

## The past, present, and future of English dialects: Quantifying convergence, divergence, and dynamic equilibrium

WARREN MAGUIRE AND APRIL MCMAHON

*University of Edinburgh*

PAUL HEGGARTY

*University of Cambridge*

DAN DEDIU

*Max-Planck-Institute for Psycholinguistics*

### ABSTRACT

This article reports on research which seeks to compare and measure the similarities between phonetic transcriptions in the analysis of relationships between varieties of English. It addresses the question of whether these varieties have been converging, diverging, or maintaining equilibrium as a result of endogenous and exogenous phonetic and phonological changes. We argue that it is only possible to identify such patterns of change by the simultaneous comparison of a wide range of varieties of a language across a data set that has not been specifically selected to highlight those changes that are believed to be important. Our analysis suggests that although there has been an obvious reduction in regional variation with the loss of traditional dialects of English and Scots, there has not been any significant convergence (or divergence) of regional accents of English in recent decades, despite the rapid spread of a number of features such as TH-fronting.

### THE PAST, PRESENT AND FUTURE OF ENGLISH DIALECTS

Trudgill (1990) made a distinction between Traditional and Mainstream dialects of English. Of the Traditional dialects, he stated (p. 5) that:

They are most easily found, as far as England is concerned, in the more remote and peripheral rural areas of the country, although some urban areas of northern and western England still have many Traditional Dialect speakers. These dialects differ very considerably from Standard English, and from each other.

We gratefully acknowledge funding from the Arts and Humanities Research Council for the project Sound Comparisons, 2005–2007 (grant no. 112229). Dan Dediú also thanks the Economic and Social Research Council for a postdoctoral fellowship award held during the preparation of this paper. We would also like to thank the reviewers of this paper for their very helpful (and extensive) comments.

Concerning the Mainstream dialects, he noted that they are primarily associated, in Britain, with the southeast of England, with urban areas, with areas where English has been introduced relatively recently (e.g., the Scottish Highlands), with younger speakers and “with middle- and upper-class speakers everywhere” (p. 6). Compared with the Traditional dialects of English, he suggested (p. 6) that “the Mainstream Modern Nonstandard Dialects differ much less from Standard English and from each other.” Trudgill (1990) defined the differences between dialects of English, Traditional and Mainstream, using carefully chosen, example phonetic and phonological features, and summarized the results of these analyses in two maps (pp. 34, 65), as well as in two trees (pp. 35, 67). Given the introductory nature of Trudgill’s account of English dialect differences, this approach is perfectly understandable, but it does raise important questions about how we determine the relationships between varieties of a language. Thus we might question why Trudgill chose the features that he did and consider whether the results would be the same if an entirely different or larger set of features were to be analyzed. And we might also wish to explore other options for visualizing the relations between varieties, because the methods Trudgill used inevitably force divisions in what is essentially a dialect continuum (as Trudgill pointed out [pp. 6–7]) and are incapable of comparing Traditional and Mainstream dialects *at the same time*.

Trudgill also considered ongoing change and the future of English dialects in Britain. Comparing English in England with English in Australia, Trudgill suggested (p. 10) that there will never be the same dialect diversity in Australia due to the recent formation of the variety and the fact that the world has changed so much since then, in terms of transport and communication. Trudgill suggested that the kind of divergence seen in earlier stages of English and in other Germanic languages is unlikely to be repeated in Australia, and the same could just as easily be said for Mainstream dialects in England too. That is not to say that Trudgill believed varieties of English, in England or elsewhere, are currently converging to the point where regional differences will disappear. Instead, Trudgill envisaged a state of affairs, already apparent in some parts of Britain such as the southeast of England, where there is a “spreading out of urban dialect speech to form new dialect areas” (p. 81). Trudgill projected a possible course of development of English dialects along these lines in a map of “future dialect areas” (p. 83) which are focused on the major urban centers in England. Although Trudgill suggested that this future development may involve a degree of convergence between varieties at the regional level—so that, for example, Tyneside and Middlesbrough English become more similar (if not the same)—he is at pains to point out that this does not mean that everyone will soon end up speaking the same throughout England (p. 84):

It is certain that, whatever the exact form of future developments will be, there will never be total uniformity across the country, because innovations will always continue to spread and recede and thus continue to produce the rich mosaic of regional variation in pronunciation which has characterized England ever since English first became its language.

Trudgill painted a vivid and complex picture of variation and change in varieties of English in Britain. Not surprisingly, his account raises as many questions as it answers, including: How different were Traditional dialects of English compared with Mainstream ones, and what degree of convergence was there in the transition between them? Is this convergence still ongoing, or is something else happening to varieties of British Englishes? Are recent and widespread changes that seem to have their source in southeast England having an effect on the relationships between varieties? And are certain regional varieties remaining more distinct than others?

This paper is an attempt to demonstrate how we might go about answering some of these questions. In particular, it describes a method for the simultaneous comparison of many varieties of English at the phonetic level and for the quantification of differences between them. Before we describe the method, we first discuss a number of studies that have highlighted the complexities of change in British Englishes, the processes and motivations behind those changes, and the interpretations of them by linguists. In so doing, we hope to clarify the patterns that lie behind Trudgill's outline of change and to set the findings of our own analysis in context so that the patterns revealed can be better understood.

#### VARIATION IN CHANGE IN BRITISH ENGLISHES

##### *Endogenous and exogenous change in Norwich*

One locality that has been characterized by significant change throughout the 20th century is Norwich English in East Anglia, as discussed, for example, in Trudgill (1999). Trudgill identified a wide range of changes in Norwich English in the second half of the 20th century and suggested that these changes have come about as a result of three different forces at work. These forces are:

1. Homogenization toward “a national mainstream,” which Trudgill identified “as the result of the influence of RP and of other varieties of English English generally” (p. 139). This homogenization mostly consists of redistribution of vowel phonemes, such that their lexical distribution more closely matches Received Pronunciation (RP) and London English (p. 137), and perhaps some vocalic phonetic changes (p. 138).
2. Homogenization between (nonstandard) varieties of English which is “bringing Norwich English into line with other regional accents” (p. 140). Changes of this type include H-dropping, TH-fronting, and labiodental pronunciation of /t/.
3. Endogenous change—changes that are absent “from RP and from neighbouring dialects” and which are “spontaneous” and “internally produced” (Trudgill, 1999:134). These changes largely affect the vocalic system (pp. 134–135).

Depending upon their relative importance, these three kinds of change are likely to have different implications for the development of Norwich English and its

relations with other varieties of the language. If homogenization toward “a national mainstream” predominates, Norwich English is likely to become less distinctive from other varieties and from RP English. If homogenization of Norwich English with other regional varieties of English predominates, a similar loss of distinctiveness may result, although things may in fact be more complex than this, as the studies discussed in the following sections indicate. If endogenous changes were to predominate, it is possible that Norwich English would diverge from other varieties of English. Much of this paper will involve discussion of how we determine the overall effects of these kinds of change.

### *Regional dialect leveling*

On a wider geographical scale, Kerswill and Williams (1999) examined the accents of three localities in England (Milton Keynes, Reading, and Hull), and suggested (p. 149) that the three varieties are, to a certain extent at least, converging “in both inventory and realisations,” while they are, at the same time, retaining features that distinguish them (especially Hull in the north). They call this process *dialect leveling*, which they define (p. 149) as “a process whereby differences between regional varieties are reduced, features which make varieties distinctive disappear, and new features emerge and are adopted by speakers over a wide geographical area.”

Kerswill (2003:223) suggested that there are two separate processes at work in dialect leveling—the first is *geographical diffusion* of linguistic features, and the second is *leveling* of variants as a result of mutual convergence between speakers of different varieties. The result of both of these processes is what Kerswill (2003:223) called, more specifically, *regional dialect leveling*. An example of geographical diffusion is the rapid spread of TH-fronting throughout Britain in the second half of the 20th century (see Kerswill, 2003:230–237, and the discussion of *Innovation diffusion in the Fens*).

Kerswill (2003) identified change in the pronunciation of the FACE lexical set in the northeast of England as a likely case of leveling. In both Tyneside and Durham, the FACE lexical set may be pronounced with the diphthong [ɪə] or the monophthong [e:] (see also the discussion of *Dialect leveling in Tyneside*). [ɪə] is highly localized whereas [e:] is characteristic of a much larger geographical area (northern England and Scotland). In both locations, [ɪə] is characteristic of (older) working-class male speech, and Kerswill suggested that this highly localized variant is subject to leveling, leading to the variant [e:], which is shared by other varieties, becoming the predominant variant. Unlike geographical diffusion, Kerswill suggests that one feature of leveling is that leveling changes establish themselves “simultaneously throughout a given region” (p. 224).

An important aspect of regional dialect leveling is that it is, as its name suggests, conceived of as a *regional* rather than a national (or indeed international) phenomenon. Although regional dialect leveling can, given the right circumstances, happen anywhere, the effects of any one instance of it are likely to be felt in a relatively small geographical area. As an example of this, Kerswill

(2003) and Torgersen and Kerswill (2004) highlighted the convergence of the entire short vowel system in two varieties of southeastern English (Ashford and Reading) which were formerly characterized by rather different pronunciations of their short vowels.

Nevertheless, many changes implicated in regional dialect leveling, both of the geographical diffusion type (e.g., TH-fronting) and of the leveling type (e.g., replacement of localized vowel pronunciations), are shared by a wide geographical spread of varieties of British English. The effects that some of these changes have when confronted with rather different, regional phonological systems, and the overall consequences of regional dialect leveling and how we might determine this are discussed below.

### *Dialect leveling in Tyneside*

Watt and Milroy (1999) and Watt (1998, 2000, 2002) examined changes in the pronunciation of the vowels in the FACE and GOAT lexical sets in Tyneside English in northeast England. They identified three variants of FACE and four of GOAT: traditional variants [ɪə] and [ʊə] (and possibly [ø:]); variants [e:] and [o:], which are characteristic of much of northern England and Scotland; and [eɪ] and [ou], which are really only characteristic of the speech of middle-class speakers. In an “apparent time” analysis of their data (for an overview, see Bailey, 2002), Watt and Milroy (1999:41) identified a “reduction in variability across two generations; localised variants either disappear or are reduced in number and younger speakers prefer mainstream, unmarked variants.” In effect, this means that highly localized [ɪə] and [ʊə] are disappearing, and pan-northern [e:] and [o:] are becoming the default pronunciations of FACE and GOAT in Tyneside English. The picture is somewhat complicated by an increased use of the (seemingly traditional) vowel [ø:] in GOAT among the younger middle-class males and the introduction of [eɪ] and [ou] which are typical of Midland and southern English accents (including RP) in middle-class speech.

Watt and Milroy (1999:25) ascribed these changes to a process of “dialect leveling” in Tyneside, whereby “variants characteristic of a larger area than the Tyneside region appear to be spreading at the expense of extremely localised variants.” They consider this dialect leveling to be a consequence of the disruption of “close-knit, localised networks which can be shown to maintain highly systematic and complex sets of socially structured linguistic norms” (Watt and Milroy, 1999:26).

The effect of this disruption is the “eradication of socially or locally marked variants” (ibid.). This leveling is not the whole story, however. Watt and Milroy (1999:42) remarked on how sharp gender differences are still maintained despite the overall reduction in variation, and note that, although these developments in Tyneside English may be part of the establishment of a leveled, *koiné* variety characteristic of northern England, this does not mean that the variety that has developed is in any sense “pan-English” (p. 43). Furthermore, some young speakers use what look like traditional forms “as a symbolic affirmation of local

identity” (p. 37). As Foulkes and Docherty (1999:13) put it, “there appears to be a tension between speakers’ desire to continue signalling loyalty to their local community by using local speech norms, and a concurrent urge to appear outward-looking or more cosmopolitan,” a tension that is summed up very nicely by Watt (1998:7) as a desire to “sound like northerners, but *modern* northerners.”

