3 Varieties and Variation

JAMES MILROY and LESLEY MILROY

1 Introduction

As Edward Sapir remarked (1921: 147), "everyone knows that language is variable." Variability in language is within everyone's experience of using and listening to language, and most people show some degree of interest in it. Despite this, however, linguistic theory has until quite recently paid relatively little attention to variation, and in many branches of inquiry languages have been treated as if they were wholly or mainly invariant entities, or as if the variability that does exist within them were unimportant, accidental, or inessential. Variability within a language or dialect and variation across languages have not been central concerns in the dominant linguistic theories of this century – Saussurean theory, American and Prague School structuralism, and Chomskyan theory. One consequence of this, to which we return below, is that linguistic theorizing has been largely based on *standardized* forms of languages, rather than on the more variable forms of naturalistic speech.

Within descriptive linguistics, the main exception to this is what can be called the *variationist paradigm*, which is based on the research methods and analytic techniques developed by William Labov (see especially Labov 1966c, 1972), on the critique of current linguistics set out by Weinreich, Labov, and Herzog (1968), and on ideas developed in several papers by Labov himself. Many important principles are set out in this work, but the most relevant to the present discussion is the principle that *variability in language is, or may be shown to be, structured*. Weinreich, Labov, and Herzog (1968) observed that linguistic scholars generally had not only focused mainly on uniform states of language, but had also equated this uniformity with structuredness. That is, they had believed that only uniform states can be structured and had tended to dismiss variability in language as unstructured or random and therefore

not worth studying. Many examples of this can be cited, and some are noticed in section 2 below. Because of this emphasis on invariance, however, linguists have often used as their subject matter "cleaned-up" or invented data, and not for the most part naturally occurring data (which is of course inclined to be variable). It should be noted that in contrast to this, the variationist paradigm is *empirical* in its methods, in that it depends on collecting naturalistic speech from real speakers and insisting on full *accountability* to the data so collected, no matter how messy some of the data may be.

A major reason for recent advances in variation studies is technological. Before tape recorders became easily available, students of spoken language had to rely on single-word citation forms and on memory. In the last 30 years, however, much attention has been devoted to collecting tape-recorded data in situations that are as "naturalistic" as possible. This advance has been of great importance in several branches of linguistic analysis, including work on interactional sociolinguistics, following the principles of John Gumperz and others, and work on conversational analysis by Sacks, Schegloff, and Jefferson, and others. The key difference between the variationist paradigm and other empirical approaches is that the former is focused on understanding variation and change in the structural parts of *language* rather than the behavior of *speakers* or the nature of speaker interaction. The activities of speakers in naturalistic settings are indeed studied, but not primarily for what they tell us about speakers or interaction between speakers; the interest is in what they tell us about varying structures of language and speakers' knowledge of these variable structures. The aim of the next section is to explore to some extent the range and depth of variation that exists in a language in terms of the different linguistic and extralinguistic dimensions in which variation is observed.

2 The Range and Depth of Variation: Some Examples

Language is inherently variable at a number of structural levels – in phonology, morphology, and syntax in particular. Phoneticians frequently point out that no two utterances of the same word by the same speaker are ever exactly alike, and it is also recognized that some variation in sound-patterns may be structured. One of these structured aspects of phonetic/phonological variation is labelled *assimilation*. For example, in a phrase such as *bacon and eggs* in British English, the final /n/ of "bacon" may be assimilated to the place of articulation of the preceding /k/ and realized as a velar rather than an alveolar nasal. This is likely to happen in relatively rapid or casual speech and is to that extent "stylistic": the same speaker may use either alternant (alveolar or velar) in a reasonably regular way according to situation and context. In

morphology and syntax also, there are many alternative ways of saying the same thing, especially in nonstandard forms of languages – for example, variation between *you were* and *you was* in London English and between *can't* and *cannot* in Newcastle speech.

The quantitative paradigm explores the regularity in linguistic variation by examining certain dimensions that are external to language itself and relating variation in these to variation in language. These dimensions are normally social; however, strictly speaking, two of the dimensions that are involved in variation are perhaps better described as "natural" dimensions. These are the dimensions of *space* and *time*, which exist independently of human society and which have been studied extensively by linguists for some centuries. Language variation in space forms the subject matter of linguistic geography, which itself includes traditional dialectology of the kind exemplified by Orton and others (1963–9). Language variation in time forms the subject matter of historical linguistics. The main advances in recent years, however, have been in the more obviously "human" dimensions of variation, that is, in *social* variation in language, and it is important to notice here that this type of variation was the first to be studied *quantitatively*, i.e., by counting variants and comparing the incidence of variants in different speakers and groups of speakers. Quantification is an essential methodological tool of the variationist paradigm, and for this reason it is sometimes called *quantitative social dialectology*.

