

Introduction

Charles Boberg, John Nerbonne and Dominic Watt

The origins of dialect variation

DIALECTOLOGY is the study of *dialect*, or regional variation in language. Dialect differences are caused by two forces operating in tandem: language change and the expansion of speech communities. Language change is of course a constant, on-going process in all speech communities. As long as communities remain small, changes are adopted or rejected by the community as a whole, or show only social differentiation. When a speech community expands sufficiently across a territory, however, the network of interpersonal communication that diffuses changes among its members is disrupted: sheer distance, or physical barriers like mountains and bodies of water, and sometimes also cultural, economic or social divisions, make it impossible for change to diffuse evenly across the entire community. Eventually, an accumulation of undiffused changes causes community members to recognize that people in other parts of the community speak a different version of their language, what we would call a dialect.

Given enough time, this process of differentiation can cause dialects to diverge to the point where they are no longer wholly mutually intelligible, in which case we begin calling them separate but historically related languages. Such divergence lies at the heart of how historical linguists conceive of the development of families of related languages, like the Indo-European languages spoken across most of Europe and the Americas today, which hypothetically began their individual existence as dialects of a common ancestral

language. In other cases, dialect differences can persist in a stable relationship for centuries, without leading to language divergence, or can decline and disappear, as the communication barriers that produced them are overcome by social or technological change. All normal languages, except those spoken in single, restricted locations, display regional variation and have always done so: accounts of dialect differences are as old as written language itself, appearing two millennia ago in Ancient Greece and China. Given its universality, dialect variation should be seen as a fundamental aspect of human language and dialectology an important branch of linguistics, the scientific study of language: a linguistics that did not include dialectology would be incomplete.

Languages vary in many ways: across time and space, as just discussed, as well as across social categories. Today, dialectology is often seen as part of a larger sub-discipline of linguistics dealing with all of these types of variation, collectively called *language variation and change* (see, e.g., Chambers and Schilling 2013, another handbook in this series). This integrative approach reflects the many ways in which these types of variation have been shown to interact, first brought into clear focus in the work of William Labov (see below). Much of the variation we observe in speech communities is in fact the synchronic manifestation of diachronic processes, or changes in progress: newer forms, before being uniformly adopted, compete for dominance with older forms, in patterns that reflect an intersection of regional and social influences. Nevertheless, as difficult as it can be to isolate regional from other types of variation, the primary focus of this book will be on regional variation.

Defining dialects

We shall begin our discussion of regional variation with just this problem, by exploring the meaning of the word *dialect*, which cannot be properly understood without reference to social variation as well. As linguistic variation arises in speech communities, it usually reflects social differences: different ways of speaking, like different ways of eating or dressing or having fun, come to be associated with groups arrayed on a socio-economic hierarchy involving wealth, power, education, ethnic or social identity and other factors. Varieties of speech associated primarily with social groups are properly called *sociolects* rather than *dialects* and are the main focus of the allied sub-discipline of *sociolinguistics*, but this type of variation also has an important place in dialectology, since regional varieties of a language -- the definition of *dialects* given above -- often develop social attributes. In particular, one variety, usually that spoken by the social, economic and political élite in a nation's capital city or other great metropolis, normally comes to be seen as the 'correct' form of the language. In many cases, this evaluation is shared not only by its speakers, who use it as a symbol and even a justification of their higher social position, but also by others in the community, who accept that their own speech is by comparison inferior, or 'incorrect'. Because of its perceived social superiority, the élite variety is promoted to the status of a regional or national 'standard' variety, which is preferred or even required in domains like broadcasting, education, government, journalism, the law, literature, liturgy and science. It often serves these functions not only in its city or region of origin but across the entire linguistic territory, at higher social levels. This establishes a nationwide diglossia between the pan-regional 'standard' variety, which comes to be seen not as a dialect but as the unmarked form of the language

itself (for instance the form taught to foreigners who want to learn the language), and the regionally restricted and socially inferior ‘dialects’, which continue to be the language of everyday life for peasants or farmers in the countryside and for factory workers and trades people in the towns and cities. Rural and urban dialects often receive distinct social evaluations. Rural dialects are frequently seen as quaint and musical, if also unsophisticated and somewhat comic, and are associated with idyllic notions of traditional country life. Urban dialects are more often seen as lazy, ignorant and linguistically and morally degenerate, since they are associated (at least in many middle-class minds) with the social problems of the working-class sections of large cities.

An amusing instantiation of this ideology can be seen in the animated adaptation of Kenneth Grahame’s children’s story *The Wind in the Willows* that was made in the 1980s for Thames Television in the U.K. Though all of the characters are animated figures of animals, the heroes of the story, Rat, Mole and Badger, speak with subtly different versions of standard British English, or ‘Received Pronunciation’; the sympathetic minor characters, like a plainspoken otter and a benign cow, have rural, West Country dialects; but the local gang of criminals, the weasels, are given working-class London (‘Cockney’) and Manchester dialects. That said, the great fool of the piece, Mr. Toad, the lord of the local manor and a sort of upper-class twit, has the poshest accent of all, reminding us that the correspondence between high-class speech and positive social attributes is not always simple or direct. Nevertheless, the fact that this is a children’s program – and a delightful and brilliantly produced one at that, it should be admitted – emphasizes the extent to which dialect ideologies are inculcated in children at a young age by schools, media and other institutions.

Even more problematic than negative attitudes about dialects is the transfer of these attitudes to the speakers themselves: people who speak what some think of as ‘lazy’ or ‘ignorant’ dialects are thought of as lazy or ignorant themselves, a stereotype that can be used to justify denying them educational, occupational or social opportunities. Conversely, speakers of standard varieties may be given unfair advantages in the same contexts, a fact that has encouraged many ambitious people from working-class social backgrounds to try to ‘improve’ their speech, often with measurable benefits. This, indeed, is the main justification for teaching standard varieties in schools, whose main purpose is to maximize the socio-economic opportunities of their students: defenders of the exalted status of standard varieties might argue that they are, in fact, democratizing (or at least meritocratizing) instruments, since they can be learned in school or by other means, thereby conferring socio-economic benefits on the ambitious and becoming a symbol of individual achievement rather than of inherited privilege. Sociolinguists have argued passionately – and correctly – that these notions are based on social prejudice rather than linguistic fact, but they have proven very difficult to dislodge from popular culture, persisting at both ends of the social spectrum (for a critical look at the concepts of ‘standard’ v. ‘dialect’ in English, see Milroy and Milroy (1999) and the contributors to Bex and Watts (1999)).

