each given a map. One subject has a map as shown in Figure 6, with a route marked. The other has no such route marked. The speaker with the route must describe this route, such that the other speaker can draw it on to their map. This is an interactive task, and to ensure that it does not consist simply of a monologue from the speaker describing the route, there are several differences between the two maps. For example, whereas one speaker has a location named *John's Arms*, the other has *Ann's Arms*. When this point on the route is reached, the difference should motivate discussion. The maps are also designed with the elicitation of different prosodic phenomena in mind. For example, this particular feature of the map may elicit contrastive stress patterns, e.g. "I haven't got ANN'S Arms, I've got JOHN'S Arms".

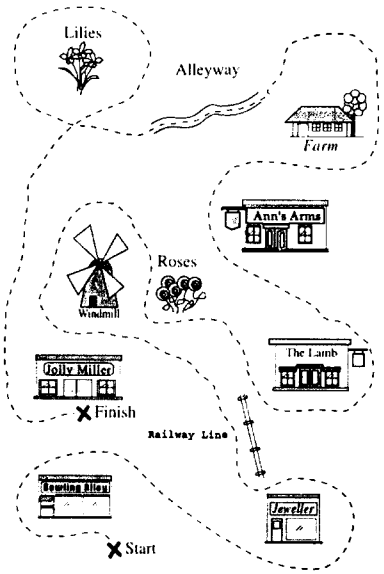


Figure 6: Example of a map used in map task

3.4.2 Free Conversation

Finally, in order to elicit spontaneous speech, there is a free conversation section. The speakers, still in pairs, are given a topic, namely the recent publicity on tobacco advertising and smoking, and are asked to discuss whether they feel it has affected them. We expect the lowest level of formality here, and as it follows the map task, and is with a friend, we hope that speakers are less intimidated by the situation so that less style-shifting will occur. We hope to supplement the information on intonation patterns that has been gathered in the rest of the interview here. One possible problem is the relative lack of control over the material: the bulk of the data may consist of simple statements for example. However, free conversation does offer interesting insights into the way in which intonational patterns vary in different speech styles, and will show whether there is any parallel to the kind of connected speech processes which affect segmental phenomena in less formal styles.

4. *Transcription System*

As already mentioned, previous studies of the intonation of individual dialects have also often employed different methods of transcribing the patterns found. This creates a further problem for those attempting comparative work. Until recently there was no widely accepted method of transcribing intonation. The development of the ToBI system in recent years has gone some way to filling this gap (see, for example, Silverman et al, 1992).

ToBI stands for 'Tones and Break Indices'. In this labelling system contours are described not in terms of pitch movements, as in the British tradition, with its rises and falls for example, but in terms of targets of H(igh) and L(ow) tones. These targets combine to form pitch accents. Intonational boundaries may also be marked by tones. ToBI was devised in the first instance for the description of American English, and it is acknowledged in the guidelines for labelling that modified variants will be necessary for the description of some other varieties. Modifications have indeed been suggested, such as Glasgow ToBI, in a paper by Mayo et al (1997). It is therefore recognised that some modification of the existing system will be necessary for the modelling of patterns in other varieties. For this project in particular, it is imperative that the system used be neutral enough to allow all the dialects to be transcribed in a way that is susceptible to common interpretation. Only then can comparative statements be made.

It appears, however, that the ToBI system is unlikely to fulfil this demand. As

Nolan and Grabe (1997) have pointed out, problems for cross-dialectal transcriptions arise particularly when it comes to the transcription of pitch movements at intonation phrase boundaries. In the ToBI system, every significant pitch event in English is transcribed with a tone, either high or low. This means that every intonation phrase boundary ends either high or low. Now, this notion is relatively unproblematic, as long as you stick to GSB, as can be seen in Figure 7.

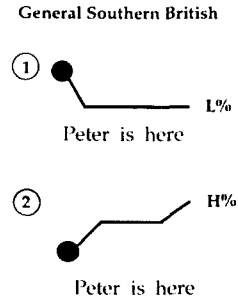


Figure 7

The two patterns in GSB can be transcribed, one with a L% boundary tone, and one with a H% boundary tone. However, as has already been noted, there is a pattern which occurs in Belfast English known as the rise plateau. This means that, in Belfast English, there are three boundary options: low, high and 'no change'. The ToBI system cannot capture the options in GSB and in Belfast English in any comparable way. Therefore in the system being used for the current project, which is based on work by Esther Grabe (1998), boundaries can be unspecified for tone. In other words, there is an extra option in our system, namely a 0% boundary, in order to cope with such patterns. The patterns found in Belfast, and the transcription of the boundary tones, are therefore as shown in Figure 8.

5. *Conclusion*

The project is in its early stages. So far data has been collected from GSB, Leeds and Belfast. Already it is clear that the materials have been successful in eliciting a variety of patterns. Figure 9 gives an example of what will be available on the forthcoming CD-ROM. At the top is a waveform, and at the bottom is a fundamental frequency trace (fundamental frequency is the acoustic measure which corresponds to

the changes in pitch which we hear). In the centre there is a labelling window, with several tiers. First there is an orthographic tier where the words spoken are labelled. The second tier is a rhythmic tier, marking stressed and accented syllables, the stressed syllable being enclosed by angled brackets. The next tier is the auditory phonetic tier, where the labels describe the perceived pitch level on the accented syllable (marked with capital letters H/M/L) relative to the immediately preceding syllable and the immediately following syllable (marked with small letters h/m/l). The IViE tier contains the transcription of the intonation, with IViE standing for *Intonational Variation in English*. Finally there is an extra tier for miscellaneous comments, such as here that the peak in the F0 contour is relatively late within the accented syllable. This tier is also used for comments on features such as errors in the reading, or hesitations, laughter etc.

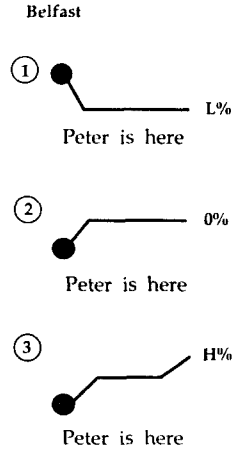


Figure 8

To sum up, this new project has several goals. The first is to provide corpus data from seven dialects of English. Since the same methodology and materials have been used for each dialect, the data will be directly comparable. Secondly, the data will be labelled for intonation patterns in the different dialects, giving both auditory phonetic information on the pitch contours and a phonological transcription of them. Thirdly, a CD-ROM of the labelled corpus will be produced and made available. This will be invaluable to researchers working on intonation and other aspects of dialectal diversity.

This paper has illustrated how the relative scarcity of work in the field so far is due to the previous lack of a suitable metholodogy on which to base comparative

studies. This is in part due to the particular difficulties posed by a study of intonation rather than segmental phenomena. However, such difficulties are not insurmountable. The approach developed for this project will hopefully serve as a model for future work in this interesting and underresearched field.

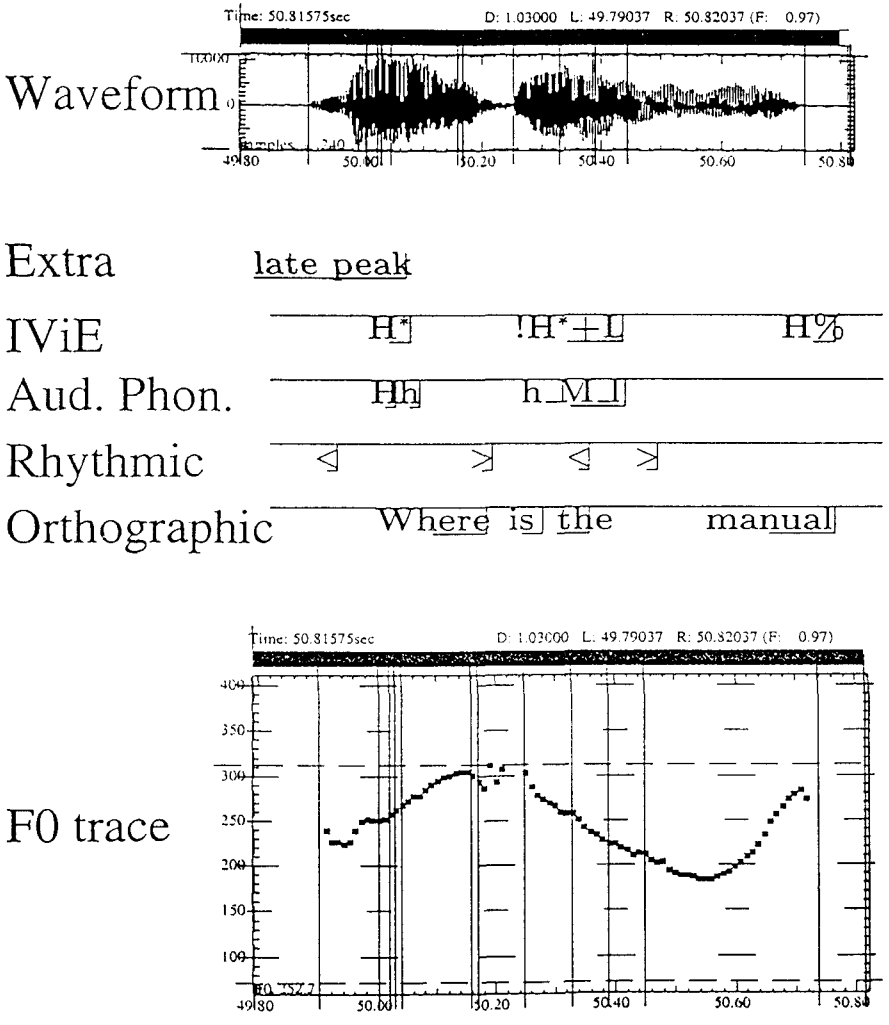


Figure 9: Example of database

NOTES

[1] One possible exception is Rohrer (1952). However, the scope of Rohrer's work is small: although seven dialects are investigated, very little data from each is examined, and the technology available was also limited.