Investigators proceed by first selecting a *variable* (for example, a sound segment such as /a/, which is observed to vary in pronunciation) and quantifying occurrences of variants of this variable in the speech of different speakers and groups of speakers. The use of quantification represents an advance in descriptive techniques, as it enables investigators to make accurate statements about fine-grained differences between groups of speakers in a community. Formerly, such statements tended to be categorical: For example, a particular usage (such as /h/-dropping in British English, for example) might have been categorized as "working-class" and the "proper" use of /h/ might have been categorized as "middle-class." In practice such usages are seldom categorical for any group of speakers. Table 3.1 shows variability in /h/ according to social class in Bradford and Norwich, England. It is clear from this that the use or non-use of initial /h/ is not categorical for any group in either city.

Table 3.1 Percentage of /h/-dropping in Bradford and Norwich (formal style) (after Chambers and Trudgill, 1980: 69, and Petyt, 1977).

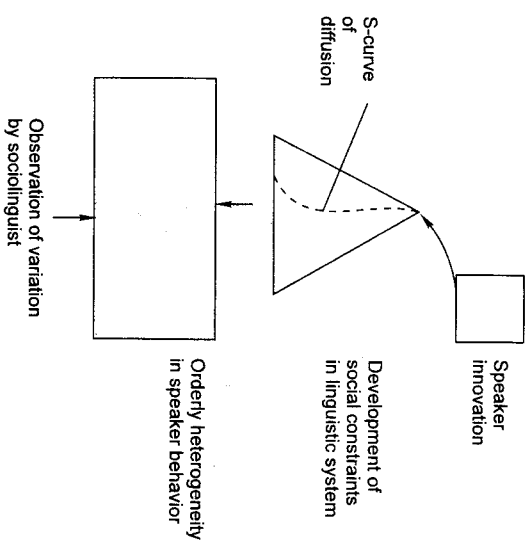|	*Bradford*	*Norwich*
Middle middle class	12	6
Lower middle class	28	14
Upper working class	67	40
Middle working class	89	60
Lower working class	93	60

The use of quantification has led to speculation as to whether human linguistic competence is in fact probabilistic (see Cedergren and Sankoff, 1974; Fasold, 1990: 249–57) but the literal claim that speakers "know" the exact quantities in which different variants should be used by them in varying situations has not been generally accepted. The main advance brought about by using quantification is methodological, not theoretical. Quantitative methods of analysis have enabled us to propose socially based explanations for aspects of language variation in time, space, and social space. Generally, they have done this by relating variation in language to variation in society and situational contexts of speech. In order to demonstrate covariation between linguistic and social categories, it is normal to identify one or more *speaker variables*. The most widely used of these is *socioeconomic class*. Other variables that are commonly used include age of speaker, sex (gender) of speaker, ethnic group of speaker, and social network (we return below to the purpose of using speaker variables in this type of research). In addition, it is usual, where possible, to recognize *contextual style* as a variable, and this variable tends to cut across or interact with the speaker variables. It is not a speaker variable in quite the same sense as the others mentioned, as variation according to social context or occasion of use (i.e., "stylistic variation") is not a characteristic of the speaker as such, but of the speaker's relationship to the resources of the language and of the situational contexts in which the speaker finds himself at different times (see further below, section 5). All normal speakers of a language exhibit stylistic variation in speech, and patterns of stylistic variation exhibited by speakers, taken together with other variables, may reveal the direction of linguistic change in progress at some particular time. It is therefore desirable in fieldwork to obtain a range of styles from informants, and it should be noted that this has important implications for fieldwork methods. In particular, certain techniques have to be used in order to elicit casual or informal styles, which informants may tend to avoid in talking to an outsider such as the fieldworker. This problem is known as the *Observer's Paradox*. There is a considerable literature on fieldwork method relevant to the observer's paradox (see especially Labov, 1972a, and L. Milroy, 1987b).

3 Speaker Variables and the Speech Community

Whereas other branches of linguistics focus on "the language" or "the dialect," quantitative sociolinguistics focuses on the speech community. This is envisaged as a sociolinguistic entity rather than a purely linguistic one. It is not supposed that all speakers in the community speak in exactly the same way or that there is some "real" or "genuine" uniform language variety that characterizes the community. It is not even necessary that the members of such a community should all speak the same language, although the most

influential quantitative work so far has focused on monolingual states of language. The speech community, according to Labov, is a locus in which speakers agree on the social meanings and evaluations of the variants used, and of course it incorporates variability in language use. In practice the speech communities studied by sociolinguists have been geographically very restricted, and this restrictedness is in itself important in the identification of the origins and diffusion of linguistic changes in progress. They are identified as happening, not in "the language" as a whole, but in some particular speech community, and the progress of these changes is then analyzed as they spread in the speech community, and possibly to other speech communities. Underlying all this is an assumption that access to speakers in present-day speech communities will bring us closer to understanding the origins of linguistic changes. It can also be argued on this basis that changes do not take place in the abstraction that we call "the language," but that they arise from the activities of speakers and then feed into the linguistic system. For this reason it has been proposed that a methodological distinction should be drawn between *innovation* and *change*. A linguistic innovation is an act of the speaker (or speakers). It may or may not become established in the linguistic system and become part of the language. If it does penetrate into the system, however, it becomes a linguistic change and will at that point display a regular structure of variation in terms of the social variables discussed above. This process is represented in figure 3.1.