Not all ‘dialects’, of course, are socially stigmatized, at least not by general consensus. Many non-standard dialects, if they lose points on the ‘status’ dimension that governs access to the most prestigious schools and jobs, gain them on the ‘solidarity’ dimension, with their speakers being perceived as friendlier, more attractive, more relaxed, funnier or more honest than speakers of the standard variety, if not more suitable

as surgeons or bank presidents. Other non-standard dialects may be generally disparaged by people outside their own region or social group but are the focus of intense local pride within it. Speakers of these dialects often have a correspondingly negative view of the standard variety and its speakers: as Fischer (1958: 56) observed half a century ago in the pioneer of sociolinguistic studies, “A variant which one man uses because he wants to seem dignified another man would reject because he did not want to seem stiff.” Still other non-standard dialects are valued even by speakers of the standard variety as genuinely beautiful or cultured, even if inappropriate for some of the domains reserved to the standard variety.

Moreover, not all regional differences are socially marked. It is easy to think of variables in North American English, for example, that appear to be purely regional, with no common perception that one variant is more correct than the other. This is often true of lexical variation, which opposes forms like *see-saw* and *teeter-totter*, both meaning a tilting board that children play on, or *cottage* and *cabin*, terms for a rural summer vacation home, or *pop* and *soda*, generic terms for non-alcoholic carbonated beverages, without social prejudice. Many regional phonological variables, too, seem to lack social symbolism: Americans as a whole have no opinion on whether it is correct to pronounce pairs of words like *cot* and *caught*, or *stock* and *stalk*, differently, as in large sections of the eastern half of the country, or the same, as in most of the western half (where opinions exist, they relate to the phonetic qualities of the vowels involved, not the presence of phonemic contrast). Grammatical variation, by contrast, is more frequently aligned with social factors: everyone in the United States, as well as in other English-speaking countries, knows that ‘double negatives’ and lack of ‘standard’ subject-verb

agreement are ‘wrong’ and that those who use them mark themselves as lower-class, a message continually reinforced by schools and other institutions. Potential interactions with social factors, then, are an important aspect of dialect study.

If national standard varieties of languages coexist in a diglossic relationship with dialects of those languages, they also, in the case of multinational languages like English, French and Spanish, coexist with other national standards. In this context, such ‘standard varieties’ are themselves ‘dialects’. In some cases, as between fellow ex-colonies like the United States and Canada or many Latin American countries, these relationships are fairly egalitarian, with national differences viewed as purely regional rather than social. In other cases, as between ex-colonies and their former colonizers, unequal sociolinguistic status can persist long after political independence. A general equality between the standard varieties of British and American English has now, after two centuries, come to be accepted by many English-speakers, including those in second-language education; for instance, few people in the United States would consider shifting toward British standards when reading the news on television or teaching the language to foreigners. This equality, however, reflects the enormous size, power and prestige of the United States, which has clearly surpassed that of Britain. By contrast, relations between standard Parisian French and the ex-colonial varieties of French spoken in Canada or other parts of the former French empire are still more hierarchical, with varieties closer (though not necessarily identical) to European French preferred in broadcasting and second-language teaching, for example. In some cases, opinion about such matters is regionally divided: while many in Spain consider *castellano*, the standard variety of Iberian Spanish based on the dialect of Castile, a global standard, Latin Americans are

less likely to accept this notion and the form of Spanish taught in the United States most commonly follows a Mexican rather than Castilian standard, for instance in failing to preserve the Castilian distinction between *s* and *z* (*casa*, ‘house’, v. *caza*, ‘hunts’).

While dialects can differ at every level of structure – phonetic, phonological, morphological, lexical, syntactic, semantic, etc. – the term *dialect* is often used in a complementary relation with another term, *accent*, whereby *dialect* means differences in grammar and lexicon, while *accent* is restricted to phonological and especially phonetic differences, such as the quality of vowel sounds (as in the exhaustive survey of ‘accents of English’ compiled by Wells 1982). This distinction takes on an important social dimension in Britain, for instance, where a three-level structure of language variation has traditionally been observed: the national élite, particularly those educated at Oxford and Cambridge Universities, spoke ‘standard’ British English with a ‘standard’ or non-regional accent known as ‘Received Pronunciation’, regardless of where they lived (at least within England -- the Celtic ‘nations’ were to some extent exempt from this standard and had their own regional standards); the urban middle class spoke ‘standard English’ with a regional ‘accent’, differing from the élite ‘standard’ only in pronunciation, especially of vowel sounds; and the working class, urban and rural, spoke regional ‘dialect’, with non-standard grammar and lexicon, which also implied a marked regional ‘accent’. These social distinctions have recently been waning, with a decline in élite use of some traditionally prestigious features now seen as unattractively snobbish and a deliberate promotion of regional accents in domains like national broadcasting (the BBC) where they were not previously accepted. Nonetheless, to a large extent this differentiation can still be heard today and might also be argued to apply increasingly to

the United States, where a sort of non-regional ‘General American English’ appears to be pushing out local speech patterns among the social élite in many regions.

A particularly problematic issue in defining *dialect* has been its taxonomic relation to the term *language*, the latter supposedly comprising a set of mutually intelligible dialects: if two people speak differently but can understand one another, they are speaking dialects of the same language; if they cannot understand one another, they are speaking different languages. It has often been pointed out that popular and even academic ideas about classifying varieties as languages or dialects reflect non-linguistic factors, like political boundaries and cultural history, as much as strictly linguistic criteria of mutual intelligibility. The stock examples in this discussion include, on the one hand, Mandarin and Cantonese, which many people think of as dialects of a single Chinese language but which are not mutually intelligible in speech (see Tang, this volume); and on the other, Hindi and Urdu, spoken in India and Pakistan respectively, which many people think of as separate languages but which are in fact largely mutually intelligible, separated more by an international boundary and by the cultural and religious affiliations of their speakers than by any marked linguistic divergence (see Deo, this volume). Europe, too, includes many instances of political boundaries creating and reinforcing ‘language’ differences across what were once gradually shifting continua of local dialects, such as those between Germany and the Netherlands (see Kürschner, this volume), Spain and Portugal (see Lipski, this volume) or parts of the former Soviet Union (see Zhobov and Alexander, this volume); Italy presents a particularly complex blend of regional ‘dialects’ and ‘languages’ that are all offshoots of Latin (see Telmon, this volume).

