The Linguist's Shoebox: Application Guide for Anthropology

David Coward
Barbara Moore
John Wimbish

Macintosh is a registered trademark of Apple Computer, Inc.

Microsoft and MS-DOS are registered trademarks, and Windows is a trademark of Microsoft Corporation.

Contents

Acknowledgments

The framework for a database of anthropological field notes is based on the work of Barbara Dix Grimes and has been incorporated with her permission.[1] John Wimbish wrote the original version of this chapter on anthropological analysis for the MS-DOS version of the Shoebox computer program.[2] David Coward updated it to make it more compatible with the Windows and Macintosh versions of Shoebox. Barbara Moore reviewed, edited, and contributed substantial material to help enhance this application guide for anthropology. Mark Pedrotti edited the authors' manuscript.

---

[1] Grimes, Barbara D. Ethnography notebook: a guide to managing cultural data on computer. Manuscript.

———. 1986. Wanted: computer programs for anthropology. *Notes on Translation* 115:7–19. Dallas: SIL International.

[2] Wimbish, John S. 1989. *Shoebox: a data management program for the field linguist*. Waxhaw, NC: SIL International.

Anthropology Databases

Overview

Goals for anthropological data collection typically include:

- learning, understanding, and reproducing appropriate patterns of behavior for day-to-day interaction in the local culture

- mapping out the cultural meanings surrounding certain words in order to better use the language and to do a better job in translation

- gathering a collection of texts that illustrate various aspects of the culture

- publishing an ethnography or ethnographic articles

Collecting appropriate anthropological data will provide the raw material needed for a good analysis of the local social structure, worldview, taboos, and many other topics that need to be understood. The information and analysis can help in making decisions about language programs or programs of change (such as community development).

There are many types of cultural material that you might obtain in a field situation. Some of the more common types are:

- daily log of observations and impressions

- kinship charts

- genealogies

- case studies and life histories

- interviews

- hortatory, procedural, and other types of texts

- mapping of the semantic range of individual lexical items[3]

This chapter focuses on a data management strategy for entering, cataloging, and analyzing cultural observations, impressions, interviews, any recorded case studies or life histories you have gathered, and any other type of text data that you might normally put in an anthropology notes journal. Other types of data will not be addressed in this chapter.[4]

Shoebox   The Linguist's Shoebox is a computer program that helps field researchers integrate various kinds of text data: lexical, cultural, grammatical, etc. It has flexible options for selecting, sorting, and displaying data. The name *Shoebox* recalls the use of shoe boxes to hold note cards on which the definitions of words were written in the days before researchers could use computers in the field.

For most linguists and anthropologists, managing data on the computer is time-consuming. They collect thousands of data items when learning a language and culture. Shoebox goes with researchers through all the stages of their field work. Because Shoebox integrates various kinds of data and makes it quickly available, field workers can spend less time on the computer and invest more time with the people interacting and learning.

To gain competence in using Shoebox, run its self-paced computer-based training, use Shoebox Help, and read *The Linguist's Shoebox: Tutorial and User's Guide*.

When you install Shoebox, there are four files that you can use for your anthropology data. They are located in the Anthro folder where Shoebox is installed, e.g., Program Files\Shoebox\Anthro.

- AnthNote.db—anthropology notes database. It contains one record (the first example in this chapter). After you have added a few entries, you can delete the sample record.

---

[3] Mapping lexical semantic ranges is discussed in *Making Dictionaries: A guide to lexicography and the Multi-Dictionary Formatter* (Coward and Grimes 1995).

[4] Shoebox does not provide a method for handling kinship charts and genealogies (except that you can type them in verbatim).

- AnthCat.db—anthropology categories (analysis) database. It contains a few empty records that correspond to topics in the Outline of Cultural Materials (OCM).

Copy the database files to an appropriate *data* folder. The corresponding database type files—AnthNote.typ and AnthCat.typ—define all the data field markers listed in this chapter. Copy them to the *user settings* folder (default: My Shoebox Settings).