Figure 3.1 Model of transition from speaker innovation to linguistic change.

4 Language Maintenance, Standardization, and Change

A basic assumption of variation studies is that "at any time we care to look at a language ...it is variable and in a state of change" (J. Milroy, 1992: 2). Sometimes change is rapid and sometimes it is slow, but there is no reason to believe that there can ever be a time when a spoken language is completely stable. It follows that the methods used for studying it should preferably recognize that languages are dynamic and not static phenomena. Traditionally, this has not been a central perception in the descriptive and comparative methodology, and language states at different times and places have often been studied as if these different states were like different physical objects that could be compared with each other as (largely invariant) wholes. It appears, however, that insofar as they are social or sociocultural phenomena, languages are subject to speaker-based processes that are initiated in social groups. When a language (such as French or English) is recognized by society as a single phenomenon, it can be assumed that it has been subject to a diachronic process of language *maintenance*. It seems to be necessary to invoke such a process in order to account for the existence of language states that are popularly perceived to be static (in reality they are not), and we give some attention to this process here.

In studies in the sociology of language the term *language maintenance* signifies the process of consciously maintaining – if necessary by government intervention – a particular form of a language in a population where there is linguistic diversity wide enough to make communication difficult; it is usually bilingual situations that are involved. In the histories of major languages, such as English and French, the process of maintenance has also been prominent – sometimes carried out by overt legislation, and sometimes in a less formal way by imposing the codified linguistic norms of elite social groups on society as a whole through education and literacy (for a discussion see Milroy and Milroy [1985] 1991; see also chapter 18). These processes of maintenance, which arise from the imposition of linguistic norms by powerful social groups, can be subsumed under the term *language standardization*. The chief linguistic consequence of standardization is a tendency to structural uniformity in a language, i.e., variability is resisted and suppressed by stigmatization of nonstandard variants. It should be noted, first, that standardization can be viewed as a diachronic process occupying an extended time-scale, and second, that it is continuously in progress, and not completed in any language except a dead one. Thus it is not correct to state (for example) that the standardization of English was completed at some particular time, such as the eighteenth century. Finally, it should further be noted that the speech communities in which quantitative sociolinguists have usually worked have been within nation states in which a standardized form of the language is considered to be a well established superordinate norm (as contrasted with pidgin situations, for example). As a consequence of this, an understanding of processes of change in such communities should ideally take account of this fact.

In the research projects carried out in Belfast by Milroy and Milroy (1975–82) the notion of language maintenance was extended to cover situations in which the pressure to maintain language states is noninstitutional. Individuals in small-scale communities do not systematically act as language planners or language maintainers, but in order to account for the survival of nonstandard or low-status varieties, noninstitutional norm enforcement of this kind must be assumed, and the effect of such norm enforcement is just as much a form of language maintenance as is overt standardization. If we wish to discriminate between the two types of maintenance, we can call the former kind institutional and noninstitutional – we can call the latter kind *vernacular maintenance*. The hypothesis followed out in the Belfast research was that community norms of language are maintained by these informal pressures, and it was further suggested that relatively localized patterns of identity marking are involved.

In nation states in which there is consciousness of a standard language, vernacular maintenance can result in conflict between two opposing norms. This emphasis on societal conflict is one of the things that differentiates this research from that of Labov, and it has obvious consequences for the characterization of the idealized "speech community" in which every speaker agrees on the evaluation of the varying norms of language. If low-prestige varieties can persist and spread within urban societies, it may be that their speakers do not evaluate variants in the way that other sectors of the community do. Thus, for example, /h/-less, rather than /h/-pronouncing norms can be seen as favored rather than stigmatized in some small-scale communities. The pattern arising is of course one of conflict rather than consensus, and this conflict pattern can be at least partially understood as arising from the conflict between status-based ideologies and solidarity-based ideologies in the community. When the latter are dominant, localized noninstitutional norms of language will tend to be preserved (see further section 7 below on "social network").

5 Extra-linguistic Variables

The main speaker variables that have been used are noted above. Their use is methodological and exploratory and not in itself explanatory. Thus it should not be assumed that to relate language variation to a social variable, such as social class, is to explain language variation as being *caused* by social class variation. There are several reasons for this caution, the chief of which is that there may be many aspects of social behavior that are not accounted for in a single social variable, and also underlying social factors that are subsumed under such a label as "social class" (such as educational level) which may

sometimes yield more precise correlations than the main composite variable (in this case social class).