Mutual intelligibility is itself a hazy concept, of course, involving not the binary distinction implied by the terms *language* and *dialect* but a cline or scale of linguistic similarity. At one end of the scale, we find cases of minimal regional difference with unrestricted mutual intelligibility, like the major national standard varieties of English or Spanish: middle-class people in London and Los Angeles, or Madrid and Mexico City, recognize clear differences in each other's speech but have very little difficulty understanding one another, if any at all. At the other end of the scale, we find complete unintelligibility, as between English and Arabic or Mandarin. In the middle, however, are many degrees of partial intelligibility. Some of these involve varieties that differ markedly from the most widely recognized international standards, such as the types of English spoken in Glasgow, Belfast, Appalachia, Jamaica, Singapore or Nigeria. Others involve closely related 'languages', such as the Scandinavian or Romance languages, which began their histories as dialects of a common ancestral language and still retain a large common grammar and vocabulary, but have since drifted far enough apart to make mutual comprehension difficult, especially in speech. In many of these cases, moreover, the partial intelligibility that does exist is not symmetrical: Danes understand Swedes better than Swedes understand Danes and Portuguese speakers can generally make out more Spanish than vice versa. Intelligibility can be affected by non-linguistic factors like education, exposure and the comparative social status and population sizes of the languages and cultures involved, as much as by purely linguistic matters like sound change or vocabulary replacement. These problematic issues will be reprised in several chapters of this book.

The origins and development of dialectology

From issues surrounding the nature and definition of dialects, let us now turn to a brief review of the history of dialectology, setting the stage for the chapters that follow. While it certainly has precedents in other places and earlier times, the modern ‘western’ tradition of dialectology began in Europe in the nineteenth century. For many of its earliest practitioners, dialect study was a hobby: an entertaining pastime for self-taught philologists with an interest in cultural history and folkways and a romantic conception of rural life, then very much in fashion (seen also, for instance, in the literature, music and painting of the period). Some early dialect collectors, for instance, were parish priests or school teachers, who had both a measure of formal education and a strong connection to the local communities they served. Henry Higgins, the fictitious dialect phonetician parodied by George Bernard Shaw in his 1913 play *Pygmalion*, though based partly on real-life phonetician and philologist Henry Sweet (1845-1912), comes off more as a gentleman of leisure with eccentric interests than as the sort of person we would recognize today as a professional academic or serious scholar. Early interest in dialect was given extra urgency by a genuine concern, not altogether unjustified, that rural culture would soon be irretrievably altered or lost as industrialization and urbanization increased. This ‘curatorial’ approach to dialect study sought to record as much of traditional rural speech as possible before it was too late, not unlike the efforts of modern linguists to record and study the thousands of indigenous and minority languages whose vitality is now threatened by digital technology and cultural globalization. Dialects also came to be seen as entertaining by the growing urban bourgeoisie of the nineteenth

century, who enjoyed tittering at rustic stereotypes presented in theaters and music halls or in novels. Many of the greatest writers of the time, like the Brontë sisters, Dickens and Hardy, filled their novels with passages of dialect, not only as a creative device, adding realism to their rural or working class characters, but as comic entertainment for their largely urban, middle-class readers (a tradition that continues today in film and television).

Over time, however, the subject also developed a more serious, academic side. As will be recounted in several subsequent chapters (see also the general accounts of the development of dialectology in Petyt 1980, Francis 1983 or Chambers and Trudgill 1998), serious academic study of dialects began in Germany, where Georg Wenker and his colleagues carried out a postal survey of dialect variation across the German-speaking territory of Europe, starting in the 1870s (Wenker *et al.* 1927-1956). Wenker asked school teachers to translate a set of 40 sentences into local dialect, as they observed it in their communities; he then collected these records and compiled them in a *Deutscher Sprachatlas*, or ‘German language atlas’, showing where each form was found and how one region differed from another. This effort was closely followed by the *Atlas linguistique de la France*, published by Jules Gilliéron in several volumes over the first decade of the 20th century. The French study was based instead on face-to-face interviews with dialect speakers carried out in the field, using a standard questionnaire administered by a trained fieldworker (Gilliéron and Edmont 1902-1920).

The ultimate goal of these projects, like many that came after them, was to produce a *dialect atlas* (see Kretzschmar, this volume): a collection of maps showing the regional distribution of linguistic variants – different words, pronunciations or

grammatical forms – over a given territory (usually the territory covered by speakers of a single language, or a subdivision of that territory). On these maps, symbols or transcriptions indicated the variants occurring in each location and lines called *isoglosses* could be drawn to divide spatial distributions of variants or mark the outer limit of a distribution; bundles of these isoglosses were taken to indicate major dialect boundaries. This aspect of dialectology is also known by the term *dialect geography*. As the name implies, dialect geographers used their maps to develop geographic interpretations of the spatial distribution of dialect forms, such as the role of barriers to communication, like mountain ranges, in preventing the diffusion of variants and thereby creating dialect divisions, or of channels of communication, like rivers and roads, in encouraging diffusion over wider areas; they also turned to information on cultural and settlement history in their efforts to explain the location of dialect boundaries.

Alongside dialect geography, an allied tradition of *dialect lexicography* also emerged in the 19th century, which involved the production of dictionaries of dialect words and phrases, with definitions, examples and usage notes, recorded in list form rather than on maps (see Van Keymeulen, this volume). At the turn of the 20th century in England, for instance, Alexander Ellis published his records of dialect pronunciation, collected two decades earlier (Ellis 1890), and the English Dialect Society produced an *English Dialect Dictionary*, compiled by Joseph Wright (1898-1905).

As the field evolved, the interests of many dialectologists expanded beyond dialect variation itself to include connections with other aspects of language study. For instance, some dialectologists became involved in a debate with linguistic historians over the nature of language change. A group of 19th-century historical linguists known as the

Neogrammarians had proposed that sound change – gradual shifts in the pronunciation of sounds found in sets of words – was a regular and exceptionless process that operated rather like the physical laws of natural science (Osthoff and Brugmann 1878).

