Part I Data Collection

1 Entering the Community
Fieldwork

CRAWFORD FEAGIN

While the ultimate goal of sociolinguistic research is to resolve questions of linguistic importance, such as how language change comes about, nothing of that sort can be accomplished without first entering a community in order to collect data which will help provide the basis for any such answers. The central problem in collecting sociolinguistic data has been described by Labov as the Observer's Paradox: "our goal is to observe the way people use language when they are not being observed" (1972a: 61). Sociolinguistic fieldwork of all kinds, whether recorded interviews, participant observations or street-corner quizzes, must be geared to overcome this problem. In this chapter, I consider several well-established methods. I begin with a section on "Planning the Project," dealing with preliminary considerations for designing and conducting a sociolinguistic survey. The heart of the chapter, as indeed of field research, is the second section on the "Sociolinguistic Interview," the Labovian protocol for selecting informants and eliciting different styles of speech. I then consider some other elicitation methods used in sociolinguistics: participant observation and rapid and anonymous observations. While telephone surveys have been fruitful in the past (see Labov *et al*. 2006; Ash 2000), today they have limited use because of the general shift to cell phones, eliminating the use of area codes or telephone directories in identifying likely participants. Long-distance surveys today can utilize internet phone services such as Skype or other internet technologies (e.g. web-based surveys for gathering self-reports of linguistic production and/or information on linguistic perceptions and attitudes; see Schilling 2013); they will not be discussed here. Instead, I will focus on face-to-face methods.

The Handbook of Language Variation and Change, Second Edition.
Edited by J.K. Chambers and Natalie Schilling.

1 Planning the Project

Although the methods involved are presented here as if they were sequential, in practice the various phases of fieldwork and other aspects of research are cyclical, or perhaps spiral. Investigation in one area will influence what can be done in another. An interview might provide insights about the community that can be incorporated into the protocol and produce a much better interview with subsequent informants. For instance, in my work in Anniston, Alabama (Feagin 1979), one teenager mentioned a recent snowstorm, an unexpected and exciting phenomenon in that part of the world, and so in later interviews I asked the rest of the teenagers about it. As a result, I came away with excited accounts of sledding on garbage-can tops and cookie sheets, wearing improvised boots made from plastic bags, and skidding dangerously over slippery roads. My interview protocol for the older people already included questions about a tornado that had hit Anniston 20 years before; the snowstorm provided similarly dramatic stories from an incident in the recent memories of the teenagers.

Similarly, sometimes in the course of an interview, investigators might discover an unexpected grammatical form or phonological realization. They must be attentive and flexible in order to pursue the newly discovered linguistic feature for that community.

As an aid to planning, a small-scale pilot project along the general lines of the main research will indicate more precisely what might be feasible goals and procedures. A larger consideration is that collecting data is only an intermediate goal. The ultimate goal is linguistic.

The hypothesis that motivates the project will influence how to go about collecting the data. Again, in my own work in Anniston, I hypothesized that over the three-and-a-half centuries of close contact, African-American speech would have influenced European-American grammar in the South. I therefore set out to elicit data from the white community that was parallel to Labov's African-American data from Harlem (Labov *et al*. 1968; Labov 1972b). Even though it turned out in large part that my hypothesis was not correct, nonetheless it was important to try to get parallel data so that a comparison would be possible.

An important guideline for fieldworkers at the planning stage is that a close analysis of a small amount of data is better than an unfinished grandiose project. With that in mind, I concentrated on the extreme generations (teenagers and grandparents) and extreme social classes (local working class and upper class), and the older rural working class (with no younger counterpart). More than that I could not handle, though ideally I would have liked to include the middle class and the middle aged, not to mention the local African-American community. However, examining only the two urban classes plus the older rural working class, using adolescents and grandparents in the city and elderly people from the country, and keeping the sample balanced in terms of gender, I was able to see change progressing through the community.

A rule of thumb in disciplines that require fieldwork is that one third of the project time will be spent in fieldwork, one third in analysis, and the final third in writing up the work. Though far from scientific, this rule provides an effective reminder of the point that time required for analysis and writing increases in a ratio of about 2:1 for each hour of data elicitation.

Competent fieldworkers have included a wide range of personality types. Because fieldwork requires face-to-face interaction, it is usually assumed that gregarious persons do best, and it seems likely that they would have an advantage, at least in getting started. Shy people might find this sort of work excruciating, especially in the beginning. However, shy people have sometimes proven highly successful in conducting interviews and obtaining data, for the simple reason that people often open up when talking to quiet people, perhaps because they find them unthreatening and perhaps because the lack of interruptions encourages them to speak at length (Schilling, personal communication).

1.1 Library research

Once the community has been selected for research, the next step is to get a perspective on the community itself – linguistic, demographic, and historical. Information on local speech, major industries, labor, religious institutions, communications, movement of peoples, and the historical development of the area can aid in understanding local society.

A survey of previous linguistic work must be carried out, both on the linguistic aspects you intend to study and on any previous research concerning the local language variety. Earlier work on the local variety, regardless of its quality, can be useful for time depth or for pinpointing interesting problems.

First-hand accounts of fieldwork can be found in Labov (1966), Feagin (1979), Milroy (1980), Dayton (1996), and Eckert (2000) for linguistics, and in Whyte (1943, 1984) and Liebow (1967) for ethnography. Such personal accounts are rarely published, but dissertations often include them in chapters on methodology. More general discussions may be found in Labov (1972a, 1984), Wolfram and Fasold (1974), Milroy (1987), Romaine (1980), Baugh (1993), Milroy and Gordon (2003), Di Paolo and Yaeger-Dror (2011), and Schilling (2013). For sociolinguistic fieldwork in non-Western societies where the investigator is clearly an outsider, see Albó (1970), Harvey (1992), Wald (1973), and Bowern (2008). Obviously, a different set of problems arises when the fieldworker is a foreigner, of different ethnicity, and not a native speaker of the language. While addressed to researchers doing basic linguistic fieldwork (rather than sociolinguistic research) in non-Western languages (frequently in remote areas), Samarin (1967) provides an overview of linguistic fieldwork, though now somewhat dated. Bowern (2008) is a more recent resource.

