

## References

- King, Brian. 2011. Language, sexuality and place: The view from cyberspace. *Gender and Language* 5: 1–30.  
[www.dmlp.org/legal-guide/pennsylvania/pennsylvania-recording-law](http://www.dmlp.org/legal-guide/pennsylvania/pennsylvania-recording-law) (accessed 2 December 2013).

## Further reading

- Johnstone, Barbara. 2000. *Qualitative Methods in Sociolinguistics*. Oxford: Oxford University Press.  
King, Brian. 2011. Language, sexuality and place: The view from cyberspace. *Gender and Language* 5: 1–30.  
Thieberger, Nicholas. 2011. *The Oxford Handbook of Linguistic Fieldwork*. Oxford: Oxford University Press. Especially Chapters 19–21 by Keran Rice, Paul Newman and Monica Macaulay.

## 4 Sampling techniques and gaining access to speakers

Once you know what you want to find out, you have to ask yourself a very important question: will you be able to access the data you need to answer your question? If you plan to collect new data rather than using a corpus, accessing speakers and finding a way to sample the community are crucial.

### Gaining access to speakers and entering the community

Gaining access to participants and entering a community can be quite a challenge. You should always have several backup plans in place and, ideally, test out several ways into the community. Always stay active and keep the ball rolling. Never sit and wait for that one contact who promised to get back to you. Explore other options in the meantime.

#### THE SPEECH COMMUNITY

What exactly a speech community is, is much debated. Some definitions focus on linguistic criteria, others focus on social aspects. If this becomes an issue in your research, consult some of these sources to follow up on the debate: Labov (1972), Buchholz (1999), Patrick (2002), Schilling (2013).

There are several ways to contact speakers in a specific community. If you are a member of that community yourself, you can use your insider status to contact individual speakers: a friend, a family member or a former workmate. If you are not a member of the community in question but you know members in that community, you may be able to contact them.

- 1 A friend or acquaintance. The 'friend of a friend' technique (Milroy and Gordon 2003: 32) has proven to be very useful in data collection. You ask your friend to introduce you to people s/he knows in the community of interest. Ideally, they will in turn introduce you to more people.
- 2 An official community 'broker'. The term 'broker' is often used to describe people with official community status such as priests, teachers and community leaders. Brokers can facilitate community access tremendously as they have very wide networks. In fact, in some cases it may be impossible to do research without these people's approval. Religious leaders or teachers are essential gate-keepers to certain communities. Nevertheless, there can be hazards in working with a broker. Because brokers may hold positions of social respect, they may be associated with higher degrees of formality, and this association may be extended to the researcher. People with official community status such as priests, teachers and community leaders are also more likely to introduce you to people within their immediate social networks and these may be predominantly speakers who use standard speech styles. So be aware of the kind of networks brokers are introducing you to and try to branch out if this doesn't fully match your goals.
- 3 If you do not have contacts in a community, you will have to find a way to access speakers in that community yourself. Find a contact by talking to, writing or e-mailing a member of the community you want to study. Some points of contact are particularly worth exploring as they have well-established networks in place that you can use, for example schools, churches, societies or internet interest groups (the 'brokers' mentioned above). Search the internet to find initial points of contact. If you have a clearly delimited task, e.g. filling in a short survey, contacting complete strangers may be appropriate. Places where people gather to spend their free time or social network sites offer good opportunities to contact people. Alternatively, you can become a member of a community yourself. For example, if you're interested in speech and interaction patterns among members of a sports team or a book club, join a sports team or a book club. Volunteer in the community! Miriam got a lot of exposure to very vernacular speech by working in a kitchen in Berquia, and Nick Wilson's rugby data (Chapters 1–2) was collected in the team Nick joined for his fieldwork.

Whichever way you enter the community, always remember that the way you enter may shape your role there. You should have a simple and honest story about who you are and what you do. Also, try not to get too reliant on and identified with one specific group or network within a community. Explore other groups as well. This will enable you to get a wider picture of the linguistic situation in a speech community. Eckert (2000) recommends a useful strategy that keeps you moving between networks: construct a random sample of community members and return to the list at intervals to make contacts with new groups.