As the methodology is exploratory, it is also open-ended. It is not necessarily the case that all language variation can be accounted for by relating it to social variation, and no one has claimed to be able to do this. It is clearly likely that other factors are involved, including linguistic constraints (Weinreich, Labov, and Herzog, 1968), and a start has been made on investigating conversational or discoursal constraints on variation (Milroy, Milroy, and Docherty, 1994). Criticisms of sociolinguistic method (e.g., Cameron, 1990) on the grounds that it claims to account socially for all linguistic variation are therefore otiose. It is also likely in any study that there will be a residue of apparently random variation which is difficult to account for using the methods of quantitative sociolinguistics.

Of the social variables that are commonly used, two at least are composite (or complex) variables, in that they are calculated by reference to a number of indicators. These are socioeconomic (social) class and social network. Quantitative measurements of social class depend on such indicators as income, trade or profession, and educational level, while social network depends on indicators of density and multiplexity in a speaker's social relationships. Certain other social variables, such as age and sex of speaker, are mathematically simpler in that they do not depend on multiple indicators and do not need to be calculated in the form of numerical scores (this does not of course imply that correlations with age and sex are simple to interpret; see chapters 8 and 9). Whereas these simplex variables are verifiable from observation at the data collection stage, there can be and has been dispute about how the complex variables (especially social class) are to be conceptualized, calculated, and interpreted in specific investigations. The most controversial social variable is socioeconomic class.

Labov's (1966c) study in New York City proceeded by measuring covariation of language with variation in social class membership, and the social measurements used were imported from sociology. They depended, moreover, on a particular social theory associated with the work of Talcott Parsons (1952), which uses the concept of *stratificational* social class. This involves classifying individuals in a hierarchy of class groupings based on the idea of a continuum from highest to lowest, and is the most familiar way of treating social class in Western countries. However, there are other theories of social class, such as those associated with Marx, which are not stratificational, but which use a process model of class. Social class is seen as emanating from economic factors, such as the means of production and distribution, and resulting in two broad groupings in society – the proletariat and the bourgeoisie. Whereas the stratificational model results in a consensus view of society, in which there is general agreement within the hierarchy, the Marxist model plainly emphasizes conflict between the different interest groups. This difference in social models is reflected in the consensus-based and conflict-based models of the speech community that were mentioned above. In proce-
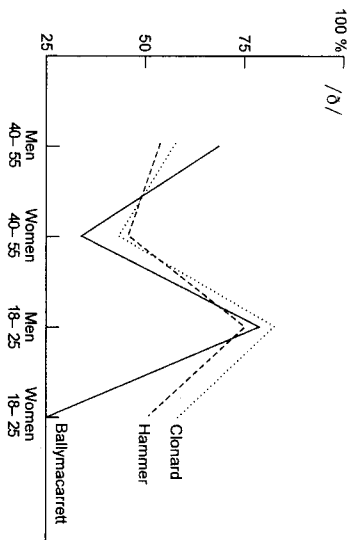
dural terms, however, a stratificational model is much more readily adapted to quantitative use than is a process model (see further below).

Social class has been by far the most widely used social variable, and it appears that this emphasis on social class is not confined to modern sociolinguistics. It is quite prominent in work on the descriptive history of English, for example, and it is usual for lay people to assume that it is the most important social category. However, since Labov's New York study, it has become the central social variable in sociolinguistic research, in that results obtained from work on other variables (particularly gender) are interpreted in terms of social class or the closely associated notion of prestige. In the Labov methodology, the direction of style-shifting (toward "careful" style) corresponds to upward movement in the social hierarchy, and is interpreted in terms of it (this correspondence is further discussed and interpreted by Bell (1984)). It also affects the interpretation of gender difference in speech. Thus the fact that females tend to speak more "carefully" than males has been interpreted as arising from a desire on their part to acquire social prestige through their speech, as they could not traditionally acquire this through career success – as males could. This type of conclusion can be objected to on various grounds, but what is at issue here is the centrality of social class, not in the quantitative methodology itself, but in the *interpretation* of the results of that methodology. In what follows, the quantitative method is taken for granted as valid, and the emphasis is on the interpretation of sociolinguistic patterns arrived at in terms of speaker variables. We first consider gender.

6 Gender

Variation according to gender appears to be universal and, in terms of style, the tendency appears (in Western societies at least) to be always in the same direction. Females tend toward the careful end of the continuum and males toward the casual end. Similarly, it can be said that females favor "prestige" norms and males vernacular norms. This is something of a paradox because, although sociologists and anthropologists normally recognize that virtually all societies have accorded higher status and greater power to males than to females (for a useful survey see Giddens, 1989), it is females who tend to use higher status variants of language. There is also another factor involved, however, which may not have to do with prestige or "carefulness" of speech: This is that males appear to favor more localized variants, which carry some kind of identity-based social meaning in the local community, whereas females identify more with supra-local variants in speech. It should be noted that gender variation in speech is not necessarily evident to the casual observer. Normally, both sexes use the same variants, but in different quantities, and the differences are fine-grained; therefore, they can normally be demonstrated only by quantitative means.

