6 Language Variation

As we have seen in previous chapters, languages vary in many ways. One way of characterizing certain variations is to say that speakers of a particular language sometimes speak different dialects of that language. Although I have already noted how difficult it is to define *dialect*, we may still find it useful to use the term in our work in sociolinguistics, and even to extend its use from studies of regional variation to those of social variation. In this way it would be possible to talk about both *regional dialects* and *social dialects* of a language. Just as a regional dialect marks off the residents of one region from those of other regions, a social dialect would be a variety associated with a specific social class or group which marks off that class or group from other classes or groups. However, if this further differentiation of varieties is to be successful, it will require us to be able to find linguistic features which are associated with differences in classes or groups and, of course, to define what we mean by these latter terms.

Sociolinguists today are generally more concerned with social variation in language than with regional variation. However, if we are to gain a sound understanding of the various procedures used in studies of social variation, we should look at least briefly at previous work in regional dialectology. That work points the way to understanding how recent investigations have proceeded as they have. Studies of social variation in language grew out of studies of regional variation. It was largely in order to widen the limits and repair the flaws that were perceived to exist in the latter that investigators turned their attention to social variation in language. As we will see, there may still be certain limitations in investigating such variation but they are of a different kind. It is also important to note that even if there are limitations to this kind of work, many sociolinguists regard it as being essentially what sociolinguistics is – or should be – all about (see p. 14). In this view the study of language variation tells us important things about languages and how they change. This chapter and the two that follow deal with such matters.

Regional Variation

The mapping of dialects on a regional basis has had a long history in linguistics (see Petyt, 1980, Chambers and Trudgill, 1998, and Wakelin, 1977). In fact, it is a well-established part of the study of how languages change over time, i.e., of *diachronic* or *historical linguistics*. Traditionally, *dialect geography*, as this area of linguistic study is known, has employed assumptions and methods drawn from historical linguistics, and many of its results have been used to confirm findings drawn from other historical sources, e.g., archeological findings, population studies, written records. In this view languages differentiate internally as speakers distance themselves from one another over time and space; the changes result in the creation of dialects of the languages. Over sufficient time, the resulting dialects become new languages as speakers of the resulting varieties become unintelligible to one another. So Latin became French in France, Spanish in Spain, Italian in Italy, and so on.

In this model of language change and dialect differentiation, it should always be possible to relate any variation found within a language to the two factors of time and distance alone; e.g., the British and American varieties, or dialects, of English are separated by well over two centuries of political independence and by the Atlantic Ocean; Northumbrian and Cockney English are nearly 300 miles and many centuries apart. In each case, linguists working in this tradition try to explain any differences they find with models familiar to the historical linguist, models which incorporate such concepts as the 'family tree' (Latin has 'branched' into French, Spanish, and Italian), phonemic 'split' (English /f/ and /v/ are now distinctive phonemes whereas once they were phonetic variants, or allophones, of a single phoneme) or phonemic 'coalescence' (English *ea* and *ee* spellings, as in *beat* and *beet*, once designated different pronunciations), the 'comparative method' of reconstruction (English *knave* and German *Knabe* come from the same source), and 'internal reconstruction' (though *mouse* and *mice* now have different vowel sounds, this was not always the case).

Dialect geographers have traditionally attempted to reproduce their findings on maps in what they call *dialect atlases*. They try to show the geographical boundaries of the distribution of a particular linguistic feature by drawing a line on a map. Such a line is called an *isogloss*: on one side of the line people say something one way, e.g., pronounce *bath* with the first vowel of *father*, and on the other side they use some other pronunciation, e.g., the vowel of *cat*. Quite often, when the boundaries for different linguistic features are mapped in this way the isoglosses show a considerable amount of criss-crossing. On occasion, though, a number coincide; i.e., there is a *bundle of isoglosses*. Such a bundle is often said to mark a *dialect boundary*. One such bundle crosses the south of France from east to west approximately at the 45th parallel (Grenoble to Bordeaux) with words like *chandelle*, *chanter*, and *chaud* beginning with a *sh*

sound to the north and a *k* sound to the south. Quite often, that dialect boundary coincides with some geographical or political factor, e.g., a mountain ridge, a river, or the boundary of an old principality or diocese. Isoglosses can also show that a particular set of linguistic features appears to be spreading from one location, a *focal area*, into neighboring locations. In the 1930s and 1940s Boston and Charleston were the two focal areas for the temporary spread of *r*-lessness in the eastern United States. Alternatively, a particular area, a *relic area*, may show characteristics of being unaffected by changes spreading out from one or more neighboring areas. Places like London and Boston are obviously focal areas; places like Martha's Vineyard – it remained *r*-pronouncing in the 1930s and 1940s even as Boston dropped the pronunciation – in New England and Devon in the extreme southwest of England are relic areas. Wolfram (2004) calls the dialect of such an area a *remnant dialect* and, in doing so, reminds us that not everything in such a dialect is a relic of the past for such areas also have their own innovations. Huntley, a rural enclave in Aberdeenshire, Scotland, where Marshall worked (2003, 2004) is also a relic area (see p. 218).

The Rhenish Fan is one of the best-known sets of isoglosses in Europe, setting off Low German to the north from High German to the south. The set comprises the modern *reflexes* (i.e., results) of the pre-Germanic stop consonants *p, *t, and *k. These have remained stops [p,t,k] in Low German but have become the fricatives [f,s,x] in High German (i.e., Modern Standard German), giving variant forms for 'make' [makən], [maxən]; 'that' [dat], [das]; 'village' [dorp], [dorf]; and 'I' [ik], [ix]. Across most of Germany these isoglosses run virtually together from just north of Berlin in an east–west direction until they reach the Rhine. At that point they 'fan', as in figure 6.1. Each area within the fan has a different incidence of stops and fricatives in these words, e.g., Düsseldorf has [ix], [makən], [dorp], and [dat], and Trier has [ix], [maxən], [dorf], and [dat]. The boundaries within the fan coincide with old ecclesiastical and political boundaries. The change of stops to fricatives, called the Second German Consonant Shift, appears to have spread along the Rhine from the south of Germany to the north. Political and ecclesiastical frontiers along the Rhine were important in that spread as were centers like Cologne and Trier. The area covered by the fan itself is sometimes called a *transition area* (in this case, between Low and High German) through which a change is progressing in contrast to either a focal or relic area.

Very often the isoglosses for individual phonological features do not coincide with one another to give us clearly demarcated dialect areas. For example, in England the isogloss that separates *stood* or *come* pronounced with [ʊ] rather than [ʌ] runs roughly east and west (with [ʊ] to the north). It intersects the isogloss that separates *farm* pronounced with or without the [r], which runs roughly northwest to southeast (with [r] to the west, except for pockets of [r] pronunciation in the West Midlands and Northeast). This gives us the four distinct areas illustrated in figure 6.2: [r] and [ʊ]; [r] and [ʌ]; Ø, i.e., nothing and [ʊ]; and Ø and [ʌ]. These two quite different distributions, i.e., the 'criss-cross' pattern, are