Success when making contacts depends a lot on common-sense considerations. Be sincere, polite and interested and do what you can to prepare for meetings and making new friends. Find out about your participants' interests: local football teams, community events, history and most importantly how they talk about their language (for example, do they consider it a language or a dialect). Find out what you can about local values, norms and community structure (see Chapter 11 on 'emic' categories) so you can conduct your research in an atmosphere of mutual respect.

### **Samples and sampling techniques**

How you gain access to participants is intimately tied up with the type of sample you aim for, i.e. the people you collect data from and on which you base your analysis. Below we will outline several types of samples, and the techniques that are usually used to collect data from such a sample. Your decision about sample size and sampling technique depends very much on your research question and the kind of community you're interested in.

#### ***Random sampling versus proportionate stratified random sampling***

Truly random sampling is hardly ever used in sociolinguistics. It may be an appropriate sampling method when the goal of a study is to provide a description of variation in a city or a town. When selecting your speakers, you will have to use a principled method of selection that gives everyone living there an equal chance of being chosen as a participant in your study. Such a selection method is normally based on some form of existing list of the relevant population (for example telephone or electoral registers). Participants are then selected randomly from those lists. Finding truly representative lists and ensuring selected participants take part in the study is an art in itself. There is also no guarantee that you will tap into speakers in all categories you're interested in. This is particularly the case if segments of the population are unevenly represented, e.g. young people on electoral rolls.

This is where a proportionate stratified random sample has an advantage. Such a sample still uses random selection techniques but the population is divided into strata, i.e. groups of individuals that may be important to the study. A random sample is then collected from each stratum and combined with the samples from other strata to form the full sample. Tagliamonte (2006: 23) proposes that as a minimum requirement a sample should be representative on the basis of age, sex, social class and/or educational level, since these variables prove over and over to be significant in studies of urban speech communities (see Chapter 11 on 'etic' categories). This may be too ambitious for a student project but it's something to be aware of.

### RAPID AND ANONYMOUS SURVEYS

These are brief, short-contact anonymous surveys that usually feature randomised speaker selection. Speakers are normally unaware that their speech is the focus of study, and we find out relatively little about the speakers we're polling (except for characteristics such as sex and age that – by and large – can be detected visually). Surveys are conducted in public places such as department stores, streets, coffee shops, etc. Eliciting data in this manner can be done through observation, that is, the researcher places him-/herself in a location where observation of language can take place, e.g. at a service counter or a table near the cash register in a café. Alternatively, the researcher can prompt speech. Gardner-Chloros' (1991) study of codeswitching in a Strassburg department store is an example of the passive technique, while Labov's classic study of language use in three socially stratified department stores in New York City (Labov 1972: 43–69) is an example of prompted speech. Labov randomly selected employees and asked questions that elicited the phrase *fourth floor* twice, which gave four tokens of the variable (V). Usually, audio recording does not take place, partly to avoid ethical issues. This technique requires the researcher to be quick and do several things at the same time: assess the social characteristics of the speaker; make an 'on-the-spot' analysis of the speech sample; note down what was heard. This works well if the variables selected for study are highly targeted and salient, e.g. your prompt always elicits the word *street* and you are interested in whether people say [stɹi:t] or [stri:t] (see Labov 1984).

### Judgement sampling

Judgement sampling is frequently used in sociolinguistics. First, you identify the kinds of speakers you're interested in (e.g. male and female adolescents at a mid-tier-class high school). Then, you collect data from a number of speakers in each group: males, females; possibly also different age groups and ethnicities, if applicable. You should try to capture the basic demographics of whatever population you decide to investigate. This may mean investigating locally important groupings (e.g. nerds, jocks, townies, etc.) in addition to or instead of macro-social categories, such as speaker sex and social class. You'll need to conduct background research and observe for a while before making decisions about categorisations. Also find out about people's attitudes and sense of identity to help you in making decisions about community groupings. Making decisions about social categories is not as straightforward as one may think – Schilling (2013: 46–54) problematises

some social categories. As we argue above, it is a good idea to create a data grid (see Chapter 2) and fill the cells, either (a) randomly, (b) by your own judgement or (c) by advancing from speaker to speaker through a social network. Keep in mind, though, that from a statistical point of view a judgement sample is not truly representative as it was not selected randomly. However, linguistic behaviour is much more uniform than other social behaviour (e.g. people's fitness routines, their holiday preferences, etc.), so even a judgement sample allows a certain degree of generalisability (see Schilling 2013: 33).