Figure 3.2 Deletion of medial /ð/ in words of the type *mother, together.*

100 %
75
50
25

/ð/

Men 40–55 Women 40–55 Men 18–25 Women 18–25
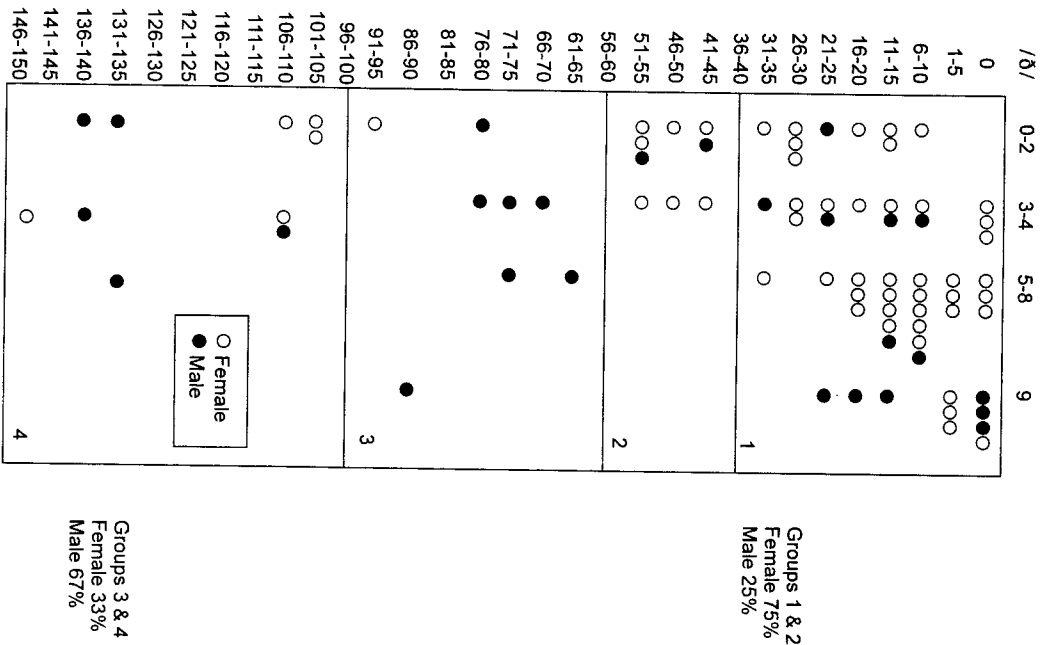
Ballymacarrett
Hanmer
Clonard

The inner-city Belfast study (Milroy and Milroy, 1978 etc.) dispensed with social class as a variable and concentrated first on variation according to age and sex differences. It demonstrated that, within the same social class or stratum, gender difference was always present and almost always moved in the same direction. Figure 3.2 demonstrates a clear gender difference in the variable [ð] (presence or absence of the medial consonant in words of the type: *mother, bother, together*) in two generations of speakers, with virtually no difference according to age.