It is worth pointing out at this stage that the developments Watt and Milroy identified in FACE and GOAT are only the latest stages in a long history of change. Examination of traditional dialect data for the northeast of England from the *Survey of English Dialects* (Orton, Sanderson, & Widdowson, 1962–1971, henceforth *SED*) and from Rydland (1998) suggested that both of these lexical sets have been subject to considerable reorganization in terms of the lexical distribution of phonemes—or, in the words of Trudgill (1999:136–137), “lexical redistribution” or “dedialectalisation.” Figure 1 illustrates the considerable lexical redistribution that fed into the patterns identified by Watt and Milroy.

This radical reorganization of the lexical incidence of phonemes was not random. The highly localized lexical distribution of phonemes in the traditional dialects has been transformed into one which is shared over a much wider part of the English-speaking world and which is much closer to standard varieties of English (see Maguire, 2009). Certainly this looks like an instance of dialect leveling, although it is not clear that this is the same kind of process as the change from [ɪə] to [e:], which does not affect the lexical incidence of phonemes in the variety.

Thus in Tyneside English, we have, for these features at least, a long history of change, involving lexical redistribution of phonemes, reduction in highly localized variants, adoption of pan-regional variants and of (perhaps standard) variants from further afield, combined with reinterpretation of local features for the projection of particular identities. These changes have happened at different periods and need not all have had the same influence on the relationships between Tyneside English and other varieties of English. How we might more fully come to understand the nature of these changing relationships is discussed further in *Consequences of Change*.

### *Innovation diffusion in the Fens*

Britain (2005) examined the effects of the diffusion of changes from southeast England into the Fenlands of eastern England. Although he suggested (p. 999) that “Within the core of the southeast, it is probably true that regional dialects are losing their distinctiveness as the dialect mixing becomes ever more intense—thanks to very high levels of social and geographical mobility in the area,” he questions whether this loss of distinctiveness is inevitable in peripheral dialect areas such as the Fens. Although a wide range of consonantal and vocalic changes are spreading to this area, Britain suggested that the effects of these changes might be rather different when they are introduced to areas characterized by phonological patterns not found in the southeast of England.

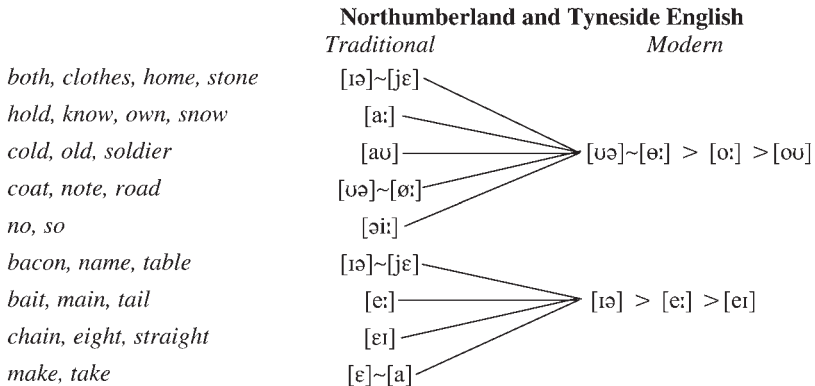


FIGURE 1. The development of the FACE and GOAT lexical sets in northeast England.

In order to illustrate this difference, Britain examined innovations in four consonantal variables and five vocalic variables in three Fenland locations (Spalding in the west, Terringtons in the east, and Wisbech/March in the central Fens). He found that all of these innovations except STRUT-fronting are found in the Fens, but the three locations typically differ in the extent to which they are present. More strikingly, Britain found that the innovations produce rather different results in the three locations depending upon pre-existing local phonological patterns. For example, L-vocalization is much less advanced in the east where clear /l/ was a traditional feature of the dialect (even though it is not for young speakers today). Similarly, open onset and monophthongal pronunciations of the PRICE vowel, though found everywhere, are only found before voiceless consonants in the central and eastern locations because these areas (but not Spalding in the west) are characterized by a “Canadian Raising” type alternation. The result of this innovation is that in the west [aɪ] ~ [a:] can occur everywhere, but in the center and east, the difference between the variant before voiceless consonants (typically [əɪ]) and the variant before voiced consonants and morpheme boundaries (where [aɪ] ~ [a:] is possible) has been accentuated.

But what are the consequences of these changes for relationships between the Fenlands varieties and between all of these and varieties in the southeast of England? Britain observed that changes in the consonantal system such as TH-fronting may well have no overall effect on the similarities of different varieties because they all shared [θ] before the change and all share [f] after it. It might be argued that the varieties in the Fens are becoming more similar to varieties in southeast England as a result, but even these varieties once had [θ] (see Kerswill, 2003, for evidence from the *SED*), so it could be argued that the Fens are simply catching up with a change that does not affect overall relations between varieties in the long term. Although all three Fenland varieties have introduced the same kinds of changes to their vowel systems, these changes have affected each differently, depending upon the pre-existing phonological patterns in the variety. The continued (and increased) distinctions between the prevoiced and prevoiceless

variants of PRICE and of the MOWN and MOAN subsets of GOAT in the central and eastern varieties suggest to Britain that “innovations don’t just overwhelm and obliterate local varieties, but, engaging in contact with them, often produce compromise outcomes, showing evidence of the interaction of the innovation with the traditional local form” (Britain, 2005:1017). Because the varieties (particularly the eastern and western ones) also differ across a whole range of phonological and morphosyntactic variables, it appears to be the case that “at the end of the twentieth century, extensive diffusion coexisted alongside local and regional differentiation” (Britain 2005:1017).

### *Innovation and maintenance in Glasgow*

Stuart-Smith (2003) and Stuart-Smith, Timmins, and Tweedie (2006, 2007) examined the incorporation of phonological features of nonlocal origin alongside the maintenance of traditional dialect features in the urban dialect of Glasgow in Scotland. Although they identify a number of features that seem to have their origin outside of Glasgow—for example TH-fronting and “Cockney-style” L-vocalization (in words such as *feel* and *milk*)—these are used alongside traditional dialect features such as monophthongal /ʌ/ in the MOUTH lexical set and traditional “Scots” L-vocalization (in words such as *all* and *hold*).

Stuart-Smith et al. found that younger working-class speakers combine use of features of exogenous origin with continued (and in some cases increased) use of traditional Scots features. In light of this complex pattern of variation and change in Glasgow, Stuart-Smith et al. (2007:248) commented that:

Taking Glaswegian in the overall U.K. context, we might think that this is dialect levelling in its broadest sense (Kerswill, 2003), and in time, this may be the outcome. But there are also important differences, both in the ideological processes involved ... and linguistically, for example, in the healthy retention of local non-standard variants ... initial analysis of Scots lexical variation as represented by Scots /ʌ/ in the OUT set (*oot* for *out*), also reveals strong maintenance of the vernacular in the same speakers ... and we suspect that similar results could follow from analysis of further vowels.

A number of issues are identified as important in this mixing of new and traditional features, including projection of a young, local, working-class Glaswegian identity and a rejection of middle-class practices and values. This involves “healthy” retention of localized phonetic and phonological patterns alongside reinterpretation of patterns of (probable) exogenous origin as manifestations of localized speech and is giving rise to an urban variety for which the terms “Scots” and “Scottish Standard English” hardly seem appropriate.

### *Consequences of change*

It is clear from the previous discussion that changes in British English have been far-reaching and complex. But it is less clear what effect these changes have had



on the relations between varieties of English. Despite these varieties sharing many changes, they have not all become the same. Although there is some evidence of convergence, particularly at the regional level, it remains to be seen whether this is so throughout Britain. As was discussed, the kinds and proportions of different changes will have an important influence on the relations between varieties. If endogenous changes predominate, divergence is likely, assuming that the same changes do not spread to other varieties. If, on the other hand, the same exogenous changes affect different varieties, convergence might be the outcome if those changes predominate. But as the studies discussed indicate, this need not be the case: diffusing changes such as TH-fronting may simply replace one set of patterns with another without any net convergence; the effects of exogenous change may only be at the regional level; and incoming changes may interact with existing phonological patterns to give rather different outcomes in different locations. On the other hand, the leveling of localized patterns at different locations is undoubtedly going to affect the relationships between varieties.

As was discussed, Kerswill (2003) suggested that the exogenous changes involved in regional dialect leveling may lead to convergence between varieties. In particular, he noted (p. 239) that:

The overall picture supports the view that regional dialect levelling is widespread in Britain. Few researchers have been able to demonstrate its opposite—divergence or diversification—in local varieties.

Although Kerswill contrasted regional dialect leveling with divergence, there may be another scenario that Kerswill does not allude to here, one that involves neither convergence nor divergence—to use the language of local government elections, no overall change. Is it possible for there to be neither convergence or divergence, but a maintenance of the *status quo*? That is, might varieties, through a combination of endogenous and exogenous changes, remain relatively equidistant from each other, even though the details change considerably? The case studies discussed suggest that this might be possible (although we cannot tell for sure). If this is indeed the case, this would be a major factor in the maintenance of dialect distinctiveness in languages without the emergence of distinct new languages. We label this kind of change *dynamic equilibrium*.

So although exogenous changes may contribute toward convergence, and although endogenous changes may contribute toward divergence, it is the overall effect of a range of changes on a range of varieties which affects the degree of similarity between them. It is possible that various kinds of change may combine and “cancel each other out,” at least in terms of the overall relationships between varieties. A key point that arises from the preceding discussion is that we cannot diagnose divergence, convergence, or dynamic equilibrium through the examination of a small number of changes, because other changes may be of an entirely different sort. It is obviously not possible to identify changes in relationships between varieties without comparing more than one variety. In

order to determine the nature and direction of changes in relationship between varieties, we suggest that the following steps are necessary:

1. A wide range of features must be compared at the same time, and these features should not be selected simply because we judge them to be important; in other words, the true significance of particular changes can only be determined in the context of other patterns of variation and change.
2. A wide range of geographical varieties need to be compared with each other so that we can determine whether change is regional, national, or international.
3. Data that allow us to assess change in “real” or “apparent time” should be included, otherwise we will only be able to reveal a synchronic picture of the relationships between varieties.
4. The global similarity of different varieties to each other should be quantified so that an objective assessment of the degree of convergence, divergence, or no overall change can be properly appreciated.

Thus we argue that to determine whether divergence or convergence (or neither) have taken place it is essential that we develop methodologies that look beyond a single geographical location and a particular set of features that may or may not be important for defining overall relationships between varieties. In this way, we suggest, we will be able to assess the extent to which Trudgill’s model of change, from Traditional dialects, to Mainstream accents, to urban-focused varieties, is a realistic representation of the development of British Englishes. In the rest of this paper, we outline a method for doing precisely this and examine what it can tell us about changing relationships between varieties.

#### A METHOD FOR COMPARING PHONETIC SIMILARITY

In Heggarty, McMahon, and McMahon (2005), McMahon and McMahon (2005), and McMahon, Heggarty, McMahon, and Maguire (2007), a new method for comparing the phonetic similarity of different varieties of the same language and of different languages was introduced. This method shares with other dialectometric methods a concern with quantifying the similarities and differences between varieties in an objective way, but it differs from methods such as that described in Nerbonne, Heeringa, and Kleiweg (1999) and Nerbonne and Heeringa (2001) in that it prioritizes linguistic accountability over computational simplicity. This method measures the similarity of (segmental) phonetic transcriptions of a list of cognate words. A full description of this method can be found elsewhere;<sup>1</sup> in summary, this approach has four steps:

1. Phonetic transcription of a list of cognate words in the varieties being compared.
2. Comparison via a slot-matching mechanism that uses a node form to inform us as to which parts of each transcription should be compared with which.
3. Application of a phonetic similarity algorithm, which compares the phonetic transcriptions for each variety with each other and computes their similarity

based on an assessment of the relative distance between segments in multidimensional phonetic space.