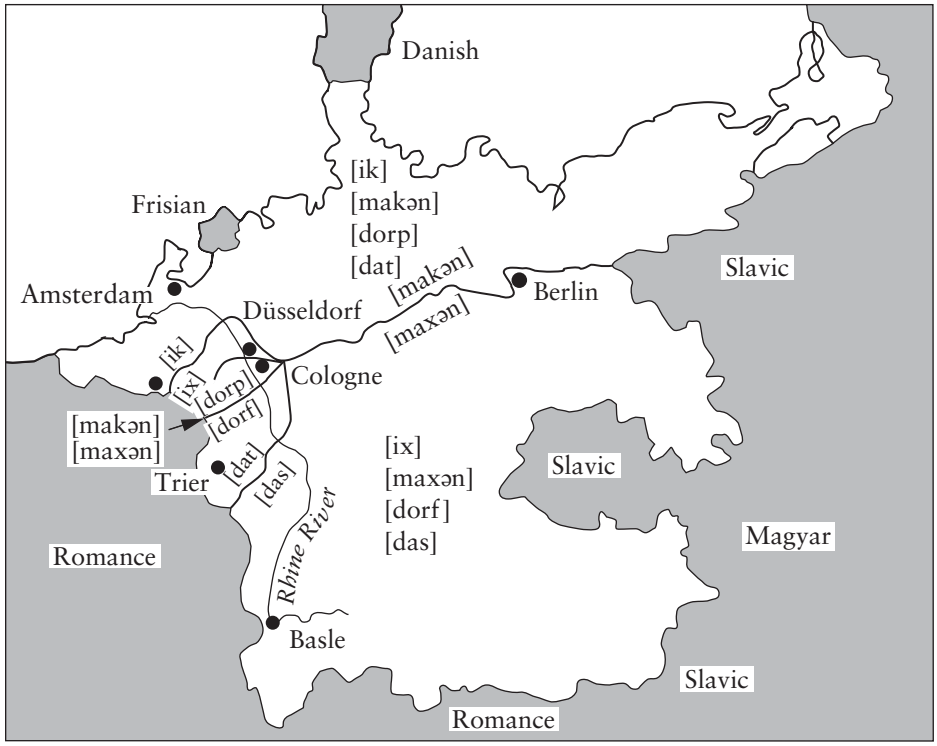

Figure 6.1 The Rhenish Fan

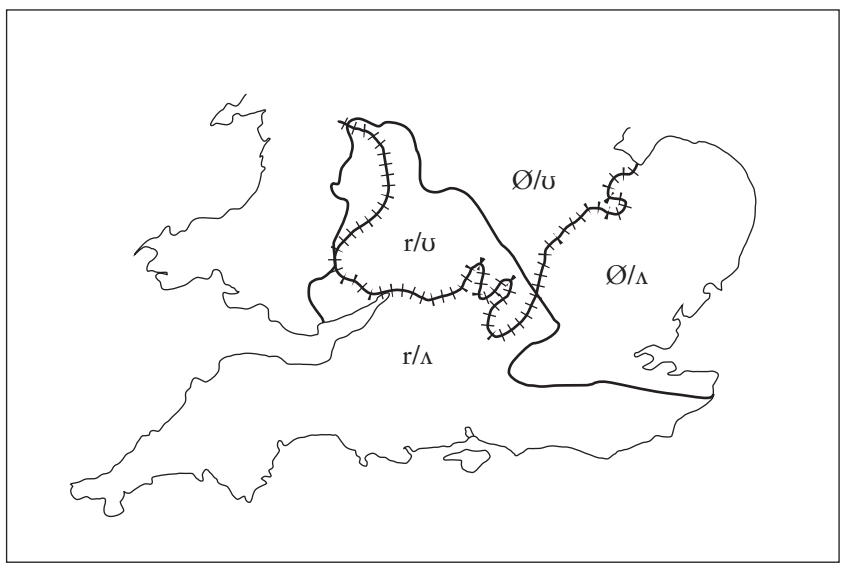

Figure 6.2 Intersecting isoglosses

just about impossible to explain using traditional 'family-tree' type models of linguistic change. We should also note that the [ʊ] and [r] pronunciations are 'retreating' before those with [ʌ] and Ø, which are more and more associated with the standard variety of the language. Isoglosses do cross and bundles of them are rare. It is consequently extremely difficult to determine boundaries between dialects in this way and dialectologists acknowledge this fact. The postulated dialect areas show considerable internal variation and the actual areas proposed are often based on only a few key items (or linguistic variables in our terminology). Consequently, as Le Page (1997, p. 18) says, 'the dialect areas outlined by the isoglosses on the maps were artifacts of the geographer; they had to be matched against such stereotypes as "southern dialect" or "Alemmanic" or "langue d'oc," concepts which often related in the minds of outsiders to just one or two variables characterizing a complete, discrete system.' Hudson (1996, p. 39) draws a somewhat negative conclusion: 'isoglosses need not delimit varieties, except in the trivial sense where varieties each consist of just one item; and if we cannot rely on isoglosses to delimit varieties, what can we use?' However, if we look at Trudgill's book (1999) on English dialects we can see the many positive results of this kind of work.

Dialect-atlas type approaches such as Trudgill's take a particular linguistic feature, which we will soon call a *linguistic variable* (see following section), and show its distribution geographically. They also attempt to relate that distribution to the historical development of the language, both internally, i.e., linguistically, and externally, i.e., politically, socially, and culturally.

Because dialect studies grew out of historical studies of languages, it should also come as no surprise that they have focused almost exclusively on rural areas. Rural areas were regarded as 'conservative' in the sense that they were seen to preserve 'older' forms of the languages under investigation. Urban areas were acknowledged to be innovative, unstable linguistically, and difficult to approach using existing survey techniques. When the occasional approach was made, it was biased toward finding the most conservative variety of urban speech. Ignoring towns and cities may be defensible in an agrarian-based society; however, it is hardly defensible in the heavily urbanizing societies of today's world as the only way to study the language variation that exists there. An alternative approach is called for.

One basic assumption in dialect geography is that regional dialects are really quite easy to sample: just find one or two people in the particular location you wish to investigate, people who are preferably elderly and untraveled, interview them, and ask them how they pronounce particular words, refer to particular objects, and phrase particular kinds of utterances. A sampling of such people from various locations throughout a wide geographical area will allow the dialect geographer to show where particular sounds, forms, and expressions are used, and where boundaries can be drawn around these so that area A may be described as an area in which linguistic feature X occurs (or is used) whereas area B has no instances of that feature. If there are sufficient differences between the

linguistic features employed in areas A and B, then we may say that we actually have two dialects, A and B, of the particular language in question.

While this kind of study of regional varieties of languages has a long and respected history, it also has serious limitations. As I have said, it tends to ignore densely populated areas, specifically large sprawling urban areas, because of the complexities of both sampling and data evaluation. The selection of informants also tends not to be very well controlled, often reflecting no more than the judgment of the person collecting the data that a particular individual is 'representative' of the area being sampled. It certainly lacks the kind of scientific rigor that sociologists have come to insist on in sampling any population. For example, the informants chosen for the *Linguistic Atlas of the United States and Canada* were of three types (Kurath, 1939, p. 44), chosen as follows:

Type I: Little formal education, little reading, and restricted social contacts
Type II: Better formal education (usually high school) and/or wider reading and social contacts
Type III: Superior education (usually college), cultured background, wide reading, and/or extensive social contacts

Each of these three types was then sub-categorized as follows:

Type A: Aged, and/or regarded by the field worker as old-fashioned
Type B: Middle-aged or younger, and/or regarded by the field worker as more modern

We should also note that it was the field worker for the *Atlas* who decided exactly where each informant fitted in the above scheme of things. A certain circularity is obviously involved: the *Atlas* studies were intended partly to find out how speech related to social class, but speech was itself used as one of the criteria for assigning membership in a social class. The field worker alone judged whether a particular informant should be used in the study, and Type IA informants were particularly prized as being most representative of local speech.

In England, the Survey of English Dialects carried out between 1950 and 1961 with informants from 313 localities in England and Wales employed similar criteria (Orton et al., 1978, p. 3):

> The selection of informants was made with especial care. The fieldworkers were instructed to seek out elderly men and women – more often men, since women seemed in general to encourage the social upgrading of the speech of their families – who were themselves of the place and both of whose parents were preferably natives also. They were to be over 60 years of age, with good mouths, teeth and hearing and of the class of agricultural workers who would be familiar with the subject matter of the questionnaire and capable of responding perceptively and authoritatively.

Typically, both informants and field workers were male. As Coates (2004, pp. 10–11) says, 'Dialectology . . . marginalized women speakers. Traditional dialectologists defined the true vernacular in terms of male informants, and organised their questionnaires around what was seen as the man's world.'

Dialect-atlas studies attempted to relate variation in language to settlement history and tended to ignore social-class factors. There was some recognition of the latter, but it was relatively small owing to the inadequate systems of social classification that were employed in most investigations. However, it is still possible to make some observations. For example, in the southern and south Midlands dialects of the United States, a form such as *you-all* is found in use among all social classes, whereas *I might could* and *a apple* are found in use only among speakers in the low and middle classes, and *I seed* and *fistes*, *postes*, and *costes* (as plurals of *fist*, *post*, and *cost*) are found in use only among speakers from the low class.

Since most of us realize that it is not only where you come from that affects your speech but also your social and cultural background, age, gender, race, occupation, and group loyalty, the traditional bias toward geographic origin alone now appears to be a serious weakness. Then, too, the overriding model of language change and differentiation is an extremely static one, and one that is reinforced, rather than questioned, by the types of data selected for analysis. Speakers from different regions certainly interact with one another; dialect breaks or boundaries are not 'clean'; and change can be said to be 'regular' only if you are prepared to categorize certain kinds of irregularities as exceptions, relics, borrowings, 'minor' variations, and so on. Furthermore, the varieties of a language spoken within large gatherings of people in towns and cities must influence what happens to other varieties of that language: to attempt to discuss the history of English, French, or Italian while ignoring the influences of London, Paris, or Florence would seem to be something like attempting to produce *Hamlet* without the prince!