1.2 Ethnography

Along with gathering linguistic data, it is important to study the community itself in situ. While material collected from library research must not be overlooked if it is available, the researcher in the field must begin by observing the physical layout of the place, who lives where, who associates with whom, and in what situations particular people associate with each other. While this type of research can be seen in Fischer (1958) and more elaborately in Labov (1963), subsequent studies have become more sophisticated and more detailed, culminating in Eckert's intricate study of a suburban Detroit high school (Eckert 2000). It is through a thorough knowledge of both the structure and dynamics of the local community that the patterning and social meanings of language variation and change in the speech community can be fully understood. While some linguists have criticized sociocultural investigations as outside the competence of linguists who are not specialists in sociology or anthropology (Bailey 1996), the only way some aspects of language behavior can be understood and analyzed is through such an undertaking.

It was through such a study that Labov was able to show that younger people on the island of Martha's Vineyard who had decided to remain on the island after their high school years were picking up the fishermen's pronunciation of (ay) and (aw), regardless of their social class, while those who had decided to leave the island for further education and employment were shifting toward mainland speech norms (Labov 1963). Similarly, Eckert (2000) was able to show that the social division between "jocks" (middle class) and "burnouts" (working class) in suburban high schools played a role in transmitting urban Detroit features into suburban teenage speech. See Eckert (2000: Chapter 3) for a valuable account of the process of studying the ethnography of a community.

1.3 Linguistic variables

In a quantitative study of linguistic variation, acquaintance with previous work and perhaps a pilot study should help to narrow the focus of the project. In practical terms, however, this does not always take place right at the beginning. What needs to be isolated before analysis can begin, and preferably before data-gathering begins, is a selection of linguistic variables to be studied. As with fieldwork more generally, though, the process is iterative, and it may turn out that the variables one originally sets out to study are not of great sociolinguistic interest, and more important features may be revealed as fieldwork progresses.

The linguistic variable, a concept originating with Labov (1963, 1966), is a linguistic entity which varies according to social parameters (age, sex, social class, ethnicity), stylistic parameters (casual, careful, formal), and/or linguistic parameters (segmental, suprasegmental). Usually the social and stylistic variation will be coordinated in some way, so that the casual speech of an accountant will be similar to the formal speech of a plumber – though that remains to be seen in the course of the investigation.

The linguistic variable can be found at all linguistic levels: most common are phonological, such as, for example, (r) might be realized as [ɹ] or as [ə] in a community which has been r-less and is becoming r-ful; morphophonological as in (ing), the English present participle marker which has two common pronunciations, standard [ɪŋ] and casual [ɪn]; morphological as in the realization of the past tense form of *dive* either as *dived* or as *dove*; syntactic as in the realization of negated *be* variously as *ain't*, *isn't*, *'s not*, *is not*; or lexical as in the use of either *hero* or *grinder* as the word to designate a particular kind of sandwich. The most frequently studied variables are phonological and morphological.

The main criterion for determining the set of variants of a single variable is that the referential meaning must be unchanged regardless of which variant occurs. (This can present a problem when dealing with grammar, as pointed out by Lavandera (1978) and Romaine (1981).) The selection of one variant from the set will generally be motivated by either social or stylistic considerations. See Wolfram (1993) and Guy (1993) for discussions of some of the problems connected with settling on the variable(s) to be investigated.

1.4 Recording equipment

To name particular types of recording equipment would not be useful, because technology changes so rapidly. However, it is crucial that researchers use recording equipment meeting the technical specifications needed to produce sound of high enough quality for potential acoustic phonetic analysis (whether or not this is the immediate goal of the study) as well as high-quality external microphones. Some types of equipment have abiding advantages. For example, the lavaliere (lapel) microphone improves the quality of the sound and minimizes the speaker's attention to the recording mechanism. Also crucial is selection of recording location. Clearly, quiet locations are better than noisy ones; however, sometimes fieldworkers must sacrifice sound quality in favor of enhanced interactional quality when quiet locations that are comfortable to interviewees are not available. In addition, researchers should be aware that some types of noise that wreak havoc on audio recordings are practically unnoticeable to the untrained ear – for example, the noises emitted by electronic equipment (including computers), kitchen appliances, and espresso machines (despite the many other advantages of conducting interviews in comfortable public locations like coffee shops). The reader is referred to Cieri (2011) for excellent, detailed advice on selecting locations for interviews and choosing microphones and recording equipment. See also Schilling-Estes (2007) for good discussion and advice about videotaping interviews.

The main point is to get the best equipment possible given the practical constraint of expense. Recording fidelity is the primary consideration, and after that come ease of use, flexibility, weight, and other factors. Field recordings can be useful for many years, for purposes unplanned. In my case, tape recordings intended only for a study of grammar have since been used for work on phonology, both using impressionistic phonetic transcription and computer-assisted vowel analysis.

1.5 Institutional Review Board approval

Before heading off to the field, it is necessary to fulfill the requirements of the IRB – the Institutional Review Board – also called Ethical Review Board (ERB) or Independent Ethics Committee (IEC) at your institution and/or granting agency. In the US IRB approval is required for all research involving human subjects. Each institution has its own requirements, so it is advisable to obtain and complete IRB approval forms early and allow enough time for your project to be reviewed and accepted. A crucial component of the approval processes involves preparing an Informed Consent Form which will have to be signed by each study participant (or legal guardian, in the case of children under 18).