[2] This project is supported by ESRC award R0002237145 to F. Nolan and E. Grabe.

[3] The term *General Southern British* is used to describe the pronunciation which serves as a prestige norm in South East England, but which is not closely tied to upper class groups, unlike RP, as traditionally defined.

[4] The last two terms are used for the native varieties of English spoken by second generation immigrants to Britain, whose parents are of Jamaican or Gujerati extraction.

REFERENCES

Cruttenden, Alan. 1997. *Intonation*. Second edition. Cambridge: Cambridge University Press.

Cruttenden, Alan. forthcoming. The deaccenting of given information: a cognitive universal? To appear in Bernini, G. (ed) *The Pragmatic Organisation of Discourse*. Berlin: Mouton de Gruyter.

Cruttenden, Alan and Coward, Louise H. 1997. *The Intonation of Mancunian English. Final Report – February 1997.* Unpublished report. University of Manchester.

Grabe, Esther. 1998. *Comparative Intonational Phonology: English and German.* PhD thesis. MPI Series in Psycholinguistics 7, Nijmegen, The Netherlands.

Local, John. 1986. Patterns and problems in a study of Tyneside intonation. In Catherine Johns-Lewis (ed) *Intonation in Discourse*, pp. 181-97. Beckenham: Croom Helm.

Low, Ee-Ling. 1998. *Prosodic Prominence in Singapore English*. Unpublished Ph.D. thesis, University of Cambridge (for further details, consult http://www.cus.cam.ac.uk/~ell1000/).

Lowry, Orla. 1997. *Intonation Patterns in Northern Irish English.* Unpublished M.Phil. thesis, University of Cambridge.

Mayo, Catherine, Aylett, Matthew, and Ladd, D. Robert. 1997. Prosodic transcription of Glasgow English: an evaluation study of GlaToBI. In Antonis Botinis, Georgios Kouroupetroglou, and George Carayannis (eds) *Intonation: Theory, Models and Applications. Proceedings of the ESCA workshop on Intonation*, September 18-20, 1997, Athens, Greece.

Nolan, Francis and Grabe, Esther. 1997. Can 'ToBI' transcribe intonational variation in British English? In Antonis Botinis, Georgios Kouroupetroglou, and George Carayannis (eds) *Intonation: Theory, Models and Applications. Proceedings of the ESCA workshop on Intonation*, September 18-20, 1997. Athens, Greece.

Pellowe, John and Jones, Val. 1978. On intonational variability in Tyneside speech. In Peter Trudgill (ed) *Sociolinguistic Patterns in British English*, pp. 101-21. London: Edward Arnold.

Rohrer, Fritz. 1952. *Untersuchungen zur Intonation der Dialekte von Dorset, Gloucester, Westmorland, Northumberland, Yorkshire, Lincoln und Norfolk*. Gutenberg, Lachen: A. Kessler.

Silverman, Kim, Beckman, Mary E., Pitrelli, John, Ostendorf, Mari, Wightman, Colin, Price, Patti, Pierrehumbert, Janet and Hirschberg, Julia. 1992. ToBI: a standard for labelling English prosody. In *Proceedings, Second International Conference on Spoken Language Processing* 2, pp. 867-70. Banff, Canada.

Wells, John C. 1982. *Accents of English*. Cambridge: Cambridge University Press.

For further details of the IViE project, and regular updates on progress, please consult our website at: ***http://www.mml.cam.ac.uk/ling/intoproj.HTM***

Address
Kimberley Farrar, Esther Grabe, and Francis Nolan
Department of Linguistics
University of Cambridge
Sidgwick Avenue
Cambridge CB3 9DA
England
<kip1003@cam.ac.uk>

The First SuRE Moves: Early Steps Towards a Large Dialect Project

Paul Kerswill, Carmen Llamas and Clive Upton

Abstract

Two factors have led to a sea-change taking place in the field of dialectology during recent times. The first is a considerable broadening of research aims to include models of the diffusion of changes through both geographical space (geographical spread) and social space (permeating different social groups at different times). The second is the very recent use of digital technology.

This paper argues that the time has come for a new survey of British and Irish English to be instituted, taking account of new orientations and methodologies. Firstly we propose continuous recording, to document the simplest facts of language variation over a wide geographical area in a way that will be useful to future linguists. Further, in a more elaborately structured enquiry of more restricted time-span, it should be possible to provide clear and detailed outcomes directly relating to current issues concerning the diffusion of language change.

In order to undertake the proposed large-scale survey of regional English, data must be obtained which are analysable on three levels of variation: phonological, grammatical and lexical. This paper outlines a new method of data elicitation which will prove to be a workable, effective and relatively simple way of obtaining data which allow for a complete picture of regional variation found throughout the British Isles at the turn of the Millennium and onwards.

1. *Introduction*

The 20th century has seen unprecedented changes in the regional speech of the British Isles. The industrialisation of Britain in the 18th and 19th centuries led to the

establishment of new, urban varieties of English in all its cities and large towns. It was recognition of the upheavals caused by the rapid pace of change which led Joseph Wright (1905: vii) to assert that by 1925 it would be 'quite impossible to get together sufficient pure dialect material to enable any one to give even a mere outline of the phonology of our dialects as they existed at the close of the nineteenth century'. The varieties created were in some sense a compromise between the dialects of the people who migrated there from the respective hinterlands and from further afield. However, in a time of unprecedented geographical and social mobility, this process of the development of 'compromise' dialects seems to be even stronger at the close of our present century. Not only are there now very few speakers of genuine rural dialects, but the overwhelmingly urban speech heard in the streets today is being 'levelled' in the direction of a set of relatively uniform varieties which have a clearly identifiable regional flavour but which are relatively difficult to pin down to a more specific locality.

It is the argument of the authors of this paper that it is now time to begin a two-pronged assault on a deficiency which exists in our knowledge of the contemporary situation. Firstly, and at bottom, we propose a process of continuous principled recording to document the simplest facts of language variation over a wide geographical area, in a way that will be useful to future generations of linguists. Further, in a more elaborately structured enquiry of more restricted time-span, it should be possible to provide clear and detailed outcomes directly relating to current issues concerning the diffusion of language change.

The previous network surveys of regional speech in Britain are the Survey of English Dialects (Orton and Dieth 1962-71), the Survey of Anglo-Welsh Dialects (Parry 1977, 1979), the Linguistic Survey of Scotland (Mather and Speitel 1975), and the Tape-Recorded Survey of Hiberno-English Speech (Barry 1981). To these can be added recent and ongoing endeavours of restricted theoretical scope or geographical spread, such as the Survey of British Dialect Grammar (Cheshire et al. 1989), Kerswill and Williams' (1997) work on the role of adolescents in dialect levelling, and the Tyneside and Derby study (Docherty et al. 1997). Each of inestimable value in its own way, these widespread or more focussed projects are of different times, have different aims, and employ different methodologies: together they do not amount to a cohesive record of the dialects.

To the fact of the fragmentary nature of our holdings of information can be added the fact that dialectology has quite recently undergone two sea-changes. Firstly, there has been a considerable broadening of its research aims to include models of the diffusion of changes through both geographical space (geographical spread) and social

space (permeating different social groups at different times) (Butters 1997). Dialectologists have used their data to model types of language change in a way entirely compatible with Labovian sociolinguistics (Bailey, Guy & Wikle 1993); indeed, it has recently been argued that dialectology must be regarded as part of sociolinguistics (Butters 1997: 11). There has also been the use of models of geographical diffusion to account for patterns of spread (Trudgill 1974; Upton 1995; Hernandez-Campoy 1996).

Secondly, there have been enormous technological advances in the recording, storage, sorting, and retrieval of data. Quite early use of such technology is exemplified by the Computer Developed Linguistic Atlas of England (Viereck 1991; Ramisch 1997), a project which has entered some of the Basic Material (Orton et al. 1962-71) of the SED onto a database allowing sophisticated displays of individual linguistic features and, importantly, displays showing overall differences between locations measured in terms of a large number of features. It is now finding further expression in the work of Elmer and Rudin (1997) and Schiltz (1997). Such scholarship as theirs shows what can be done with material not designed with computerisation in mind, and hints at the possibilities for enquiries which presume upon the digitised storage of speech signals, the tagging of speech samples for automatic retrieval, and the facility of remote retrieval of stored data.

Large-scale surveys of regional speech in this country have not before had the benefit of full computerization. The obvious UK point of reference, the British National Corpus (1994), did not intend to achieve any degree of representativeness as a sample of regional speech, and a more directed approach is needed.

2. *SuRE: The New Project*

The project being embarked upon, the Survey of Regional English (SuRE), will create from the outset a large computer-held corpus whose form will be guided by the need for it to be the object of analytical work addressing current research questions, those to do with levelling, while at the same time being sufficiently broad to allow to be addressed research questions which may arise in the future.

Having regard to issues of current and continuing theoretical linguistic interest, and to the technologies available to us, active steps are under way to put in place a double-banked project. Funding is being sought by a Leeds/Sheffield/Reading axis intensively to survey the speech of a planned network of British and Irish localities. In the meantime, a doctoral research grant has been secured, part of the brief of the holder

being to advance thinking on that method and, in particular, to generate ideas for rapid data-collection which can be implemented both within and beyond a planned network and limited time-frame.

The 'double-banked' nature of our approach is fundamental (and here it seems reasonable to follow Orton and Wright's Introduction to *A Word Geography of England* by venturing into military metaphor). We see the need in the most detailed way possible to reconnoitre in strength the dialects as they exist at the turn of the Millennium: for this reason a 'heavy brigade' approach is required, targetting a finite set of localities with a comprehensive, tightly structured campaign. Such a foray is logistically demanding, however, and cannot be sustained indefinitely. For this reason, a 'light brigade' action is envisaged: this can be sustained indefinitely, with field linguists super-adding a simple technique of elicitation to any methodology which they are employing, in any locality and at any time, so that a bank of data accumulates. This nugget of comparable data, properly tagged, will be the kernel of all collecting, growing to an open-ended resource capable of analyses known and as yet unforeseen.

