An up-to-date metadata catalogue is a key component of your documentary corpus. Even if you have chosen a file-naming convention that captures a lot of information, you still need to keep a digital catalogue. Initially you are likely to be creating metadata in your paper notebooks, jotting down the name of a speaker or the object being photographed, the date, the location, and so on, and then keying this information into your catalogue at the earliest opportunity. Ask your archive what metadata they require for deposit, and you can also add other information that is important to your particular situation.

4.3.3.1 Relational databases

A relational database is a good choice for a catalogue, and in this section we discuss how one can be conceptualized (but we refer readers to other sources for the details of using any particular database management system (DBMS), like Filemaker Pro, MS Access, or OpenOffice.org Base). If you do not know how to build a relational database, you can still keep well-organized metadata in a spreadsheet, but you p. 106 should be aware that a spreadsheet has some limitations. For instance, you will ↳ need to enter identical information multiple times, and there is no easy mechanism for enforcing consistency, whereas a DBMS can constrain entries to a fixed list, using dropdown menus to assist in data entry. For these reasons you may wish to make the leap from a spreadsheet to a relational database when you have time to learn how to use one (Harrington 2009 is an excellent introductory guide), or adapt an existing relational database to suit your needs. As with all digital files, periodically backing up your catalogue to a readable text format is good insurance against losing your catalogue in the event that someday the DBMS software is no longer supported.

Relational databases provide ways of linking related but conceptually different kinds of data, such as information regarding recordings, transcripts, and people that are part of your fieldwork project, as shown in Fig. 4.5. In your database, each *record* relates to an *item* in your collection. An item can be whatever you select it to be, perhaps a recording session that lasted two hours of which there are photos and video and audio, all of which can be summarized in one database record. On the other hand, you may want to list each recording as an item, each with a unique name and with a catalogue record devoted to it. This would allow you to find information at a level lower than the session. Sample catalogue files in various formats are available from archiving projects (see e.g. Arbil,[17] IMDI,[18] PARADISEC[19]) to help you decide. We hope that in the near future we will have access to more user-friendly metadata entry tools such as Saymore[20] and Fieldhelper,[21] which promise to use drag-and-drop functions for metadata entry.

Everyone has different ways of working, but we all need to keep track of some basic kinds of information, such as information about media recordings, about people, about transcripts and texts, and about lexica. A DBMS stores similar kinds of information together in *tables*, and then establishes *relationships* between tables so that you should (in theory) never need to enter the same piece of information twice, thus saving time and eliminating chances of typos. The links between tables in a DBMS work by using *keys*, a unique identifier for each record in the database.

To illustrate this, consider Fig. 4.5. A metadata database will have a table that keeps track of your **recordings**. You will want to store a number of different pieces of information about each one, e.g. the file name, the date recorded, the equipment used, the location of the recording, a summary description of the contents, and the length. At the same time, the DBMS can also establish a unique identifier, or key, for each recording (here, Recording_ID).

To understand how relations between tables work, imagine you also want to keep track of your **transcripts**, p. 107 which are related to, but separate from, the recordings ↳ they transcribe. You can keep information about your transcripts, like the start date and the status (in progress, finished, etc.), in another table. Notice in Fig. 4.6 that the transcripts table not only contains the unique identifier, or key, for each transcript, but also