Anthropology notes database

The *anthropology notes* database is where you enter the data that you collect. Here are the recommended data field markers:

\date *The date the data was actually observed or elicited*. This is normally the record marker. To allow sorting, enter the dates as YYYY-MM-DD (e.g., **1998-03-23**). Sorting by date keeps the events or observations in chronological order (like a diary or journal). If several notes are written for the same day, the date should be followed by the letters *a*, *b*, *c*, etc. (e.g., **1998-03-23b**). If a day has more than 26 entries (not common), use *za*, *zb*, *zc*, etc.

\de *Date entered*. It is not always possible to enter data the same day we collect it. Use this field to indicate the date the data is *actually entered* on the computer (or when it was written down in a data notebook). This distinction is important because the longer we wait to transcribe our notes or write down an incident we participated in or heard about, the more unreliable the facts become. Later on, we might find that we could possibly explain a discrepancy in our data by the fact that it was not written up until several days (or even weeks) after the event was observed. Enter the date using the same format as the **\date** field or enter the time difference, e.g., **d02** (two days later), **d15**, **w03** (w = weeks), **y03** (y = years).

\wthr *The day's weather*. You might be able to draw conclusions concerning the weather patterns of the area and its relation to behavior and cultural events, if you faithfully enter the data over an extended time period. If several entries are written for a given day, you only need to fill in this field in the first entry.

\rscr *Researcher*. If more than one person is involved in data collection, the team should use this field to specify which

member made the entry. This allows data from multiple researchers to be stored in one database, and yet to be viewed separately, if needed. Generally, initials are adequate. Set up a Shoebox range set for this field to ensure consistency.

\type *Type of data*. Use this field to specify whether the data for this entry comes from an observation, reported information, an impression, an informal conversation you had or overheard, a more formal interview, a book or article, or some other data type. It can help you evaluate the validity of your data. Abbreviations can be used. Set up a range set to ensure consistency.

\loc *Location*. Where were you when you collected the data? In your home, office, courtyard, a house or garden, the town square? Location can affect the type of information you are getting. Depending on how you enter the locations, you might be able to set up a range set.

\srce *Source*. This field contains the name of the person(s) from whom the data for this entry is obtained. If the data comes from observation, the source is yourself (or spouse or colleague). If the data comes from a book or article, cite the author. For conversations or interviews, the person who gave the information is the source. Ideally, you should maintain a biographical sketch of each of these persons, including age, sex, marital status, kinship group, place of birth and childhood, social standing, etc. You might want to enter the information in a biographical database. You could use **\srce** as the record marker and set up a Shoebox data link.

\part *Participants in the data gathering*. Who else was present during a conversation or interview? Enter the names of those who were present, were involved in the conversation, or listened to the interview in this field. People often adjust their information to fit their audience. Certain information might not be shared (or might be shared differently) when certain people are present. Therefore, kinship data can be relevant (e.g., certain topics are not discussed before certain relatives). For comments on biographical data, see the description of the **\srce** field.

\data *The actual data in a fully expanded form*. The data can be a single sentence or a description that takes several pages. Be as complete as memory, notes, and time allow. What did you see,

hear, understand, learn, etc.? Often you will only take brief notes at the time, whether at a ceremony or during an interview. Or you might only have time to scribble some notes down after the event or when the person giving the data has left. The full description needs to be written down as soon as possible while your memory is fresh (see the description of the **\de** field). Divide long descriptions into multiple entries, so that you can use the **\anth** fields to retrieve only the relevant "chunk" of data, not the entire description.

\anth *Anthropology category*. The categories (or key words) that you select to catalog the data in the entry. Many researchers use the *Outline of Cultural Materials* (OCM) codes—often referred to as the HRAF codes—for this field. The OCM is a compilation of over 625 categories for organizing cultural data that is very useful for guiding research and avoiding "holes" in your data. OCM is recommended—certainly worth learning how to use—but each of us will probably find the need to add certain categories of our own. Shoebox can easily handle this. Set up a data link from this field to the **\anth** field in the analysis database. This will ensure consistency when entering data and let you jump to the topic record when compiling the analysis.