1.6 Self-presentation of the fieldworker

Having selected the community and investigated the locale, culture, and speech, and having the approval of the IRB, the investigator finally has to actually go there and find people to talk to. This is a rather stressful position to be in, from all accounts. Eckert (2000) describes the nightmares she had before beginning her work in the Detroit suburbs. Entering any community carries with it certain responsibilities for respecting the privacy and customs of local people. Most often, this is not a great problem because researchers tend to investigate cultures with which they have some personal familiarity. It is a much greater problem, obviously, in a culture and language that is not native to the investigator. In these situations, Samarin (1967: 19) recommends that the researcher undertake meticulous planning to deal with the pressures, being aware of the problems that might arise and arranging for breaks in order to get away from the locale from time to time.

Often, cultural alienation is not a factor. My own fieldwork, for instance, took place in my home town, where I had lived until I was 15, and where both my mother and grandfather had grown up. My role there, while conducting fieldwork between 1969 and 1973 and again in 1990 and 1991, was both as a visitor in the town, staying with my grandparents, and as a researcher working on my dissertation, carrying out interviews. On my side of town I was known to the people I interviewed as a friend's granddaughter or cousin, but on the other side of town I was a complete stranger doing research. I told people that I was working on a book on growing up in the town, and how it was changing over time, especially for the teenagers. I said I wanted to record speech in the interests of accuracy, so I would get the dialog right. As a former resident with kinship ties in the town, I attended church with my family, visited friends, and took my grandmother to her club meetings. I also attended revival meetings and visited a church on the other side of town, which helped me learn about the life and culture outside of my own experience and to meet older people who were members of the church I visited.

I was careful to dress suitably according to local custom, always wearing a skirt and stockings to interview older people and to attend classes at the high school, but sometimes wearing blue jeans and sitting on the floor when interviewing teenagers, explaining that I needed to watch the level on the tape recorder while we were talking. In this way I was showing respect to my elders and solidarity with the younger group. With teenagers, I generally took along sodas and chips, which helped make the interview less formal, though the crunch of potato chips sometimes can be heard on the recording.

In reporting on his research in a small town in North Carolina, Hazen (2000) explains that before beginning his fieldwork he had married a woman from the community, which gave him entree. However, as a native of suburban Detroit, he was not as well acquainted with the culture as he might have liked, though this also allowed him to assume the role of a student of that culture and ask questions that only an outsider could ask.

Albó (1970) describes in detail his entry into rural communities in Andean Bolivia where his identity as a priest proved advantageous. He was sometimes asked to bless houses, which gave him an opportunity to observe the living standards of the families. This contributed to his understanding of the degree of modernization of the household, giving insight into the relationship between the borrowing of linguistic forms and of material culture. It also gave him opportunities to line up interviews. Similarly, Harvey (1992), whose research was in Southern Peru after it became a dangerous area for outsiders, was considered the adopted daughter of a local family, which gave her a place in the community, allowing her to observe both language and culture.

Both Whyte (1943) and Liebow (1967) emphasize that it is never possible to completely fit in, nor is it necessary or even advisable. As white middle-class men carrying out ethnographic research among working-class men, one group white and the other black, they report that they were able to lower the barriers between their subjects and themselves but not to remove them. Liebow uses the image of the chain-link fence: you can see through it, but it remains a barrier. The researcher can become a friend, and even find a role in the community, but skin color, class affiliation, speech, or education may all set the investigator apart, which may of course result in less than ideal conditions for collecting maximally naturalistic speech data and maximally informed community understandings but which can also serve as a protection in some situations.

2 The Sociolinguistic Interview

The classic method of sociolinguistic research is the one-on-one recorded conversational interview (Labov 1972a, 1984; Wolfram and Fasold 1974). Recording has the obvious advantage of permanency, so that it is possible to return to the recording again and again, either for clarification or for further research. A second major advantage is that the recording permits the researcher to fulfill the Principle of Accountability (Labov 1972c: 72), so that all occurrences as well as non-occurrences of the variable in question can be identified and accounted for. In this way statistical manipulations of the data can show whether the occurrence of a variant is happenstance or patterned, and, if patterned, to what degree in contrast to its occurrence in the speech of others of varying social characteristics – age, sex, social class, ethnicity – and across speech styles. This, then, is the primary method of quantitative sociolinguistics.

Variations on this classic sociolinguistic interview approach include interviewing two or more speakers together (Feagin 1979), or even breaking down the one-on-one interview structure with pairs of interviewers. Labov used group interviews in his work with Harlem street gangs (Labov *et al*. 1968; Labov 1972b), with one lavaliere microphone per person, and a multitrack recorder, while Wolfram, Hazen and Schilling-Estes (1999) reported great success in using pairs or even teams of interviewers in their work in Ocracoke, North Carolina. These variations on the one-on-one interview are intended to reduce the formality of the interview, turning it into a more natural social event.

The sociolinguistic interview – regardless of the variations on it – does carry some disadvantages. The interview as a speech event is a special genre (Wolfson 1976), so the naturalness and certainly the informality of the recorded speech can be called into question, regardless of efforts to make the speaker feel comfortable with the situation. The use of lavaliere microphones may remove the microphone from view, but the recording device is always there. However, despite our fears that recorded sociolinguistic interviews may be less than fully "natural," Eckert (2000) reminds us that speakers are adaptable in both research and non-research contexts, and they can just as readily shape their speech to naturally fit a conversational sociolinguistic interview as any other speech event in which they find themselves in the course of their daily interactions. Furthermore, recording devices are becoming less obtrusive as technology continues to advance, and non-research-related recording situations are more commonplace.

The interview method works best for frequently occurring variables, especially phonological and morphological, and certain syntactic forms, such as negation. But many syntactic structures, including interrogatives, double modals, and special auxiliaries such as perfective *done*, do not occur frequently enough in interviews to provide sufficient data for analysis. Moreover, the interview is problematic for discourse studies and ethnomethodology (Briggs 1986).