Systematic sound changes, gradually transforming all of the instances of a given sound simultaneously, were held to be responsible for the linguistic diversification of speech communities. This process had given rise, over thousands of years, to the families of “genetically” related languages and their distinct branches observable in contemporary Europe, particularly the Indo-European family and its Italic, Celtic, Germanic, Slavic and other offshoots, as mentioned above. For example, the initial /p/ sound of the hypothetical ancestral Indo-European language, evident in Latin words like *pater*, *pe(di)s* and *piscis*, had become an /f/ in the Germanic languages, producing German *Vater*, *Fuß* and *Fisch*, or English *father*, *foot* and *fish*, against French *père*, *pied* and *poisson* or Italian *padre*, *piede* and *pesce*, among dozens of other examples. Other linguists saw Neogrammarian theory as an extreme view, which idealized the process of sound change and ignored a great deal of contradictory evidence.

As dialectologists began their survey work, the Neogrammarians initially hoped that the collection of data on traditional rural dialects, free of the complicating factors of urban speech communities and standardized literary languages, would prove their theory right, by showing systematic and regular application of sound changes. When these data began to be analyzed, however, dialectologists found that they often revealed glaring exceptions to the hoped-for patterns of regular change. In some villages, a mixture of changed and unchanged forms was found, suggesting that some changes, at least, were irregular, affecting some instances of a sound but not others, and that the basic unit of

phonological change was the word, not the phoneme, or sound. A classic illustration comes from Dutch: Kloeke (1927) found that local forms of *house* and *mouse*, which both had long /u:/ in Proto-Germanic and should have followed parallel developments according to Neogrammarian theory, displayed different sounds in some Dutch towns, a direct contradiction of the regularity of sound change (see Bloomfield (1933: 328-331) for an influential discussion of these data).

In response to the Neogrammarian dictum that sound change is regular and suffers no exceptions, dialectologists therefore advanced their own, opposite slogan, that “every word has its own history,” apparently denying any sort of regularity to sound change. In its French form, *chaque mot a son histoire*, this view is usually attributed to Gilliéron (see Gilliéron and Roques (1912)), but it goes further back to Hugo Schuchardt in the 19th century and perhaps as far as Grimm (1819: XIV), who says, “...*jedes Wort hat seine Geschichte und lebt sein eigenes Leben*” (‘every word has its history and lives its own life’). The dispute over the regularity of sound change produced a deep cleft between what would become two separate traditions of linguistic thought. The dialectologists accused the Neogrammarians of ignoring the complexity of actual data in their efforts to attain higher levels of generalization and theoretical abstraction, while the Neogrammarians accused the dialectologists of obsessing over minutiae and variability for their own sake, like stamp collectors, without addressing questions of broader scientific significance. This rift is still observable today, in the division between formal theoretical linguistics, which is in some ways the heir of Neogrammarian philology (with other important influences, like the work of Saussure), and the field of language variation and change, including much of modern historical linguistics, sociolinguistics and

dialectology, which carries on the more skeptical or at least more empirical and data-oriented viewpoint of the 19th-century dialectologists.

While formal theoretical approaches to the study of language, such as the structuralism of the mid-20th century and the generative school of the late 20th century, came to dominate modern academic linguistics, especially in eastern North America, a robust tradition of work on language variation and change, including dialectology, also continued to thrive, even if it was increasingly sidelined in many prestigious linguistics programs at major universities. By the 1930s, the French method of interviewing dialect speakers in the field and making meticulous records of their speech that could later be transformed into maps was extended to North America by Hans Kurath, who produced a *Linguistic Atlas of New England*, intended to be the first of several regional dialect atlas projects that would eventually cover the entire continent (Kurath *et al.* 1939-1943). This project was sadly never completed but has nevertheless produced a great deal of data and a tradition of work that continues today. In addition to the original New England atlas, the major published atlases of American English now cover the Middle and South Atlantic states (McDavid *et al.* 1980), the Gulf states (Pederson, McDaniel and Adams 1986-1993) and the Upper Midwest (Allen 1973-1976). Following World War II, the *Survey of English Dialects* published maps of dialect variation across England (Orton and Dieth 1962-1971; see also Upton and Widdowson 1996), and in the 1960s a second major American dialectology project, the *Dictionary of American Regional English*, was undertaken, which is now complete (Cassidy and Hall 1985-2012; see also Carver 1987). Traditional dialect survey work has also continued in many other countries across the globe, as attested in the chapters of Section 3 of this volume.

Dialectology received a new stimulus in the 1960s from the work of William Labov in the closely allied field of *sociolinguistics*, which investigates relationships between linguistic variation and social structure and identity. One of the main concerns of early dialectology, as mentioned above, had been the effects of urbanization, mass education and other forms of social change on traditional rural dialects, which were feared to be disappearing. A priority of many dialectologists was therefore to collect and study records of these dialects before they were lost. The best exemplars of traditional dialects were thought to be older rural males with minimal formal education and long family histories in the region, who were consequently favored as informants. Comparatively little interest was taken in other types of speakers, who were seen as less representative of ‘pure dialect’, or in cities, which were seen to offer nothing more than chaotic mixtures of modified regional dialects brought in by migrants from the surrounding countryside, or working-class urban varieties that were seen as linguistically and morally corrupt. By the 1960s, these assumptions no longer seemed justified. A new generation of sociolinguists sought to base their descriptions and theories of linguistic variation on the speech of the majority of the population. In the United States, Britain and other western nations, this majority now lived in cities, where it comprised not only old men of local stock but women, young people, recent migrants and a wide range of social classes and ethnic groups, including those who spoke varieties stigmatized as debased, indolent and ugly.

When Labov began to study urban speech communities, starting with New York City, he found that they displayed not the chaotic dialect mixture dismissed by some dialectologists as uninformative but *orderly heterogeneity*, a pattern in which the

probability of occurrence of competing linguistic variants, such as standard and non-standard pronunciations or grammatical forms, depends on a complex yet systematic interplay of many different factors (Labov 1972). These factors included social attributes of speakers, like age, sex and social class, as well as speech style, or the social context of speech (the identities of interlocutors, the setting and topic of conversation, etc.).