All of this is not to say that this kind of individual and social variation has gone unnoticed in linguistics. Linguists have long been aware of variation in the use of language: individuals do speak one way on one occasion and other ways on other occasions, and this kind of variation can be seen to occur within even the most localized groups. Such variation is often ascribed to *dialect mixture*, i.e., the existence in one locality of two or more dialects which allow a speaker or speakers to draw now on one dialect and then on the other. An alternative explanation is *free variation*, i.e., random 'meaningless' variation of no significance. However, no one has ever devised a suitable theory to explain either dialect mixture or free variation, and the latter turns out not to be so free after all because close analyses generally reveal that complex linguistic and social factors appear to explain much of the variation.

There have been some recent developments in this kind of work which hold promise for future discoveries. They result largely from our growing ability to process and analyze large quantities of linguistic data. One, for example, is

Kretzschmar's work on the Linguistic Atlas of the Middle and South Atlantic States (LAMSAS). He shows (1996) how it is possible to use quantitative methods to demonstrate the probability of occurrence of specific words or sounds in specific areas. Another quantitative survey (Labov et al., 2005) used a very simple sampling technique to survey the whole of North American English in order to produce the *Atlas of North American English* (ANAE), a study of all the cities on the continent with populations of over fifty thousand. This study showed that 'regional dialects are getting stronger and more diverse as language change is continuing and that the structural divisions between them are very sharp, with very tight bundling of the isoglosses' (p. 348). In still another approach to dialects, this one focusing on how a specific dialect emerged, Lane (2000) used a variety of economic, demographic, and social data from 3,797 residents of Thyborøn, Denmark, covering the years 1890–1996, to reveal how the local dialect 'is the result of a constant situation that led to the formation of a new dialect as a result of massive in-migration . . . a new system created largely out of materials selected from competing systems in contact and from innovations that indexed the new local linguistic community' (p. 287). It was clearly another triumph for an aspiration to achieve a local identity. We can see a similar emphasis on using traditional dialect materials to help us account for current language varieties in recent writings on new Englishes (see Gordon et al., 2004, Hickey, 2004, and Trudgill, 2004).

I have deliberately not focused at length on dialect geography since it is not one of our major concerns and I have already (in chapter 2) considered certain issues related to dialects. However, dialect geography does raise a number of issues which are important to our concerns. One is the kind of variation that we should try to account for in language. Another has to do with sampling the population among which we believe there is variation. Still another is the collection, analysis, and treatment of the data that we consider relevant. And, finally, there are the overriding issues of what implications there are in our findings for theoretical matters concerning the nature of language, variation in language, the language-learning and language-using abilities of human beings, and the processes involved in language change. It is to these issues that I will now turn, and in doing so, focus on social rather than regional variation in language. The major conceptual tool for investigation of such variation will be the 'linguistic variable.'

The Linguistic Variable

The investigation of social dialects has required the development of an array of techniques quite different from those used in dialect geography. Many of these derive from the pioneering work of Labov, who, along with other sociolinguists,

has attempted to describe how language varies in any community and to draw conclusions from that variation not only for linguistic theory but also sometimes for the conduct of everyday life, e.g., suggestions as to how educators should view linguistic variation (see chapter 14). As we will see, investigators now pay serious attention to such matters as stating hypotheses, sampling, the statistical treatment of data, drawing conclusions, and relating these conclusions to such matters as the inherent nature of language, the processes of language acquisition and language change, and the social functions of variation.

Possibly the greatest contribution has been in the development of the use of the 'linguistic variable,' the basic conceptual tool necessary to do this kind of work (see Wolfram, 1991). As I have just indicated, variation has long been of interest to linguists, but the use of the linguistic variable has added a new dimension to linguistic investigations. Although not all linguists find the concept useful in their work, it has nevertheless compelled most of its severest critics to reconsider just what it is they are theorizing about when they talk of 'language,' of a speaker's 'knowledge' of language, and of the relationship between such knowledge and actual 'use.'

A *linguistic variable* is a linguistic item which has identifiable variants. For example, words like *singing* and *fishing* are sometimes pronounced as *singin'* and *fishin'*. The final sound in these words may be called the linguistic variable (ng) with its two variants [ŋ] in *singing* and [n] in *singin'*. Another example of a linguistic variable can be seen in words like *farm* and *far*. These words are sometimes given *r*-less pronunciations; in this case we have the linguistic variable (r) with two variants [r] and Ø (pronounced 'zero'). Still another example involves the vowel in a word like *bend*. That vowel is sometimes nasalized and sometimes it is not; sometimes too the amounts of nasalization are noticeably different. In this case we have the linguistic variable (e) and a number of variants, [ɛ], $[\tilde{\varepsilon}]^1, \ldots, [\tilde{\varepsilon}]^n$; here the superscripts ^{1} to ^{n} are used to indicate the degree of nasalization observed to occur. We might, for example, find two or even three distinct quantities of nasalization.

There are at least two basically different kinds of variation. One is of the kind (ng) with its variants [ŋ] or [n], or (th) with its variants [θ], [t], or [f], as in *with* pronounced as *with*, *wit*, or *wif*. In this first case the concern is with which quite clearly distinct variant is used, with, of course, the possibility of Ø, the zero variant. The other kind of variation is the kind you find above in (e): $[\tilde{\varepsilon}]^1, \ldots, [\tilde{\varepsilon}]^n$, when it is the quantity of nasalization, rather than its presence or absence, which is important. How can you best quantify nasalization when the phenomenon is actually a continuous one? The same issue occurs with quantifying variation in other vowel variables: quantifying their relative frontness or backness, tenseness or laxness, and rounding or unrounding. Moreover, more than one dimension may be involved, e.g., amount of nasalization *and* frontness or backness. In such cases usually some kind of weighting formula is devised, and when the data are treated it is these weights that are used in any calculations, not just the ones and zeros that we can use in the case of (ng): [ŋ] or [n], where [ŋ] = 1 and [n] = 0.

Linguists who have studied variation in this way have used a number of linguistic variables. The (ng) variable has been widely used; Labov (2006, p. 259) says it 'has been found to have the greatest generality over the English-speaking world, and has been the subject of the most fruitful study.' The (r) variable mentioned above has also been much used. Other useful variables are the (h) variable in words like *house* and *hospital*, i.e., (h): [h] or Ø; the (t) variable in *bet* and *better*, i.e., (t): [t] or [ˀ]; the (th) and (dh) variables in *thin* and *they*, i.e., (th): [θ] or [t] and (dh): [ð] or [d]; the (l) variable in French in *il*, i.e., (l): [l] or Ø; and variables like the final (t) and (d) in words like *test* and *told*, i.e., their presence or absence. Vowel variables used have included the vowel (e) in words like *pen* and *men*; the (o) in *dog*, *caught*, and *coffee*; the (e) in *beg*; the (a) in *back*, *bag*, *bad*, and *half*; and the (u) in *pull*.

Studies of variation employing the linguistic variable are not confined solely to phonological matters. Investigators have looked at the (s) of the third-person singular, as in *he talks*, i.e., its presence or absence; the occurrence or nonocurrence of *be* (and of its various inflected forms) in sentences such as *He's happy*, *He be happy*, and *He happy*; the occurrence (actually, virtual nonoccurrence) of the negative particle *ne* in French; various aspects of the phenomenon of multiple negation in English, e.g., *He don't mean no harm to nobody*; and the beginnings of English relative clauses, as in *She is the girl who(m) I praised*, *She is the girl that I praised*, and *She is the girl I praised*.

To see how individual researchers choose variables, we can look briefly at three studies. In a major part of his work in New York City, Labov (1966) chose five phonological variables: the (th) variable, the initial consonant in words like *thin* and *three*; the (dh) variable, the initial consonant in words like *there* and *then*; the (r) variable, *r*-pronunciation in words like *farm* and *far*; the (a) variable, the pronunciation of the vowel in words like *bad* and *back*; and the (o) variable, the pronunciation of the vowel in words like *dog* and *caught*. We should note that some of these have discrete variants, e.g., (r): [r] or Ø, whereas others require the investigator to quantify the variants because the variation is a continuous phenomenon, e.g., the (a) variable, where there can be both raising and retraction of the vowel, i.e., a pronunciation made higher and further back in the mouth, and, of course, in some environments nasalization too.