2.1 Selecting speakers

The earliest community-based research in sociolinguistics, Labov's work in Martha's Vineyard (1963), used a judgment sample, selecting subjects to fill preselected social categories, all locally born and raised adults and teenagers. His categories crisscrossed geographic area, profession, and ethnicity. It is interesting that in this early study gender was not considered a separate social variable, though only men were used for acoustic analysis. In his New York study a few years later, Labov was able to base his subject selection on a previous random survey by the Mobilization for Youth, a project of the School of Social Work at Columbia University, which had conducted a random-sample survey of the Lower East Side. Labov used their demographic data to select natives of the area or people who had arrived by age five, as well as people from across a range of social strata. This was, then, a stratified random sample in that it selected a stratified sample from what had originally been a random sample. In his third major project (Labov *et al*. 1968; Labov 1972b), Labov worked with teenage boys who were members of street gangs. This represents an early – possibly the earliest – study of language variation through social networks.

Trudgill (1974: 20–30), who followed soon after with a study of Norwich, England, relied on a quasi-random sample taken from four ward voter registration lists. The names from the voter lists were chosen randomly, but the wards were not random but were selected "so that they had, between them, social and economic characteristics that were, on average, the same as those of the city as a whole" (1974: 22).

My own work in Anniston, Alabama, was based on a judgment sample, filling pre-selected cells on a number of criteria. First, speakers were chosen because they were native-born or had arrived by age five. Second, preference was given to those whose parents were from the area. Though I did not know of the literature on networks at the time, I often selected subjects who were "a friend of a friend," using the resources of my family and their acquaintances for contacts. I began with friends of a younger cousin, then moved on to friends of my grandparents. Later, when I wanted to work in another section of town, I began with a home economics teacher who turned out to be an acquaintance and an admirer of my grandfather. She welcomed me into her classes where I was able to observe, and in some cases (with permission) to record the students and make appointments with them for interviews. Twenty years later, in 1990, I followed the same procedures to find teenage subjects on both sides of town. Luckily enough, the new home economics teacher said that if the earlier one, who had been her own teacher, had let me visit her class, it was all right with her. The now-retired home economics teacher was still in touch with the students I had interviewed 20 years before, and through her I was able to find those students again, most of whom still lived in the area.

When the Milroys were selecting informants in Belfast (Milroy 1980; Milroy and Milroy 1985), they were forced to rely on the "friend of a friend" method for contacts because of the sectarian problems in the city, and especially in the working-class neighborhoods in which they intended to conduct their research. Their methods auspiciously introduced the concept of the network to sociolinguists. (See Milroy and Llamas, this volume, on social networks.)

Generally, researchers must use common sense to select subjects not by some pre-ordained "social-science" formula but according to the prevailing conditions of the setting they are working in, as well as their research goals. Thus, in selected subjects for their study of Ocracoke, Wolfram *et al*. (1999) chose ancestral islanders whose families had been on the island for at least several generations, because the purpose of the study was to recover, as far as possible, the traditional dialect that was rapidly eroding in the face of incursion into the island community by tourists and new residents from a range of dialect areas. In a quite different vein, Eckert (2000) selected high school students of opposing ideologies and styles, known as "burnouts" and "jocks," basically working-class and upper-middle-class adolescents, because she was studying the dynamics of adolescent speech and culture in the school setting.

One danger with selecting informants by pre-selected categories is that results can be self-fulfilling or circular. For a more general community study, Horvath (1985) gathered speech data from a stratified judgment sample in Sydney, Australia, and analyzed it using principal components analysis, a statistical technique which grouped speakers into clusters according to their linguistic similarities, and in that way revealed what the sociolinguistic groupings of Sydney were, based entirely on speech, rather than on preconceived notions about class membership, sex, or other social groupings.

Except for studies that take a special interest in the language of children (as with Roberts, this volume), it is better to avoid speakers younger than adolescents, since there is the possibility of confounding phonological or grammatical development with local variation.

The two genders must be kept fairly even numerically in order to prevent a confounding of gender differences with other distinctions. Many studies have demonstrated gender differences in language, beginning with Fischer's (1958) study of (ing) which showed that boys in a small New England town were more likely to use the [ɪn] variant than girls.

Attention must also be given to social class (as in Ash, this volume), as well as its interaction with age and gender. The older members of any class usually have the most conservative phonology; teenage working-class boys and girls are often the leaders in innovation, with certain items being more characteristic of one gender than the other. Eckert (2000) elaborates a striking example of highly innovative teenagers who show gender- and social group-based differentiation in their usage patterns for new linguistic features. In regard to grammar, the higher classes will usually use a local variety of the standard; the older members of the working class will maintain older forms which have become nonstandard and which may be obsolete in other places, while the younger speakers may still use those forms, but may also show innovative forms. For example, an older working-class woman in Anniston used *clim* as the past participle of *climb*, a form which existed in seventeenth- and eighteenth-century English but which since has become obsolete.

Ethnicity often provides a striking correlate with linguistic variation. Wolfram *et al*. (1999) and Rickford (1985) have shown that African Americans and European Americans living together on isolated islands, of the same socioeconomic background, education and age, show consistent differences in their speech, both in phonology (on the Outer Banks) and in grammar (Sea Islands and Outer Banks).

2.2 *Sample size*

The next question to be resolved is how many speakers are needed. The question depends most directly on the number of independent variables. If you are interested in comparing the speech of working-class men and women of the same age, say, 30 years old, then you have subjects in only two cells: 30-year-old women and 30-year-old men. If you expand the study to include men and women of 60 as well, the number of cells doubles to four. If you expand to include both working-class and middle-class subjects, it doubles again to eight cells. Obviously, each cell must be filled with enough subjects to provide confident generalizations about the social group.