At the moment the envisaged characteristics of the 'heavy brigade' project are as follows. Firstly, a network of pre-defined localities in England, Wales, Scotland, Northern Ireland and the Republic of Ireland will be targetted. Locality distribution will be determined essentially by density of population along lines set by the Survey of British Dialect Grammar (Cheshire et al. 1989: 190), a project which made use of a classification of locations in Britain according to, among other variables, the degree of urbanisation (CURDS Functional Regions framework, see Champion et al. 1987). In principle, this will allow every resident to have an equal chance of selection, and will reflect the urban bias of much of the British Isles. Geographical models of diffusion can be applied to the data deriving from the network. Subsets of localities will be subject either to fuller or less detailed sociolinguistic investigations.

Secondly, a standard set of items will be elicited, with some variation allowed to enable known regionally-significant variables to be investigated. The methodology for this aspect of the collection, which is currently being designed and trialled, is now outlined.

3. *SuRE: towards a new methodology*

In order to undertake a large-scale survey of regional variation in contemporary spoken British English, data must be obtained which can be analysed on three levels

of possible variation; phonological, grammatical and lexical. Although it is difficult to combine the three levels, to discount any would be to obtain an incomplete picture of the regional linguistic variation found in the British Isles at the turn of the Millennium. The phonological, grammatical and lexical data must be comparable across the localities to be studied, permitting quantitative analyses of the different levels of regional and social variation.

As the SuRE project is collaborative in nature, the problem of combining the three levels of analysis into a single data elicitation method which will be satisfactory to all interested parties is considerable. The problem is further compounded by the necessity of any data elicitation technique to be relatively quick and easy to administer. Researchers must be able to apply the methodology to their fieldwork with the minimum of prior preparation or administration superfluous to their particular fieldwork needs. With these considerations and underlying difficulties in mind, a new method of data elicitation and collection is proposed. After refinement, it is anticipated that this method will prove to be an effective and relatively simple way of gaining data which are analysable on a number of different levels, and which will be usable by researchers whatever their particular research interest.

The primary aim of the SuRE interview is to obtain informal speech from the informant from which an analysis can be made on one, or more than one, level. This being the case, the somewhat formal context of the fieldworker asking set questions to elicit lexis or grammar in an extremely lengthy interview, as in the SED (Orton et al. 1962-71), would be entirely inappropriate. Similarly, methods involving the fieldworker asking questions to elicit involving personal narratives (cf. Labov 1972), or allowing the informants to converse in pairs on topics of their choosing (cf. Docherty et al. 1997; Llamas 1998) would also be unsuitable, as the possibility of obtaining any comparable data on lexical variation would be almost completely removed.

It is necessary, then, to find a way of combining informal conversation, from which data for phonological and, to some extent, grammatical analyses are obtained, with information on lexical variation. With this in mind it is intended that the fieldworker will 'lead' a conversation around linguistic domains, ideally with socially paired informants, permitting interaction to be more like a conversation and less like an interview. The fieldworker prompts discussion about lexical items used in the given area, encouraging the informants to discuss their 'dialect' words; how they are used and what connotations they have. Whilst producing relatively casual conversation, this means of eliciting data yields valuable information on knowledge and use of lexical items, as well as revealing possible age or gender variation in lexis

within a given dialect. In the course of the conversation, how much the speakers are actually aware of variation, as well as attitudinal information on lexis and dialect, are also revealed.

4. *Sense Relation Network Sheets*

4.1 *SRNs: Visual Design and Content Design*

However, although it is successful in obtaining informal speech, simply *talking* about lexical variation does not yield comparable or quantifiable lexical data. In order to allow the information on lexical items to be comparable across the network of localities to be studied, and to give a somewhat flexible structure to the interview, Sense Relation Network sheets (SRNs) have been designed. An example of one of the SRNs, three of which form the core of the SuRE interview, is given in Figure 1 (see Appendix). Both the visual design and the content design of the SRNs are inspired by the idea that there exists a 'web of words', or a series of interconnected networks which define, delimit and store linguistic expressions in the mind (Aitchison 1994, 1996).

Visually, a network is designed as shown in the example given in Figure 1. The language domain of *Feelings, Actions & States* is broken down into subdivisions which form the network. Standard notion words are listed, and space is provided for the informant to insert a dialectal partial synonym for the standard notion word. The visual design of the SRN is aimed at making it as visually pleasing and unthreatening as possible, so the informant will actually *want* to complete it.

In terms of content, the SRNs are built around domains of language, much like the grouping of questions by subject matter in the SED questionnaire. Standard notion words are given as pointers, since interviews which use indirect elicitation techniques are much more time-consuming than those which use direct ones. Also, with an indirect question, the interaction may feel more like an interview or a test of some sort, rather than a conversation, and this may have the affect of increasing the formality of the speech of the informant. In the SRNs, the presence of an *indirect* prompt would considerably increase the density of the written input, which may result in a negative reaction from the informants. Through trialling and revision, although none of the original domains have been entirely lost, the number of SRNs has been drastically cut from eight to three, reducing both the time needed by informants to complete them and the time necessary to conduct the interview.

The SRNs then, as well as being a visual network, rather than a list of

questions, represent the interrelated network of paradigmatic and syntagmatic sense relations in which linguistic expressions from similar semantic fields define and delimit each other's meaning. They also represent the sense relation of partial synonymy, which the dialectal variant holds with the standard notion. Additionally, in time they will represent a geographical sense relation network of dialectal variation of partial synonyms found throughout the British Isles.

4.2 *SRNs: Technique of Administration, and Data Yielded*

A crucial part of the new methodology used for the SuRE interview involves the actual administration of the SRNs and the conduct of the interview. Informants are given the SRNs some five days in advance of the interview. This allows them time to consider the words they use, and lessens drastically the possibility of the mind going blank if an informant is called upon to give an immediate response to something which is not often consciously thought about. Also, if the informants are aware of what is going to be discussed, the feeling of somehow being tested may be lessened considerably. It is important that the interview is enjoyable and unthreatening, in order to ensure the ready recruiting of informants and to maximise the possibility of gaining access to their least overtly careful or monitored speech style.

The method of basing an informal recorded conversation on the SRNs allows the fieldworker to secure the written record of the informant's responses on the SRNs, which are retained by the fieldworker, and also to secure the backup of the recorded spoken version of the responses for pronunciation purposes. The spoken recording of the responses also acts as a safeguard against mis-spellings, which may indeed prove to be an interesting research exercise in itself. In the course of the interview, other lexical items not given on the SRNs may also be revealed, with informants becoming aware that they use a particular word when they hear someone else use it, or with informants using dialectal variants without necessarily being aware they are doing so. The amount of lexical data obtained through the SRNs is considerable. In terms of obtaining phonological and grammatical data, informal speech is produced through the use of the SRNs. Informants are willing to talk at length about lexis, and about attitudes towards lexical items and awareness of variation which also yields a mass of attitudinal data.

5. *Identification Questionnaire*

Combined with the SRNs an Identification Questionnaire (IdQ) is included in the

SuRE interview. An example of the questions posed in the IdQ is given in Figure 2 (see Appendix). The complete IdQ comprises 15 questions, whose primary aim is to act as a safety net: the questions posed elicit a relatively extended response, should the informant respond minimally to the SRNs. The questions on the IdQ are designed to combine informal speech with valuable information on people's attitudes towards language and identity (Le Page & Tabouret-Keller 1985), existence and awareness of age and gender differences (Kerswill 1996; Kerswill and Williams 1997; J. Milroy, L. Milroy and S. Hartley 1994; Trudgill 1974), and rudimentary ideas on density of networks (Milroy 1987). They may also elicit short personal narratives, and information on people's perception of language areas and boundaries (Preston 1988). As communities and boundaries are often symbolic, it is difficult to impose a definition of speech community onto a geographical area and a group of people, even when an investigator is a native of the geographical area to be studied. The similarities and differences which define and delimit communities are often not a matter for objective assessment, but are largely subjective, involving feelings and existing in the minds of the members of the community (Cohen 1985: 21). There is no reason why the topical content of the interview should not be of use in this regard, with the fieldworker tapping the natural resource of the informant for information on language, area, boundaries and attitudes found in Britain at the turn of the Millennium. Attitudinal information from the individual informant is yielded through the use of the IdQ. This attitudinal information will give comparable data across regions of Britain and may reveal differing regional attitudes towards areas and dialects.

There also exists the possibility of scoring the responses given in the IdQ in terms of positive, negative and neutral answers, as in Labov's Martha's Vineyard study (1972). Such quantification can be used to discover whether any correlation exists between attitudes to the area and linguistic or non-linguistic variables, again considered both in terms of the individual or social groups of informants within a given region and in terms of differing attitudes towards regional areas as a whole, on a national scale. The questions on the IdQ are also given to the informants prior to the interview, and are included with the SRNs in a small interview pack.

Ultimately included in the interview may be a word or sentence list to be read by the informant to facilitate observation of stylistic variation, and for control of phonological features. Also, a more formal grammatical element may be included along the lines of an adapted questionnaire similar to that used by Cheshire et al. (1989) for the Survey of British Dialect Grammar.

6. *Conclusion*

Two elements of SuRE are envisaged: firstly, a detailed survey designed to provide up-to-date information over a wide area, informing the linguistic community on current issues; secondly, accumulation of an uncomplicated bank of consistently collected data, rolling forward, built by and available to anyone committed to its creation.

At present the methodology envisaged for the data elicitation involves an informal interview with socially paired informants, providing the data for phonological and, to some extent, grammatical analyses. This interview is centred on the three SRNs, which will also yield data on lexical variation, and the IdQ, which will provide attitudinal and ethnographic information. This, then, can be seen as the core element of the SuRE interview, which can be expanded for the purposes of the 'heavy brigade' project, i.e. the detailed survey. A reduced version of this core will act as the 'light brigade', i.e. the bank of consistently collected data, if the whole proves to be too lengthy or complex to administer.