### Social networks and communities of practice

Our assumption so far has been that you want to conduct a study of a speech community. However, there are other frameworks for viewing multiple speakers: social networks and communities of practice (see Dodsworth 2013 for a thorough comparison of these frameworks). They differ from the speech community framework in various respects, for example in how many speakers are considered, the methods used to study them and the assumptions researchers make about social interaction and language. In the speech community framework, composite data is usually investigated and compared across demographic categories. Thinking about sampling is therefore of utmost importance. It's a good approach to use when you are interested in the global distribution of a variable across several external categories within a city or a town.

However, if your focus is highly specific or if you want to find out more about how language is socially embedded within a community, work informed by the traditions of ethnography (see Chapter 5) may be more suited to your research question. Few sociolinguists devote the time for a genuine ethnography, but the practice of immersing yourself in a community for some period of time has been used for studying networks and communities of practice. We will introduce these here briefly. If you work in one of these traditions, it is almost inevitable that you will collect naturally occurring, spontaneously produced speech (see Chapter 6).

The network approach focuses on some pre-existing social network, rather than abstract social categories such as social class or age. Social network analysis finds out who people interact with and how these patterns of interaction relate to language; it is particularly good at shedding light on how features spread or don't spread from person to person and network to network. Researchers attach themselves to such a network and often move from contact to contact using the 'friend of a friend' technique. As you get to know more people in a network, you may become a participant observer within this network of friends and acquaintances (so the work shifts subtly into a more ethnographic mode). What makes such a study a social network study, rather than just a convenient way to collect data, is that some aspect of the structure of the network will also be investigated. Often, an attempt is made to quantify, or at least visualise, this structure. Here are some examples.

Mackeigan and Muth (2006a, b) conducted a social network analysis of Tzotzil Mayan colour terms. They investigated a social network of weavers in Chiapas, Mexico. Some of the weavers collect thread from a central location and then go back into their communities and pass on the thread to other people. Mackeigan and Muth have shown how certain forms of colour terms seem to be transmitted within particular groups within this network. The details of this shall not concern us here, but we'd like to point out the complexity of the network, shown in Figure 4.1.

When attempting to correlate linguistic data with network structure, some measure will have to be found that allows researchers to do this. If the number of people is limited and the linguistic data is not too complex, visual representation may suffice (see Mackeigan and Muth 2006b: 33). However, if a statistical analysis is to be conducted, you can assign a social network score to every individual. Such a score is often a reflection of how dense and multiplex a person's social network lies are. For example, Milroy assigned participants network scores from zero to five and included these in her statistical analyses (Milroy 1987: 141–142). Such a score may be based on participant observation, interview data or a survey.

A social network is still a structure that is, to a certain extent, objective rather than subjective. Sociolinguists are very much interested in how individuals see the world from an even more local perspective. We have, during the last two decades, made increasing use of a construct that allows exactly this: the community of practice (Eckert 1989; Wenger 1998).

Have another look at Figure 4.1: individuals 1, 2, 3, 6, 7, 8, 9, 10 and 13 seem to have dense network connections; so do individuals 27, 28, 50, 60 and 61. To find out whether they are a community of practice we would have to show that there is mutual engagement between all community members, that they're working towards a jointly negotiated enterprise and that there is a shared repertoire. These are the three essential criteria by which communities of practice are defined. Members in a community of practice all interact with each other and they do this by orienting to shared norms and attitudes. For example, a group of close friends may be a community of practice but people who all happen to come together in a carriage on a commuter train are probably not.