How many subjects should fill each cell? The simple answer is: the more the better. In practice, sociolinguistic analysis requires isolating and classifying dozens and sometimes hundreds of tokens from each subject. It bears little resemblance to the sampling carried out in many kinds of social sciences for the purposes of opinion polls or voter preferences. As a rule of thumb, five persons per cell is often adequate, assuming the cells are well-defined in terms of local social categories (Guy 1980). I followed this rule in my Anniston study, where cells consisted of the independent variables of age/sex/social class/locale (urban/rural); so, for instance, I had to locate and interview at least five older rural working-class male informants.

2.3 Interview protocols and questionnaires

There are two main types of sociolinguistic interviews. The most influential one, modeled on Labov's work, uses a set of questions to elicit as much free conversation as possible, with some reading tasks designed to elicit a range of styles. Another way of going about it is simply to let the conversation flow (Briggs 1986; Hazen 2000). This more open-ended type of interviewing is intended to reduce the distance between interviewer and subject, making the interaction more naturalistic.

For the more structured interview, protocols may be found in appendices of several reports (Labov 1966; Feagin 1979; Labov 1984; Horvath 1985; Wolfram *et al*. 1999, to name a few). The chief goal is to obtain large quantities of speech that is as relaxed and naturalistic as possible; often, too, researchers will design protocols to sample other speech styles as well, for example reading styles considered to be more "formal" than spoken conversational speech. Some researchers, however, have considered the conceptualization of style as a unidimensional "formal–informal" continuum to be problematic (see Schilling, this volume), and so will focus solely on conversation rather than including readings as well.

Sociolinguistic interviews usually begin by asking subjects about themselves – year and place of birth, parents' birthplace, schooling (speaker's and parents'), occupation (their own or that of their parents or spouse). Questions like these often yield a relatively formal or self-conscious speaking style, known as Interview Style, as will discussion of school or the workplace (see Sankoff and Laberge 1978). Such questions invite self-conscious responses by asking the subjects to reflect on their histories and their accomplishments. However, in some circumstances, asking about school activities may elicit informal and spontaneous speech, for example, if directed to subjects deeply and personally involved in those activities. Thus Eckert's teenaged subjects become very animated when talking about activities, groups, and characters in their school, as did mine (Eckert 2000; Feagin 1979). This distinction is crucial in planning the interview protocol, since it is not really topic per se that correlates with degree of self-consciousness but rather extent and type of involvement with the topic area. People tend to be least self-conscious when talking about subjects with which they are intimately involved, while the most self-conscious speech comes from asking people to talk about their credentials.

In the opening section on demographics, asking the subjects to list the houses they have lived in can lead to a discussion of the neighborhood where the speakers grew up, and that can lead to discussing childhood friends and describing rules for various games, jump-rope rhymes, and so on. Here the speaker will probably switch to a less formal, more conversational style. It is difficult to monitor one's speech when recalling and reciting such rhymes as "Fatty, fatty two-by-four, can't get through the bathroom door."

Asking the subjects about their first dates or how they met their spouses sometimes elicits a flood of speech, at least in the European-American context. Labov's best known question has to do with the danger of death: "Have you ever been in a situation where you were in serious danger of being killed, where you thought to yourself, *This is it* . . . What happened?" (Labov 1972a: 113).

While sometimes this elicits an outstanding narrative, it seems to work better in New York City than anywhere else. My speakers in Alabama, asked the same question, generally responded, after a pause, "No." Others have had similar experiences – Trudgill (1974) in Norwich, England, Chambers (1980) in Toronto, and Milroy (1980) in Belfast. In Anniston, after the danger-of-death question proved unsuccessful, I discovered that the question "Have you ever heard of anybody seeing a ghost around here?" often elicited long elaborate narratives of local mayhem and murder from older working-class speakers. Similarly, with his Canadian subjects in Toronto, Chambers discovered he could elicit passionate speech by saying, "People keep saying we're getting more and more American. Do you think that's true?"

The interview, obviously, must be adjusted for local conditions. Familiarity with local customs helps develop questions such as "When did you get your first gun?" in the southern United States, or "What were you doing when that tornado hit back in 1954?" There is no simple formula for eliciting relatively unmonitored, casual styles. The best advice is for researchers to know their regions, especially the tensions in the community, when planning the interview protocol.

2.4 More formal styles: Reading passages, word lists, minimal pairs

The use of written materials in the interview protocol depends on the focus of the research. Presenting subjects with a reading passage, word list, and minimal pair list can certainly be useful for research oriented toward phonology, because the researcher can ensure that the same words, involving particular phonological contrasts or certain variables in particular contexts, are recorded for every subject. In studying syntax, having the speaker read sentences while being recorded can produce valuable results, if they are used to elicit judgments on grammaticality or acceptability. The speakers can be asked who would use such a sentence, even if they themselves would not. If reading is a problem, as it often is for the oldest rural subjects either through poor eyesight or through illiteracy, having subjects repeat sentences read by the interviewer can also be a source of information. Wolfram and Fasold (1974) discuss repetition tasks and some of the information they can yield. In my own work, I started out using word lists and sentences, but dropped them, since I was concentrating on grammar alone. However, judgments on sentences proved to be useful, as ancillary evidence. Now that I am using the same recordings to work on phonology, I am very much aware that it would have been helpful to have kept the word list to observe style shifting, and to get an idea of what might be considered more self-conscious speech.

Word lists and reading passages that have been used successfully may be found in the appendix to Labov (1966), Trudgill (1974), and elsewhere. See also Labov (1984) for a description of various field experiments and references to their use. Each community and each set of variables requires its own materials, but looking at previous models can be helpful.