Knowledge of the early periods of each stage of SED shows that a final viable project is different in detail from that which is initially envisaged, and this may prove the case for the SuRE project. But what we are decided upon is that we should make every effort to follow the example of our predecessors in principled data gathering. It would be a sad reflection on us if we were now to continue to rely on the ever more dated findings of those who preceded us, without setting in place the coherent collection of data which can in some measure at least be compared with earlier findings, and which will just as importantly provide data for the instruction of those who will follow.

REFERENCES

Aitchison, Jean. 1994. *Words in the Mind: An Introduction to the Mental Lexicon.* 2nd edn. Oxford: Blackwell.

Aitchison, Jean. 1996. *Reith Lectures.* B.B.C. Radio 4.

Bailey, Guy, Tom Wikle, Jan Tillery & Lori Sand. 1993. Some patterns of linguistic diffusion. *Language Variation and Change* 5: 359-90.

Barry, Michael V. 1984. The English Language in Ireland. In R.W. Bailey and M. Gorlach (eds) *English as a World Language* 84-133. Cambridge: Cambridge University Press.

British National Corpus. 1994. Longman, Chamber-Harrap, OUP, British Library, Unit for Computer Research, Lancaster University.

Butters, Ronald. 1997. Dialectology and sociolinguistic theory. In Alan Thomas ed. *Issues and Methods in Dialectology* 1-13. Bangor: Department of Linguistics, University of Wales, Bangor.

Champion, A.G., A.E. Green, D.W. Owen, D.J. Ellin & M.G. Coombes. 1987. *Changing Places: Britain's Demographic, Economic and Social Complexion.* London: Edward Arnold.

Cheshire, Jenny, Viv Edwards & Pam Whittle. 1989. Urban British Dialect Grammar: the question of dialect-levelling. *English World Wide* 10: 185-225.

Cohen, Antony, P. 1985. *The Symbolic Construction of Community.* Chichester: Ellis Horwood.

Docherty, Gerard J., Paul Foulkes, James Milroy, Lesley Milroy & David Walshaw. 1997. Descriptive adequacy in phonology: variationist perspective. *Journal of Linguistics* 33: 275-310.

Elmer, Willy & Ernst Rudin. 1997. The Survey of English Dialects as an electronic database for research in areal and variationist linguistics. In Alan Thomas (ed) *Issues and Methods in Dialectology* 234-46. Bangor: Department of Linguistics, University of Wales, Bangor.

Hernandez-Campoy, Juan Manuel. 1996. *Modelos de difusión geografica de las innovaciones sociolingüísiticas en los acentos del Reino Unido.* Unpublished Ph.D. thesis. Universidad de Murcia.

Kerswill, Paul. 1996. Children, adolescents and language change. *Language Variation and Change* 8: 177-202.

Kerswill, Paul & Ann Williams. 1997. Creating a new town koine: children and language change in Milton Keynes. In *Reading Working Papers in Linguistics* 3: 205-58. Department of Linguistic Science, University of Reading.

Labov, William. 1972. *Sociolinguistic Patterns.* Philadelphia: University of Pennsylvania Press.

Le Page, Robert B., & Andrée Tabouret-Keller. 1985. *Acts of Identity: Creole-based Approaches to Language and Ethnicity.* Cambridge: Cambridge University Press.

Llamas, Carmen. 1998. Language variation and innovation in Middlesbrough: a pilot study. In *Working Papers in Linguistics vol. 6.* Department of Linguistics and Phonetics, University of Leeds.

Mather, J.Y., H.H. Speitel & G.W. Leslie. 1975. *The Linguistic Atlas of Scotland.* 3 vols. London: Croom Helm.

Milroy, James, Lesley Milroy & Sue Hartley. 1994. Local and supra-local change in British

English: The case of glottalisation. *English World Wide* 15:1 1-33.

Milroy, Lesley. 1987. *Language and Social Networks.* Oxford: Blackwell.

Orton Harold et al. 1962-71. *Survey of English Dialects.* Introduction and 4 vols. in 3 parts. Leeds: E.J. Arnold.

Orton, Harold & Nathalia Wright. 1974. *A Word Geography of England.* London: Seminar Press.

Parry, David (ed). 1977, 1979. *The Survey of Anglo-Welsh Dialects.* 2 vols. Swansea: University College of Swansea.

Preston, Dennis R. 1988. Methods in the Study of Dialect Perceptions. In Alan R.Thomas (ed) *Methods in Dialectology* 373-95. Clevedon: Multilingual Matters.

Ramisch, Heinrich. 1997. Dialectological and cartological features of the *Computer Developed Linguistic Atlas of England.* In Alan Thomas (ed) *Issues and Methods in Dialectology* 224-33. Bangor: Department of Linguistics, University of Wales, Bangor.

Schiltz, Guillaume. 1997. A dialectometric analysis of the dialects in the Grand-Duchy of Luxemburg. in Alan Thomas (ed) *Issues and Methods in Dialectology* 94-108. Bangor: Department of Linguistics, University of Wales, Bangor.

Trudgill, Peter. 1974. *The Social Differentiation of English in Norwich.* Cambridge: Cambridge University Press.

Upton, Clive. 1995. Mixing and fudging in Midland and southern dialects of England: the *cup* and *foot* vowels. In Jack Windsor Lewis (ed) *Studies in General English Phonetics: Essays in Honour of Professor J.D. O'Connor.* London: Routledge.

Viereck, Wolfgang. 1991. *Computer Developed Linguistic Atlas of England 1.* Tübingen: Niemeyer.

Wright, Joseph. 1905. *The English Dialect Grammar.* Oxford: Clarendon Press.

Addresses

Paul Kerswill
Department of Linguistic Science
University of Reading
Whiteknights
PO Box 218
Reading RG6 6AA
England
<p.e.kerswill@reading.ac.uk>

Carmen Llamas and Clive Upton
School of English
University of Leeds
Leeds LS2 9JT
England
<engml@leeds.ac.uk>
<c.s.upton@leeds.ac.uk>

APPENDIX

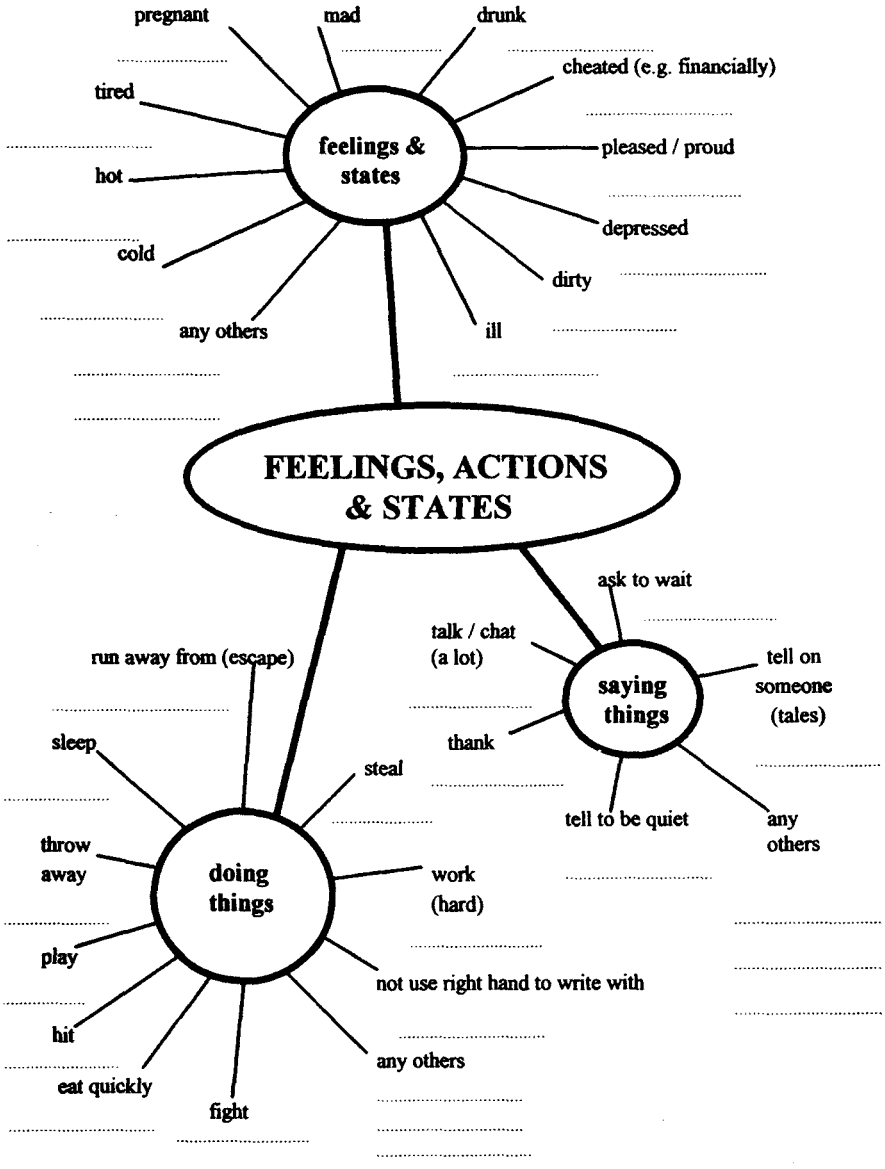


Figure 1: Sense Relation Network sheet

Your Language

- What accent would you say you had, and do you like it?
- Have you ever been in a situation where you've deliberately changed the way you talk? If so, why?
- Do you think there's a difference between how males and females speak here?
- Where, geographically, would you say people stop talking the same as you and start sounding different?

Your Area

- If you were watching a regional news programme, what places would you expect to hear news from?
- What image or description of your home town would you give to someone who didn't know it?
- If you could, would you change where you came from? Why/why not?
- What do you think the best and worst things are about growing up and living in your home town?

Figure 2: Identification Questionnaire (example questions)

The Future of Dialectology

William A. Kretzschmar, Jr.