Labov's focus on correlations between linguistic and social variables and on shifting frequencies of variants has earned this type of sociolinguistics the names *correlational* or *quantitative* or *variationist sociolinguistics*. Some have called it *urban dialectology* (e.g., Chambers and Trudgill 1998: 54), though of course the sort of variation that Labov and others have studied in major cities can also be found in small towns and rural communities (as Labov himself did on the island of Martha's Vineyard, off the Massachusetts coast), if on a smaller scale reflective of their narrower range of social diversity. Once the focus of interest shifts from a curatorial mission to preserve obsolescent traditional speech varieties to a more objective interest in how language reflects social identity, systematic variability can be found in any speech community; subsequent studies in British cities like Norwich (Trudgill 1974a) and Glasgow (Macaulay 1977) clearly demonstrated that this property of speech communities was not unique to New York City or to the United States. Moreover, communities can be compared simultaneously in the regional and social dimensions: they differ both one from another and within themselves, so that regional comparisons have to take local, community-internal variability into account. For instance, regional divergence may be greater at certain social levels, or among particular ethnic groups. This hybrid approach

has been called *social* or *socio-dialectology* (e.g., by Rona 1976; see also Kristiansen, this volume).

Labov's contributions to modern dialectology go beyond shifting the focus to cities. He also pioneered the use of acoustic phonetic analysis to make detailed and reliable measures of vowel quality (Labov, Yaeger and Steiner 1972; see also Thomas, this volume), which could be used to track the progress and distribution of sound changes that were continually modifying the pronunciation of urban dialects in contemporary American English. This work produced a new hybrid subfield normally called *socio-phonetics*; despite this label, the variation measured and analyzed by these techniques was as much regional as social.

Finally and perhaps most importantly, following on the insights of several predecessors (Weinreich 1954; Moulton 1960, 1962), Labov sought to re-establish connections between dialect study and theoretical linguistics, particularly structural phonemics (see Gordon, this volume). By framing his investigations of linguistic variation in terms of major questions of linguistic theory, such as resolving the Neogrammarian controversy discussed above (Labov 1981), or explaining what kinds of linguistic elements typically get transferred between dialects in contact situations and what kinds need to be learned by children from their parents (Labov 2007), Labov hoped to end what he saw as the intellectual isolation of dialectology, thereby augmenting its scientific value and stature. At the most fundamental level, he argued persuasively that the development of linguistic theory should not be divorced from the close study of data on how language is actually used by real people in real communicative contexts, and must give a satisfactory account of the variability found in these data, rather than

dismissing it as optional rule application. He further believed that the study of dialect, in turn, could profit from new insights provided by reference to concepts and questions in general theoretical linguistics, as shown by Chambers (1973) and discussed here by Hinskens (this volume). Though not all dialectologists today feel the need to engage with questions of general linguistic theory, Labov's work has illuminated many opportunities for those who do; it has transformed the modern study of language variation and change, producing a whole new tradition of dialect study best represented by the *Atlas of North American English* (Labov, Ash and Boberg 2006), the journal *Language Variation and Change*, and the annual *New Ways of Analyzing Variation (NWAV)* conference, now a major venue for the latest research in dialectology as well as sociolinguistics.

Modern dialectology has seen other advances as well. Where traditional dialectologists had to draw their maps by hand, a laborious and necessarily selective process, today's practitioners benefit from a growing array of computer cartography tools, which support new insights into the regional distribution of dialect variants (see Rabanus, this volume; also Wikle 1997; Lameli, Kehrein and Rabanus 2010). Another group of scholars has been concerned with developing objective measures of dialect difference that rise above anecdotal accounts and avoid selective analysis, a subfield known as *dialectometry*, which has incorporated sophisticated and rigorous quantitative methods from other social sciences and from statistics and computer engineering (Goebel and Nerbonne and Wieling, this volume; also Goebel 1982; Kretzschmar 1996; Boberg 2005; Nerbonne and Kleiweg 2007; Nerbonne and Kretzschmar 2006; Nerbonne 2009; Grieve, Speelman and Geeraerts 2011).

Still other dialectologists have focused on the nature of the borders between dialect regions (Watt, Llamas and Johnson 2010; Watt and Llamas 2014). These have often been observed to be ‘fuzzy’, involving *transition zones* in which the features of neighboring dialects are commingled, with a gradual shift from one dialect to the next, rather than sharp, with a sudden and easily perceptible change in speech at a specific location. Related to these topics are studies of the rise of dialect continua across stretches of terrain (Heeringa and Nerbonne 2001), of contact between dialects (Trudgill 1986), of the transitional forms that arise from this contact (Britain 1997; Chambers and Trudgill 1998), of the spatial diffusion of linguistic elements from one dialect to another (Trudgill 1974b; Callary 1975; Bailey *et al.* 1993; Auer, Hinskens and Kerswill 2005) and of the rise of new dialects created by migration and dialect mixture (Kerswill and Williams 2005).

Another recent trend in dialect study has been to turn from production to perception, by examining what ordinary people think about the dialect diversity that surrounds them (Preston, this volume; Preston and Long 1999); how dialect differences interfere with cross-dialectal intelligibility (Labov and Ash 1997); or how listeners categorize speakers by dialect region and which features these categorizations rely on (Clopper and Pisoni 2004).

Finally, dialectology, like all fields of study, has not been immune to the recent influence of the Internet, which presents new opportunities (and challenges) for data collection and analysis, dramatically increasing both the quantity of data available to researchers and the speed at which these data can be collected and analyzed (e.g., Eisenstein, O’Connor, Smith and Xing 2010; Grieve, Asnaghi and Ruetten 2013).

Technological advances have contributed to the creation and analysis of large, searchable corpora of data, as discussed by Szmrecsanyi and Anderwald (this volume), allowing conclusions about variation and change to be drawn from ever larger sets of data. In particular, internet searches make possible the rapid collection of vast quantities of data on regional variation in ordinary language – most commonly written language but also speech – as opposed to language deliberately collected for the purposes of study, a potentially transformative change that minimizes the gap between dialectology and the variation it attempts to study.