Trudgill (1974) also chose certain phonological variables in his study of the speech of Norwich: three consonant variables and thirteen vowel variables. The consonant variables were the (h) in *happy* and *home*, the (ng) in *walking* and *running*, and the (t) in *bet* and *better*. In the first two cases only the presence or absence of *h*-pronunciation and the [ŋ] versus [n] realizations of (ng) were of concern to Trudgill. In the last there were four variants of (t) to consider: an aspirated variant; an unaspirated one; a glottalized one; and a glottal stop. These variants were ordered, with the first two combined and weighted as being least marked as nonstandard, the third as more marked, and the last, the glottal stop, as definitely marked as nonstandard. The thirteen vowel variables were the vowels used in words such as *bad*, *name*, *path*, *tell*, *here*, *hair*, *ride*,

Exploration 6.1: The (ng) variable

As I have said, the (ng) or /ŋ/ variable is the most generally noticeable phonological variable throughout the English-speaking world. This task requires you to do some 'field work.' Devise a way of finding out instances of the use of (ng). You may want to listen to pop songs, recorded interviews, talk shows, casual conversations, etc. You may want to attempt to elicit *-ing* forms through some kind of task you ask people to perform. The key is to access unmonitored speech, i.e., talk that is focused on 'content' rather than on 'form,' as well as more conscious varieties. After you have collected some data and analyzed what you have, try to figure out how you might improve your results if you were to repeat the task. (You could then repeat it to see what progress you made.) You can be sure that none of the research findings reported in this chapter and in the following two came from first attempts but rather from studies that were preceded by one or more such 'pilot' studies.

bird, *top*, *know*, *boat*, *boot*, and *tune*. Most of these had more than two variants, so weighting, i.e., some imposed quantification, was again required to differentiate the least preferred varieties, i.e., the most nonstandard, from the most preferred variety, i.e., the most standard.

The Detroit study (Shuy et al., 1968) focused on the use of three variables: one phonological variable and two grammatical variables. The phonological variable was the realization of a vowel plus a following nasal consonant as a nasalized vowel, e.g., *bin* realized as [bĩ] rather than [bɪn]. The grammatical variables were multiple negation, which I have already mentioned, and pronominal apposition, e.g., *That guy, he don't care*. In a further study of Detroit speech, Wolfram (1969) considered certain other linguistic variables. These included the pronunciation of final consonant clusters, i.e., combinations of final consonants in words like *test*, *wasp*, and *left*, *th* in words like *tooth* and *nothing*, final stops in words like *good* and *shed*, and *r*-pronouncing in words like *sister* and *pair*. So far as grammatical variables were concerned, Wolfram looked at matters such as *he talk/talks*, *two year/years*, *she nice/she's nice*, *he's ready/he ready/he be ready*, and multiple negation as in *He ain't got none neither*.

This brief sample indicates some of the range of variables that have been investigated. The important fact to remember is that a linguistic variable is an item in the structure of a language, an item that has alternate realizations, as one speaker realizes it one way and another a different way, or the same speaker realizes it differently on different occasions. For example, one speaker may say *singing* most of the time whereas another prefers *singin'*, but the first is likely

to say *singin'* on occasion just as the second may be found to use the occasional *singing*. What might be interesting is any relationship we find between these habits and either (or both) the social class to which each speaker belongs or the circumstances which bring about one pronunciation rather than the other.

Labov (1972b) has also distinguished among what he calls indicators, markers, and stereotypes. An *indicator* is a linguistic variable to which little or no social import is attached. Only a linguistically trained observer is aware of indicators. For example, some speakers in North America distinguish the vowels in *cot* and *caught* and others do not. Whether one distinguishes the vowels or not carries little or no social significance. On the other hand, a *marker* does carry with it social significance. In fact, markers may be potent carriers of social information. You do not always have to drop every *g*, i.e., always say *singing* as *singin'*. Labov says that 'we observe listeners reacting in a discrete way. Up to a certain point they do not perceive the speaker "dropping his g's" at all; beyond a certain point, they perceive him as always doing so' (p. 226). G-dropping is a marker everywhere English is spoken. People are aware of markers, and the distribution of markers is clearly related to social groupings and to styles of speaking. Pronouncing *car* and *cart* in New York City in their *r*-less varieties marks you as using a type of pronunciation associated with lower-class speech in that city. New Yorkers are conscious of this fact and may vary their use of *r* according to circumstances. A *stereotype* is a popular and, therefore, conscious characterization of the speech of a particular group: New York *boid* for *bird* or *Toitytoid Street* for *33rd Street*; Texas 'drawling' or *Howdy Pardner*; a Northumbrian *Wot-cher* (What cheer?) greeting; the British use of *chap*; or a Bostonian's *Pahk the cah in Hahvahd Yahd*. Often such stereotypes are stigmatized everywhere, and in at least one recorded case (see Judges 12: 4–6) a stereotypical pronunciation of *shibboleth* had fatal consequences. A stereotype need not conform to reality; rather, it offers people a rough and ready categorization with all the attendant problems of such categorizations. Studies of variation tend therefore to focus on describing the distributions of linguistic variables which are markers; they may explain how stereotypes arise, but they merely note indicators. (See Johnstone, 2004, for a discussion of stereotypes in Pittsburgh speech.)

Social Variation

Once we have identified the linguistic variable as our basic working tool, the next task becomes one of employing that tool in an effort to see how linguistic variation relates to social variation. An early study of linguistic variation by Gumperz (1958), but one cast in a 'modern' mold, shows some of the intricacies involved in trying to relate linguistic variation to social variation. Because the society he was studying is rigidly stratified on the basis of caste membership, the

problems are considerably fewer than those encountered in such cities as New York, Detroit, or even Norwich, but they are still present. Gumperz shows how rather small differences in speech can effectively distinguish sub-groups in society from one another in a study of linguistic usage in the village of Khalapur, eighty miles north of Delhi in India. The social structure of the village is determined by Hindu caste membership with Brahmans at the top, then Rajputs (warriors), Vaishyas (merchants), and several groups of artisans and laborers lower down. At the bottom are three untouchable castes: Chamars (landless laborers), Jatia Chamars (leather workers and shoe makers), and Bhangis (sweepers). The latter are restricted to living in certain neighborhoods and have less freedom to move in the village than do members of the upper castes. Ten percent of the population are not Hindus but Muslims; they are outside the caste system.

So far as language is concerned, certain characteristics of the Khalapur village dialect are clear markers of social-group membership. For example, Bhangis do not make certain phonological contrasts that speakers of all the other castes make. Chamars and Jatia Chamars also lack certain phonological contrasts made by all others, and some, in attempting to make such a contrast, actually *hyper-correct*; that is, they over-extend a particular usage in trying to emulate others. Jatia Chamars have a characteristic pronunciation of words that end in [æ] in all other village varieties. Each of the three untouchable castes therefore has speech characteristics that clearly set it off both from the other two untouchable castes and also from the touchable castes in the village. Muslim speech resembles that of the touchable classes.

An anomaly is that the variety of village speech spoken by the lowest caste, the Bhangis, is closest to the dialect of the region in which Khalapur is situated. This fact constrains members of the upper castes in their use of the regional dialect since using it would make them sound like untouchables. In their linguistic usage therefore they are forced to innovate away from the regional variety. Since untouchables apparently try to emulate the touchables, the direction of innovation for all groups in Khalapur is away from the regional variety with the innovations prompted, of course, by different needs: the touchables' need to signal their clear distinction from the untouchables, and the untouchables' attempt to reduce that distinction as much as possible. This study quite clearly shows a direct relationship between linguistic variation and caste membership. If we know certain things about one, we can predict certain things about the other. It is just such connections or correlations that interest sociolinguists working with the linguistic variable.

The next task becomes one of collecting data concerning the variants of a linguistic variable in such a way that we can draw certain conclusions about the social distribution of these variants. To draw such conclusions, we must be able to relate the variants in some way to quantifiable factors in society, e.g., social-class membership, gender, age, ethnicity, and so on. As we will see, there are numerous difficulties in attempting this task, but considerable progress has been made in overcoming them, particularly as studies have built on those that

have gone before in such a way as to strengthen the quality of the work done in this area of sociolinguistics.

While it is fairly easy to relate the occurrences of such variants to factors like gender and age, relating them to factors like race and ethnicity is somewhat more troublesome since these are much more subjective in nature and less easily quantifiable. The most complicated factor of all is social-class membership, if we consider 'social class' to be a useful concept to apply in stratifying society – and few indeed would deny its relevance! However, we must be cautious in any claims we make, particularly if we attempt regional or historical comparisons. The social-class system of England in the 1950s was different from what it is today and, presumably, it will be different again in another half century, and all these class systems were and are different from those existing contemporaneously in New York, Brazil, Japan, etc.