Since then, clear patterns of gender differentiation have been demonstrated in a number of studies, so much so that it can be suggested that gender difference may be prior to class difference in driving linguistic variation and change. In an important study, Horvath (1985) has regraphed some of Labov's New York City data in terms of gender difference instead of social class. Figure 3.3 shows her results. In the graph the lower social classes are on the left (0–2) progressing upwards to the right to the upper middle class (9) at the extreme right of the diagram. It shows that, although there is certainly an effect of class, sex of speaker accounts for the distribution more satisfactorily than class. In the top half of the graph females are dominant, and in the bottom half males dominate. The one upper middle-class male, Nathan B., who appears in the bottom half of the graph has been discussed by Labov (1966) as anomalous. In a gender-based interpretation this individual is no longer anomalous: he is converging on the male norm rather than the class norm.
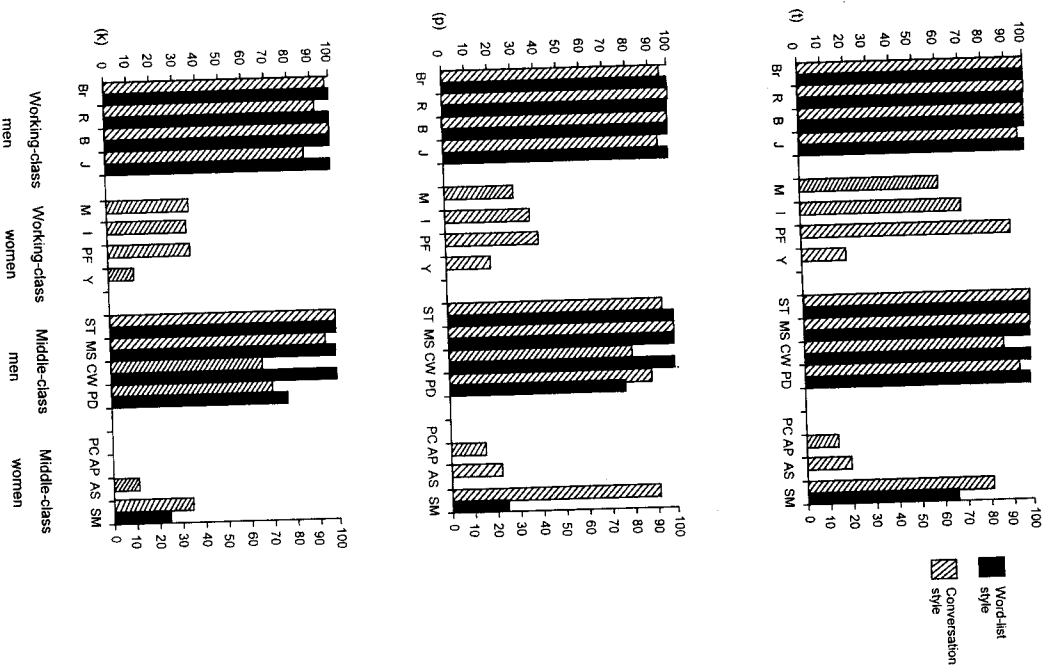
Other studies that suggest the priority of gender over class have been carried out in Newcastle upon Tyne under the direction of L. Milroy (for a summary see L. Milroy, 1992). Rigg (1987) studied glottalization of /p/, /t/, and /k/ in medial and word-final positions (as in *pepper, butter, flicker, what, top, lock*), and her findings are shown in figure 3.4 and table 3.2.

Figure 3.3 /ð/ in New York City (after Labov, 1966, and Horvath, 1985).

/ð/   0-2   3-4   5-8   9

0
1-5
6-10
11-15
16-20
21-25
26-30
31-35
36-40
41-45
46-50
51-55
56-60
61-65
66-70
71-75
76-80
81-85
86-90
91-95
96-100
101-105
106-110
111-115
116-120
121-125
126-130
131-135
136-140
141-145
146-150

O Female
● Male

Groups 1 & 2
Female 75%
Male 25%

Groups 3 & 4
Female 33%
Male 67%

1   2   3   4

The most obvious finding here is that glottalization is sex-marked rather than class-marked (the effect of class is quite trivial). In the conversational style shown in table 3.2, the overall female scores do not overlap at all with the

Figure 3.4 Percentage of glottalized variants of /p/, /t/, and /k/ in word-medial position in two speech styles of 16 Tynesiders.

male scores. Hartley's (1992) study reported in Milroy, Milroy, and Hartley (1994) demonstrates further that, within the glottalization pattern, males lead in glottal reinforcement whereas, surprisingly, females lead in use of the glottal stop. Other studies of glottalization, such as Mees (1987) and

Table 3.2 Percentage of glottalized variants of three voiceless stops in spontaneous speech of 16 Tynesiders.

	Working class			Middle class		
	/p/	/t/	/k/	/p/	/t/	/k/
Male	99.5	97.0	94.5	96.5	91.0	80.5
Female	60.0	31.0	28.0	27.0	32.5	11.0

Kingsmore (1994), also associate the glottal stop with female usage. As it is hardly feasible to explain these findings in terms of prestige, it has been suggested that glottal reinforcement is favored by males as a traditional localized pattern, whereas the glottal stop is spreading at a supra-local level. Comments by Wells (1982) and others to the effect that the glottal stop is beginning to enter the prestige accent, Received Pronunciation, would seem to support an interpretation based on the idea of supra-local diffusion. Before leaving the subject of gender differentiation, we should further note that the priority of gender over class is also suggested by various studies in the Arab world (Alahdal, 1989; Jawad, 1987; Jabeur, 1987). Thus it is no longer clear that gender-marking in language should be interpreted in terms of class, status, or prestige as prior categories. It may be that female norms for some reason become the prestige norms in the course of time, and hence that gender differentiation is an important driving force in linguistic change, independently of class (for a different view see Labov, 1990).

7 Social Network

Social network was developed as a quantitative speaker variable by L. Milroy ([1980] 1987) as part of the Belfast inner-city study. The main methodological difference between network (as used here) and other variables that have been examined is that it is based, not on comparisons between groups of speakers, but on relationships contracted by individual speakers with other individuals. It is assumed that all individuals are embedded in networks of personal ties. Furthermore (following a considerable body of social anthropological research), it is argued that when these ties are strong, they can act as normenforcement mechanisms. The idea of social network was adopted from social science research as a means of accounting for patterns of vernacular maintenance over time. In the context of the Belfast research (and indeed more generally) this is an important issue, as stigmatized and low-status forms of language tend to persist despite strong pressure from "legitimized" norms.