The present and future state of dialectology

Studies of diffusion, which includes the spread of features not only from one region to another but from social group to another, give rise to the question of whether the advent of mass education, personal mobility and instant communication in modern, industrialized nations threatens the very survival of dialects, echoing the original concerns of 19th-century dialectologists (Britain 2009; Kristiansen 1998). Insofar as many traditional rural and non-standard urban dialects are now declining or disappearing, this implies a gradual contraction of the subject matter of dialectology, which might suggest a rather pessimistic future for the field. On the other hand, many older dialects are sustained by a strong force of local identity that prevents their decline. Even in fully industrialized or post-industrial countries like Britain and Germany, some distinctive regional dialects, like those in the North of England (e.g. Tyneside, Yorkshire or Lancashire English) or the South of Germany (e.g. Bavarian or Swabian German),

continue to be spoken enthusiastically by millions of people and show little sign of disappearing anytime soon, even if their features are constantly modified by contact with non-local speech. In other cases, like that of Denmark, which saw traditional dialects virtually disappear over the 20th century (Kristensen and Thelander 1984; Pedersen 2003), local identity manifests itself today in subtler forms of variation, sustaining small regional differences in an otherwise homogeneous supra-regional or national type of speech that diffuses from cosmopolitan centers (Kristiansen 1998). Boberg (2005, 2008, 2010) observes a similarly fine-grained yet tenacious regional differentiation, which we might call micro-variation, in Canadian English, which is otherwise reputed to be remarkably homogenous over the country's vast territory.

Moreover, if we look beyond traditional dialects, we find a proliferation of new dialects constantly emerging. This is true in many large urban centers, for example, whose populations are increasingly diverse: distinct ethnic and cultural subgroups in cities like London (Cheshire et al. 2011), Berlin (Wiese 2009), Stockholm (Kotsinas 1988) and Copenhagen (Quist 2000) mark their emerging social identities in linguistic as well as other ways, though intra-community variation of this kind is strictly speaking more a concern of sociolinguistics than of dialectology.

An even more important source of new dialects, however, is the most consequential linguistic development in the world today: the rise of English as a global language. This phenomenon, which has generated a whole new subfield of language variation and change called *World Englishes*, with its own conferences and journals, fundamentally involves the creation and development of dozens of new dialects of English (Crystal 2003; Schneider 2007; Kachru, Kachru and Nelson 2009; Melchers and

Shaw 2013). These were spoken at first in countries with historical ties to the former British Empire, where many native yet distinctive varieties of English are to be found, usually in multilingual settings (e.g., India, Ghana, Nigeria, Hong Kong and Singapore; the ‘outer circle’ of Kachru (1985)). Today, they are increasingly flourishing in countries with no such connection, where English has no historical status but is now used as a lingua franca for intercultural and international communication (Kachru’s ‘expanding circle’), especially in domains such as advertising, digital media, diplomacy, high technology, international commerce, international sport, popular entertainment, post-secondary education, scientific research and of course tourism. In these contexts, new, non-native varieties of English, which may become native varieties in the future, exhibit distinctive features that reflect the native languages with which they coexist.

In Europe, for instance, where English has become the *de facto* working language of the European Union and is widely learned as a second language in primary school by children across the continent, the point at which future generations will regard Dutch English, Swedish English, even German English as legitimate dialects of English, spoken by bilingual populations, seems less remote every year (after all, Irish, Scottish and Welsh English, now universally accepted as dialects of English, also began as second-language varieties, spoken by Britain’s Celtic populations). Assuming this scenario persists (which should not be taken for granted), the evolution of Russian English, Chinese English and Japanese English may not be far behind. Once these originally second-language varieties become semi-native, mutually-intelligible regional types of English, they enter the legitimate domain of dialectology; from this perspective, the future of the field looks bright (if only for English dialectology!).

In short, the field of dialectology has grown and adapted in many ways and continues to respond to a changing environment today. While it no longer holds the central position in linguistic science that it enjoyed in the 19th century, it nevertheless remains a dynamic and relevant sub-discipline that continues to produce new scholarly work and attract new generations of students. From its origins in Europe it has now spread across the globe, with dialect studies available or in progress on languages spoken in every region of the world, as seen in Section 3 of this volume. Yet, despite all of these changes, most dialectologists today continue to focus on the central questions that gave rise to the field over a century ago:

1. How do languages vary across the territories in which they are spoken?
2. What are the common patterns in this variation, including the linguistic constraints that govern it, viewed across different languages?
3. How do settlement history, topography, social patterns, urbanization and other non-linguistic factors explain the spatial distribution of linguistic features?
4. What is the nature of the transitions or boundaries between spatial distributions?
5. How do innovative features spread across new territory?
6. Is regional variation receding, stabilizing or increasing over time?

Despite their long history, all of these questions remain relevant today, as they are addressed with the new methods described above and with new data, both from new communities and from previously studied communities that continue to change.

Rationale and plan of this book

Given the recent expansion and diversification of dialectological scholarship reviewed above, contemporary students of dialect at all levels of expertise now face a significant challenge. They must keep up with technical and theoretical advances in a wide range of different sub-disciplines, as well as with a constantly growing body of data on dialect variation in a wide range of languages. Moreover, as the demands of assimilating all of this new information grow heavier, it becomes more difficult – yet no less important – to maintain an intellectual connection between contemporary research and the scholarly achievements of the past. Foundational work should always be taught and re-taught as an underpinning for modern research, but also critically re-evaluated in light of new information and alternative, innovative thinking.

In light of these challenges, the field has become far too large for even the most senior and widely-experienced scholar to have more than a passing acquaintance with all of its various sub-divisions, let alone for the junior scholar or beginning student who wishes to progress beyond the surveys available in short introductory monographs suitable for undergraduate courses. Yet those with an interest in developing a broad knowledge of dialectology, or in having access to such knowledge on an as-needed basis, have been faced with making their own surveys of a very large and in some cases inaccessible corpus of materials, which is simply not available in many libraries. Dialect atlases in particular are expensive and space-consuming luxuries found only in large or specialized collections, while much of the original work on non-English dialects, particularly that written before the late 20th century, has not been translated and cannot be read by most English-speaking students even where it is available.