2.5 *During the interview*

In conducting the sociolinguistic interview, it is important for researchers to give interviewees plenty of space to elaborate on topics of interest and to not waste too much time recording their own voices. Perhaps the most embarrassing moment for novice fieldworkers is the discovery, on listening to interviews they have made, that their own contributions limited what the subject might have offered by interjecting friendly asides or interrupting the flow of the subject's conversation. The resulting interviews sometimes preserve hard evidence of misguided sociability. Nonetheless, as Milroy (1987) notes, it is important to remember that interviews are exchanges, and interviewers do have to make contributions to get quality conversation in return. Keeping the attention and interest of the speaker during the interview is obviously important, and that makes it hard for the researcher to limit back-channelling. It is natural to respond to what the speaker says, to offer your own opinions and to bring up parallel experiences. And whereas sometimes interviewers can get a bit carried away, providing a reasonable amount of co-conversation can be valuable indeed. Breaking my self-imposed silence in a second interview with one of my subjects, comparing notes with the speaker on some experiences we shared, I discovered that the speaker's phonology and grammar altered at that point, with more local vowels – more breaking and shifting – and nonstandard grammar where there had been little or none before.

Thus, while controlling the inclination to take the floor, the interviewer must provide signs of involvement – both verbal and nonverbal (for example, maintaining eye contact, if culturally appropriate) – at the same time keeping a watchful eye on the recording equipment and a dutiful ear on the production of the desired variables.

2.6 *After the interview*

Whether or not to provide monetary compensation to informants is subject to debate (as in Whyte 1984: 361–365). While I have never paid speakers for interviews, others have and do. This may be a community-specific issue. Researchers are often graduate students working on doctoral dissertations – unpaid or poorly paid themselves, so that most of them rely on an exchange of services, such as giving rides, if the researcher has a car, helping with schoolwork, or writing letters, as did Dayton (1996). As Whyte points out (1984) paying speakers can change the nature of the enterprise, even compromising the possibility of further research by making it much too expensive for others following after.

As noted above, another very important matter must be addressed before leaving the speaker: The person interviewed must sign an Informed Consent Form, indicating their understanding of the basic purpose of the research project and recording and their permission for the interview to be used for research purposes. The wording must be approved by the researcher's Institutional Review Board and any other relevant organizations or agencies (for example, appropriate school officials, if recording teenagers in a school).

Finally, as detailed in Wolfram, this volume, most sociolinguists feel strongly that they must give back to their communities of study, partly in exchange for community members' having shared with them their voices, life stories and life experiences, and partly because, no matter what community members have given them, sociolinguistic researchers feel a scientific and ethical obligation to share the linguistic knowledge they have gained through community studies with as wide an audience as possible, including not only academic audiences but also research communities and the general public.

2.7 *Ethics*

Surreptitious recordings, made by planting a recording device where it will capture ambient conversations without the knowledge or consent of the participants, are often illegal and are considered unethical – and pointless – by the vast majority of sociolinguists. In their favor, of course, is the elimination of the Observer's Paradox, but in purely practical terms, apart from ethics, sound quality is usually so poor that it is a waste of time, and discovery by the community can lead to serious repercussions. The legal aspects of surreptitious recordings have been discussed by Larmouth *et al*. (1992), who review state and federal laws of the United States, defend the use of such recordings, and illustrate their points with examples of real or possible situations and their legal outcomes.

Harvey (1992) made covert recordings of drunken speech because it was central to her research, and she states that, while she found it distasteful, she would do it again (1992: 80). She considers surreptitious recordings as no more unethical than researchers not being entirely open about their research agenda with speakers, as in my telling speakers that I was interested in what it was like growing up in Anniston, Alabama, rather than saying outright that I was interested in their grammar.

Most researchers consider that surreptitious recording violates the privacy of the subjects. Even in open recording, it is usually necessary to respect the privacy of subjects by disguising their identities. Some researchers use alpha-numeric codes for speakers, but a better system is to use pseudonyms that preserve clues to ethnic background and other essential traits, so that someone with a German name would be given a German pseudonym, and the same style of naming. Using carefully constructed pseudonyms rather than mysterious codes renders analysts' jobs easier and also results in more readable text. Recorded discussions of illegal activities or private matters should be treated as confidential, regardless of the informant's attitude toward such things at the time.

3 Participant Observation

Because the effect of recording on the interview can never be completely eradicated and because interviews are entirely unsuitable for obtaining certain kinds of data, participant observation has come to be seen as a complementary method of data collection in variationist sociolinguistics. This entails living and participating in the community in some function other than as a linguist, while at the same time observing and noting particular types of linguistic data. Such observations are frequently used to supplement material collected from interviews, as by Labov *et al*. (1968) and Feagin (1979), but they can also be used as the primary source of data, as in Rickford (1975), Mishoe and Montgomery (1994), and Dayton (1996).

Participant observation is especially useful for studying infrequent grammatical items such as questions, modals, and particles, where recorded interviews will not capture these forms. Either the discourse constraints are such that the question/answer format or the extended narrative of the interview do not allow the forms, or the forms are too rare to make an interview worthwhile. For such variables, participant observation becomes necessary. It is crucial to remember that both participation and observation are crucial: The researcher must immerse him- or herself in the community as far as possible while at the same time maintaining some measure of outside, "observer" status.

One of the best discussions of the rationale for using participant observation as well as one of the most complete descriptions of this method as employed in variationist sociolinguistics is found in Dayton (1996: Chapter 2). Here Dayton relates how she, a white woman, became a member of an African-American working-class community in Philadelphia. She first lived in that neighborhood for two years simply as a graduate student, not participating in the life there. Then she lived as a participant observer for four and a half more years, becoming a block chairman, organizing clean-ups, volleyball games, and generally entering into the local African-American life in that block.

The participant observer studying forms not likely to surface in sociolinguistic interviews will write down their data rather than make audio recordings. Dayton managed to write down most of the data for her study within an hour of hearing it. She seldom attempted to store and remember more than three items at a time. Mishoe and Montgomery (1994), who collected their corpus of double modals through participant observation, report that they wrote items down within a minute of hearing them.