This has always been difficult to explain. Social network analysis thus provides a methodology for studying the interaction between patterns of maintenance and patterns of change. A basic assumption is that, in order to understand how language changes are adopted by communities, we must also take account of patterns of resistance to change.

In his study of Harlem street-gangs and subsequently in Philadelphia, Labov (1972a; Labov and Harris, 1986) used typically ingenious methods to measure personal relationships, which are similar to (and indeed a precursor of) social network analysis. The main difference from social network analysis is that Labov's studies are of bounded groups, whereas social networks are in principle open-ended and anchored on the individual. Furthermore, Labov's research agenda does not specifically refer to a maintenance/change hypothesis and is not primarily concerned with accounting for language maintenance.

The Belfast research operationalized the network strength of individual speakers, which are based on the notions of *density* and *multiplexity*. A maximally dense network is one in which everyone knows everyone else, and a multiplex relationship is one in which A interacts with B in more than one capacity (for example, as workmate and friend). Statistical analysis suggested that the network variable was capable of accounting for certain patterns of linguistic variation, and further that the network variable interacted with the variables of gender and age. Extensive discussions of the use of statistical analysis in sociolinguistics are provided by Fasold (1984) and L. Milroy (1987b).

Quantitative social network analysis has been used in other urban monolingual situations, notably by Bortoni-Ricardo (1985) in Brazil. Here the method was successful in revealing patterns of adaptation to the urban dialect by rural migrants. Other examples of sociolinguistic applications of network analysis are Schmidt (1985: Australian Aboriginal adolescents), Lippi-Green (1989: an Alpine rural community in Austria), V. Edwards (1986: black adolescents in England), and W. Edwards (1990: black Detroit speakers). A network-based approach is also very suitable for analyzing situations of bilingualism, language contact, and language shift. The classic network-based study of language shift is Gal (1979: Hungarian/German-speaking peasant workers in Austria). More recently the method has been used to study language use and language shift in the Newcastle Chinese community (see Li, Milroy, and Pong, 1992; Li, 1994). Attempts have also been made to project the idea of social network on to past states of language (Van der Wurff, 1990). Whereas the maintenance/change model provides a framework in which past language changes may be discussed in an illuminating way, it is dubious whether the social networks of individuals who are no longer accessible to systematic observation can be adequately reconstructed.

Social network and social class

Social network and social class represent different orders of generalization about social organization. Class accounts for the hierarchical structure of

society (arising from inequalities of wealth and power), whereas network deals with the dimension of solidarity at the level of the individual and his or her everyday contacts. An attempt has been made to link the two concepts together in a sociolinguistic model by using the notion of *weak* network ties (Milroy and Milroy, 1992).

It is evident that close-knit solidarity ties are characteristic of lower and higher social groups, and that, in the middle sectors of society, social network density and multiplexity tend to be weak. A process model of social class, such as Thomas Højrup's theory of *life-modes*, suggests that different kinds of social network structure do not occur accidentally, but "fall out" naturally from different life-modes, such as those of the self-employed, of wage-earners (both poor and relatively affluent), and of professionals. A high proportion of close-knit ties on the one hand, and of loose-knit ties on the other are consequent upon the life-modes which themselves are constitutive of distinct classes. In this way, different kinds of social network can be linked to the wider organization of society, and it is suggested that these links can be explicated by considering the properties of weak as well as strong ties (for details, see Milroy and Milroy, 1992).

8   The Sociolinguistic Variable

Critics of sociolinguistics have had much to say about the social variables discussed above, but much less about the idea of the linguistic variable. This is a relatively old concept in linguistics, most familiar in the idea of the *phoneme*, which typically manifests itself in the form of variants known as allophones. The sociolinguistic variable is also manifested in the form of variants. It differs from the phoneme, however, in that the focus is on social variation rather than exclusively on intra-linguistic variation. Thus the range of a sociolinguistic variable does not normally correspond to that of a phoneme, as different social values may overlap with different phonemes. The nonidentity of the sociolinguistic variable with the phoneme is not always sufficiently emphasized by investigators.

In the foregoing it has been assumed, without comment, that sociolinguistic variables are usually phonological elements. In practice, this has often been so, but the principle underlying the method is more general than this. What is important is that variants of a variable should demonstrably be variants of the same underlying linguistic element. At higher levels of linguistic organization (particularly syntax) it is difficult to meet this condition, as it is often not clear that two syntactic variants (for example, active and passive sentence forms) have the same meaning and distribution in the language system. This difficulty is discussed by (among others) Romaine (1984), Lavandera (1978), L. Milroy (1987b). It is less often pointed out that comparable problems may also

emerge at very particularistic levels of subphonemic organization. For example, in British English, the glottal stop (for /t/) occurs in different positions within words: medial, final, and in some dialects syllable-initially. However, not only does the likelihood of the glottal stop differ in these different positions; the variants that it alternates with may also differ in the different positions. Furthermore, the social meaning attached by the community to these variants may also vary according to where they occur in words or syllables. Therefore a correct quantitative statement depends on isolating environments in which we can be sure that we are dealing with variants of the same sociolinguistic phenomenon. If therefore we regard glottalization as a "variable," we must acknowledge that it is a complex variable that contains a number of subvariables within it, and it is possible that these subvariables will display different (even contrasting) patterns.