The present volume therefore seeks to provide both experienced practitioners and their apprentices with an overview of the field of dialectology -- past and present -- comprising three main aspects of the topic: principal theoretical approaches; methodological traditions; and sets of data. This trio of topics provides the main organizational basis for the book, reflected in the three sections of the Table of Contents, each comprising twelve chapters. Because dialectology is and always has been fundamentally a data-driven field, committed to empirical investigation more than to theoretical speculation, or rather to basing the development of theory firmly on competently collected and analyzed sets of data, the methods of dialectology and the data of dialectology are just as important to any review of the field as the various aspects of dialectology theory. Accordingly, while the book's first section gives a detailed account of the historical and contemporary development of dialectological thinking, including crucial concepts like the dialect dictionary, the dialect atlas and the various interfaces with other areas of linguistics and non-linguistic sciences discussed above, the second section is concerned entirely with methodological matters -- how dialect data are collected and analyzed -- and the third with the data themselves, illustrated with descriptive overviews of dialect variation in the world's most widely spoken languages and language families, particularly those that have produced the richest traditions of dialect study. Primary editorial responsibility for each of these sections was assigned to one of the three co-editors of the volume: Watt oversaw the section on theory, Nerbonne that on method and Boberg that on data. Beyond this co-written general introduction, each section editor provides a specialized introduction to his section, which introduces and discusses the chapters it contains in more detail than is possible or appropriate here.

Works cited

- Allen, Harold B. 1973-1976. *The Linguistic Atlas of the Upper Midwest*. 3 vols.
Minneapolis: University of Minnesota Press.
- Auer, Peter, Frans Hinskens and Paul Kerswill (eds). 2005. *Dialect Change: Convergence and Divergence in European Languages*. Cambridge, UK: Cambridge University Press.
- Bailey, Guy, Tom Wikle, Jan Tillery and Lori Sand. 1993. Some patterns of linguistic diffusion. *Language Variation and Change* 5: 359-390.
- Bex, Tony, and Richard J. Watts (eds). 1999. *Standard English: The Widening Debate*. London: Routledge.
- Bloomfield, Leonard. 1933. *Language*. Chicago: University of Chicago Press.
- Boberg, Charles. 2005. The North American Regional Vocabulary Survey: Renewing the study of lexical variation in North American English. *American Speech* 80/1: 22-60.
- Boberg, Charles. 2008. Regional phonetic differentiation in Standard Canadian English. *Journal of English Linguistics* 36/2: 129-154.
- Boberg, Charles. 2010. *The English Language in Canada: Status, History and Comparative Analysis*. Cambridge, UK: Cambridge University Press.
- Britain, David. 1997. Dialect contact and phonological reallocation: 'Canadian Raising' in the English Fens. *Language in Society* 26: 15-46.

- Britain, David. 2009. One foot in the grave? Dialect death, dialect contact, and dialect birth in England. *International Journal of the Sociology of Language* 2009: 196-197.
- Callary, Robert E. 1975. Phonological change and the development of an urban dialect in Illinois. *Language in Society* 4: 155-169.
- Carver, Craig M. 1987. *American Regional Dialects: A Word Geography*. Ann Arbor, MI: University of Michigan Press.
- Cassidy, Frederic Gomes, and Joan Houston Hall (eds). 1985-2012. *Dictionary of American Regional English*. Cambridge, MA: Harvard University Press.
- Chambers, J.K. 1973. Canadian raising. *Canadian Journal of Linguistics* 18: 113-35.
- Chambers, J. K., and Natalie Schilling (eds). 2013. *The Handbook of Language Variation and Change, 2nd Edition*. Wiley-Blackwell.
- Chambers, J.K., and Peter Trudgill. 1998. *Dialectology*, 2nd ed. Cambridge, UK: Cambridge University Press.
- Cheshire, Jenny, Paul Kerswill, Sue Fox and Eivind Torgersen. 2011. Contact, the feature pool and the speech community: The emergence of Multicultural London English. *Journal of Sociolinguistics* 15/2: 151-196.
- Clopper, Cynthia G., and David B. Pisoni. 2004. Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics* 32: 111-140.
- Crystal, David. 2003. *English as a Global Language*. Cambridge, UK: Cambridge University Press.

- Eisenstein, Jacob, Brendan O'Connor, Noah A. Smith and Eric P. Xing. 2010. A latent variable model for geographic lexical variation. *Proceedings of Empirical Methods on Natural Language Processing 2010*: 1277-1287.
- Ellis, Alexander John. 1890. *English Dialects, Their Sounds and Homes: Being an Abridgment of the Author's Existing Phonology of English Dialects*. London: K. Paul, Trench, Trübner & Co.
- Fischer, John L. 1958. Social influences on the choice of a linguistic variant. *Word* 14: 47-56.
- Francis, W. Nelson. 1983. *Dialectology: An Introduction*. London: Longman.
- Gilliéron, Jules, and Edmond Edmont. 1902-1920. *Atlas linguistique de la France*. Paris: Champion.
- Gilliéron, Jules, and Mario Roques. 1912. *Étude de géographie linguistique*. Paris: Champion.
- Goebel, Hans. 1982. *Dialektometrie; Prinzipien und Methoden des Einsatzes der numerischen Taxonomie im Bereich der Dialektgeographie*. Vienna: Verlag der Österreichischen Akademie der Wissenschaften.
- Grieve, Jack, Costanza Asnaghi and Tom Ruetten. 2013. Site-restricted web searches for data collection in regional dialectology. *American Speech* 88/4: 413-440.
- Grieve, Jack, Dirk Speelman and Dirk Geeraerts. 2011. A statistical method for the identification and aggregation of regional linguistic variation. *Language Variation and Change* 23: 193-221.
- Grimm, Jacob. 1819. *Deutsche Grammatik, 1. Theil*. Göttingen: Dieterich.

- Heeringa, Wilbert, and John Nerbonne. 2001. Dialect areas and dialect continua. *Language Variation and Change* 13: 375-400.
- Kachru, Braj B. 1985. Standards, codification, and sociolinguistic realism: The English language in the outer circle. In Randolph Quirk and Henry Widdowson (eds), *English in the World: Teaching and Learning of Language and Literature* (Cambridge, U.K.: Cambridge University Press), pp. 11-30.
- Kachru, Braj B., Yamuna Kachru and Cecil L. Nelson (eds). 2009. *The Handbook of World Englishes*. Malden, MA: Blackwell Publishing Limited.
- Kerswill, Paul, and Ann Williams. 2005. New towns and koineization: Linguistic and social correlates. *Linguistics* 43: 1023-1048.
- Kloeke, G.G. 1927. *De Hollandsche expansie in de zestiende en zeventiende eeuw en haar weerspiegeling in de hedendaagsche Nederlandsche dialecten*. Den Haag: Martinus Nijhoff.
- Kotsinas, Ulla-Britt. 1988. Immigrant children's Swedish—a new variety? *Journal of Multilingual and Multicultural Development* 9/1-2: 129-140.
- Kretzschmar, William A, Jr. 1996. Quantitative areal analysis of dialect features. *Language Variation and Change* 8: 13-39.
- Kristensen, Kjeld, and Mats Thelander. 1984. On dialect levelling in Denmark and Sweden. *Folia Linguistica* 18(1-2): 223-246.
- Kristiansen, Tore. 1998. The role of standard ideology in the disappearance of the traditional Danish dialects. *Folia Linguistica* 32/1-2: 115-129.
- Kurath, Hans, Miles Hanley, Bernard Bloch and Guy S. Loman, Jr. 1939-1943. *Linguistic Atlas of New England*. Providence: Brown University Press.

- Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Labov, William. 1981. Resolving the Neogrammarian controversy. *Language* 57: 267-308.
- Labov, William. 2007. Transmission and diffusion. *Language* 83/2: 344-387.
- Labov, William, Malcah Yaeger and Richard Steiner. 1972. *A Quantitative Study of Sound Change in Progress*. Philadelphia: U.S. Regional Survey.
- Labov, William, and Sharon Ash. 1997. Understanding Birmingham. In Cynthia Bernstein, Thomas Nunnally and Robin Sabino (eds), *Language Variety in the South Revisited* (Tuscaloosa, AL: University of Alabama Press), 508-573.
- Labov, William, Sharon Ash and Charles Boberg. 2006. *The Atlas of North American English: Phonetics, Phonology and Sound Change*. Berlin: Mouton de Gruyter.
- Lameli, Alfred, Roland Kehrein and Stefan Rabanus (eds). 2010. *Language and Space: An International Handbook of Linguistic Variation. Vol. 2: Language Mapping*. Berlin: Walter de Gruyter.
- Macaulay, Ronald K.S. 1977. *Language, Social Class, and Education: A Glasgow Study*. Edinburgh: Edinburgh University Press.
- McDavid, Raven I., et al. 1980. *Linguistic Atlas of the Middle and South Atlantic States*. Chicago: University of Chicago Press.
- Melchers, Gunnel, and Philip Shaw. 2013. *World Englishes*. London: Routledge.
- Milroy, James, and Lesley Milroy. 1999. *Authority in Language: Investigating Standard English*. London: Routledge.

- Moulton, William G. 1960. The short vowel systems of Northern Switzerland: A study in structural dialectology. *Word* 16: 155-182.
- Moulton, William G. 1962. Dialect geography and the concept of phonological space. *Word* 18: 23-32.
- Nerbonne, John. 2009. Data-driven dialectology. *Language and Linguistics Compass* 3/1: 175-198.
- Nerbonne, John, and Peter Kleiweg. 2007. Toward a dialectological yardstick. *Journal of Quantitative Linguistics* 14: 148-166.
- Nerbonne, John, and William A. Kretzschmar, Jr. 2006. Progress in dialectometry: Toward explanation. *Literary and Linguistic Computing* 21: 387-397.
- Orton, Harold, and Eugen Dieth. 1962-1971. *Survey of English Dialects: Basic Materials*. Leeds: E. J. Arnold & Son.
- Osthoff, Hermann, and Karl Brugmann. 1878. *Morphologische Untersuchungen auf dem Gebiete der indogermanischen Sprachen*. Leipzig: S. Hirzel.
- Pedersen, Inge Lise. 2003. Traditional dialects of Danish and the de-dialectalization 1900-2000. *International Journal of the Sociology of Language* 159: 9–28.
- Pederson, Lee, Susan L. McDaniel and Carol M. Adams (eds). 1986-1993. *Linguistic Atlas of the Gulf States*. 7 vols. Athens, GA: University of Georgia Press.
- Petyt, K.M. 1980. *The Study of Dialect*. London: Andre Deutsch.
- Preston, Dennis R., and Daniel Long. 1999. *Handbook of Perceptual Dialectology*. Amsterdam: Benjamins.
- Quist, Pia. 2000. Ny københavnsk ‘multietnolekt’: Om sprogbrug blandt unge i sprogligt og kulturelt heterogene miljøer [A new Copenhagen ‘multi-ethnolect’:

- Language use among adolescents in linguistically and culturally heterogeneous settings]. *Danske Talesprog* 1: 143–211.
- Rona, José Pedro. 1976. The social dimension of dialectology. *International Journal of the Sociology of Language* 9: 7-22.
- Schneider, Edgar W. 2007. *Postcolonial English : Varieties around the World*. Cambridge, U.K.: Cambridge University Press.
- Trudgill, Peter. 1974a. *The Social Differentiation of English in Norwich*. Cambridge, UK: Cambridge University Press.
- Trudgill, Peter. 1974b. Linguistic change and diffusion: Description and explanation in sociolinguistic dialect geography. *Language in Society* 3: 215-246.
- Trudgill, Peter. 1986. *Dialects in Contact*. Oxford, UK: Blackwell.
- Upton, Clive, and J.D.A. Widdowson. 1996. *An Atlas of English Dialects*. Oxford, UK: Oxford University Press.
- Watt, Dominic, and Carmen Llamas (eds). 2014. *Borders and Identities*. Edinburgh: Edinburgh University Press.
- Watt, Dominic, Carmen Llamas and Daniel Ezra Johnson. 2010. Levels of linguistic accommodation across a national border. *Journal of English Linguistics* 38: 270-289.
- Weinreich, Uriel. 1954. Is a structural dialectology possible? *Word* 10: 388-400.
- Wenker, Georg, Ferdinand Wrede, Walther Mitzka and Bernhard Martin. 1927-1956. *Der Sprachatlas des deutschen Reichs*. Marburg: Elwert.
- Wells, J.C. 1982. *Accents of English*. Cambridge, U.K.: Cambridge University Press.
- Wiese, Heike. 2009. Grammatical innovation in multiethnic urban Europe: New linguistic practices among adolescents. *Lingua* 119: 782–806.

- Wikle, Tom. 1997. Quantitative mapping techniques for displaying language variation and change. In Cynthia Bernstein, Thomas E. Nunnally and Robin Sabino (eds.), *Language Variety in the South Revisited* (Tuscaloosa, AL.: University of Alabama Press), 417-433.
- Wright, Joseph (ed.). 1898-1905. *The English Dialect Dictionary, 6 vols.* Oxford: Oxford University Press.