4. Representation of the resultant distance matrix using trees and networks, in particular the phylogenetic software program NeighborNet (see Huson & Bryant, 2006; for applications to linguistic data, see Heggarty et al., 2005; McMahan & McMahan, 2005; McMahan et al., 2007; and *Investigating multidimensional relationships* below).

This method is particularly well suited to analyses of the sort desired here because it is capable of comparing numerous varieties of a language at the same time (e.g., RP with Tyneside, RP with Liverpool, RP with Glasgow, Tyneside with Liverpool, Tyneside with Glasgow, Liverpool with Glasgow). The resulting distance matrix (expressed in percentage terms for each possible pairwise comparison between the varieties) can be interrogated in various ways to reveal the (changing) relationships between varieties. In the following sections, we discuss a number of issues that were not covered in the previous papers, because they are the result of further developments to the method, developments which should help us to answer the questions posed at the beginning of this article.

### *The word list*

As the preceding discussion indicates, it has been our aim to collect data from a wide range of geographical locations throughout the British Isles (and beyond, although we do not report on that here) to determine the relationships between varieties of English (and its Germanic relatives). As was noted in the introduction to this article, it is essential that such a comparison avoids preselection of features, because this would give us at best a limited insight into the relationships between the varieties, and at worst leads us into a kind of self-fulfilling prophecy whereby we find that the features selected for analysis are exactly those that appear to be important.

Because the method requires the analysis of strictly comparable data, and because strictly comparable data are all too often hard to come by for modern varieties of English,<sup>2</sup> it was decided to collect these data from scratch. Given that we hoped to cover a wide range of varieties of English and other Germanic languages, and given that our method requires the comparison of (Germanic) cognates, the most straightforward and feasible way of getting the data we required was by using a word list. A representative phonetic transcription of this word list was required for each geographical data location.

*The word list wish list.* In using a word list for comparative purposes, it is essential that the words on that list have not been selected to give us the “best” results because, prior to analysis, we cannot be sure which features or, more likely, combinations of features are most significant in determining the overall relationships between varieties (cf. the discussion of Trudgill’s analysis of English dialects).

This does not mean, however, that the members of the word list were selected at random. Two (related) aspects of our methodology demand that only certain kinds of words may be compared in the first instance:

1. The comparison mechanism requires that all of the words compared are cognates in all of the varieties analyzed.
2. As it is our intention to develop a method that is applicable beyond varieties of English—so that, for example, (varieties of) English can be compared to (varieties of) German—these cognates must be shared by English and its Germanic relatives (see McMahan & McMahan, 2005, and McMahan et al., 2007, for further details).

Therefore, it follows that all words in our cognate list stem from the inherited Germanic lexicon; indeed their original Proto-Germanic forms stand as a “node-form” through which we can match up their modern phonetic reflexes precisely, to ensure that we are comparing “like with like” between all the present-day varieties (see McMahan et al., 2007:120).

Because our method depends crucially on the comparison of cognates across a wide range of Germanic varieties (including regional dialects of many languages), this considerably restricts the number of possible cognates, because not all are attested in every variety. Additionally, because we seek to measure the phonetic similarity of these cognates in different Germanic varieties, we have sought to avoid cases where differences have arisen as a result of morphophonemic or morphological differences, in Proto-Germanic or subsequently. Thus a word such as *water* is excluded because the Scandinavian languages reflect the Proto-Germanic stem *\*watnan* whereas the West Germanic languages reflect the Proto-Germanic stem *\*watar* (Orel, 2003:451). The result is a reduced population of cognates that may be sampled for our word list. The full word list is reproduced in Table 1.

This desire for the widest possible applicability of the method across Germanic cognates was an important determiner of the choice of words for our word list, and one that acted without reference or bias toward features traditionally favored for

TABLE 1. *The word list*

all	cow	fight	hear	ice	mother	oak	salt	swear	two
ash	daughter	fish	heart	in	mouse	one	see	tear	warm
bath	day	five	holy	is	mouth	open	seven	ten	wash
better	drink	foot	home	knee	nail	out	sharp	thing	what
bite	ear	four	honey	lamb	naked	oven	sit	thorn	white
blood	earth	full	horn	leaf	name	over	six	three	wind
bone	eat	good	hound	liver	needle	quick	snow	thunder	wool
brother	eight	goose	house	long	new	rain	sore	toe	word
calf	eye	green	hundred	mid	night	red	stone	tongue	yard
cold	fast	hand	hunger	milk	nine	right	stool	tooth	year
corn	father	head	I	moon	north	ring	storm	top	young

categorizing varieties of English (and other Germanic languages), so that preselection of features was not an issue. Within the constraint of needing to compare cognates, we adjusted the particular selection in order to try to bring the frequency of particular phonological features in our word list close to being representative of the English lexicon. This was done by comparing the frequency of selected phonological features (in particular initial consonants and stressed vowel phonemes) in our word list with a standard word list—the Thorndike and Lorge (1944) list of the 1,000 most frequent words in a range of English texts from both sides of the Atlantic. This word list was chosen for comparison because it is entirely independent from our own and contains a wide selection of fairly basic (often Germanic) vocabulary. The frequency of particular features in our word list and in the Thorndike and Lorge word list are compared in Table 2 and Figure 2.

The comparison of the percentages of each of Wells' lexical sets (Wells, 1982) reveals that the two word lists match well.<sup>3</sup> The discrepancies that do exist may well be attributable largely to the etymological restrictions placed on our word list (entirely Germanic in origin, with no French loans). Similarly, Figure 2 reveals that the frequency of initial consonant phonemes in the two word lists is roughly comparable, with some obvious differences (e.g., /p/ and /h/) being the result of our word list containing only words of Germanic origin.

Thus despite their different origins and intentions, these two word lists are very similar in terms of the frequency of stressed vowel and initial consonant phonemes, which suggests that our word list is a reasonable representation of the sound patterns of English. Furthermore, the *quid pro quo* for any minor discrepancies is the major boon of being able to extend our comparisons across all Germanic languages and dialects (see Heggarty, McMahan, & Maguire, forthcoming).

TABLE 2. *The percentage of Wells' lexical sets in the two word lists*

LexSet	SC	TL	LexSet	SC	TL
CHOICE	.0	.9	FOOT	3.6	1.7
CURE	.0	.5	NEAR	3.6	1.0
commA	.0	.0	DRESS	4.5	12.0
PALM	.9	.2	FACE	5.4	8.3
SQUARE	.9	1.7	GOOSE	5.4	4.0
LOT	1.8	3.6	MOUTH	5.4	3.8
CLOTH	1.8	2.1	NORTH	5.4	2.2
NURSE	1.8	2.8	FLEECE	6.4	8.0
FORCE	1.8	1.6	lettER	8.1	6.9
happY	1.8	3.8	GOAT	9.1	6.8
TRAP	2.7	6.1	PRICE	9.1	8.1
BATH	2.7	2.2	STRUT	10.0	6.9
THOUGHT	2.7	2.3	KIT	11.8	10.8
START	2.7	2.4	Rhoticity	24.3	19.1

SC = Sound Comparisons; TL = Thorndike and Lorge.

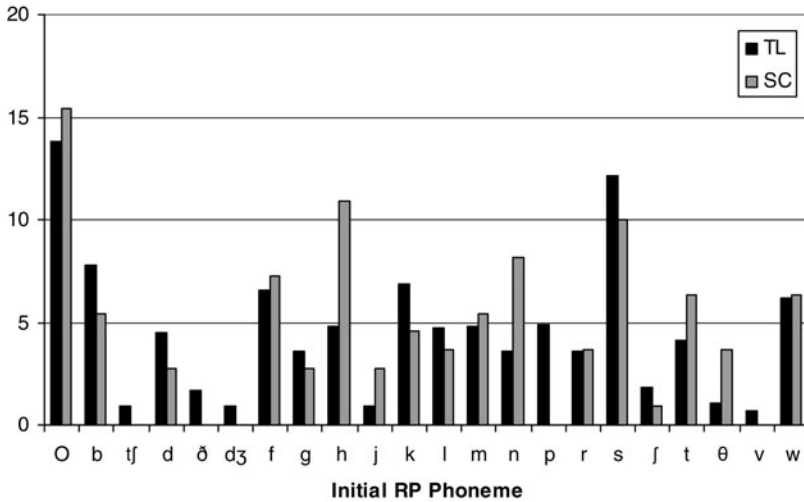


FIGURE 2. Comparison of the frequency of RP initial phonemes in the Sound Comparisons (SC) and Thorndike and Lorge (TL) word lists.

#### *Data collection and transcription*

Following considerable success, albeit with a much smaller sample, in the pilot project reported in McMahon and McMahon (2005) and McMahon et al. (2007), it was initially felt that the best way to get data for a range of varieties of English was to approach experts in particular dialects and ask them to provide us with transcriptions that would be representative of the working-class accents found at particular locations. This approach provided us with data for London, Norwich/Norfolk, Sheffield, Liverpool, Glasgow, Shetland, and Dublin.<sup>4</sup> Although these data were very much what we required, and are included in our analysis, there were a number of issues with this approach to data collection that meant that we moved away from it in favor of collection of primary data ourselves. Firstly, and perhaps most importantly, it proved difficult to find experts in a wide enough range of varieties of English to give us a sizable sample. In many cases, much was known about particular features in a given accent but, pending further research, it proved difficult to find experts who were confident in providing accurate transcriptions of our full word list in particular accents. This was even more of a problem in out of the way areas where research on modern accents (as opposed to traditional dialect research from the middle of the 20th century) has still to be done in any form. Secondly, the transcriptions received in this way were, despite the best efforts of the contributors, rather diverse in their interpretation of our transcription requirements, such that some transcriptions were more detailed and others were more normalized. The result of this was that all of these transcriptions had to be given a second pass by our research team to iron out “transcriber isoglosses” and to standardize details and methods of transcription. This was not always easy, despite being done in discussion with

the contributors, and it was felt that this process added an extra level of subjectivity which might be avoided were other methods of data collection to be employed.

Because of these issues, it was decided that data should be collected directly (although the data we had received from intermediary experts were not rejected). Audio recordings were made of native speakers of varieties of interest reading the word list. Although it was intended that these speakers should be representative of their respective varieties, lack of research in some areas meant that their representativeness or otherwise had to be gauged by the fieldworker.

Once the data had been collected, a narrow segmental phonetic transcription was made by a single specialist in English dialectology and phonetics (the lead author of this paper) to avoid transcriber isoglosses.<sup>5</sup> Details of the transcription procedure and example transcriptions can be found on the Sound Comparisons website.<sup>6</sup>

It is clear that readings of word lists are never going to be the same as vernacular speech, but by the same token, nor are they a necessarily inaccurate reflection of (a subset of) the linguistic patterns present in the speech community. Some examples from our data indicate the kinds of vernacular patterns that, perhaps unexpectedly, surfaced in readings of word lists by working-class native speakers. With reference to H-dropping, for example, speaker JB (F, born 1962) from North Devon dropped [h] twice in the 12 H-initial words, whereas speaker SA (M, born 1968) from Rossendale in Lancashire dropped [h] in all H-initial words. Speaker SW (M, born 1988) from Tyneside likewise produced a high number of highly vernacular forms in his reading of the word list, including seven tokens (out of nine) with [f] for historical /θ/, [ä] in *all* and *father*, and [øə] in *snow* and *toe*. Many other speakers produced equally localized pronunciations, suggesting that for many speakers reading need not necessarily suppress the vernacular—see Stuart-Smith et al. (2007:241–242, 247) for discussion of a similar finding. In any case, because we are comparing like with like, any effect of reading on the kind of speech produced should not affect the overall results.