Two more aspects are of particular importance in the community of practice approach. First, the community of practice is a useful concept to bridge global and local concerns and to do quantitative as well as qualitative work, i.e. qualitative sociolinguistics or discourse analysis (Holmes and Meyerhoff 1999). You may describe the community of practice and simultaneously locate it in the wider social context which gives it meaning and distinctiveness. This may mean considering larger social categories, such as social class and gender, or working with categories that make sense to informants and are not imposed by the researcher.

Second, the community of practice approach focuses on linguistic (and other) practice, repertoires and values in a small community of speakers, on the social meanings of variables and on how they are used to build up speaker identity.

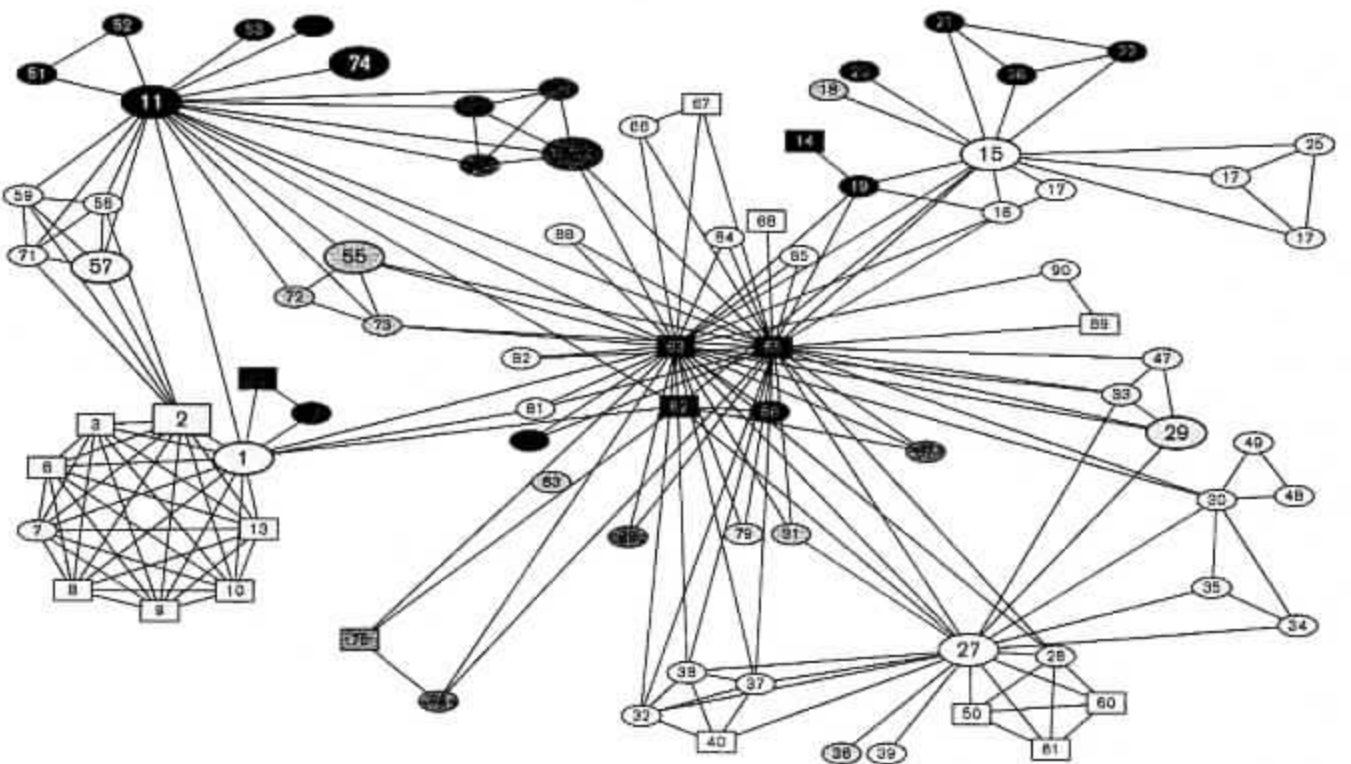


Figure 4.1 Thread passing network

Source: Mackeigan and Muth (2006a), unpublished figure reproduced with permission