This problem of the correct input to the variable has not been widely discussed in the literature. It seems to be most prominent in what have been called "divergent dialect" studies, where the range of variation encountered is very large (for a discussion see J. Milroy, 1992: 68–75). It also happens in these studies that some salient variables do not occur frequently enough for quantification throughout the whole range (for treatment of such a variable see Milroy and Harris, 1980). It should also be noted that for reasons of time (and probably also in principle) investigators cannot quantify all the variation that exists in a speech community. Selection of representative variables depends on the skill of the investigators, and nonquantitative description is also necessary for a reasonably comprehensive account.

In the final section, we are concerned with the relevance of variation studies to our conception of what constitutes a language, a dialect, or a variety.

9 Languages and Dialects as Physical Entities

As we have noted, variation studies have led scholars to question the definition of "a language" and what kind of object a language is. Linguists have generally relied on a working assumption that there exists a structured and stable entity which we can call a language or a dialect of a language. This can be accessed or described in internal structural terms, e.g., as having a "phonology," "grammar," and "lexis," without reference to society – i.e., independently of the speakers who use it in their speech communities. As noted above (see section 1), it has also been usual to treat this entity as having an invariant underlying structure. Social dialectology has called into question the discreteness of these entities that we call languages, and seeks to contribute to a clearer understanding of what we actually mean when we say that we are describing a "language." In order to characterize a "language" or any quasi-discrete variety of a language, we need to invoke sociopolitical criteria in addition to structural linguistic criteria.

Sociolinguists (e.g., Downes, 1984; Chambers and Trudgill, 1980) commonly point out that boundaries between languages cannot be wholly determined in terms of structural difference or mutual (in)comprehensibility. Several Scandinavian languages, for example, are mutually comprehensible to a great extent and some dialects of English are not readily comprehensible to speakers of other dialects. There are many areas of the world in which variability within and between languages is very great, and some situations in which speakers may not be entirely certain as to which language they are speaking (see Grace, 1990, 1992, for comments on blurring of distinctions between certain Melanesian languages in speaker usage). Similarly, there are many situations in which two or more languages are mixed. Finally, there are rapidly changing situations, especially in the genesis and development of pidgin and creole languages, in which younger generations may use markedly different varieties of the language from those of older speakers.

From a variationist point of view, a language is a dynamic phenomenon. It is appropriate to liken languages to relatively fluid and variable physical states, and to use process models rather than product or static models in describing them.

It can be suggested that discreteness of individual *languages* is not inherent in the nature of *language* as a structural phenomenon: This apparent discreteness is socially or sociopolitically imposed. French is a "language" not merely because it has a linguistic structure that differentiates it from other languages and which is peculiarly "French," but also because its structures are recognized, prescribed, imposed, and agreed within a particular nation-state (and certain other areas formerly influenced by this nation-state). Separateness of languages is therefore largely the result of social and political processes, and among these processes language standardization is particularly important. Our tendency to think of languages as discrete phenomena is partly conditioned by the existence of standard languages, such as standard English and standard French.

This lack of discreteness in real language states is an important matter in the study of the histories of languages. In studies of language change, there are many examples of the tendency to regard a language as a physical entity. Yet it never seems to have been possible to specify *purely in terms of language structure* the precise point in history at which one language "became" another language. Therefore, just as it is difficult to specify discreteness of languages in space, so it is also difficult to differentiate them in time. The point in history at which one language becomes another may have more to do with political history than with linguistic differentiation.

A final point, which arises from the above discussion, is that variation studies have in many ways blurred the Saussurean distinction between synchronic and diachronic linguistics. Mesthrie (1992) has coined the useful term "panchronic" to describe such an orientation. As sociolinguists study speech communities at a single point in time, their analytic work is primarily synchronic, and their quantitative statements are synchronic statements. How-

ever, the paradigm has subtly altered the relationship between historical and other forms of linguistics, in that variation in time is grouped together with variation in space and social space as one aspect of linguistic variation. Similarly, as we have noted, process models of language in society have begun to have some impact. It is clear, however, that insofar as it is concerned with linguistic change, Labov (1994) considers the theoretical content of this work to be chiefly a contribution to historical linguistics. While this is certainly Labov's theoretical position, it is not necessarily a view shared by all sociolinguists. Equally important, as we have tried to show, is the development of an integrated account of variation encompassing not only dimensions of time and geographical space, but the various dimensions of social space, such as gender, generation, status, and network structure, discussed in earlier sections of this article.