#### *Incorporating social variation—the subvarieties*

Much sociolinguistic research makes it clear that an analysis of the structured variation in the speech community is necessary to identify change in “apparent time” (see any of the studies referred to in the introductory section of this article, for example). Given that we wished to cover a wide range of geographical varieties of English in the most economical way possible (using a word list), it is also clear that the same depth of analysis of social variation in each speech community would not be possible. Nevertheless, we wished to capture diachronic variation to a certain extent, because it is precisely this kind of variation that gives us insights into the changing relationships between varieties. We therefore decided to gather, in many of our regional locations, not just one but multiple recordings, categorized in broad subvarieties, to provide some insight into apparent time change in modern English. We have labeled these subvarieties “Typical,” “Traditional,” and “Emergent,” each of which is defined in turn.

*The Typical subvariety.* Although there is undoubtedly considerable variation among speakers at any given location, it was assumed that relative to other locations it is possible to find speakers who are typical, in terms of their speech patterns, of their speech community. For our purposes, we defined such speakers as locally born and resident, working-class, and aged between 30 and 55 years. Where only one representative speaker was recorded in the locality, a phonetic transcription of that speaker's word list reading was used in our comparison; when more than one speaker was recorded at a location, a single representative speaker was transcribed or, in cases where there was noticeable variation between speakers, a composite transcription was made that included the most typical pronunciation form of each word. We consider our Typical subvarieties to be roughly equivalent to Trudgill's Mainstream dialects.

*The Traditional subvariety.* The first author of this paper grew up in rural County Tyrone in Northern Ireland, has lived most of his adult life in Newcastle Upon Tyne in northeast England, and currently works in the University of Edinburgh in Scotland's capital city. In all three of these (rather different) locations, he has regularly encountered (and continues to encounter) traditional dialect pronunciations of a sort familiar from traditional dialect surveys but which are often peripheral to modern sociolinguistic studies. These are, as might be expected, most typical of (and sometimes exclusive to) working-class, (older) male speech. Obvious examples include [ɔːl] 'old' and [ɛt] 'eat' in Tyrone, [niːt] 'night' and [ʊt] 'out' in Newcastle, and [ɡɪd] 'good' and [meːə] 'more' in Edinburgh.<sup>7</sup> But these are precisely the kinds of pronunciations that are so deeply embedded in the vernacular that they (almost) never surface in word list readings. Nevertheless, it is clear that these kinds of pronunciations form a part of the speech of speakers in many locations across the British Isles, and indeed are particularly salient. In addition, it may be assumed, given that these kind of forms have a long and consistent recorded history, that they represent survivals of older speech patterns which were once more typical of local working-class populations. As such, and because the data available to us from traditional dialect surveys such as the *SED* are not compatible with our word list or transcription methods, it was decided that an attempt would be made to gather, where possible, traditional dialect pronunciations of this sort at many of the locations surveyed in this project. However, because traditional dialect pronunciations are typically suppressed in the reading of word lists (see, for example, Harris, 1985:243), alternative means of elicitation were required in order to gather these data. To overcome this problem, the following approach was adopted:

1. Speakers who were likely to have good fluency in traditional dialect pronunciations were sought out (usually older working-class males).
2. These speakers were engaged in general conversation to set them at ease and to gauge how "traditional" their speech was.
3. They were asked to read the word list without any prompting as to what was required—most did so in a "nontraditional" accent (but see the following discussion).



4. Speakers were then asked if they had “broader” pronunciations for these words, and an obvious test example was used (e.g., “Do you have a ‘broader’ pronunciation of the word ‘old’?”); most of these speakers responded well to this approach.
5. Assuming that the nature of the task was understood, the speakers were asked to read the word list again, but using these broad, local alternative pronunciations.
6. If speakers missed anything obvious, the interviewer returned to these words and asked if any other pronunciations were known; this usually was met with replies such as “Oh aye, ....”<sup>8</sup>
7. In a few instances, traditional pronunciations known to the interviewer from previous descriptions were not elicited, and the speaker was asked directly (after the preceding methods had been exhausted), “Have you ever heard anything like ...?”. It was usually immediately obvious whether or not the speaker was familiar with the pronunciation in mind.

This approach is similar in many respects to methods used in traditional dialect studies (e.g., Mather & Speitel, 1986; Rydland, 1982), but note also the adherence to this pattern of questioning (but see the following), designed to avoid unnecessary priming of the speakers.

The confidence some speakers had in identifying their “broader,” “local” pronunciations was often striking. One speaker from northern County Antrim produced, without prompting, doublets of many words, exemplifying his more standard Ulster English pronunciations and his traditional Ulster Scots forms, despite denying that he spoke “Ulster Scots.” Pairs produced included [bɪlðə] ~ [bɪiðə] ‘brother’, [k<sup>h</sup>ɐY] ~ [k<sup>h</sup>ɛ:] ‘cow’, [fäðə] ~ [fɛːðə] ‘father’, [fʏt] ~ [fɪt] ‘foot’, [hɛ:d] ~ [hid] ‘head’, [hɔːm] ~ [heːm] ‘home’, [hɐys] ~ [hɪs] ‘house’, [hʌndrɛt] ~ [hʌnə] ‘hundred’, [wʌn] ~ [jɛn] ‘one’, [jü:ɪd] ~ [jɔːd] ‘yard’. Another speaker, from Hawick in the Scottish Borders, read the word list only in her broadest Scots and seemed reluctant to read it any other way. Perhaps most instructive of all, and typical of many speakers who have comfortable command of traditional dialect pronunciations alongside more standard ones, was an older male speaker from the West End of Newcastle Upon Tyne, who, when asked to read the word list, said the following:

[Speaker]: *Well I'll say them what I ... you know, I'm not being broad or nowt, I'll not say them me eh ... I'll just say the way I talk [t<sup>h</sup>ä:k].*

[Interviewer]: *Just the way you, just the way you*

[Speaker]: *The way I would talk [t<sup>h</sup>ä:k] to me mates. I mean to talk [t<sup>h</sup>ɔ:k] to you, I would say all this [ɔ:l ɪs]. If I was talking [t<sup>h</sup>ä:kʔŋ] to me mates I would say all [ä:l].*

[Interviewer]: *OK.*

[Speaker]: *You know what I mean?*<sup>9</sup>

This speaker then proceeded to produce a range of traditional Tyneside dialect pronunciations and engaged in a running commentary on what was possible and what was not (e.g., *Now, sometimes I would say ...*), blurring the line between

reading the word list as he might read aloud generally and reading it as he might pronounce the words in more informal circumstances.

Thanks to this method it was possible to compile a word list that represented the traditional dialect end of the spectrum for many locations. We consider this Traditional subvariety to be equivalent to Trudgill's Traditional dialect type. Note that these pronunciations are best thought of as "potential" rather than actual (see Mather & Speitel, 1986:xii, for a discussion of this issue), for although it was the intention to collect forms still used by the speakers, no attempt was made to gauge how frequent these variants were in any particular individual's speech. Nor is it our assumption that this version of the word list in some sense represents the "actual" way these speakers talk; rather it represents one part of a continuum of variation in the speech community, a part which may well represent the final stages of the move away from traditional dialect in Britain and elsewhere.

*The Emergent subvariety.* As one of the aims of this research is to determine whether recent changes in British English have affected the relationships between varieties, it was deemed necessary to sample the speech of younger speakers at our locations where this proved possible. These younger speakers, either male or female, were working-class, local in birth and residence, and aged between 16 and 25 years. Note that these speakers were not chosen on linguistic grounds, so that in some cases (e.g., Berwick) their speech did not differ dramatically from their older counterparts, whereas in others (e.g., Edinburgh) they did. The subvariety sampled from these younger speakers has been labeled Emergent, even if in practice it transpired to be fairly similar to the Typical subvariety for the same location, precisely so the two can then be compared. We consider our Emergent subvarieties to be roughly equivalent to Trudgill's urban-focused incipient varieties of English.

Table 3 summarizes the number of speakers recorded for each subvariety at each location in the British Isles. Note that in some instances, especially where informants were few in number, a single speaker may have provided both Typical and Traditional versions of the word list. The geographical locations of these data points are indicated in Figure 3.

In the following section, we examine the results of the comparison of these varieties. The inclusion of more than one subvariety for many locations means that it is possible to examine the differential relationships between them and, in so doing, to begin to assess whether the relationships between varieties of English are changing in apparent time.

## RESULTS

The result of the comparison between these varieties of English is a matrix expressing the percentage similarity (or conversely, difference) between each pairwise comparison. These similarity matrices (see McMahan et al., 2007:126, for an example) require further interrogation and manipulation for the patterns

TABLE 3. *Speakers and subvarieties per location*

Variety	Typical	Traditional	Emergent
Belfast	2		
Berwick	4	2	2
Black Country	1		
Bristol	7		
Buckie		1	
Buxton	3		1
Coldstream	5	3	
Cornhill	6	2	
Dublin	transcript	transcript	transcript
Edinburgh	5	2	1
Fermanagh		1	
Glasgow	1 + transcript	transcript	transcript
Hawick	3	3	
Holy Island		1	
Lewis	1		
Liverpool	2 + transcript	transcript	
London	transcript	transcript	transcript
Longtown	1		
Manchester	1		
Middlesbrough	1		1
Morley	1		
Morpeth	1	1	
North Antrim	1	1	
North Devon	10	2	4
Norwich/Norfolk	transcript	transcript	transcript
Renfrewshire	1	1	
Rhymney	10	2	
Rossendale	1		
RP	1		
Sheffield	transcript	transcript	2
Shetland		transcript	
Somerset			1
SSE	3		
Stoke	1		
Tyneside	5	3	4
Tyrone	6	1	

“Transcript” refers to phonetic transcriptions provided by experts in the dialects concerned; see text for details.

within them to be revealed. In *Similarity of varieties to RP*, we analyze one dimension of the data—the similarity of the Typical, Traditional, and Emergent varieties to RP. In *Investigating multidimensional relationships*, we analyze the multidimensional relationships that hold between the varieties more fully. As was discussed in McMahon et al. (2007:130–133 in particular), one method of representing the complexities in such matrices in a visually interpretable way without oversimplifying the data is to use network-type phylogenetic analysis programs, which draw trees where the relationships in the data are treelike, but draw networks when the relationships between varieties are more complex. In particular, the NeighborNet algorithm, part of the phylogenetic software suite

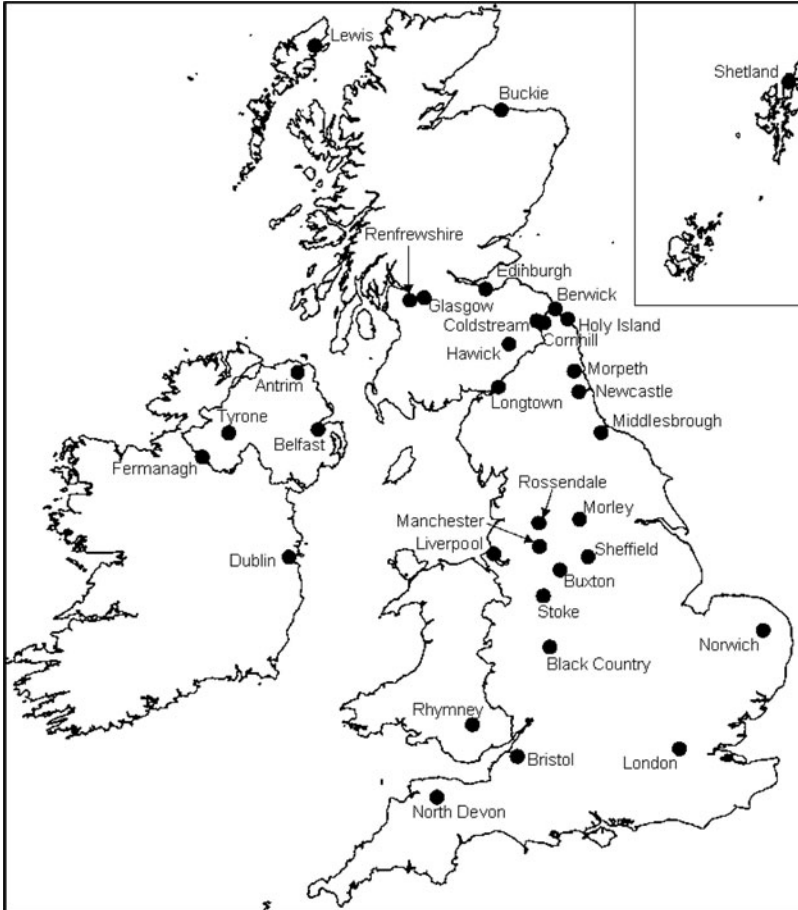


FIGURE 3. Localities surveyed in the British Isles.