Abstract

We rely on our teachers and admire their achievements in the first generation of dialectology, and now must decide how to proceed into the next generation. Language variation is not going away, and in order to deal with it most effectively we should consider three things: 1) our theoretical position; 2) our collection, encoding, and analysis of data; and 3) our presentation of data and results. In theory we should recognize that we are doing something different from our generativist and structuralist colleagues, and should ally ourselves with other linguists doing empirical research. We should make best use of emerging technologies in our data handling, and we should improve on first-generation procedures through use of discourse interviews, execution of valid survey research methods, and collection of speech perception as well as speech production data. We should create SGML-tagged transcripts of interviews. We should attempt to analyze our data as we go, with emphasis on quantitative methods but without reliance on any one best analytical method. Finally, we should make every effort to share our data and results with the public, which is badly in need of better information about their language.

Contemporary work in dialectology should always reflect the achievements of our teachers in the first generation of work on Linguistic Atlases. My usual work concerns the American Linguistic Atlas Project, most of whose interviews were carried out in the 1930s and 1940s, though some field work continued long after and some is even going on today. My students execute new projects on contemporary speech, but even then the past informs their best work and results (e.g., Johnson 1996). Thus it has been altogether appropriate for us to celebrate Harold Orton at this conference. The monumental labor on English dialectology of Orton and his

colleagues in England, and of Hans Kurath, Raven McDavid, and their colleagues in America, underlies what we do now. We begin with their example, and however we make our way forward we do so in respect of their monumental work.

My topic today, however, is the future of dialectology, and that kind of prediction is an unaccustomed task for me; I have therefore turned for help to our most impressive and popular contemporary vision of the future, Star Trek. I could hardly come before this audience without new field work, and the Internet provided a convenient portal to the 23rd and 24th centuries, at least as far as the Starfleet Library. I am sorry to report that the future of dialectology in the 24th century is bleak indeed. While language contact situations still do occur despite the use of a machine-based universal translator, as shown in the accented Galactic Standard English of Ensign Chekov, dialects seem almost to have disappeared. As late as the 23rd century (in the original Star Trek series), remnants of an American Southern accent still remained in the speech of Dr. McCoy, but these seem to have disappeared in the Next Generation. Scots and Irish accents seem to have become an occupational dialect for engineers like Mr. Scott, or Miles O'Brien in the Next Generation, apparently optional but preferred for job advancement. And astoundingly, what should have been dialectal French from the Next Generation's Captain Picard somehow turned into the Captain's Conservative RP – which indicates either the long-term success of the language teaching unit of Oxford University Press or some *truly* bizarre sociopolitical development in France. For dialectologists, the world of Star Trek's 24th century will not have much to offer, at least professionally; in Peter Trudgill's words, 'A world where everyone spoke the same language [or dialect] could be a very dull and stagnant place' (Trudgill 1975: 16), even among the stars.

The Star Trek Syndrome, as I name this dialect-empty prediction for my students, has its adherents in our own era. Joseph Wright felt a sense of urgency about recording English dialect features because 'pure dialect speech is rapidly disappearing from our midst' (1898: v), a sentiment shared by American dialectologists at the time and still by many people today. For instance, the 1995 annual meeting of the American Dialect Society featured a special session on 'Endangered Dialects,' and many of the papers at that session commented on 'receding' dialects, or dialects threatened with eradication. Perhaps a great many people wish that language variation *would* go away, as shown in the recent American controversy over Ebonics. However, change or even loss of dialects need not lead to the Star Trek Syndrome. William Labov has written recently about

> . . . the increasing diversity of American English [which is]

> the main finding of our research [at the Penn Linguistics Laboratory], one that violates the most commonsense expectation of how language works and is supposed to work. In spite of the intense exposure of the American population to a national media with a convergent network standard of pronunciation, sound change continues actively in all urban dialects that have been studied, so that the local accents of Boston, New York, Philadelphia, Atlanta, Buffalo, Detroit, Chicago, and San Francisco are more different from each other than at any time in the past. . . . Though the first findings dealt with sound change in Eastern cities, it is now clear that it is equally true of Northern, Western, and Southern dialects.
>
> (1997: 508)

This is good news for dialectologists, even if it does not feel like commonsense for some people (cf. Kretzschmar 1997). Language variation remains strong; however much Standard English may be preferred in the media and the schools, it has not driven out varieties that belong to local communities. We see the truth of this on a larger scale in the innumerable English voices that we come to hear as English continues its progress as *the* world language. The fact is that particular dialects change and come and go, just as particular languages change and come and go, so that what is constant is the presence of dialects and languages, not the fixed existence of any particular dialect or language. Indeed, there is evidence that salient features of many a dialect that we have taken for granted, like American Southern, are actually of recent vintage (Bailey and Ross 1992, Bailey 1997). From the situation as we observe it in the 20th century, so long as we get outside of the classrooms and the suburbs and the new towns and the newspaper columns, there seems to be little danger of progress towards the Star Trek Syndrome – and no threat to our job security.

The question of the future of dialectology, then, is not whether there will *be* any language variation, but rather what we should do about it. Whatever else we choose to do, there will be no replacement for having highly-trained field workers talk to speakers as part of a planned survey. This is a distinguishing characteristic of dialectology, as our field may be separated from the essentially structuralist approach of Labovian sociolinguistics or from the formal logic of generative linguistics (more about these soon). The original impulse of Wenker and Gilliéron in the first wave of Atlas studies was sound: there is great value in finding out what real people actually

say. And that takes real labor, as many here know all too well. Dialectology is not the kind of thing that one can do solely from a comfortable study, and stories from the field are legion. My students still laugh, for instance, when I tell them about the American field worker, Guy Lowman, who was chased from a Virginia farmhouse at the point of a pitchfork for daring to ask the woman of the house about names for male farm animals. We can look forward to yet more stories in the future. However, to go along with the basic necessity of field work, we have new technology to consider, and we form a party in the continually changing flow of ideas about dialects, and more generally about linguistics, and we need to take account of technology and theory as we imagine our future. We need to consider where we stand with regard to three major points: 1) our theoretical stance; 2) our collection, encoding, and analysis of data; and 3) our presentation of our results. I would like to address each of these points in turn.

The first thing to say about theory for dialectologists in the future is that we should claim one. In my own training with Raven McDavid, explicit talk about linguistic theory was minimal. For instance, when I once seriously raised the issue of why an isogloss should have been drawn where Raven drew it, all he replied was that that was where it ought to go. When American dialectologists have turned their minds to theory in years past, as for instance in Davis's work on diafeatures (1973), the prevailing climate did not really allow such thinking to get very far. Everybody much preferred to talk about 'methods' – and so the word 'methods' is part of the title for the triennial conference in dialectology, the International Conference on Methods in Dialectology. This is not to say, of course, that Orton, Kurath, and McDavid did not actually have any theory, only that the theory was most often left unstated. Orton and Dieth were explicitly interested in modern dialects as evidence for study of historical varieties of English, as well as in modern regional differences (Orton, Sanderson, and Widdowson 1978: Introduction). Raven McDavid confirmed this affiliation with historical linguistics when he wrote that 'Dialect geography is a venerable subfield within the new science of linguistics, and is basically a branch of historical linguistics' (McDavid et al. 1986:390). That said, the practical problems of doing the work were the most important consideration for our teachers.

Once could say that this concentration on methods and historical linguistics insulated dialectology from the mid-century ferment in structural linguistics and the somewhat later battles over the transformational paradigm. Historical linguistics has remained in large part a traditional discipline, the only area of linguistics where lists of works cited still typically include many titles from the nineteenth century. One might also say that dialectology was not insulated but isolated by that view, rendered

marginal, because it did not or could not marshal all of its tremendous weight of evidence to participate in the linguistics wars of the 1950s and thereafter. Whether or not this was the cause, we do in fact find ourselves to be isolated and marginalized, perhaps more in America than here in England or in Europe. My only regret among the successes of the previous generation of dialectologists is this apparent retreat from the main theater in linguistics, to my way of thinking without effectively having defended dialectology from attacks of irrelevance (as by Chomsky) or of faulty and inappropriate procedures (as by some sociolinguists, notably C. J. Bailey).

The sort of defense in theory that might have been mounted has been outlined by Lee Pederson, who has written that dialectology is

> a logically ordered and systematic approach that begins with common sense, proceeds through deductive cycles, and concludes in enumeration. It conducts research in a geographic context, but its research concerns a few words of a language, not the language itself and its universe of discourse. . . . American dialectologists, for example, concentrate on sorting and counting components – American English synonyms, morphs, and phones. They are not concerned with the identification of new linguistic classes, semantic, grammatical, and phonological sets established according to the scientific method. . . . In word geography, [deduction] concerns the engagement of target forms. It takes them first as contrastive lexical sets, and then carries the work forward through segmentation of self-evident morphemes, phonemes, allophones, and distinctive features, according to the needs of a descriptive problem. Taken this way, deductive word geography studies only classes and components of phonological words as they characterize speakers classified according to geographical place and analyzed according to social factors.
>
> (1995: 35-36)

What a concise description this is, and how well it captures the essentially French, Gilliéronian tradition of dialectology that always has informed the American Atlas underneath the German layer of isoglosses and dialect areas (cf. Kretzschmar 1995). I completely agree with Pederson that 'common sense itself suggests the importance of these considerations . . . because American word geographers have so far given them little attention, "disdaining as they do to turn their minds to such simple things . . ."'(35). Simple in Pederson's formulation, but crucial to express and to

defend as a way of studying language.