This technique has certain advantages over the recorded interview in that the researcher becomes an insider, in so far as possible, and can in this way overcome the Observer's Paradox. In order to do this the researcher must reach the point of understanding the communicative and interactional norms of the speech community and participating in the informal social ties and exchange relationships that hold the community together (Dayton 1996: 71).

In the course of her study, Dayton collected 3,610 tokens of African-American tense/mood/aspect markers (Dayton 1996: 55), probably the largest corpus of these grammatical forms. Her observations also included the more general social context as well as the linguistic context of the use of these markers.

The drawback of participant observation is that researchers cannot write down all the tokens of the variable they might hear. There is an inevitable selectivity in the linguistic record. The selectivity means that the data cannot be quantified, so that it is impossible to provide information on the relative frequency of the variable. In addition, there is no permanent record of the speakers, so that it is not possible to return to the source of the evidence. Here the question of accuracy and reliability naturally arises. Counterbalancing that, it permits the study of rare forms, otherwise undocumentable. And the perceptual saliency of the items can abet the accuracy of the observations. In another context, Wolfram suggests that socially marked items are the most transparent differences, and as such they rank high on a "continuum of linguistic trustworthiness" (Wolfram 1990: 125; similarly Dayton 1996: 68–80).

4 Rapid and Anonymous Observations

While participant observation is a very time-consuming and labor-intensive way to overcome the Observer's Paradox, another, faster technique is "rapid and anonymous observation," first described by Labov (1966, 1972c). By this method, the variable under study is embedded in the answer to a question that can be posed to strangers. Labov, in a famous example, asked sales clerks in department stores, "Where are the women's shoes?" The respondents replied, "Fourth floor." What Labov was interested in was the pronunciation of (r) in the words *fourth* and *floor*. Labov selected a range of stores, from luxury (Saks Fifth Avenue) to bargain basement (Kleins), and was able to confirm that sales clerks tend to speak in a manner that reflects the clientele. The clerks at Saks were r-ful as are upper-middle-class New Yorkers, while those at Kleins were r-less, like working-class New Yorkers. Labov was able to capture 528 tokens of *fourth floor* from 264 subjects in approximately 6.5 hours.

The simplicity of this study has encouraged replications of it in New York and many other places, either studying (r) or other variables. For example, in some communities, the question "Excuse me. Could you tell me what time it is?" (at the right time of day) will produce many tokens of *five* or *four*. This type of study obviously sacrifices knowledge of the background of the speaker in favor of the naturalness of the speech.

5 Life after Fieldwork

Whatever methods the researcher uses, when the fieldwork is finally completed, any sense of relief evaporates rapidly as the reality of analysis of all that data dawns. Analysis, of course, moves the sociolinguist onto an entirely different level, with its own problems and its own rewards (as the following chapters in this volume make clear). The crucial first step, the fieldwork, becomes subordinated to finding, expressing and disseminating the substantive results of the project. Many sociolinguists firmly maintain that the more successful the fieldwork, the less noticeable it is in the final analysis and that fieldwork draws attention to itself mainly when the researcher has to concede that there are gaps in the data, flawed elicitations, or results that require caution in the interpretation. However, as variationists increasingly incorporate ethnographic and social constructionist viewpoints into their work, more attention is being given to how data and analyses are inevitably shaped by research methods, research contexts, and researchers themselves, and so the fieldwork process most likely will not remain quite as backgrounded as it traditionally has been. For the moment, though, the sociolinguist's prowess as fieldworker is often a private source of professional pride that only occasionally seeps into the public domain when sociolinguists gather informally at conferences and meetings. Inconspicuous it may be, fieldwork is the bedrock of the sociolinguistic enterprise, and it is crucial for novice researchers and advanced scholars to understand the methodological underpinnings of even the most theoretically sophisticated analyses.

REFERENCES

Albó, Xavier (1970) Social constraints on Cochabamba Quechua. PhD dissertation, Cornell University.

Ash, Sharon (2000) Sampling strategy for the Telsur/Atlas Project. http://www.ling.upenn.edu/phonoatlas/sampling (accessed February 5, 2013).

Bailey, Charles-James N. (1996) *Essays on Timed-based Linguistic Analysis*. Oxford: Oxford University Press.

Baugh, John (1993) Adapting dialectology: The conduct of community language studies. In Dennis Preston (ed.), *American Dialect Research*. 167–192. Philadelphia: John Benjamins.

Bowern, Claire (2008) *Linguistic Fieldwork: A Practical Guide*. New York/Basingstoke, Hampshire: Palgrave McMillan.

Briggs, Charles L. (1986) *Learning How to Ask: A Sociolinguistic Appraisal of the Role of the Interview in Social Science Research*. Cambridge: Cambridge University Press.

Cieri, Christopher (2011) Making a field recording. In Marianna Di Paolo and Malcah Yaeger-Dror (eds.), *Sociophonetics: A Student's Guide*. 24–35. New York: Routledge.

Chambers, J.K. (1980) Linguistic variation and Chomsky's "homogeneous speech community." In Murray Kinloch and A.B. House (eds.), *Papers from the Fourth Annual Meeting of the Atlantic Provinces Linguistic Association*. 1–32. Fredericton: University of New Brunswick.

Dayton, Elizabeth (1996) Grammatical categories of the verb in African-American Vernacular English. PhD dissertation, University of Pennsylvania.

Di Paolo, Marianna and Yaeger-Dror, Malcah (2011) Field methods: gathering data, creating a corpus, and reporting your work. In Marianna Di Paolo and Malcah Yaeger-Dror (eds.), *Sociophonetics: A Student's Guide*. 7–23. New York: Routledge.

Eckert, Penelope (2000) *Linguistic Variation as Social Practice: The Linguistic Construction of Social Identity in Belten High*. Oxford: Blackwell.

Feagin, Crawford (1979) *Variation and Change in Alabama English: A Sociolinguistic Study of the White Community*. Washington, DC: Georgetown University Press.