Splitstree 4 (Huson & Bryant, 2006) has proven especially suited to displaying the complex relationships between dialects and languages as encoded in similarity matrices. In *Investigating multidimensional relationships*, we again use NeighborNet, in conjunction with statistical testing, to assess the relationships between the Typical, Traditional, and Emergent subvarieties as a whole.

#### *Similarity of varieties to RP*

The mean differences, along with standard deviations, between all the Traditionals, all the Typicals, and all the Emergents in our sample from RP are given in Table 4. The most striking feature of this comparison is that the Typicals and Emergents are identical, on average, in distance from RP. Thus our data, in an apparent time interpretation, suggest that there has been no movement toward RP English in

TABLE 4. *The distance of the subvarieties from RP*

	Traditionals	Typicals	Emergents
Mean distance from RP	.171	.114	.114
Standard deviation	.059	.043	.040

regional varieties of English in recent decades. There is, however, a significant difference between the similarity of the Typicals and Emergents to RP on the one hand and the similarity of the Traditionals to RP on the other, with RP being considerably less similar to the Traditionals than the other subvarieties.<sup>10</sup> We discuss the import of these findings in *Convergence, divergence or dynamic equilibrium?* below.

#### *Investigating multidimensional relationships*

In this section, we compare the similarity of all the Traditionals, Typicals, and Emergents with each other. We visualize the relationships between the varieties in two NeighborNet networks, which allow us to overcome the difficulties inherent in the comparison of Traditional and Mainstream dialects in Trudgill (1990). The first of these networks illustrates the relationships between the Typical and Traditional subvarieties, whereas the second illustrates the relationships between the Typical and Emergent subvarieties. It is possible to combine all of the subvarieties in one network, but separating them out in this way makes the patterns in the data easier to discern because networks become harder to read when they contain many data points.

Figure 4 illustrates the relationships between the Typical and Traditional subvarieties at those locations where both of these subvarieties were sampled in the British Isles. It is important in interpreting these networks to remember that they are phenograms, which depict distance between varieties, as encoded in the similarity matrices, regardless of whether this is the result of common ancestry, contact, or parallel developments. The distance between any pair of varieties is represented by the shortest distance along the lines between them—thus this network reveals, for example, that the distance between Typical Hawick and Typical Coldstream is rather small (as we might expect, given their geographical proximity), but the difference between Traditional London and Traditional Antrim is relatively big (again not surprising given their geographical locations). Where there are “boxes” in the network (created by “reticulations” that join separate “branches”), these reflect conflicting signals in the data, because they are the result of varieties that are quite similar to each other nevertheless having different relationships with other varieties. A good example of this is the different behavior of Typical Devon and Emergent Devon in Figure 5. In this network, Typical Devon groups with Dublin, and, regardless of the reticulations in the network, is not particularly close to London. Emergent Devon, on the

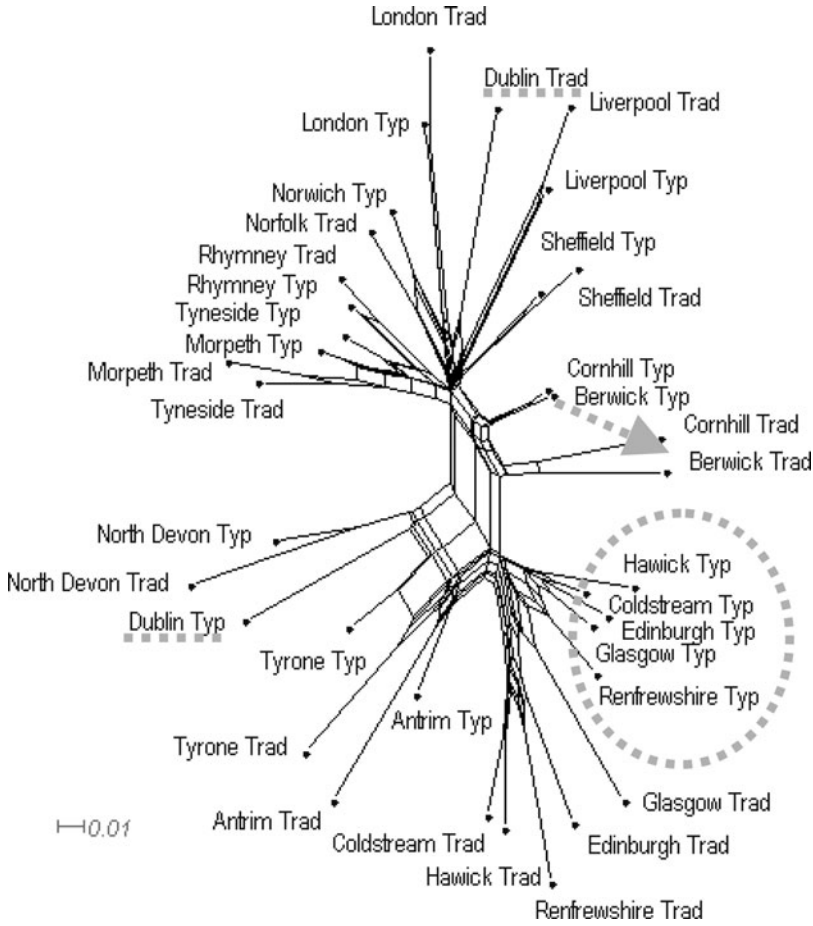


FIGURE 4. Network illustrating the relationships between the Typical and Traditional subvarieties.

other hand, groups with London and is not particularly close to Dublin, again regardless of the reticulations in the network. The fact that Typical Devon and Emergent Devon are, despite their different relationships with other varieties, rather close to each other, is indicated by the reticulation that joins them; but this reticulation does not change the relationship between Dublin and London. Note also that we can identify “splits” in the networks—a clear example can be seen in the middle of Figure 4—which suggest that the varieties on one side of the split share a significant feature or combination of features with each other but not with the varieties on the other side of the split.

Looking in detail at the network in Figure 4, it is clear that there is a major split between, on the one hand, a set of varieties made up of Typical and Traditional Devon, Typical and Traditional Scottish, Typical and Traditional Northern Irish,

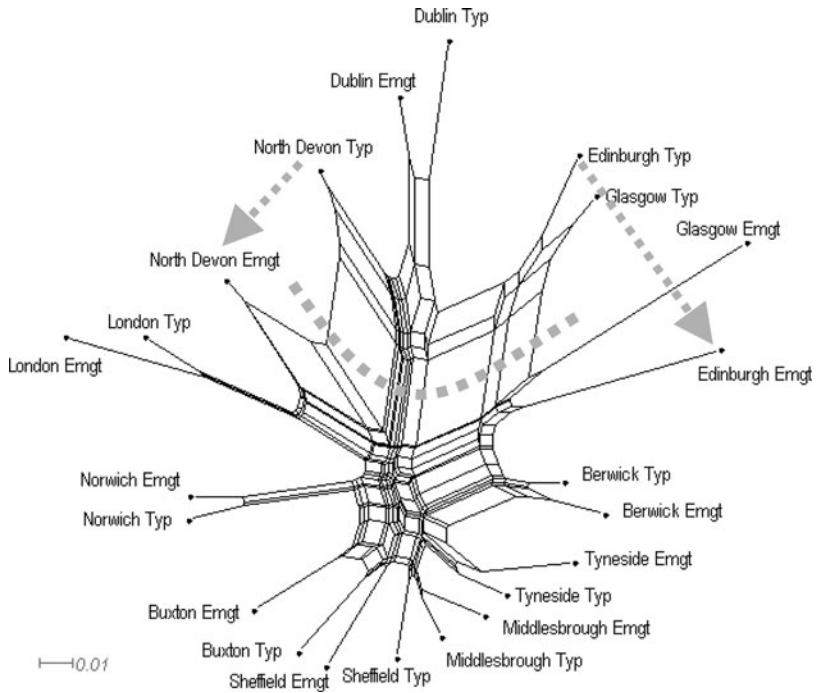


FIGURE 5. Network illustrating the relationships between the Typical and Emergent subvarieties.

and Typical Dublin, and, on the other hand, all other Typical and Traditional English varieties along with Traditional Dublin. Cornhill and Berwick (both Typical and Traditional) lie in an intermediate position. We suggest that this split is the result, at least in part, of the presence versus absence of rhoticity—the inclusion of Devon Typical and Traditional on the (otherwise) Scottish/Irish side of the split may well be a result of these subvarieties being rhotic, whereas the other English varieties are nonrhotic. The intermediate positions of Berwick and Cornhill suggest that these (fairly close) varieties share features with varieties on either side of the split, which is not surprising given their geographical position on the Scottish/English border and the fact that speakers in Cornhill (though not Berwick) were variably rhotic.

As might be expected, geographically proximal varieties tend to be closer together in the network than geographically distant ones (cf. Trudgill's comments about the artificiality of discrete division of dialects; see also Séguy, 1971). Of particular interest, given the questions raised throughout this paper, are the relationships between the Typical and Traditional subvarieties for each location. In some cases (e.g., Liverpool, London, Rhymney, and Sheffield) there is not a great deal of difference between the subvarieties, at least relative to other localities, although it is noteworthy that in each case the Traditional subvarieties

are further away from other varieties than the Typical subvarieties are. In other cases there is a clear difference between the positions of the Typicals and Traditionals. For example, there is a clear distinction between the rather homogeneous Scottish Typicals, which we might term “Scottish Englishes,” and the more diffuse Scottish Traditionals, which we might term varieties of Scots (see Corbett, McClure, & Stuart-Smith, 2003:2–4, for a discussion of this distinction). In the case of the border English varieties, Cornhill and Berwick, the Traditionals appear, given their position in the network, to share more in common with Scottish varieties than the Typical subvarieties do (cf. the findings in Glauser, 1974, on the crystallization of the Scottish/English linguistic border). Most striking of all is the difference between the positions of Dublin Typical and Dublin Traditional. Dublin Typical finds itself between the Northern Irish varieties and Devon on the rhotic side of the split, whereas Dublin Traditional lies between London and Liverpool, but is not particularly close to either of them (or anything else). An analysis of the linguistic features that lie behind these relations would be necessary to fully understand these distributions (see Figure 6 for an example of how this might be done). In this case, the position of Dublin Traditional is very likely explicable in terms of features (such as lack of rhoticity, glottal replacement, lenition) which it shares to varying degrees with Liverpool and London (see Hickey, 2005). It is precisely these features that distinguish Dublin Traditional from Dublin Typical, explaining why the two subvarieties are so dissimilar (again see Hickey, 2005).

In sum, then, the network which illustrates the relationships between the Typical and Traditional subvarieties shows a strong signal for rhoticity and reveals that the Typicals are, by and large, closer to each other than the Traditionals (see Table 5 for details).

