Researchers define speech communities in different ways, some of which focus more on internal, subjective perceptions of commonality and some of which focus more on objectively (and externally) observed patterns of commonality (cf. Labov 1972a; Corder 1973; Duranti 1997). We will use 'speech community' as a very general cover term:

> A 'speech community' is any socially meaningful grouping of speakers whose direct and indirect interactions with each other contribute to the maintenance, establishment or contestation of a social order recognizable to the speakers or the researcher.

This definition is useful because it identifies some important issues for sociolinguistic fieldwork. These include:

- Contact between speakers may be direct or indirect (i.e. we need not restrict ourselves to only people who are aware of co-membership).
- Interactions may have very different linguistic outcomes (i.e. we are not concerned with debates over whether community is constituted through consensus or competition)—these differences in outcome are the object of sociolinguistic study (i.e. sociolinguistics attempts not only to document and describe variation in language use, but to relate patterns of language to social dynamics such as the exercise of power or what it means for something to be 'innovative' or 'conservative').
- Language use can be related in an orderly and systematic manner to features of the social setting that the speakers orient to (i.e. the formal linguist's notion of 'free variation' ignores linguistic or social systematicity in the variable use of different linguistic forms).

The focus on patterns of language use often leads sociolinguists to collect their own data, and in the following sections we will discuss in more detail some of the methods used. But it is important to realize that some sociolinguistic questions can be asked and answered using freely available sources of data. The media provides an excellent source of language in use, without the access issues (see below) that sometimes go along with collecting discourse from private domains. The internet has increased enormously sociolinguists' potential datasets in the last decades. Generally, a minimal amount of social information is required about a speaker to enable sociolinguists to explore social correlates of variation (e.g. sex, approximate age, and general social class/occupation). This may not always be available for data on the internet (or other forms of media). However, if research questions don't require too much knowledge about
p. 126 who is producing the data and under ↳ what circumstances, the internet can be a useful tool for exploring some basic descriptive questions about language variation and language use (e.g. Herring 1996; Androutsopoulos 2006). Even some questions about social groups and variation can be explored by targeting subject-specific blogs or user groups. YouTube has recently increased greatly the accessibility of lesser-known languages for armchair sociolinguistic fieldwork (Wrobel forthcoming).

Having outlined some of the key concepts underlying work in sociolinguistics, we turn to more practical matters. In the next section, we look at practical issues associated with getting started. We then discuss the structure of the classic sociolinguistic interview and explain how the observer's paradox can be addressed within this kind of fieldwork methodology. We then discuss the use of group recordings as another means for addressing the observer's paradox, and finally, we discuss methods for enriching sociolinguistic fieldwork that borrow more from the anthropological tradition of participant observation.