Sociologists use a number of different scales for classifying people when they attempt to place individuals somewhere within a social system. An occupational scale may divide people into a number of categories as follows: major professionals and executives of large businesses; lesser professionals and executives of medium-sized businesses; semi-professionals; technicians and owners of small businesses; skilled workers; semi-skilled workers; and unskilled workers. An educational scale may employ the following categories: graduate or professional education; college or university degree; attendance at college or university but no degree; high school graduation; some high school education; and less than seven years of formal education. Once again though some caution is necessary: graduating from college or university in the 1950s indicated something quite different from what it does today. Income level as well as source of income are important factors in any classification system that focuses on how much money people have. Likewise, in considering where people live, investigators must concern themselves with both the type and cost of housing and its location.

In assigning individuals to social classes, investigators may use any or all of the above criteria (and others too) and assign different weights to them. Accordingly, the resulting social-class designation given to any individual may differ from study to study. We can also see how social class itself is a sociological construct; people probably do not classify themselves as members of groups defined by such criteria. Wolfram and Fasold (1974, p. 44) point out that 'there are other objective approaches [to establishing social groupings] not exclusively dependent on socio-economic ranking. . . . An investigator may look at such things as church membership, leisure-time activities, or community organizations.' They admit that such alternative approaches are not at all simple to devise but argue that a classification so obtained is probably more directly related to social class than the simple measurement of economic factors. We should note that there is a current emphasis on 'lifestyle' in classifying people, so obviously patterns of consumption of goods and appearance are important for a number of people in arriving at some kind of social classification. Coupland (2007, pp. 29–30) calls the current era 'late-modernity.' It is a time in which 'Social

life seems increasingly to come packaged as a set of lifestyle options able to be picked up and dropped, though always against a social backdrop of economic possibilities and constraints. . . . Social class . . . membership in the West is not the straitjacket that it was. Within limits, some people can make choices in their patterns of consumption and take on the social attributes of different social classes. . . . the meaning of class is shifted.'

Alternative approaches to using a somewhat simple social-class scale are, however, still rather infrequent. What we usually find is that people are assigned to social classes through the use of composite scores derived from various scales which 'measure' some of the factors mentioned above. It is also the case that the actual scales used must necessarily vary from community to community since exactly the same characteristics cannot serve to classify people in England and the United States or in New England and New Mexico, and, as I have mentioned above, any 'time' factor has its own problems. However, nearly all such scales take into account such matters as educational achievement, professional training, occupation (sometimes parental occupation too), 'blue'- or 'white'-collar work, salary or income level, source of that salary, income, or wage (this difference also being important), gender, age, residential area, race, and ethnicity. Weights are then assigned to each of these and some kind of unitary scale is devised so that individuals can be fitted into slots carrying such designations as 'upper class,' 'middle class,' 'lower working class,' and so on. Sometimes the stratifications, or gradations, are few ('upper' vs. 'middle' class), but at other times they are many ('upper middle' vs. 'middle middle' class). Most work in sociolinguistics has drawn on commonly used unitary scales of this kind to designate the social-class membership of individuals in an attempt to describe the characteristic linguistic behavior of various social classes.

In his study of linguistic variation in New York City, Labov (1966) used the three criteria of education, occupation, and income to set up ten social classes. His class 0, his lower class, had grade school education or less, were laborers, and found it difficult to make ends meet. His classes 1 to 5, his working class, had had some high school education, were blue-collar workers, but earned enough to own such things as cars. His classes 6 to 8, his lower middle class, were high school graduates and semi-professional and white-collar workers who could send their children to college. His highest class 9, his upper middle class, were well educated and professional or business-oriented. In this classification system for people in the United States about 10 percent of the population are said to be lower class, about 40 percent working class, another 40 percent lower middle class, and the remaining 10 percent fall into the upper middle class or an upper class, the latter not included in Labov's study. In his later study (2001) of variation in Philadelphia, Labov used a socio-economic index based on occupation, education, and house value.

In his study of linguistic variation in Norwich, England, Trudgill (1974) distinguishes five social classes: middle middle class (MMC), lower middle class (LMC), upper working class (UWC), middle working class (MWC), and lower

working class (LWC). Trudgill interviewed ten speakers from each of five electoral wards in Norwich plus ten school-age children from two schools. These sixty informants were then classified on six factors, each of which was scored on a six-point scale (0–5): occupation, education, income, type of housing, locality, and father's occupation. Trudgill himself decided the cut-off points among his classes. In doing so, he shows a certain circularity. His lower working class is defined as those who use certain linguistic features (e.g., *he go*) more than 80 percent of the time. Out of the total possible score of 30 on his combined scales, those scoring 6 or less fall into this category. Members of Trudgill's middle middle class always use *he goes*, and that behavior is typical of those scoring 19 or more. His study is an attempt to relate linguistic behavior to social class, but he uses linguistic behavior to assign membership in social class. What we can be sure of is that there is a difference in linguistic behavior between those at the top and bottom of Trudgill's 30-point scale, but this difference is not one that has been established completely independently because of the underlying circularity.

Shuy's Detroit study (Shuy et al., 1968) attempted to sample the speech of that city using a sample of 702 informants. Eleven field workers collected the data by means of a questionnaire over a period of ten weeks. They assigned each of their informants to a social class using three sets of criteria: amount of education, occupation, and place of residence. Each informant was ranked on a six- or seven-point scale for each set, the rankings were weighted (multiplied by 5 for education, 9 for occupation, and 6 for residence), and each informant was given a social-class placement. Four social-class designations were used: upper middle class, those with scores of 20–48; lower middle class, those with scores of 49–77; upper working class, those with scores of 78–106; and lower working class, those with scores of 107–134.

There are some serious drawbacks to using social-class designations of this kind. Bainbridge (1994, p. 4023) says:

> While sociolinguists without number have documented class-related variation in speech, hardly any of them asked themselves what social class was. They treated class as a key independent variable, with variations in speech dependent upon class variations, yet they never considered the meaning of the independent variable. In consequence, they seldom attempted anything like a theory of why class should have an impact, and even more rarely examined their measures of class to see if they were methodologically defensible.

Woolard (1985, p. 738) expresses a similar view: 'sociolinguists have often borrowed sociological concepts in an ad hoc and unreflecting fashion, not usually considering critically the implicit theoretical frameworks that are imported.' She adds (p. 739), 'However, to say that our underlying social theories are in need of examination, elaboration, or reconsideration is not to say that the work sociolinguists have done or the concepts we have employed are without merit.' (See also Horvath's comment cited on p. 11.)

Exploration 6.2: Social Class

How would you try to place individuals in the community in which you live into some kind of social-class system? What factors would you consider to be relevant? How would you weigh each of these? What class designations would seem to be appropriate? Where would you place yourself? You might also compare the scale you have devised for your community with similar scales constructed by others to find out how much agreement exists.

Chambers (2003, ch. 2) wrestles with the problem of class as a category and with what he calls the 'fuzziness' inherent in class boundaries. He admits (p. 44) that sociolinguists 'often rely on their intuitions in assigning social classes to individuals in the sample population in their studies,' but avers (p. 44) that any such 'judgment sample' made by an experienced sociolinguist familiar with the region under investigation 'carries few risks.' For Chambers it is apparently enough that the sociolinguist has an intuitive grasp of the social-class composition of the group being investigated and chooses representative (or 'prototypical') individuals using his or her own best judgment. He admits that this is not the preferred sociological method and reveals an 'abyss between the sampling methods of sociolinguistic surveys and the type of survey represented by opinion polls' (p. 45). He claims, however, that the particular sampling methods of sociolinguistic investigation have been justified by their results, adding (p. 54) that the social stratification that interests sociolinguists is 'often crystal clear.'

Chambers' view is an optimistic one. 'Class' is not a transparent concept and 'fuzziness' is ever present (see Ash, 2002, for an extended discussion). Can you clearly assign any John Doe or Jane Doe a class membership? Are the same criteria applicable to all individuals in society, e.g., to both the black and white inhabitants of northern cities in the United States, and to both recent immigrants to London and the residents of Mayfair? Is class structure the same in both the industrialized and non-industrialized parts of the same society? Do the criteria for classification apply equally to John Doe and Jane Doe? Are the different generations fairly treated? Another way of looking at John Doe is to try to specify what kinds of groups he belongs to and then relate his various uses of language to membership in these groups. The obvious disadvantage of such an approach is the lack of generalizability of the results: we might be able to say a lot about the linguistic behavior of John Doe *vis-à-vis* his membership in these groups, but we would not be able to say anything at all about anyone else's linguistic behavior. We can contrast this result with the statements we can make from using the aforementioned social-class designations: they say something about

the linguistic usage of the 'middle middle class' without assuring us that there is really such an entity as that class; nor do they guarantee us that we can ever find a typical member.