Figure 5 illustrates the relationships between the Typical and Emergent subvarieties for those location where both were sampled. In most cases, the paired Typicals and Emergents are found close together in the same part of the network, and it does not appear to be the case that the Emergents are any closer together than the Typicals (in fact in some cases, e.g., London and Berwick, they are slightly further apart). As with the network showing the relationships of the Typicals and Traditionals, this network has a noticeable split which separates the Dublin subvarieties and the Typicals from Devon, Glasgow, and Edinburgh on the one hand, from the remaining English subvarieties and the Glasgow and Edinburgh Emergents on the other. It seems likely that this rather striking pattern also reflects (at least in part) rhoticity, because the varieties on the top side of the split are consistently rhotic, whereas those below the split are nonrhotic, or are only variably so—see Romaine (1978) and Stuart-Smith (2003) for the loss of rhoticity among younger speakers in (urban) Scotland. We suggest that it is largely this change in the status of rhoticity in the southwest of England and in Scotland that accounts for the realignment of the southwest English and urban Scottish Emergent subvarieties, although an analysis of other possible contributing linguistic features is necessary to confirm this.

One simple way of assessing this is to examine the similarity relationships that hold between varieties for *each word* entered into the comparison. Figure 6 compares the similarity of each word in the comparisons of Typical Edinburgh



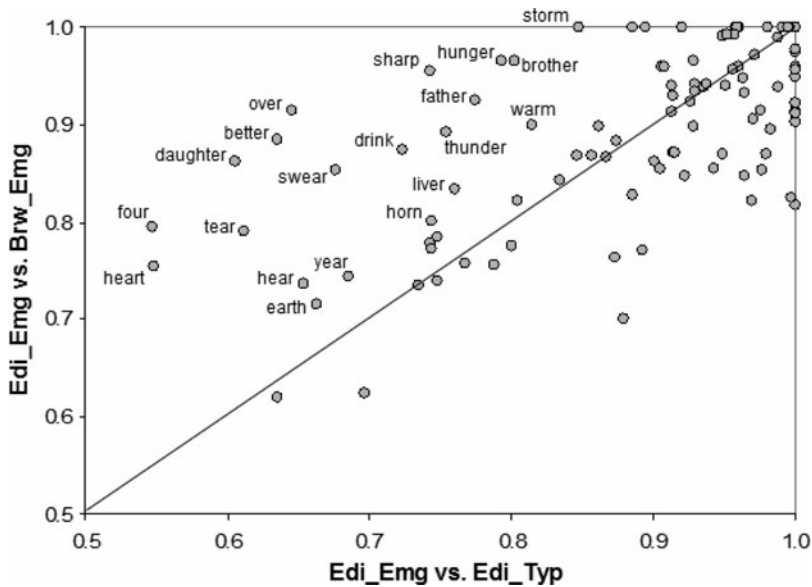


FIGURE 6. Emergent Edinburgh compared with Typical Edinburgh and Emergent Berwick.

and Emergent Edinburgh on the one hand, with Emergent Edinburgh and Emergent Berwick on the other. Comparing Emergent Edinburgh in this way, with two varieties from each side of the major split in the network, allows us to assess which words are involved in the apparent time movement of Emergent Edinburgh toward the English varieties. In this comparison, words that are more similar in the comparison of Emergent Edinburgh with Emergent Berwick than in the comparison of Emergent Edinburgh with Typical Edinburgh appear above the diagonal line, and vice versa. Higher scores on the  $x$  and  $y$  axes indicate higher similarity for the words concerned.

Although not all words are identified, an obvious pattern emerges. In many of the words where Emergent Edinburgh and Emergent Berwick are more similar, historical postvocalic /r/ was, but no longer is, present in these two Emergent subvarieties (whereas it *is* present in Typical Edinburgh). If we removed these words from the comparison, most words in Emergent Edinburgh would be more similar to Typical Edinburgh than to Emergent Berwick. This suggests that the “move” of the Scottish urban Emergents is a consequence of one major structural change: the loss of rhoticity in these varieties. We discuss this issue further in *Convergence, divergence or dynamic equilibrium?* below.

Table 5 summarizes the average distances between all the Traditionals, between all of the Typicals, and between all of the Emergents. This summary brings us to a crucial consideration for determining whether our data offer evidence of change in the relationships between varieties of English and Scots. Although the networks give us a very useful visual interpretation of the (changing) relationships between the varieties, they do not (and are not intended to) tell us what exactly

TABLE 5. *Average distances within the groups of subvarieties*

	Between Traditionals	Between Typicals	Between Emergents
Mean distance	.217	.143	.157
Standard deviation	.053	.046	.040

constitutes a statistically significant difference between varieties. As described earlier, our data consist of *distances* between speech varieties  $S_1, S_2, \dots, S_m$ , and the general type of hypothesis we wish to test is of the form “varieties  $S_1, S_2, \dots, S_n$  are more similar/more different than varieties  $S_{n+1}, \dots, S_m$ .” Therefore, we need to compare the distances between the first set of varieties,  $d(S_i, S_j)$ ,  $1 \leq i, j \leq n$ , against the distances between the second set of varieties,  $d(S_k, S_l)$ ,  $n + 1 \leq k, l \leq m$ . This is usually done using an independent samples  $t$  test (or, when more than two sets are involved, using analysis of variance [ANOVA]).

To be more explicit, we will consider in detail one of the cases described later, namely the differences between Traditionals and Typicals in the LAX case (i.e., only for those locations where at least two types of variety were recorded). To this end, we computed the distances between all 18 Traditional varieties, resulting in a matrix of 324 ( $= 18 \times 18$ ) distances, one distance for each pair of varieties, with a mean distance of .22 and a standard deviation of .053. We did the same for the 20 Typical varieties, resulting in 400 ( $= 20 \times 20$ ) distances between all pairs of such varieties, with a mean of .14 and a standard deviation of .046 (see also Figure 8). We could compare these distances using an independent samples  $t$  test, which would result in  $t(610.78) = 19.25$ ,  $p < 2.2 \cdot 10^{-16}$ , but the resulting  $p$  value would be *too liberal* because our data violate the assumption of independence of observations. This is because each Traditional variety is used to compute 17 distances (to all the other Traditionals), so that the matrix with 324 distances does not contain 324 *independent* data points, which artificially decreases  $p$  values computed by parametric tests such as the  $t$  test.

A way to address this issue is to use *randomization tests* (Edgington, 1987), which are robust to the independence assumption being violated, and to the data not being normally distributed.<sup>11</sup> For example, comparing the Traditional and Typical LAX varieties, a randomization  $t$  test (Edgington, 1987) confirms the very highly significant difference between Traditionals and Typicals, with the latter much more similar to each other than the former.

We report results for two broad cases, depending on the number of speech varieties collected at each location:

1. The STRICT case, limited to only those locations for which have recordings of all three subvarieties.
2. The LAX case, covering all locations for which have at least two subvarieties.

Using the one-way randomization ANOVA and  $t$  tests described previously, we can assess the significance of various groupings of speech varieties, and of the

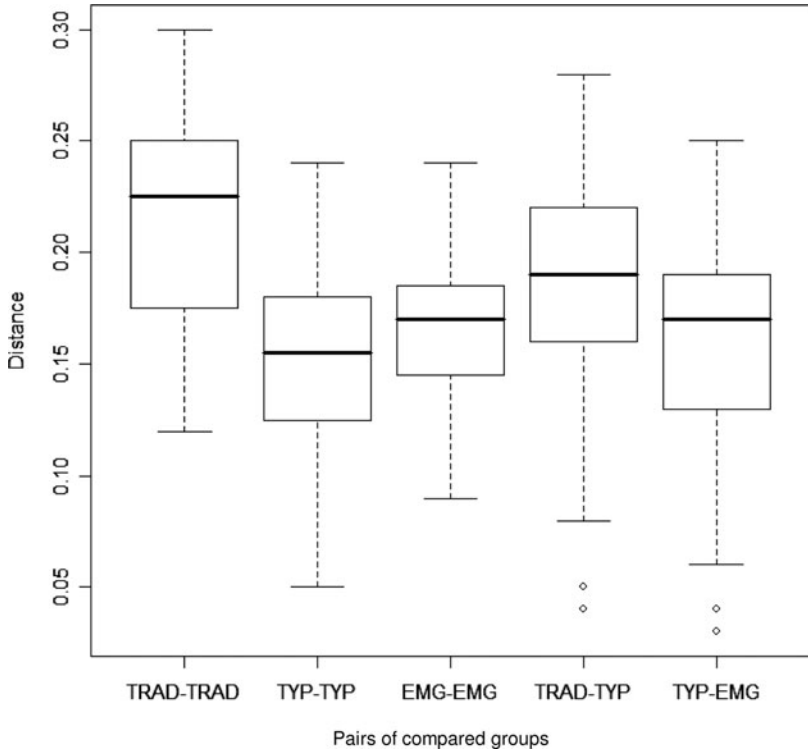


FIGURE 7. Distances between all pairs of varieties from each group (the STRICT case). TRAD = Traditional; TYP = Typical; EMG = Emergent.

differences between them. The box plots in Figures 7 and 8 show the distances for all pairwise comparisons within each of five relevant groupings of speech varieties. For the STRICT case, Figure 7 plots, from left to right, the distances between all pairs of Traditionals, all pairs of Typical, all pairs of Emergents, all pairs made up of one Traditional and one Typical and, finally, all pairs composed of one Typical and one Emergent. Likewise, Figure 8 plots the same distances, but for the LAX case.<sup>12</sup>

The five distance groups described, and presented visually in Figures 7 and 8, were compared using randomization independent sample ANOVAs. We have also used randomization independent sample *t* tests (see note 11) to compare these distance groups with each other (e.g., the distances between all Typical against the distances between all Traditionals). This allowed us to test which intergroup distances are significantly different from each other. Table 6 shows the pairs of distances that are significantly different (after multiple testing correction)<sup>13</sup> in the LAX and STRICT cases. White cells represent comparisons that are not significant comparisons, light gray cells those which are significant only in the LAX condition, and dark gray cells those significant in both the LAX and STRICT conditions (there are no cases of comparisons significant

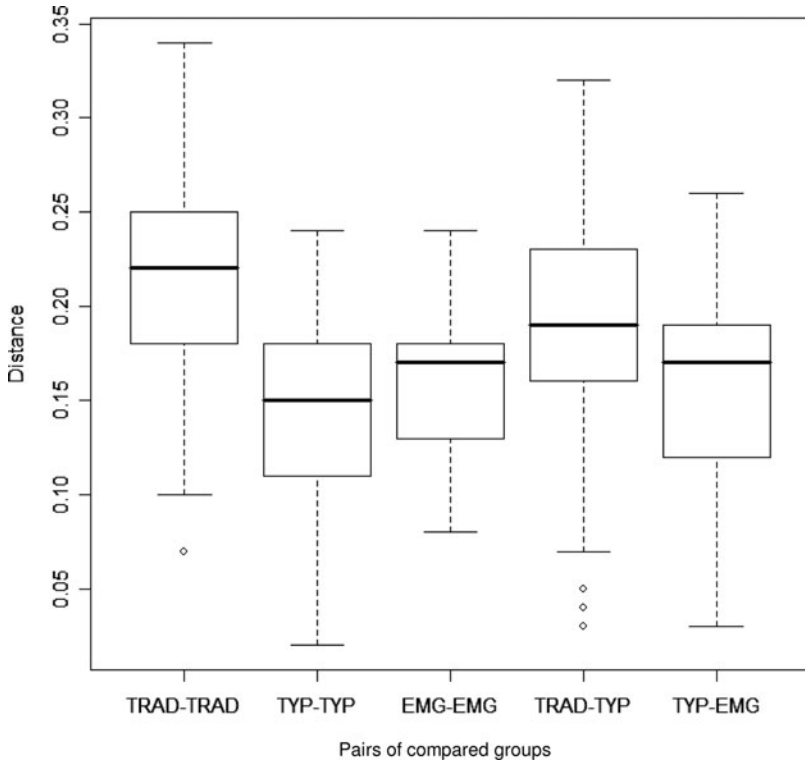


FIGURE 8. Distances between all pairs of varieties from each group (the LAX case). TRAD = Traditional; TYP = Typical; EMG = Emergent.

only in the STRICT but not in the LAX condition). The significance level used is .05, all *p* values are corrected using Holm’s (1979) procedure, and the number of random permutations were generated was 10,000.