In future, I think that we should affiliate ourselves with the emerging area of empirical linguistics. Empirical linguistics, I would argue, is an alternative to the structural and generative paradigms, and it develops from the example of Gilliéron. To illustrate the difference, let us consider the basic question of how to make a generalization about a language from each point of view. A structuralist is interested in creating a generalization about a language or dialect, often to create a dictionary or grammar, on the basis of what real speakers *say*. However, a structuralist might interview only one real speaker, or just a few speakers, in the belief that every native speaker shares in the linguistic system of the language or dialect. Some of our colleagues, the sociolinguists, follow the structuralist model of making generalizations about a group based on evidence from a small number of speakers. A generativist, alternatively, is interested in creating a generalization about a language or dialect, usually a grammar, on the basis of what a speaker *thinks*. While the generativist would agree with the structuralist that each speaker shares in the linguistic system of the language under study, and so just one or a few speakers can serve as the basis for a generalization, the underlying rules that generate real speech are the target. The generativist needs to test examples of speech against native speaker intuitions, not usually to collect and multiply examples of real speech. So, structuralists are interested in speakers' *memory* of language, while generativists may be said to be interested in speakers' *processing* of language. Both of them make general statements about language, most often from rather restricted evidence, in accord with the axiom of systematicity of the language or dialect.

The empirical linguist, on the other hand, does not necessarily assume that each native speaker shares the same linguistic system, or conversely that speakers possess only one linguistic system or inventory, and instead wants to collect great quantities of real speech from a great many speakers in order to describe what people actually say. As Pederson earlier represented the Gilliéronian point of view, the empirical linguist is interested in 'a few words of a language, not the language itself and its universe of discourse.' Empirical linguists typically employ the grammatical categories postulated by structuralists and generativists, but they test each category empirically to assess its reality in use. Empirical linguists also test the distribution of words, whether as lexical units or as they embody morphological markers or pronunciations, not necessarily as elements in a contrastive system but for themselves, to observe the dynamics of real speech by real people in samples taken from whole regions or communities. When an empirical linguist makes a generalization, it boils down large quantities of speech from many sources, as opposed

to the structural or generative prediction of the speech of the group on the basis of one or a few individuals. Clearly, some of our colleagues in sociolinguistics follow this empirical model more closely than they do the structural or generative models. Corpus linguists are also members of the empirical group, and the study of corpora has, especially here in England, proven to have a number of industrial applications. Empirical linguistics is not a replacement for structural or generative linguistics; it is just different, and there is room for all sorts.

In the future of dialectology, I hope that we will understand that what we do is not the same as what our structuralist and generativist colleagues do – and I hope that each of our students will understand how and why it is different on theoretical grounds.

Let us turn now to collection, encoding, and analysis of data. As we enter our future, we need to take best advantage of modern recording and computing resources. Technology does not determine what we do, but it does offer particular opportunities that we cannot afford to ignore. In the last generation, the great technological advance was the portable tape recorder. Dialectologists have come to use it routinely, but, as Raven McDavid has written in his last-published book review, it has never been a cure-all (1985: 60):

> I cannot be completely objective about my own experience, but for all my early ineptitude, I think I developed a large measure of cooperation from those I interviewed, but then, I knew I was undertaking something different from my usual experience, and I dearly wanted it to work out. At the beginning, I don't think the tape recorder would have made me a better field worker. For all its convenience, it cannot substitute for interest, imagination and training, and the determination to do a good job. It can improve the work of even the best field worker by providing a permanent record and picking up far more evidence than the best can record with the unaided ear, but no student should take the tape recorder out for his interviewing without previously demonstrating the ability to get along without it.

This advice certainly applies again now, as we continue to try to cope with the computer as a technological aid, and we find ourselves in the same situation with computers as McDavid described for the tape recorder. Successful integration of very human, ethnographic data collection with computer encoding and analysis will continue to be a pressing need.

Two particular aspects of data collection are not directly linked to technology but still seem clearly to be the way of the future. First is the nature of the dialect field interview. Our colleagues the sociolinguists have convinced just about everyone, including most of us, of the value of discourse-based interviews as opposed to the long-question-with-a-short-answer method. Their arguments about interview style and its effects on results are compelling. Still, we need not abandon our wish to elicit particular words or other usages. Lee Pederson has described a modified interview technique that addresses our need for specific elicitation targets in the context of a discourse-style interview (1996a, 1996b; Pederson and Madsen 1989). He offers 'a framework for a tape-recorded interview' of approximately three-hour duration, composed of four 45-minute sections, which has already received extensive use in on-going field work in the Western states. There are 360 elicitation targets, but these are embedded under a dozen topical headings that allow the field worker to introduce the general topic and to direct conversation within it, rather than to ask pointed questions. The resulting interview should be transcribed in full, with tags in the transcription to identify elicitation targets but with full opportunity to preserve the continuous speech of the speaker for analysis of verb form frequency and other discourse features (1996a: 54-59). Interviews on this model are not 'danger-of-death' performances, but they do establish a consistently informal interview style, and they should be directly comparable with the large body of sociolinguistic discourse data obtained under similar, informal interview conditions.

Another very important aspect of data collection for dialectology is the emergence since the planning of first-generation Linguistic Atlases of reliable survey research methods. This is not the time for me to recapitulate what my colleagues and I have written about at length (Kretzschmar et al. 1993, Kretzschmar and Schneider 1996), so I will not offer any detailed argument here. Suffice it to say that I believe that we have no choice but to employ valid survey research methods. We should take no short cuts. We should not follow traditional speaker selection methods just because our teachers used them. If what I have argued earlier is true, that the distinction of dialectology from competing linguistic theories is based upon large-scale, planned, empirical survey research, then we simply must adopt the methods that are accepted, even required, in the other modern empirical sciences. We need not give up anything to do so. There is no conflict between the randomization techniques of survey research and the traditional goals of dialect surveys to achieve balanced regional and social coverage. All we have to do is learn enough about modern sampling techniques to do within the accepted parameters whatever we wanted to do anyway. And we need not give up historical comparison with the first generation of Atlas studies to do so, as

Ellen Johnson's book on the Southeastern lexicon has demonstrated with her statistically-valid paired-sample survey, matching new 1990 interviews with 1930s interviews (1996). Since we can very well apply currently-accepted scientific methods to our discipline, it seems quite clear to me that nobody would or should take what we do very seriously in future if we fail to use them.

One last aspect of data collection that I think requires comment is what we consider to be 'data.' In the future we should not only collect data about the speech produced by speakers but also data about how people perceive speech features and dialects. The breakthrough book in this area is Dennis Preston's *Perceptual Dialectology* (1989), but as Preston's forthcoming *Handbook of Perceptual Dialectology* will show (in press), there has been active work in the area for over fifty years in Japan and the Netherlands. The more we know about the distribution of individual dialect features, the more we realize that they do not pattern as neatly across the land as people (including dialectologists) expect them to. We all perceive clear differences between dialects that, in the end, we cannot document except by recourse to a small number of selected diagnostic features. Our work in future must recognize the difference between speech as produced and speech as perceived, and we need to begin collecting perceptual as well as production data. Our best work will incorporate both aspects, as Macafee's paper in this volume shows.

I have already mentioned how interviews with Pederson's new framework could be transcribed, with tags to identify elicitation targets. This, I believe, will be the future of data encoding for dialectology. I have spent many years working on a database approach to encoding questionnaire responses, and I will spend more time yet trying to get all the rest of the first-generation American Atlas data into digital form in a database structure. However, tagged transcripts should be the format for digital storage of any new surveys. My earlier choice of the database structure was conditioned by the state of computer processing at the time. In the 1980s when Lee Pederson and I were building database structures, desktop computer processing was not very good or fast for text files, but was very good indeed for databases. Moreover, highly efficient storage of responses in databases went along with the shortage of mass-storage space that existed then. Pederson designed his databases to fit on floppy disks; a few years later I designed mine for 10Mb and 20Mb hard drives. Now there is no shortage of mass storage – I have 500 times as much hard-drive space as I used to have – and processors are fast enough for full-text searches in large files. And there are substantial advantages to using tagged texts. For one thing, the automatic taggers that already exist for corpus linguistics, such as the BNC automatic tagger (cf. Aston and Burnard 1997), could be applied immediately to dialect interviews, which would render

the information in them much more accessible in more ways than ever before. As early as 1989 Pederson illustrated additional uses of tagged texts [Fig. 1; Pederson and Madsen 1989: 20]. You can see how additional information is coded in Pederson's transcription. There are unique markings to identify the elicitation targets, and to label other relevant aspects of the text such as to keep the field worker's speech separate from the informant's speech. Each type of information subsequently can be extracted automatically to form concordances or different kinds of indices. Today we should all be using SGML, the standard markup language for texts, instead of Pederson's homemade tagging, but the point is the same, and SGML allows for the customized tags that we need for the special purposes of dialectology. In the 1980s dialectologists needed to be inventive to bend emerging computer technology to our own purposes; today and in the future, we need instead to be receptive to developments that are already out there, and inventive enough to see their applications for our own work.

After collection and encoding of data, we come to the issue of analysis. The future of analysis in dialectology is like the future of theory: we need some. It is not enough merely to collect and preserve data. In the past it has been an axiom of dialectology that our goal was full and fair presentation of our data. It is understandable that dialectologists have been preoccupied with display and publication because we have had such large quantities of data, in fact so much that in America most of it never has achieved publication in any better medium than microfilm. Orton did far better with *SED* (1962-71), but even *SED* analyses have come late (e.g. Orton and Wright 1974; Orton, Sanderson, and Widdowson 1978; Upton, Sanderson, and Widdowson 1987). A neat statement of our conservative position came at the end of an article called 'Inside a Linguistic Atlas,' written by Raven McDavid and the editors of the Middle and South Atlantic Atlas project, including me (McDavid et al. 1986: 404-05):

> The business of the linguistic atlas is to provide the evidence, not verdicts. It would be silly to say that we are not interested in what may come out of our materialsñnaturally we are, and we will remain in the forefront among those who interpret themñbut we cannot afford to make interpretations before we present the data or, worse, to insert interpretations into our presentation of the materials. Those of us on the inside have a responsibility to get the data out, and this we will do, in time, as clearly, fully, and objectively as possible.