Identity, in this framework, is not fixed as it is in much quantitative, variationist sociolinguistics, which works with categories such as adults, white people, middle class speakers, and, to an extent, presupposes that these categories are unified and essential. The community of practice approach challenges this categorical and homogeneous view of identity. For example, Mendoza-Denton (2008) conducts an analysis based on ethnographic research among two Latina girl gangs. Her work has resulted in a very deep understanding of the lives and oppositional social categories of the two groups. She was able to link and explain the meanings behind choices in domains as diverse as clothing colours, Spanish and English use, art, makeup, hairstyles – and that's precisely what we mean by practice, repertoires and values. Language is just one aspect of it.

In conclusion, if your interest is in a larger speech community such as a town, random sampling, stratified random sampling or a judgement sample may be best suited to your needs. If, on the other hand, your interest is in a small group of specific speakers, such as a book club or a friendship group, and you aim to uncover how variation and other practice is locally meaningful and/or how this practice is used to construct identities, immersive ethnographic methods of data collection focusing on a community of practice may be better suited. You may, of course, also decide to combine sampling techniques.

## EXERCISES

### Exercise 1

What sampling technique would you use for the general representation of an urban speech community? What sampling technique would you use if you were particularly interested in certain members of a speech community, for example the Chinese community in Auckland, New Zealand? How would your sampling techniques differ if you'd noticed that many Chinese immigrants work in retail, often running their own coffee shop or bistro?

### Exercise 2

The community of practice is a useful concept to bridge global and local concerns and to do quantitative as well as qualitative work. How could one do this in practice? Use as an example a relatively well-studied city such as London or Philadelphia and start reading articles that give you a good quantitative base and some useful ideas for further study at the local level, e.g. Cheshire, Fox, Kerswill and Torgersen (2008) for London and Labov, Rosenfelder and Fruehwald (2013) for Philadelphia. Devise a plan to study language use in a community of practice. What would you want to study? How would you go about accessing the community? And why could it be important to do such a study?

### Exercise 3

Consider the beginning of the methods section of an article by Schlee, Meyerhoff and Clark (2011), which investigates the acquisition of variation by Polish immigrant adolescents in London, England and Edinburgh, Scotland.

#### 3.1 Data collection

Our study was conducted in two high schools, one in Edinburgh and one in London, where recent immigration has led to an increase in the number of non-locally born students. We interviewed both Polish migrants and teenagers from local British families so as to have a benchmark of the local norms that the teenage migrants were exposed to most frequently. Students volunteered for the study following a presentation from the research assistant about the general nature of the tasks; they were interviewed in friendship pairs in order to facilitate the most casual atmosphere possible given the school-based setting for the interviews (Milroy and Gordon 2003: 66).

The Edinburgh sample consisted of 16 Polish migrants (8 males, 8 females) and 21 Edinburgh-born teenagers. The London sample consisted of 21 Polish migrants (8 males, 13 females) and 24 London-born teenagers. The Polish teenagers were all aged between 12 and 18 with a mean age of 14 in both the London and Edinburgh samples. The length of time that each adolescent had spent in the UK varied from seven months to five years, with an average in both cities of two and a half years in the UK. A locally born female research assistant carried out sociolinguistic interviews in both Edinburgh and London respectively. Recordings were made using the M-AUDIO Microtrack II 2-channel mobile digital recorder and SHURE headset microphones. The interviews were transcribed orthographically using ELAN (<<http://www.lit-mpievu/tools/elan/>>), resulting in a time-aligned corpus of around 200 000 words.

Source: Schlee, Meyerhoff and Clark (2011)

How was the community (most likely) accessed? What sampling technique was used? Why? What limitations might the researchers have encountered when attempting to collect a large sample of Polish adolescents? Why were English/Scottish students included as well?