Fischer, John L. (1958) Social influences on the choice of a linguistic variant. *Word* 14: 47–56.

Guy, Gregory R. (1980) Variation in the group and in the individual. In William Labov (ed.), *Locating Language in Time and Space*. 1–36. New York: Academic Press.

Guy, Gregory R. (1993) The quantitative analysis of linguistic variation. In Dennis Preston (ed.), *American Dialect Research*. 223–249. Philadelphia: John Benjamins.

Harvey, Penelope (1992) Bilingualism in the Peruvian Andes. In Deborah Cameron, Elizabeth Frazer, Penelope Harvey, M.B.H. Rampton and Kay Richardson (eds.), *Researching Language: Issues of Power and Method*. 65–89. London: Routledge.

Hazen, Kirk (2000) *Identity and Ethnicity in the Rural South: A Sociolinguistic View through Past and Present Be*. Publication of the American Dialect Society 83. Durham, NC: Duke University Press.

Horvath, Barbara (1985) *Variation in Australian English: The Sociolects of Sydney*. Cambridge: Cambridge University Press.

Labov, William (1963) The social motivation of a sound change. *Word* 19: 273–309.

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

Labov, William (1972a) Some principles of linguistic methodology. *Language in Society* 1: 97–120.

Labov, William (1972b) *Language in the Inner City*. Philadelphia: University of Pennsylvania Press.

Labov, William (1972c) *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

Labov, William (1984) Field methods of the project on linguistic change and variation. In John Baugh and Joel Sherzer (eds.), *Language in Use*. 28–53. Englewood Cliffs, NJ: Prentice-Hall.

Labov, William, Ash, Sharon, and Boberg, Charles (2006) *Atlas of North American English: Phonetics, Phonology, and Sound Change*. Berlin: Mouton de Gruyter.

Labov, William, Cohen, Paul, Robins, Clarence, and Lewis, John (1968) A study of the Non-Standard English of Negro and Puerto Rican speakers in New York City. Cooperative Research Project No. 3288. Office of Education, US Department of Health, Education, and Welfare.

Larmouth, Donald W., Murray, Thomas E., and Ross Murray, Carin (1992) *Legal and Ethical Issues in Surreptitious Recording*. Publication of the American Dialect Society, 76. Tuscaloosa: University of Alabama Press.

Lavandera, Beatriz (1978) Where does the sociolinguistic variable stop? *Language in Society* 7: 171–83.

Liebow, Elliot (1967) *Talley's Corner: A Study of Negro Streetcorner Men*. Boston, MA: Little, Brown.

Milroy, James and Milroy, Lesley (1985) Linguistic Change, Social Network and Speaker Innovation. *Journal of Linguistics* 21: 339–384.

Milroy, Lesley (1980) *Language and Social Networks*. Oxford: Basil Blackwell.

Milroy, Lesley (1987) *Observing and Analyzing Natural Language*. Oxford: Blackwell Publishers.

Milroy, Lesley and Gordon, Matthew J. (2003) *Sociolinguistics: Method and Interpretation*. Oxford: Blackwell.

Mishoe, Margaret and Montgomery, Michael (1994) The pragmatics of multiple modal variation in North and South Carolina. *American Speech* 69: 3–29.

Rickford, John (1975) Carrying the new wave into syntax: The case of Black

English BIN. In Ralph W. Fasold and Roger W. Shuy (eds.), *Analyzing Variation in English*. 162–183. Washington, DC: Georgetown University Press.

Rickford, John (1985) Ethnicity as a sociolinguistic boundary. *American Speech* 60: 99–125.

Romaine, Suzanne (1980) A critical overview of the methodology of urban British sociolinguistics. *English World-Wide* 1: 163–198.

Romaine, Suzanne (1981) On the problem of syntactic variation: A reply to Beatriz Lavandera and William Labov. *Working Papers in Sociolinguistics 82*. Austin, TX: Southwest Educational Development Laboratory.

Samarin, William (1967) *Field Linguistics: A Guide to Linguistic Field Work*. New York: Holt, Rinehart and Winston.

Sankoff, David and Laberge, Suzanne (1978) The linguistic market and the statistical explanation of variability. In David Sankoff (ed.), *Linguistic Variation: Models and Methods*. 239–250. New York: Academic Press.

Schilling, Natalie (2013) *Sociolinguistic Fieldwork*. Cambridge: Cambridge University Press.

Schilling-Estes, Natalie (2007) Sociolinguistic fieldwork. In Robert Bayley and Ceil Lucas (eds.), *Sociolinguistic Variation: Theories, Methods, and Applications*. 165–189. Cambridge: Cambridge University Press.

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

Wald, Benji (1973) Variation in the system of tense markers of Mombasa Swahili. PhD dissertation, Columbia University.

Whyte, William F. (1943 [1993]) *Street Corner Society: Structure of an Italian slum*. 4th ed. Appendix A: On the Evolution of Street Corner Society. Chicago: University of Chicago Press.

Whyte, William F. (1984) *Learning from the Field: A Guide from Experience*. Beverly Hills, CA: Sage Publications.

Wolfram, Walt (1990) Review article: Re-examining Vernacular Black English. *Language* 66: 121–133.

Wolfram, Walt (1993) Identifying and interpreting variables. In Dennis R. Preston (ed.), *American Dialect Research*. 193–221. Amsterdam: John Benjamins.

Wolfram, Walt and Fasold, Ralph (1974) *The Study of Social Dialects in American English*. Englewood Cliffs, NJ: Prentice Hall.

Wolfram, Walt, Hazen, Kirk, and Schilling-Estes, Natalie (1999) *Dialect Change and Maintenance on the Outer Banks*. Publication of the American Dialect Society 81. Tuscaloosa: University of Alabama Press.

Wolfson, Nessa (1976) Speech events and natural speech. *Language in Society* 5: 189–209.