While I agree with and am still acting upon the last statement here, the irony of this

assertion is twofold. First, the failure of dialectologists to provide analysis of their materials has actually *prevented* publication of the data. After a long series of grant applications to the National Endowment for the Humanities, it has become clear to me that the proposal reviewers are not willing to fund the Atlas project just for its intrinsic merit. They have trouble seeing what the Atlas data is good for, and increasingly so the more time passes after the interviews. The second irony is that the greatest success of the American Atlas was the benchmark set of analyses by Kurath, Atwood, and McDavid (1949, 1953, 1961), which were all based on a subset of the data and were produced before publication of any but the first American regional Atlas. These studies, particularly the *Word Geography* and the *Pronunciation of English in the Atlantic States*, caught the imagination not just of dialectologists but of a wide range of readers, from philologists to cultural geographers. They established scholarly expectations about American dialects for a generation. On the basis of this experience, dialectologists in the future ought to publish working analyses early and often!

The nature of our analysis, I believe, will also be different in the future. Lee Pederson has described an essential contradiction in the goals of dialectologists and the expectations of their audience (1995: 39):

> The [Linguistic Atlas] method carries analysis through an enumeration of features and records them in lists and/or reports them in maps. Such analytic word geography ends its work at this point in a taxonomy of observed sociolinguistic facts. But the research invariably implies more than that because planners, editors, and their critics fail to characterize the work at hand. For that reason, a reader expects an identification of dialect areas and a description of dialects within those geographic divisions in a concordance of social and linguistic facts projected across space and through time. . . . Both [Hans Kurath and Harold Allen in their association of American settlement patterns with speech areas] synthesize geographic, historical, and social facts in their reorganization of evidence in an effort to meet the unreasonable expectations of linguistic geography.

The name 'dialectology' itself raises the expectation that our job is to describe and to find the borders of dialects. In the past our teachers believed that they were conducting research in historical linguistics, and well-bounded dialects are one of the tools of study in historical linguistics, which in theory is essentially a structuralist enterprise. Yet another irony in our field is that American and some British dialectologists today

regularly question the status of dialects and their borders, for example in papers by Davis, Houck, and others about Kurath's Midland region (Carver 1987; Davis and Houck 1992; Frazer 1987, 1993, 1994; Johnson 1994), or Davis, Houck, and Upton's paper from the last Methods conference about our general failure to draw convincing dialect boundaries (1997), or Davis, Houck, and Horvath's paper in this volume – all this while our sociolinguist and structuralist colleague William Labov quite happily describes well-bounded dialect areas, including the American Midland. Now and in the future, if we dialectologists believe ourselves to be engaged in empirical linguistic study, then we should have much broader scope for analysis than our teachers did. For this reason I have renamed the 'Dialectology' course at my university to 'Language Variation,' so as to enlarge the expectations of my students. In future, dialectology may lose its name if others do what I have done, and that may not be a bad thing if it would help us to be clear about what we and others expect from our work.

As my own writing reveals (e.g. 1992, 1996a; Kretzschmar and Schneider 1996), I believe that quantitative analysis will be the hallmark of future analysis in dialectology, and also in empirical linguistics more generally. I will not belabor that point here. I would like to confirm, however, that there is room in dialectology for many kinds of research and analysis. I have mentioned perceptual dialectology as an important constituent of our work, and that branch of study verges on psychology. Those who study language and identity, following LePage and Tabouret-Keller (1985), will also find friends among the dialectologists (e.g. Lanehart 1996). Qualitative research of many kinds should exploit the ethnographic side of dialectology. Finally, as much as I have suggested that we need to break free from exclusive concentration on historical linguistics, we should in future continue to contribute to that discipline (e.g. Kretzschmar 1996b). The future of dialectology should be pluralistic in its approaches to analysis.

My last point today concerns our presentation of our data and the results of our analyses. The publication of many American Atlases has been held up for years by the shear weight of the data, and it is no more practical to produce Atlases on paper today than it has been for the last fifty years. I believe that the best solution for this problem is publication via the Internet, and to that end we have created a Web site (on a server at the Georgia editorial office, http: //us.english.uga.edu) that provides a framework for comprehensive display of Linguistic Atlas data and for visualization of the data on maps – with every map created to the user's order. The problem of editing all of that data is still present, and it will take us time to get all of it onto our Atlas Web site, but when we have done that we will have a fully interactive Atlas with comprehensive coverage of the first generation of American regional projects. I believe

that, within the foreseeable future, we will have largely fulfilled Hans Kurath's dream of an American Atlas.

Let me take a few moments to illustrate the Web site for those of you who may not have accessed it. The opening screen of the Web site is followed by a clickable map of the American Atlas regions. We have made the Middle and South Atlantic project (LAMSAS) operational, with local pilot funding for creation of the site based on prior computerization of LAMSAS data (see Kretzschmar et al. 1993). Base maps for the site have come from the US government 'Tiger' Internet site, which provides the digital images for no cost. Next comes another clickable map, this time just of the LAMSAS region, from which users can find detailed information about the people interviewed for the project. A click on any state brings up a more detailed map of the state on which are plotted the locations of all the people interviewed, and a click on any person's plot reveals a full description of their particulars (age, sex, type, educations, etc.). It also allows the user to ask what that person responded to any of the questions of the survey questionnaire. Another screen shows the non-informant-based services available: 1) browsing the data, 2) searching the data, 3) generating maps of particular items, 4) a table of phonetic symbols used for the project, and 5) utilities available for downloading, notably a TrueType font with which users may view and print Atlas phonetic symbols on their own PC-compatible computers. We have also implemented a demonstration project for linguistic survey research via email. Users may browse lists of data for any survey question, or search the database for words in which they are interested. In order for users to make maps, they first select the survey question in which they are interested, and the site produces a list of all the answers; users then select the answer that they want to map. Software underlying the Web page then produces a map to order. The key feature of the Web site is that it is an interactive resource. It is extensively cross-linked in addition to allowing the user to ask several different kinds of questions of the database. When we have more data available, it will be possible to ask questions across several projects at once. We hope to automate several of the quantitative analytical methods that we have developed, so that those can be run by users online in real time. The Web is the research tool of the future, and we have it now.

Technology will certainly be important to us in presentation of our data and results, as in all aspects of data collection, encoding, and analysis, but there is another equally important point to make. I believe that we need to accept as central to our purposes the goal of informing the public, not just the scholarly community, about the facts of language variation, especially as that information can affect education and public policy. And technology can help us to achieve that goal, too. The existing

Linguistic Atlas Web site, even given the limited area for which its software is fully activated, has become the most accessible source of information about regional American English for the general public (see also http://www.ling.upenn.edu/phono_atlas/home.html, Labov's site, which is more oriented to professionals; it now offers sound samples). The scholarly works that I have mentioned have been largely technical in nature, best suited for specialists. The *Dictionary of American Regional English* (Cassidy, Hall, et al. 1985-) is much more accessible to the general public in its content, and it has sold thousands of copies to libraries, but unlike the Atlas Web site it is not available in the millions of homes in America and abroad that can access the Internet. An earlier version of the Atlas Web site averaged over 10,000 'hits' per month during its peak of use, a very large number of them from non-academic addresses. I regularly get email from high-school students who ask for help with class papers, and so far I have been able to answer every one. Many of the speech patterns documented by the first American Atlases still exist, so public and expert users can look up words and pronunciations that they have noticed in their own speech. Speakers of different ages and social circumstances, in cities as well as rural areas, participated in the American interviews, including a number of African Americans, so the Atlas surveys include a wide spectrum of American speech – and thus a great many Americans can use the Atlas Web site to find out about American words that are still relevant to them and which reflect their American cultural heritage. We hope to do better still with the site in future, now that we know that it has become such an active public resource. We have added to our new site more explanatory information keyed to the needs of our non-specialist audience. We also hope to help the public to ask appropriate questions of our data, and to allow them to ask questions using a natural-language interface in an expert system built with Artificial Intelligence tools (cf Rochester and Kretzschmar 1998). We have felt a little like the recent popular movie about a man who built a baseball diamond in an Iowa cornfield, whose credo was 'If you build it, they will come.' We built it, and they came, and because they came we now see that we have to build it better for them.

The need for this public access has been made clear to us all in the debate over Ebonics this past year. The extent of misinformation about language variation that has been displayed by politicians, pundits, and even in newspaper editorials by the normally sensible American public, has been nothing short of astounding. In large part the misinformation results from the overselling of Standard English in our educational system, the same reason that my students and many more educated people believe in the Star Trek Syndrome. The fact that a great many people have visited our Web site is a hopeful sign, a measure of the public's willingness to seek out new

information about something on which they already have definite opinions.

This curiosity on the part of the public, and the fact of extensive variation in contemporary English, both show that the future of dialectology is not bleak, no matter what things look like on Star Trek. The foundations of study laid down by our teachers in the first generation of Linguistic Atlas work still serve us well, and we have a great many technological and other kinds of improvements that we can make as we enter our own Next Generation. The most important element for our future, among all of these factors, is us ourselves. As Raven McDavid said, nothing can 'substitute for interest, imagination and training, and the determination to do a good job.' We will succeed in the next century, and beyond, to the degree that we are determined, that we can interest and train ourselves and our students, and that we have the imagination to see the way forward. I believe that we will make Harold Orton and all of our teachers proud.

REFERENCES

Aston, Guy, and Lou Burnard. 1997. *The BNC Handbook.* Edinburgh: Edinburgh University Press.

Atwood, E. Bagby. 1953. *A Survey of Verb Forms in the Eastern United States.* Ann Arbor: University of Michigan Press.

Bailey, Guy. 1997. When did Southern American English begin? In Edgar Schneider (ed) *Englishes Around the World 1*, pp. 255-75. Amsterdam: Benjamins.

Bailey, Guy, and Garry Ross. 1992. The evolution of a vernacular. In Matti Rissanen, et al. (eds) *History of Englishes*, pp. 519-31. Berlin: Mouton de Gruyter.

Carver, Craig. 1987. *American Regional English.* Ann Arbor: University of Michigan Press.

Cassidy, Frederic, Joan Hall, et al. 1985-. *Dictionary of American Regional English.* Cambridge, MA: Belknap Press.