## References

- Buchholz, Mary. 1999. Why be normal? Language and identity practice in a community of need girls. *Language in Society* 28: 203–223.
- Cheshire, Jenny, Sue Fox, Paul Kerswill and Eivind Torgersen. 2008. Ethnicity, friendship network and social practices as the motor of dialect change: Linguistic innovation in London. *Sociolinguistics* 22: 1–23.
- Dodsworth, Robin. 2013. Speech communities, social networks, and communities of practice. In Janet Holmes and Kirk Hazen (eds) *Research Methods in Sociolinguistics*. Oxford: Wiley Blackwell, 262–275.
- Eckert, Penelope. 1989. *Jocks and Burnouts: Social Categories and Identity in the High School*. New York: Teachers College Press.
- Eckert, Penelope. 2000. *Linguistic Variation as Social Practice*. Malden, MA: Blackwell.
- Gardner-Chironis, Penelope. 1991. *Language Selection and Switching in Strasbourg*. Oxford: Clarendon Press.
- Holmes, Janet and Miriam Meyerhoff. 1999. The community of practice: Theories and methodologies in language and gender research. *Language in Society* 28: 173–183.
- Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia, PA: University of Pennsylvania Press.
- Labov, William. 1984. Research methods of the project on linguistic change and variation. In John Baugh and Joel Sherzer (eds) *Language in Use: Readings in Sociolinguistics*. Englewood Cliffs, NJ: Prentice Hall, 28–53.
- Labov, William, Ingrid Rosenfelder and Josef Fisielwald. 2013. One hundred years of sound change in Philadelphia: Linear incrementation, reversal, and reanalysis. *Language* 89: 30–65.
- Mackeligan, Terri and Stephen Q. Muir. 2005a. The structures of conceptual change. XXVI *International Sunbelt Social Network Conference*, Vancouver BC, April 2006.
- Mackeligan, Terri and Stephen Q. Muir. 2006b. A grammatical network of Tzotzil Mayan colour terms. In Carolee P. Biggam and Christian J. Key (eds) *Progress in Colour Studies: Vol. 1 Language and Culture*. Amsterdam: John Benjamins, 25–36.
- Mendoza-Denton, Norma. 2008. *Homegirls: Language and Cultural Practice among Latina Youth Gangs*. Malden, MA: Blackwell.
- Milroy, Lesley. 1987. *Language and Social Networks*. 2nd edition. Oxford: Blackwell.
- Milroy, Lesley and Matthew Gordon. 2003. *Sociolinguistics: Method and Interpretation*. Malden and Oxford: Blackwell.
- Patrick, Peter. 2002. The speech community. In JK. Chambers, Peter Trudgill and Natalie Schilling-Estes (eds) *The Handbook of Language Variation and Change*. Malden, MA: Blackwell, 573–597.
- Schilling, Natalie. 2013. *Sociolinguistic Fieldwork*. Cambridge: Cambridge University Press.
- Schieff, Erik, Miriam Meyerhoff and Lynn Clark. 2011. Teenagers' acquisition of variation: A comparison of locally-born and migrant teens' realisation of English (ing) in Edinburgh and London. *English World-Wide* 32: 206–236.
- Tagliamonte, Sali A. 2006. *Analysing Sociolinguistic Variation*. Cambridge: Cambridge University Press.
- Wenger, Etienne. 1998. *Communities of Practice*. Cambridge and New York: Cambridge University Press.

## Further reading

- Dodsworth, Robin. 2013. Speech communities, social networks, and communities of practice. In Janet Holmes and Kirk Hazen (eds) *Research Methods in Sociolinguistics*. Oxford: Wiley Blackwell, 262–275.
- Eckert, Penelope. 1989. *Jocks and Burnouts: Social Categories and Identity in the High School*. New York: Teachers College Press.
- Holmes, Janet and Miriam Meyerhoff. 1999. The community of practice: Theories and methodologies in language and gender research. *Language in Society* 28: 173–183.
- Milroy, Lesley and Matthew Gordon. 2003. *Sociolinguistics: Method and Interpretation*. Malden and Oxford: Blackwell.
- Schilling, Natalie. 2013. *Sociolinguistic Fieldwork*. Cambridge: Cambridge University Press.