One of the major problems in talking about social class is that social space is multi-dimensional whereas systems of social classification are one-dimensional. As we have seen, at any particular moment, an individual locates himself or herself in social space according to the factors that are relevant to him or her at that moment. While he or she may indeed have certain feelings about being a member of the lower middle class, at any moment it might be more important to be female, or to be a member of a particular church or ethnic group, or to be an in-patient in a hospital, or to be a sister-in-law. That is, creating an identity, role-playing, networking, etc. may be far more important than a certain social-class membership. This is the reason why such concepts as 'social network' and 'communities of practice' are attractive to some investigators. Sometimes, too, experience tells the investigator that social class is not a factor in a particular situation and that something else is more important. For example, Rickford's work (1986) on language variation in a non-American, East Indian sugar-estate community in Cane Walk, Guyana showed him that using a social-class based model of the community would be inappropriate. What was needed was a conflict model, one that recognized schisms, struggles, and clashes on certain issues. It was a somewhat similar perspective that Mendoza-Denton (2008) brought to her work among rival Latina groups in a California school where the main issue was Norteña-Sureña rivalry.

The work of Labov, Trudgill, and others tries to describe the speech characteristics of members of social groups, that is, various *sociolects*. Traditionally, linguists have been interested in *idiolects*, the speech characteristics and linguistic behavior of individuals. They have also maintained that, once free variation is taken into account, an idiolect is highly representative of the linguistic behavior of all the speakers of that language. In fact, that is usually the approach linguists adopt in studying an exotic language: they find a speaker who is willing to serve as an *informant*, and they attempt to describe that speaker's language using appropriate *field methods*, i.e., elicitation techniques. They usually show little hesitation in generalizing their statements about that speaker's linguistic behavior to all speakers of the language. Sociolects, however, are statements about group norms arrived at through counting and averaging. To the extent that the groups are real, that is, that the members actually feel that they do belong to a group, a sociolect has validity; to the extent that they are not, it is just an artifact. In the extremely complex societies in which most of us live, there must always be some question as to the reality of any kind of social grouping: each of us experiences society differently, multiple-group membership is normal, and both change and stability seem to be natural conditions of our existence. We must therefore exercise a certain caution about interpreting any claims made about 'lower working-class speech,' 'upper middle-class speech,' or the speech of any other social group designated with a class label – or any label for that matter.

Distinguishing among social classes in complex modern urban societies is probably becoming more and more difficult. We are far removed from the caste system described by Gumperz (1958) in his village of Khalapur in India, or the clearly differentiated societies so often described by anthropologists. We are also considerably distanced from the rural societies favored by dialect geographers. Cities like New York and London continue to change, and some would argue that the process of change itself has actually speeded up – certainly the process of social change has. If such is the case, the very usefulness of 'social class' as a concept that should be employed in trying to explain the distribution of particular kinds of behavior, linguistic or otherwise, may need rethinking.

It was for reasons not unlike these that Milroy (1987a) preferred to explore network relationships and the possible connection of these to linguistic variation, rather than to use the concept of 'social class.' In her work, Milroy found that it was the network of relationships that an individual belonged to that exerted the most powerful and interesting influences on that individual's linguistic behavior. When the group of speakers being investigated shows little variation in social class, however that is defined, a study of the network of social relationships within the group may allow you to discover how particular linguistic usages can be related to the frequency and density of certain kinds of contacts among speakers. Network relationships, however, tend to be unique in a way that social-class categories are not. That is, no two networks are alike, and network structures vary from place to place and group to group, e.g., in Belfast and Boston, or among Jamaican immigrants to London and Old Etonians. But whom a person associates with regularly may be more 'real' than any feeling he or she has of belonging to this or that social class. I will have more to say in the following chapter about this use of network structure in the study of linguistic variation. We will also see how the concept of 'communities of practice' will be helpful in understanding differences in language behavior.

Data Collection and Analysis

Once an investigator has made some decision concerning which social variables must be taken into account and has formed a hypothesis about a possible relationship between social and linguistic variation, the next task becomes one of collecting data that will either confirm or refute that hypothesis. In sociolinguistics, this task has two basic dimensions: devising some kind of plan for collecting relevant data, and then collecting such data from a representative sample of speakers. As we will see, neither task is an easy one.

An immediate problem is one that I have previously referred to (p. 18) as the 'observer's paradox.' How can you obtain objective data from the real world

without injecting your own self into the data and thereby confounding the results before you even begin? How can you be sure that the data you have collected are uncontaminated by the process of investigation itself? This is a basic scientific quandary, particularly observable in the social sciences where, in almost every possible situation, there is one variable that cannot be controlled in every possible way, namely, the observer/recorder/analyst/investigator/theorist himself or herself. If language varies as much as it does, the presence of an observer will have some effect on that variation. How can we minimize this effect? Even data recorded by remote means, e.g., by hidden cameras and sound recorders, may not be entirely 'clean' and will require us to address additional ethical issues which severely limit what we can do and which we would be extremely unwise to disregard. We know, too, that observations vary from observer to observer and that we must confront the issue of the reliability of any observations that we make. Sociolinguists are aware that there are several serious issues here, and, as we will see, they have attempted to deal with them.

The usual kind of data collection device is a questionnaire designed to elicit data illustrative of the use of the variable or variables that are being investigated. Since experience has shown that the different variants of a variable occur in different circumstances, the questionnaire must be designed to elicit data in a variety of circumstances. Many studies have made a four-fold distinction in categorizing those circumstances: (1) a casual situation, with sub-categories such as speech outside the formal interview, or conversation with a third party (i.e., not the person doing the interviewing), or responses to general questions, or recall of childhood rhymes, or the narration of a story about feeling one's life to be in peril; (2) an interview situation; (3) the reading aloud of a story; and (4) the reading aloud of lists of words and of pairs of words like *den* and *then*. A questionnaire which elicits these various kinds of linguistic behaviors will cover very casual speech (the casual situation), more formal speech (the interview situation), and the most formal speech of all (the different reading tasks). A person who says *shootin'* when explaining how he at some time felt himself to be in mortal danger may well read the same word presented on a list as *shooting*, and someone who pronounces *caught* and *court* as homophones during an interview may well distinguish them in some way when the words appear in contrast with each other on a list of pairs of words.

In his work in New York City, Labov (1966) investigated both careful and casual speech. His four types of careful speech, from most to least careful, were: reading lists of close pairs (e.g., *den* and *then*), reading lists of words, reading a prose passage, and participating in a formal interview. His five types of casual speech came from situations such as speech outside the formal interview, conversation with a third party, responses to questions, telling childhood rhymes, and recounting an incident which might have proved fatal. This classification gave Labov a total of nine *contextual styles* for analysis in his work. He also insisted that each style had to be accompanied by appropriate *channel cues*. In particular, casual style had to be marked by such cues. These cues involved changes

of speech pitch, volume, and rate of breathing, and perhaps such things as outbursts of laughter. Labov regarded speech not accompanied by one or more of these cues as formal rather than spontaneous and casual. Such cues most often accompanied either the subject's breaking away from the topic of the recorded interview to deal with some situation in the immediate environment, e.g., a family interruption, or a change of topic, particularly a change brought about by Labov's asking subjects to talk about a narrow escape from death. Labov also included what he called a *subjective reaction test* in his questionnaire, requiring subjects to react to taped samples of speech containing the five variables he was concerned with in his study. In this way he was able to compare what informants said about their own and others' usage with their actual usage, note differences between the two, and hypothesize about the consequences for such matters as linguistic change. Mendoza-Denton (2008, pp. 222–5) questions the 'naturalness' of such interview-derived data and also the usefulness of the *danger of death question*. She says that in her work using the latter would have been an 'outright *faux pas* . . . highly suspicious to gang members . . . very personal, and only to be told to trusted friends.' However, she does admit that 'the sociolinguistic interview paradigm . . . has yielded replicable results that allow us to contextualize variation in a broader context.'

Trudgill's questionnaire (1974, pp. 195–201) required his subjects to answer certain questions, e.g., 'What different parts of Norwich have you lived in?' and 'Which schools did you go to?' It also required that subjects read word lists aloud 'as naturally as you can,' and later 'as rapidly as you can,' and also to read pairs of words. The word lists contained words like *paper*, *baker*, *silly*, *you*, *avoid*, and *girl*; the pairs lists contained pairs like *boot–boat*, *hair–here*, *bust–burst*, *daze–days*, and *moon–moan*. Questions about individual local words were also asked, e.g., 'Do you know what a "dwile" is?' and questions about Norwich itself, e.g., 'What do you think of Norwich as a place to live?' Trudgill then asked his subjects to read aloud a short story 'as naturally as you can,' to make judgments about Norwich speech ('Do you like the way people in Norwich speak?'), to listen to certain pronunciations of words and judge whether or not they themselves used those pronunciations, and to judge ten paired sets of words in order to 'tick which way you think is correct, and then . . . underline the way you say it yourself, either if it's the same or different.'