Davis, Lawrence M. 1973. The diafeature: an approach to structural dialectology. *Journal of English Linguistics* 7: 1-20.

Davis, Lawrence M., and Charles L. Houck. 1992. Is there a Midland dialect area? – Again. *American Speech* 67: 61-70.

Davis, Lawrence M., Charles L. Houck and Clive Upton. 1997. The question of dialect boundaries: the SED and the American atlases. In Alan Thomas (ed) *Issues and Methods in Dialectology*, pp. 271-83. Bangor: Linguistics Department, University of Wales

Bangor.

Frazer, Timothy. 1987. *Midland Illinois Dialect Patterns*. PADS 73. Tuscaloosa: University of Alabama Press.

Frazer, Timothy. 1993. *'Heartland' English: Variation and Transition in the American Middle West*. Tuscaloosa: University of Alabama Press.

Frazer, Timothy. 1994. On transition areas and the 'Midland' dialect: a reply to Davis and Houck. *American Speech* 69: 430-35.

Johnson, Ellen. 1994. Yet again: the Midland dialect. *American Speech* 69: 419-29.

Johnson, Ellen. 1996. *Lexical Change and Variation in the Southeastern United States 1930-1990*. Tuscaloosa: University of Alabama Press.

Kretzschmar, William A., Jr. 1992. Isoglosses and predictive modeling. *American Speech* 67: 227-49.

Kretzschmar, William A., Jr. 1995. Dialectology and sociolinguistics: same coin, different currency. *Language Sciences* 17: 271-82.

Kretzschmar, William A., Jr. 1996a. Quantitative areal analysis of dialect features. *Language Variation and Change* 8: 13-39.

Kretzschmar, William A., Jr. 1996b. Foundations of American English. In Edgar Schneider (ed) *Focus on the USA*, pp. 25-50. Philadelphia: John Benjamins.

Kretzschmar, William A., Jr. 1997. American English for the 21st century. In *Englishes Around the World* 1, edited by Edgar Schneider, 307-24. Amsterdam: Benjamins.

Kretzschmar, William A., Jr., and Edgar Schneider. 1996. *Introduction to Quantitative Analysis of Linguistic Survey Data: An Atlas by the Numbers*. Thousand Oaks, CA: Sage.

Kretzschmar, William A., Jr., et al. 1993. *Handbook of the Linguistic Atlas of the Middle and South Atlantic States*. Chicago: University of Chicago Press.

Kurath, Hans. 1949. *A Word Geography of the Eastern United States*. Ann Arbor: University of Michigan Press.

Kurath, Hans, and Raven I. McDavid, Jr. 1961. *The Pronunciation of English in the Atlantic States*. Ann Arbor: University of Michigan Press.

Labov, William, and Sharon Ash. 1997. Understanding Birmingham. In C. Bernstein, T. Nunnally, and R. Sabino (eds) *Language Variety in the South Revisited*, pp. 508-73. Tuscaloosa: University of Alabama Press.

Lanehart, Sonja. 1996. The Language of Identity. *Journal of English Linguistics* 24:322-33.

LePage, Robert, and Andree Tabouret-Keller. 1985. *Acts of Identity*. Cambridge: Cambridge University Press.

McDavid, Raven I., Jr. 1985. Review of Edward Ives, *The Tape-Recorded Interview*. *Journal*

of English Linguistics 18: 54-63.

McDavid, Raven I., Jr., et al. 1986. Inside a linguistic atlas. *Proceedings of the American Philosophical Society* 130: 390-405.

Orton, Harold, et al. (eds). 1962-71. *Survey of English Dialects.* Intro. and 4 vols. in 3 pts. Leeds: Arnold.

Orton, Harold, Stewart Sanderson, and John Widdowson (eds). 1978. *The Linguistic Atlas of England.* London: Croom Helm.

Orton, Harold, and Nathalia Wright. 1974. *A Word Geography of England.* London: Seminar.

Pederson, Lee. 1995. Elements of word geography. *Journal of English Linguistics* 23: 33-46.

Pederson, Lee. 1996a. LAWCU project worksheets. *Journal of English Linguistics* 24: 52-60.

Pederson, Lee. 1996b. LAMR/LAWS and the main chance. *Journal of English Linguistics* 24: 234-49.

Pederson, Lee, and Michael Madsen. 1989. Linguistic geography in Wyoming. *Journal of English Linguistics* 22: 17-24.

Preston, Dennis. 1989. *Perceptual Dialectology.* Dordrecht: Foris.

Preston, Dennis, ed. in press. *Handbook of Perceptual Dialectology.* Amsterdam: John Benjamins..

Rochester, Eric, and William A. Kretzschmar, Jr. 1998. Intuitive interfaces for the retrieval of linguistic data, ACH/ALLC, Debrecen (Hungary).

Trudgill, Peter. 1975. *Accent, Dialect, and the School.* London: Arnold.

Upton, Clive, Stewart Sanderson, and John Widdowson. 1987. *Word Maps.* London: Croom Helm.

Wright, Joseph. 1898. *English Dialect Dictionary.* Vol. 1. London: Frowde.

URLs:

Labov's site: http://www.ling.upenn.edu/phono_atlas/home.html.

Linguistic Atlas site: http: //us.english.uga.edu.

Star Trek site: http://www.startrek.com (the official site; there are others which include sound samples.)

Address
William A. Kretzschmar, Jr.
Department of English
Park Hall
University of Georgia
Athens
GA 30602-6205
USA
<billk@atlas.uga.edu>

Also published *by Leeds Studies in English* are the occasional series:

LEEDS TEXTS AND MONOGRAPHS
(ISSN 0075-8574)
and
LTM Medieval Drama Facsimiles
(ISSN 0143-0335)

The most recent volumes in the **LTM** series are:
The Gawain Country: Essays on the Topography of Middle English Poetry by R.W.V. Elliot (1984) 165pp.
Staging the Chester Cycle, Lectures given on the Occasion of the Production of the Cycle at Leeds in 1983, edited by David Mills (1985) vii + 123pp.
Thomas Campion: De Puluerea Coniuratione (On the Gunpowder Plot), edited by David Lindley with translation and additional notes by Robin Sowerby (1987) v + 123pp.
Úr Dölum til Dala: Guðbrandur Vigfússon Centenary Essays, edited by Rory McTurk and Andrew Wawn (1989) x + 327pp.
P. J. Cosijn: Notes on 'Beowulf', introduced, translated and annotated by Rolf H. Bremmer Jr, Jan van den Berg and David F. Johnson (1991) xxxvi + 120pp.
A Study and Edition of Selected Middle English Sermons by V. M. O'Mara (1994) xi + 245pp.
Concepts of National Identity in the Middle Ages, edited by Simon Forde, Lesley Johnson and Alan V. Murray (1995) viii + 213pp.
The Old English Life of St Nicholas with the Old English Life of St Giles by E. M. Treharne (1997) viii + 218pp.
And in the **Medieval Drama Facsimile** series:
The York Play, a facsimile of BL Additional MS 35290, with a facsimile of the *Ordo Paginarum* section of the A/Y Memorandum Book, edited by Richard Beadle and Peter Meredith (1983) lxi + 551pp., 5 colour plates.
The Chester Mystery Cycle, a facsimile of BL Harley MS 2124, edited by David Mills (1984) xxii + 288pp.

Full details of these series and of past numbers of *Leeds Studies in English*, including prices and availability, can be obtained from:

The Secretary,
Leeds Studies in English,
School of English,
University of Leeds,
Leeds LS2 9JT, U.K.

All orders should be placed with the Secretary at the above address.
Those with standing orders for any series (*Leeds Studies in English*, *LTM* or the Drama Facsimiles) are entitled to a 25% discount on all publications.

Also published *by Leeds Studies in English* are the occasional series:

LEEDS TEXTS AND MONOGRAPHS
(ISSN 0075-8574)
and
LTM Medieval Drama Facsimiles
(ISSN 0143-0335)

The most recent volumes in the **LTM** series are:
The Gawain Country: Essays on the Topography of Middle English Poetry by R.W.V. Elliot (1984) 165pp.
Staging the Chester Cycle, Lectures given on the Occasion of the Production of the Cycle at Leeds in 1983, edited by David Mills (1985) vii + 123pp.
Thomas Campion: De Puluerea Coniuratione (On the Gunpowder Plot), edited by David Lindley with translation and additional notes by Robin Sowerby (1987) v + 123pp.
Úr Dölum til Dala: Guðbrandur Vigfússon Centenary Essays, edited by Rory McTurk and Andrew Wawn (1989) x + 327pp.
P. J. Cosijn: Notes on 'Beowulf', introduced, translated and annotated by Rolf H. Bremmer Jr, Jan van den Berg and David F. Johnson (1991) xxxvi + 120pp.
A Study and Edition of Selected Middle English Sermons by V. M. O'Mara (1994) xi + 245pp.
Concepts of National Identity in the Middle Ages, edited by Simon Forde, Lesley Johnson and Alan V. Murray (1995) viii + 213pp.
The Old English Life of St Nicholas with the Old English Life of St Giles by E. M. Treharne (1997) viii + 218pp.
And in the **Medieval Drama Facsimile** series:
The York Play, a facsimile of BL Additional MS 35290, with a facsimile of the *Ordo Paginarum* section of the A/Y Memorandum Book, edited by Richard Beadle and Peter Meredith (1983) lxi + 551pp., 5 colour plates.
The Chester Mystery Cycle, a facsimile of BL Harley MS 2124, edited by David Mills (1984) xxii + 288pp.

Full details of these series and of past numbers of *Leeds Studies in English*, including prices and availability, can be obtained from:

The Secretary,
Leeds Studies in English,
School of English,
University of Leeds,
Leeds LS2 9JT, U.K.

All orders should be placed with the Secretary at the above address. Those with standing orders for any series (*Leeds Studies in English*, *LTM* or the Drama Facsimiles) are entitled to a 25% discount on all publications.