TABLE 6. Significant corrected comparisons between pairs of varieties

	TRAD-TRAD	TRAD-TYP	TRAD-EMG	TYP-TYP	TYP-EMG	EMG-EMG
TRAD-TRAD	-					
TRAD-TYP		-				
TRAD-EMG			-			
TYP-TYP				-		
TYP-EMG					-	
EMG-EMG						-

TRAD = Traditional; TYP = Typical; EMG = Emergent

As can be seen, in both the LAX and STRICT conditions the difference between the Typical (TYP-TYP) and Emergent (EMG-EMG) was not significant. This means that, as the networks and box plots indicate, the Emergents are *not* significantly more or less similar to each other than are the Typicals, whereas the Typicals and Emergents *are* significantly more similar to each other than are the Traditional.

In the next section, we discuss the significance of our findings in light of the models of variation and change discussed in the introductory section of this article.

#### CONVERGENCE, DIVERGENCE, OR DYNAMIC EQUILIBRIUM?

This paper examines the changing relationships between varieties of English in the British Isles as a result of endogenous and exogenous changes. We have suggested that although certain developments may be indicative of change in a particular direction, we cannot tell whether they are, in fact, leading to overall change in relationships between varieties if we only look at them in isolation (either changes in individual linguistic features or changes in individual varieties). Rather, we have argued, we must use methods that examine not individual features or individual varieties in isolation but the behavior of a wide range of linguistic features in a wide range of different varieties *at the same time*. One such method has been introduced here. Our method first rates the similarity of phonetic transcriptions of a standard word list in pairwise comparisons of varieties from across the British Isles. This produces a similarity matrix that captures the complex relationships that exist between these varieties. Using up to three subvarieties from many locations, this method has revealed five key patterns:

1. Rhoticity is a very important feature for defining the relationships between varieties of English (whereas it only counts as but one feature among many in Trudgill's analysis, for example).
2. There has been a highly significant convergence in the transition from traditional dialects to modern accents of English (as Trudgill hypothesizes).
3. This is mirrored in the relationship of the Traditional to RP, although the effect is less marked.
4. There has been no corresponding (nor indeed significant) convergence between modern working-class varieties of English as used by middle-aged speakers (Typicals) and those used by younger speakers (Emergents).
5. Likewise, there is no evidence that the Emergents are more (or less) similar to RP than the Typical.

Thus our methodology gives us different answers, depending upon which period of development (in apparent time) we look at in the history of English. Items 2 and 3 involve definite convergence. Items 4 and 5, on the other hand, do not.

#### *Lack of divergence*

Perhaps not surprisingly, our data and methods reveal little evidence for divergence among varieties of English in Britain and Ireland. This finding accords with the

comment by Kerswill (2003:239) that few researchers have been able to demonstrate divergence in dialects of English in the British Isles. Although the Emergent subvarieties are very slightly less similar to each other than the Typical subvarieties, this difference is not significant, and it is likely that it has other causes. It is possible, for example, that it is a consequence of the analysis containing fewer Emergent subvarieties from a wide range of geographical locations and more Typical subvarieties from a denser geographical network; or it might be that our younger speakers were less likely to suppress the vernacular when reading, in a similar fashion to the younger speakers analyzed in Stuart-Smith et al. (2007).

### *The transition from Traditional to Typical*

Our data reveal that the Traditional subvarieties are significantly more heterogeneous than the Typical and Emergent subvarieties, and that the Traditionals are much further from RP than the other subvarieties. An apparent time interpretation of this situation suggests that there has been a marked convergence of varieties of English at some point in the past, in the transition from traditional English and Scots dialects to modern accents of English. This accords with the analysis of English dialects in Trudgill (1990). Two processes that might explain this change were identified earlier: geographical (innovation) diffusion and leveling. Leveling in particular seems likely to lead to convergence between varieties because it involves the reduction in precisely those features that make them most distinctive. Furthermore, as Trudgill (1999) pointed out, exogenous changes can have different sources, such that some changes have the effect of making varieties more similar to the “national mainstream” (including RP), whereas others bring varieties into line with other regional accents. However, the results of these two processes can look rather similar (and this is perhaps why Trudgill does not use the term “standardization”). This is particularly so with leveling changes, which eradicate highly localized variants (by their nature nonstandard), leaving features that are shared by a range of varieties (which may ultimately be of standard origin). So, for example, the word *old* may be pronounced as [ɑ̃] and [õd] in northeast Scotland, as [ɑ:d] and [old] in Tyneside English, and as [ø̃ɹl] and [old] in Tyrone English. Clearly [ɑ̃], [ɑ:d], and [ø̃ɹl] are highly localized and hence may be subject to leveling, leaving [õd], [old], and [old], which, not coincidentally, are nearer to RP English [əʊld] phonetically and phonologically. Even if the leveling change happens without any (further) input directly from the standard variety itself, the result may look like standardization because divergent regional varieties are already likely to share forms that have their origin in more standard patterns of speech. This may well explain why our Traditional subvarieties are further from RP than the Typical and Emergents, and this difference need not indicate that speakers have moved toward RP English directly.

Despite this complication, it is clear that here we are dealing with a process like that described by Kerswill and Williams (1999). The differences between varieties

are being reduced (cf. Figure 4), and markedly local linguistic patterns are being replaced by patterns found in wider geographical areas. Although the term “regional dialect leveling” seems suitable for this change, an alternative term, “dialect death” (see Britain, 2009), might be more appropriate, given the wholesale restructuring of the lexical incidence of phonemes and change at all linguistic levels that underlies this convergence. The term “regional dialect leveling” typically applies to changes in regional accents of English in the second half of the 20th century, whereas the abandonment of traditional dialect by the vast majority of the population was a change that was already well under way in the second half of the 19th century, if the evidence in Ellis (1889), Wright (1905), and the *SED* is anything to go by. Although traditional dialect pronunciations do still survive, for some speakers at least (see *The ‘Traditional’ sub-variety*), a number of factors suggest that the transition from traditional dialect to modern varieties of English is not really an ongoing change for the vast majority of the population and has not been for some time. First, only a small percentage of the population appears to have much command of traditional dialect pronunciations, and this was true even in the middle of the 20th century and before, given the length that surveys such as the *SED* went to identify suitable informants (see Orton, 1962, for details). Second, our own methods for eliciting traditional dialect pronunciations make no claim as to how commonly used these pronunciations are, and this is also the case with traditional dialect surveys such as the *SED* or Ellis (1889). It could well be that these traditional dialect pronunciations are (and were) a minor (if salient) part of the repertoire of even these specially selected speakers, so that the differences between them and more “typical” speakers of modern varieties of English represent a considerable difference in apparent time.

#### *Comparing Typical and Emergent subvarieties*

When we compare the Typical and Emergent subvarieties, we find that there is no significant evidence for overall convergence or divergence. This is despite the fact that most of our Emergent varieties are characterized, to one degree or another, by many exogenous changes suspected of originating in the southeast of England. That the introduction of patterns of exogenous origin is possible without convergence was discussed in above. The fact that there seems to be no evidence of convergence might suggest that leveling has not been a major factor in their development in recent years, even if it was an important factor in the transition from traditional dialects to modern varieties of English. This is not to say that our Emergent subvarieties are the same as the equivalent Typical ones. Although they are obviously quite similar, they are different, for all of the reasons discussed in the first part of this article. Things have not stood still in recent decades. Thus we find evidence for continued change without overall change in similarity or difference—precisely the situation that we introduced the term “dynamic equilibrium” to describe.

But might it be the case that the overall similarity of the Typicals and the Emergents is obscuring changes in the relationships between varieties at the

regional level? After all, the process of regional dialect leveling, as defined by Kerswill (2003), assumes this as part of its definition, and Trudgill (1990) suggested that it is at the regional level that focusing on urban varieties is occurring. In order to answer this question fully, a much larger sample of (sub) varieties would be needed from a variety of British regions. Nevertheless, some insight can be gained by examining the changing behavior of varieties in two regions in England—the northeast and the south. If we compare the three varieties in the northeast of England for which we have both Typical and Emergent data (see Table 7), there is still no evidence for convergence between them (cf. the comments in Trudgill, 1990:81–82). If anything, the Emergents are less similar than the Typicals, although the small number of varieties involved means that the significance of this pattern is impossible to gauge—see *Lack of divergence* above for some suggested reasons for this difference.

A somewhat different pattern emerges when we compare the varieties from the south (admittedly a much larger “region”). Table 8 indicates that although the London and Norwich Emergents are somewhat less similar than the London and Norwich Typicals, Devon Emergent is rather more similar to both the London and Norwich Emergents than Devon Typical is to the London and Norwich Typicals.

Although it is again difficult to gauge the significance of these results given the small number of varieties being compared, the fact that Devon Emergent seems to be acting in a different way from the others is intriguing. One of the features that Devon Emergent shares with London and Norwich but which Devon Typical does not is that it is a largely nonrhotic variety. It seems, on further analysis, that it is this loss of rhoticity that is responsible for the increased similarity of the variety to other southern varieties. In Table 9, the similarity of the southern varieties to each other is calculated *without* including those words which contained historical coda /r/, hence factoring rhoticity out of the equation. As can be seen, there is absolutely no difference in the similarity of Devon Typical and Devon Emergent with the other southern varieties once this is done.

Thus like Edinburgh, the apparent change in similarity between Devon and other varieties appears to be a symptom of one major phonological change—the loss of rhoticity. Whether or not this should be interpreted as a case convergence is debatable—it may, like the spread of TH-fronting, simply represent an example of varieties “falling into step” with others in the diffusion of a change, resulting in a re-establishment of the equilibrium.

Thus there is no evidence in our data (which is admittedly limited in this respect) for convergence at the regional level. Further data will need to be analyzed to

TABLE 7. *The similarities of northeastern varieties*

	Berwick—Tyneside	Berwick—Middlesbrough	Tyneside—Middlesbrough
Typicals	.92	.92	.94
Emergents	.90	.90	.91



TABLE 8. *The similarities of southern varieties*

	London—Norwich	London—Devon	Norwich—Devon
Typicals	.87	.83	.84
Emergents	.84	.87	.87

TABLE 9. *The effect of loss of rhoticity*

	Devon—London	Devon—Norwich
Typicals	.89	.89
Emergents	.89	.89

determine whether this holds true for English in Britain more generally and to evaluate the reality of Trudgill's future dialect areas.

#### AND THE FUTURE

We began this paper by examining the change in English dialects described in Trudgill (1990). The data we have analyzed here support Trudgill's scenario, at least to an extent—there has been a significant convergence in the transition from traditional dialects to modern varieties of English. But as Trudgill, Kerswill, Watt and Milroy, Britain, and Stuart-Smith et al. demonstrated, change, both endogenous and exogenous, is complex and can have unexpected results. Varieties that share exogenous changes *may* converge, but they need not, depending upon the type of change, the pre-existing system that the changes interact with, and the geographical relations of the varieties involved. The result is that varieties change, often in similar ways, but can remain just as distinct as they ever were—that is, they are in a state of dynamic equilibrium.