The other part of the linguist's task is sampling: finding a representative group of speakers. The conclusions we draw about the behavior of any group are only as good as the sample on which we base our conclusions. If we choose the sample badly, we cannot generalize beyond the actual group that comprised the sample. If we intend to make claims about the characteristics of a population, we must either assess every member of that population for those characteristics or sample the whole population in some way. Sampling a population so as to generalize concerning its characteristics requires considerable skill. A genuine sample drawn from the population must be thoroughly representative and completely unbiased. All parts of the population must be adequately represented, and

no part should be overrepresented or underrepresented, thereby creating bias of some kind. Fortunately, according to Sankoff (1974, p. 22):

> A speech community sample need not include the large number of individuals usually required for other kinds of behavioural surveys. If people within a speech community indeed understand each other with a high degree of efficiency, this tends to place a limit on the extent of possible variation, and imposes a regularity (necessary for effective communication) not found to the same extent in other kinds of social behaviour. The literature as well as our own experience would suggest that, even for quite complex speech communities, samples of more than about 150 individuals tend to be redundant, bringing increasing data handling problems with diminishing analytical returns. . . . It is crucial, however, that the sample be well chosen, and representative of all social sub-segments about which one wishes to generalize.

The best sample of all is a *random sample*. In a random sample everyone in the population to be sampled has an equal chance of being selected. In contrast, in a *judgment sample* the investigator chooses the subjects according to a set of criteria, e.g., age, gender, social class, education, and so on. Sometimes, too, it is the investigator who judges each of these categories, e.g., to which social class a subject belongs. A judgment sample is obviously less adequate than a random sample. However, it is the kind of sample preferred in most sociolinguistic studies (see Chambers, 2003, pp. 44–5).

In sampling the speech of the Lower East Side in New York City, Labov did not use a completely random sample because such a sample would have produced subjects who were not native to the area, e.g., immigrants from abroad and elsewhere in the United States. He used the sampling data from a previous survey that had been made by Mobilization for Youth, a random sample which used a thousand informants. Labov's own sample size was eighty-nine. He employed a *stratified sample*, i.e., one chosen for specific characteristics, from that survey. He also wanted to be sure that he had representatives of certain groups which he believed to exist on the Lower East Side. When he could not, for various reasons, interview some of the subjects chosen in the sample, he tried to find out by telephoning the missing subjects if his actual sample had been made unrepresentative by their absence. He was able to contact about half of his missing subjects in this way and, on the basis of these brief telephone conversations, he decided that his actual sample was unbiased and was typical of the total population he was interested in surveying.

The Detroit study (Shuy et al., 1968) initially collected data from 702 informants in the city. However, the data used for the actual analysis came from only thirty-six informants chosen from this much larger number. In selecting these thirty-six, the investigators wanted to be sure that each informant used had been a resident of Detroit for at least ten years, was 'representative,' had given a successful interview, and had provided an adequate amount of taped material for analysis. In other words, to any initial biases that might have been created

in choosing the first set of 702 informants was added the possibility of still further bias by choosing non-randomly from the data that had become available. This is not to suggest that any such biases vitiate the results: they do not appear to do so. Rather, it is to point out that the kinds of concerns sociolinguists have about data and sources of data have not necessarily been the same as those of statisticians.

Wolfram (1969) chose forty-eight black informants from those interviewed in the Detroit study. These informants were evenly divided into four social classes used in that study. Each group of twelve was further divided into three age groups: four informants in the 10–12 age group, four in the 14–17 age group, and four in the 30–55 age group. Wolfram also selected twelve white informants from the highest social class in the Detroit project, again by age and sex. Wolfram's study therefore used a total of sixty informants: twenty-four (twelve white and twelve black) from the upper middle class and thirty-six who were black and were members of the working classes, with equal numbers in each such class. Such a sample is very obviously highly stratified in nature.

It is actually possible to use a very small sample from a very large area and get good results. For their *Atlas of North American English* Labov and his co-workers sampled all North American cities with populations over 50,000. Labov (2006, p. 396) reports that they did this through a telephone survey: 'Names were selected from telephone directories, selecting by preference clusters of family names representing the majority ethnic groups in the area. The first two persons who answered the telephone and said that they had grown up in the city from the age of four or earlier, were accepted as representing that city (four or six persons for the largest cities). A total of 762 subjects were interviewed.' The investigators were very pleased with the results of this sampling procedure for the ANAE.

Investigations may also have a 'time' dimension to them because one purpose – increasingly becoming the major purpose – of such studies is trying to understand language change. They may be *apparent-time* studies in which the subjects are grouped by age, e.g., people in their 20s, 40s, 60s, etc. Any differences found in their behavior may then be associated with changes that are occurring in the language. *Real-time* studies elicit the same kind of data after an interval of say 10, 20, or 30 years. If the same informants are involved, this would be in a *panel* study; if different people are used it would be in a *trend* study. Obviously, real-time studies are difficult to do. The study of the Queen's English (see pp. 44–5) is one such study, but she was the sole panel member. (The British Granada Television series *Seven Up* directed by Michael Apted is a panel study but its focus is social rather than linguistic.) The study (see p. 205) that replicated Labov's work on Martha's Vineyard was a real-time trend study. As we will see in the following pages, most studies of change in progress are apparent-time studies for reasons which should now be obvious.

Studies employing the linguistic variable are essentially correlational in nature: that is, they attempt to show how the variants of a linguistic variable are related to social variation in much the same way that we can show how

children's ages, heights, and weights are related to one another. However, a word of caution is necessary. It is quite possible for two characteristics in a population to covary without one being the cause of the other. If A and B appear to be related, it may be because either A causes B or B causes A. However, it is also possible that some third factor C causes both A and B. The relationship could even be a chance one.

To avoid the problems just mentioned, we must distinguish between *dependent variables* and *independent variables*. The linguistic variable is a dependent variable, the one we are interested in. We want to see what happens to language when we look at it in relation to some factor we can manipulate, the independent variable, e.g., social class, age, gender, ethnicity, and so on: as one of these changes, what happens to language? As Chambers (2003, p. 26) expresses it, '*Socially significant linguistic variation requires correlation*: the dependent (linguistic) variable must change when some independent variable changes. It also requires that the change be orderly: the dependent variable must stratify the subjects in ways that are socially or stylistically coherent.'

This kind of sociolinguistic investigation is often called *quantitative sociolinguistics* and it is, as I have indicated previously, for some sociolinguists the 'heart of sociolinguistics' (Chambers, 2003, p. xix). Quantitative studies must therefore be statistically sound if they are to be useful. Investigators must be prepared to employ proper statistical procedures not only in their sampling but also in the treatment of the data they collect and in testing the various hypotheses they formulate. They must be sure that what they are doing is both valid and reliable. In our case the issue of validity, that is, whether or not the sociolinguist is really measuring what he or she is claiming to be measuring, hardly ever arises. Such work certainly meets Lepper's criterion (2000, p. 173): 'the researcher must show that what is being described is accurately "named" – that is, that the research process has accurately represented a phenomenon which is recognizable to the scientific community being addressed.' However, the issue of reliability, that is, how objective and consistent the measurements of the actual linguistic data are, is a real and pressing one. There are well-known procedures for making sure that the data we gather have been collected reliably, and there are approved tests of that reliability. However, in some sociolinguistic investigations little attention is paid to this issue. We are simply presented with sums of informants' responses or averages of one kind or another, and given little or no account of how confident we can be concerning the actual items summed or averaged. If only one person collected the data, how consistent was that person in the actual collection? If two or more were involved, how consistently and uniformly did they employ whatever criteria they were using? Bailey and Tillery (2004, pp. 27–8) have identified a cluster of such issues, e.g., the effects of different interviewers, elicitation strategies, sampling procedures, and analytical strategies, and pointed out that these can produce significant effects on the data that are collected and, consequently, on any results that are reported. Therefore, there may still be room for improving the reliability of our results.

Serious empirical studies also require experimental hypotheses to be stated *before* the data are collected, and suitable tests to be chosen to decide whether these hypotheses are confirmed or not and with what degree of confidence. Such tests often require that the data collected be treated quantifiably so that the variation in the actual distribution of the various responses is taken into account in each category. It is not enough just to calculate simple means or averages. The *standard deviation*, that is, the actual distribution of the various measurements around those means, is critical if we wish to compare different means. This is a standard procedure. Peculiar kinds of variations around means, e.g., skewing or bi-modal distributions, must also be noted.

Petyt (1980, pp. 188–90) points out how the kinds of figures that sociolinguists use in their tables may be misleading in a very serious way. Sociolinguists stratify society into sub-groups, the members of which are measured in certain ways, and then these measurements are pooled. Individual variation is eliminated. Hudson (1996, p. 181) offers a similar criticism, declaring that such pooling

> loses too much information which may be important. Information about the use of individual variants is lost when they are merged into variable scores, and information about the speech of individuals is also lost if these are included in group averages. At each stage the method imposes a structure on the data which may be more rigid than was inherent in the data, and to that extent distorts the results – discrete boundaries are imposed on non-discrete phonetic parameters, artificial orderings are used for variants which are related in more than one way, and speakers are assigned to discrete groups when they are actually related to each other in more complex ways.

Petyt (p. 189) provides the data given in table 6.1. These data come from an investigation of *h*-dropping in West Yorkshire, and the table shows the means for five sub-groups, i.e., social classes. Petyt points out that, if the range of variation within each sub-group is also acknowledged to be of consequence, there is a considerable overlap among the performances of individuals, so that 'it is not the case that this continuum can be divided in such a way that the members of each social class fall within a certain range, and members of other classes fall outside this.' He indicates the range of individual scores in table 6.2, and adds that 'in the case of Classes II and V the bracketed figures indicate what

Table 6.1 *H*-dropping: means for five social groups

I	*II*	*III*	*IV*	*V*
96	64	43	21	17

Source: Petyt (1980, p. 189)

Table 6.2 *H*-dropping: within-group ranges for five social groups

I	*II*	*III*	*IV*	*V*
81–100	7–100 (40–100)	2–100	0–86	0–80 (0–37)

Source: Petyt (1980, p. 189)

the range would have been had there not in each case been one individual whose speech was markedly "status incongruent." If these two individuals had not formed part of the sample the figures would look more "regular," but there would still not be "discrete groups which are relatively unified in their linguistic behavior." ' There is considerable overlap.

It is quite obvious that if we look only at means in such a case we are tempted to say one thing, whereas if we consider the distribution of responses within each class we may draw some other conclusion. The overriding issue is that there are approved procedures to help investigators to decide how far they can be confident that any differences that they observe to exist among the various classes, that is, among the various means, are due to something other than errors in measurement or peculiarities of distribution. Such procedures require an investigator not only to calculate the means for each class, but also to assess the amount of variation in the responses within each class, and then to test pairs of differences of means among the classes using a procedure which will indicate what likelihood there is that any difference found occurs by chance, e.g., one chance in twenty.

Most social scientists employing statistical procedures regard this last *level of significance* as a suitable test of a hypothesis. In other words, unless their statistical procedures indicate that the same results would occur by chance in fewer than one case in twenty, they will *not* say that two groups differ in some respect or on a particular characteristic; that is, they insist that their claims be significant at what they call the 0.05 level of significance. We are also much more likely to find two means to be significantly different if they are obtained from averaging a large number of observations than a small number.

Figure 6.3 provides a further illustration of the problems inherent in comparing populations in this way. The two groups A and B that are compared there for characteristic X produce different mean scores and that difference may even be statistically significant. However, there is an enormous overlap among individuals in groups A and B – the shaded area. The majority of individuals in the two groups overlap in their X-ness, whatever that may be. It would be very unsafe indeed to make claims about the X-ness of the next person you meet from either group A or group B.

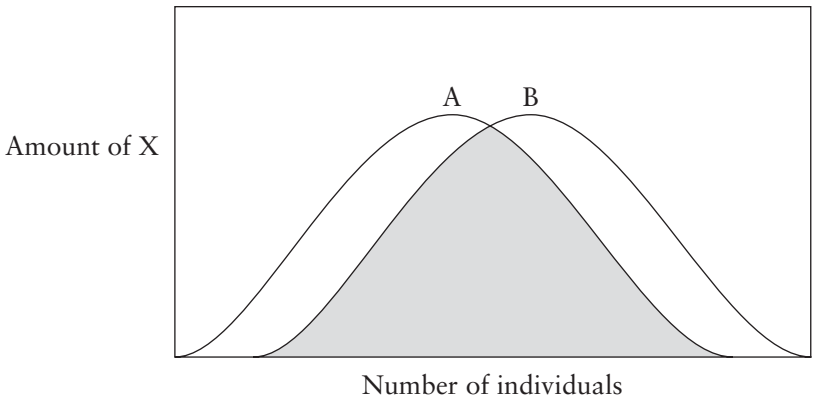

Figure 6.3 X for two groups, A and B

Whenever you look at results reported by sociolinguists, you should keep in mind the above-mentioned issues concerning the formulation of hypotheses and the collection, analysis, and interpretation of data. Statisticians certainly keep them in mind in assessing the claims they make. In examining individual sociolinguistic investigations, therefore, you must ask what exactly are the hypotheses; how reliable are the methods used for collecting the data; what is the actual significance of results that are reported on a simple graph or histogram; and what do the findings tell us about the initial hypotheses. Anyone who attempts to do serious work in sociolinguistics must address himself or herself to such issues. Likewise, anyone who wishes to draw conclusions about either the structure of language or how language varies within groups or between groups must ask the same questions of data that someone presents to support this or that conclusion. Some sociolinguists have tended not to be very rigorous in their statistical treatments, but this has not stopped them from drawing very strong conclusions, which seem 'obvious' and 'interesting' to them; whether these conclusions are 'significant' in the sense of having met an appropriate statistical test of a well-stated hypothesis may not be a concern.

Milroy and Gordon (2003, p. 168) are aware that there may be problems here. However, they ask: 'should we equate failure to achieve statistical significance with sociolinguistic irrelevance?' Their answer is that 'statistical tests, like all quantitative procedures are tools to provide insight into patterning in variation. They must be used critically.' Labov himself (1969, p. 731) has stated that statistical tests are not always necessary: 'We are not dealing here with effects which are so erratic or marginal that statistical tests are required to determine whether or not they might have been produced by chance.' Dealing with a critic of Labov's work, Milroy (1992, p. 78) says:

Exploration 6.3: Interpreting Results

Percentage of non-RP forms for three consonants

	(ng):[n]	*(t):[ʔ]*	*(h):∅*
MMC	31	41	6
LMC	42	62	14
UWC	87	89	40
MWC	95	92	59
LWC	100	94	61

Source: Trudgill (1995, p. 36)

Given information of the kind contained in the above table, and no supporting explanations of how each mean (expressed as a percentage) was determined, what can you say with confidence about the linguistic behavior of members of each of the groups mentioned and about the reported differences in behavior between the various groups? Can you be more confident of some conclusions than of others? You might consider questions such as the following: How important are the differences among the various working classes (UWC, MWC, and LWC)? Between the middle classes (MMC and LMC)? Between the middle classes as a whole and the working classes as a whole? Between the LMC and the UWC? How does performance on the (h) variable appear to be different from performance on the (ng) and (t) variables in all social classes? Is there any difference in performance on the (ng) and (t) variables? What other kinds of information would you require to strengthen any conclusions you would wish to draw?

(Before you begin this task you should be quite sure that you understand all of the information contained in the table. For example, what does the table tell you about the likelihood of finding *singing* pronounced as *singin'* in the MMC, *butter* pronounced as *bu'er* in the UWC, and *house* pronounced as *'ouse* in the LWC?)

> It is not surprising that an anti-quantitative linguist should advocate confirmatory statistical testing, but it is very important to understand the proposition put forward here is simply wrong. If Labov's interpretations were suspect (and of course they are not), this would not arise from the fact that he failed to test for significance. There was no reason for him to do so because the claims he wished to make were quite simple . . . and because in his analysis the same patterns were repeated for every variable studied.

According to Milroy, since this kind of sociolinguistic inquiry is 'exploratory' in nature, it can be likewise 'exploratory' in its quantitative approach. Labov's most recent work (2001) is still exploratory in nature but it is also extremely sophisticated in its sampling, data collection, and hypothesis-testing. Sociolinguists now make increasing use of VARBRUL, a set of computer programs specifically designed to deal with the kinds of problems encountered in studies of variation (see Bayley, 2002, particularly pp. 124–34).

Further Reading

Chambers (2003), Chambers et al. (2002), Lucas and Bayley (2007), Milroy and Gordon (2003), and Tagliamonte (2006) are good introductions to issues discussed in this and the following two chapters. Linn (1998) is also useful. Wolfram and Schilling-Estes (2005) and Wolfram and Ward (2006) focus on American English.

See Sankoff (1978, 1985) for the use of statistics in linguistics.