So what does the future hold for British Englishes? It seems unlikely to us that differences of the order seen between the traditional dialects will arise in the foreseeable future (cf. Trudgill's comments about Australian English). Further data and analysis are required to determine whether varieties at the regional level will remain distinct or are destined to converge to a degree, perhaps focusing on the varieties of major urban areas, as suggested by Trudgill (1990). This paper is only a first attempt to address this complex issue in a quantitative fashion. We have begun by analyzing similarities in segmental phonology but, as Meyerhoff and Niedzielski (2003) demonstrated, it is desirable to analyze similarity across multiple linguistic levels (which may or may not act in similar ways). Likewise, we have only analyzed a rather limited range of speakers which, we hope, gives us some insight into change (or lack of it) in apparent time. It is equally

desirable that this analysis should be broadened to include strictly comparable samples of speakers at each location, something which we have begun to do at two urban locations, Newcastle and Edinburgh. Nevertheless, quantitative methods of this sort have much to tell us about language variation and change, and it is only through holistic quantification between varieties in both social and geographical space that we can hope to determine whether varieties are converging, diverging, or coexisting in a state of dynamic equilibrium.

#### NOTES

1. For more information, see Heggarty et al. (2005), McMahon and McMahon (2005), McMahon et al. (2007), Nerbonne et al. (1999), and Nerbonne and Heeringa (2001). Information is also available at [www.languagesandpeoples.com/MethodsPhonetics.htm](http://www.languagesandpeoples.com/MethodsPhonetics.htm).

2. But not for traditional dialects, thanks to surveys such as the *SED*. We have not (at this time) made use of these data, however, because they do not include modern accents of English, and present problems of strict compatibility in terms of transcription practices.

3. Note that the figure for Rhocitiy is the percentage of words that belong to NURSE, NEAR, SQUARE, START, NORTH, FORCE, CURE, and letter.

4. Thanks to Eivind Torgersen, Peter Trudgill, Mark Jones, Patrick Honeybone and Kevin Watson, Jane Stuart-Smith, Robert McColl Miller, and Raymond Hickey, respectively. Peter Trudgill provided Typical and Emergent transcriptions for Norwich and a Traditional transcription for rural Norfolk. See *Incorporating social variation – the sub-varieties* for an explanation of these terms.

5. The first author also standardized the transcriptions received from other contributors in accord with his own transcription practices.

6. <http://soundcomparisons.com/>.

7. Note that for many such speakers, these sorts of pronunciations go well beyond a few lexicalized survivals, perhaps used to signal local identity, in the way many residents of Newcastle Upon Tyne might refer to *Newcastle* or *Newcastle United*, as the [t<sup>h</sup>u:n] (*Toon/Town*), but never refer to other towns in this way (see Beal, 2000:349).

8. It follows that the fieldworker must be aware, in advance, of the kinds of traditional pronunciations to be expected in each area.

9. Note the speaker's use of [t<sup>h</sup>ä:k] when referring to his "mates", as opposed to [t<sup>h</sup>ɔ:k] in reference to the interviewer (the lead author). The lead author can confirm that this speaker did quite a bit of [t<sup>h</sup>ä:R?n] to him too, however!

10. A randomization independent samples *t* test of TRAD-RP versus TYP + EMG-RP finds that the difference is very highly significant ( $p = .00019$ ), with TRAD-RP (mean .171) > TYP + EMG-RP (mean .114). For details of this procedure, see *Investigating multidimensional relationships*.

11. First, for the original data (e.g., the allocation of speech varieties to Traditional and Typical categories) an appropriate statistic (or summary value) is computed and stored. As an example, for a one-way randomization ANOVA, such a statistic is  $\Sigma_i(T_i^2/n_i)$ , where  $T_i$  is the total and  $n_i$  is the number of observations in group  $i$  (Edgington, 1987:71–74). Afterward, the algorithm randomly permutes the elements of the groups (e.g., randomly allocating varieties to the Traditional or Typical categories) and recomputes the value of the statistic for this new configuration. Finally, after a large number (in our case, 10,000) of such randomized values have been computed, the "original" value is compared to their distribution and, if an extreme (very low or high), then the original structure of the data was probably not due to random sampling (for details see Edgington, 1987). Moreover, the  $p$  value associated with rejecting the null hypothesis is simply the proportion of permuted values more extreme than the original value and can be judged in relation to the standard alpha levels .05 or .01.

12. The box plots in Figures 7 and 8 represent the distribution of distances between pairs of varieties, one belonging to the first type and the second to the other type (e.g., the fourth box plot shows the distribution of distances between all possible pairs composed of one Traditional variety and one Typical variety).

13. Given that conducting many statistical tests of the same type inflates the chance of obtaining a significant result simply by chance, we corrected our  $p$  values for multiple testing using Holm's (1979) method.

## REFERENCES

- Bailey, Guy. (2002). Real and apparent time. In J Chambers, P Trudgill, & N Schilling-Estes (eds), *The handbook of language variation and change*. Malden, MA: Blackwell Publishing. 312–332.
- Beal, Joan. (2000). From George Ridley to viz: Popular literature in Tyneside English. *Language and Literature* 9(4):343–359.
- Britain, David. (2005). Innovation diffusion: ‘Estuary English’ and local dialect differentiation: the survival of Fenland Englishes. *Linguistics* 43(5):995–1022.
- . (2009). One foot in the grave? Dialect death, dialect contact and dialect birth in England. In D Britain, R Vandekerckhove, & W Jongenburger (eds), *Dialect death in Europe?* Special issue of *International Journal of the Sociology of Language* 196/197:121–155.
- Corbet, John, McClure, J. Derrick, & Stuart-Smith, Jane (eds). (2003). *The Edinburgh companion to Scots*. Edinburgh: Edinburgh University Press.
- Edington, Eugene. (1987). *Randomization tests*. New York: Marcel Dekker.
- Ellis, Alexander. (1889). *On Early English pronunciation, part V: the existing phonology of English dialects compared with that of West Saxon*. New York: Greenwood Press.
- Foulkes, Paul, & Docherty, Gerard (eds). (1999). *Urban voices: Accent studies in the British Isles*. London: Arnold.
- Glauser, Beat. (1974). *The Scottish-English linguistic border*. Bern: Francke.
- Harris, John. (1985). *Phonological variation and change: Studies in Hiberno-English*. Cambridge: Cambridge University Press.
- Heggarty, Paul, McMahon, April, & McMahon, Robert. (2005). From phonetic similarity to dialect classification: A principled approach. In N Delbecque, J. van der Auwera, & D Geeraerts (eds), *Perspectives on variation*. Amsterdam: Mouton de Gruyter. 43–91.
- Heggarty, Paul, McMahon, April, & Maguire, Warren. (forthcoming). Splits or waves? Trees or webs? How divergence measures and network analysis can unravel language histories. In J Steele, E Cochrane, and P Jordan (eds), *Philosophical Transactions of the Royal Society (B)* special issue on *Cultural and Linguistic Diversity: Evolutionary Approaches*.
- Hickey, Raymond. (2005). *Dublin English: Evolution and change*. Amsterdam: John Benjamins.
- Holm, Sture. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* 6:65–70.
- Huson, Daniel, & Bryant, David. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* 23(2):254–267.
- Kerswill, Paul. (2003). Dialect levelling and geographical diffusion in British English. In D Britain & J Cheshire (eds), *Social dialectology: In honour of Peter Trudgill*. Amsterdam: John Benjamins Publishing Company. 223–243.
- Kerswill, Paul, & Williams, Ann. (1999). Dialect levelling: Change and continuity in Milton Keynes, Reading and Hull. In P Foulkes & G Docherty (eds), *Urban voices: Accent studies in the British Isles*. London: Arnold. 141–162.
- Maguire, Warren. (2009). Quantifying dialect similarity by comparison of the lexical distribution of phonemes. *International Journal of Humanities and Arts Computing* 2(1–2):261–277.
- Mather, James, & Speitel, Hans. (1986). *The linguistic atlas of Scotland, Scots section, vol. 3, phonology*. Beckenham: Croom Helm.
- McMahon, April, Heggarty, Paul, McMahon, Robert, & Maguire, Warren. (2007). The sound patterns of Englishes: Representing phonetic similarity. *English Language and Linguistics* 11(1):13–142.
- McMahon, April, & McMahon, Robert. (2005). *Language classification by numbers*. Oxford: Oxford University Press.
- Meyerhoff, Miriam, & Niedzielski, Nancy. (2003). The globalisation of vernacular variation. *Journal of Sociolinguistics* 7(4):534–555.
- Nerbonne, John, Heeringa, Wilbert, & Kleiweg, Peter. (1999). Edit distance and dialect proximity. In D Sankoff & J Kruskal (eds), *Time warps, string edits and macromolecules: The theory and practice of sequence comparison*. Stanford: CSLI Press. v–xv.
- Nerbonne, John, & Heeringa, Wilbert. (2001). Computational comparison and classification of dialects. *Dialectologia et Geolinguistica* 9:69–83.
- Orel, Vladimir. (2003). *A handbook of Germanic etymology*. Leiden and Boston: Brill.
- Orton, Harold. (1962). *Survey of English dialects (A): Introduction*. Leeds: E. J. Arnold and Son Ltd.
- Orton, Harold, Sanderson, Stewart, & Widdowson, John (eds). (1962–1971). *Survey of English dialects (B): The basic material*. Leeds: Arnold & Son.
- Romaine, Suzanne. (1978). Post-vocalic /r/ in Scottish English: Sound change in progress? In P. Trudgill (ed.), *Sociolinguistic patterns in British English*. London: Edward Arnold. 144–157.

- Rydland, Kurt. (1982). *Vowel systems and lexical-phonemic patterns in South-east Cumbria*. Bergen: Steens Offset.
- . (1998). *The Orton corpus: A dictionary of Northumbrian pronunciation, 1928–1939*. Oslo: Novus Press.
- Séguy, Jean. (1971). La relation entre la distance spatiale et la distance lexicale. *Revue de Linguistique Romane* 35:335–357.
- Stuart-Smith, Jane. (2003). The phonology of modern urban Scots. In J. Corbet, J. McClure, & J. Stuart-Smith (eds), *The Edinburgh companion to Scots*. Edinburgh: Edinburgh University Press. 110–137.
- Stuart-Smith, Jane, Timmins, Claire, & Tweedie, Fiona. (2006). Conservation and innovation in a traditional dialect. *English World-Wide* 27(1):71–87.
- . (2007). ‘Talkin’ Jockney’? Variation and change in Glaswegian accent. *Journal of Sociolinguistics* 11(2):221–260.
- Thorndike, Robert, & Lorge, Irving. (1944). *The teacher’s word book of 30,000 words*. New York: Bureau of Publications, Teachers College, Columbia University.
- Torgersen, Eivind, & Kerswill, Paul. (2004). Internal and external motivation in phonetic change: Dialect levelling outcomes for an English vowel shift. *Journal of Sociolinguistics* 8(1):23–53.
- Trudgill, Peter. (1990). *The dialects of England*. Oxford: Blackwell.
- . (1999). Norwich: Endogenous and exogenous linguistic change. In P. Foulkes & G. Docherty (eds), *Urban voices: Accent studies in the British Isles*. London: Arnold. 124–140.
- Watt, Dominic. (1998). *Variation and change in the vowel system of Tyneside English*. Ph.D. thesis, University of Newcastle.
- . (2000). Phonetic parallels between the close-mid vowels of Tyneside English: Are they internally or externally motivated? *Language Variation and Change* 12(1):69–101.
- . (2002). ‘I don’t speak with a Geordie accent, I speak, like, the Northern accent’: Contact induced levelling in the Tyneside vowel system. *Journal of Sociolinguistics* 6(1):44–63.
- Watt, Dominic, & Milroy, Lesley. (1999). Patterns of variation and change in three Newcastle vowels: Is this dialect levelling? In P. Foulkes & G. Docherty (eds), *Urban voices: Accent studies in the British Isles*. London: Arnold. 25–47.
- Wells, John. (1982). *Accents of English*. Cambridge: Cambridge University Press.
- Wright, Joseph. (1905). *The English dialect grammar*. Oxford: Henry Frowde.