\mtrl *Related materials not on the computer*. Use this field to catalog non-text information, such as the location of photographs, drawings, audio and video recordings, artifacts, published and unpublished material on the current topic.

\hypo *Hypothesis by the researcher*. The data might suggest certain hypotheses that need to be checked. For example, the data might suggest that "Property is owned by men, but domestic animals are owned by women" (true among some groups in Brazil); or "Only men can have personal fetishes" (true for some groups in Africa). Once a hypothesis is stated, you can be on the lookout for other data that either supports or refutes the hypothesis.

\q *Questions for further investigation*. Use this field to list data to be collected later. As you write up the data, you might find that there is certain other information that you need. For example, if the entry is discussing the market, possible questions might be, "What do the men sell? What do women sell? Do outsiders sell anything not sold by local people?"

\cf   *Confer, compare*. Use this field to cross-reference other entries that you consider *closely* related. Enter the date of the other entry. Set up a data link from this field to the **\date** field. This will ensure consistency when entering data and let you jump to related data when compiling the analysis.

\nt   *Notes*. Use this field for any notes or comments pertaining to the entry. It can be used to qualify the data.

\dt   *Date last edited*. This field contains the date the journal entry was last edited. Select it as the date stamp, so that Shoebox will update it automatically.

The following Shoebox record is an example of a typical entry in an anthropology notes database. It is an abbreviated entry about the researcher's first trip to a village market in West Africa.

\date	**1997-02-26a**
\de	1997-02-28
\wthr	*sunny, cloudless, 35°C*
\rscr	bjm
\type	**observation**
\loc	marketplace
\loc	Tehini
\srce	self
\part	Lisunami, Élizabet
\data	This morning I went to the market with Lisunami and Élizabet. The marketplace is situated… There were few actual stalls; most people had their goods spread on the ground… It seemed like most of the vendors were women, but certain objects were sold only by men: bicycle parts, tobacco… I saw so many different things for sale: tomatoes, onions, herbs and spices, chickens (live), fried breads and other cooked foods…, clothing, cloth, cosmetic items… Some of the vendors do not look like they are local. They have both different physical features and dress differently and sometimes are selling things no one else has; i.e., one man in a turban with heavy eye-makeup was selling medicines, potions, and charms. I began to ask what things cost, getting names for things, buying a few things.
\anth	**443 Marketing**
\anth	**252 Food preparation**
\anth	**263 Condiments**
\anth	**…**
\mtrl	diagram of market (in materials file #1), photos of vegetables I bought (photo file #1) …
\hypo	There are certain items that only men sell, only women sell, only outsiders sell. Most of the buyers and sellers are from the town or surrounding villages.
\q	What do women sell? Men sell? Where do the outsiders come from? What outlying villages do buyers and sellers come from?
\cf	**1996-10-05**
\nt	
\dt	28/Feb/1997

Comments about the data fields

Some important things to note concerning the data fields:

- The locations are entered in two separate **\loc** fields. This makes it possible for Shoebox to index both **marketplace** and **Tehini** (if you sort the entries by location).

  The same is true of the **\anth** fields. Traditionally, this cataloging information has been strung together in one field:

\anth	**443 Marketing, 252 Food preparation, 263 Condiments**
\anth	**<443><252><263>**

  Entering the anthropology codes like this only allows Shoebox to index the first one. It is *much* better to list each code in a separate **\anth** field.

- The participants are entered in a single **\part** field. It not as critical to enter each participant in a separate field. If you need to focus on a particular participant, you can use a filter or the Find command.

- Notice that the **\data** field in this example has several '…' in it. This is simply to shorten the example for this chapter. It should be restated that if a data entry becomes too long, it becomes unwieldy and difficult to analyze. The number of relevant topics becomes unmanageable and finding the sought-for comment or observation in a large data field requires excessive reading.

  In such cases, it is far better to divide a long, detailed observation or interview into smaller "chunks" (these can even be arbitrary breaks at first). The chunks of data can be categorized more precisely (because you are dealing with a smaller scope). This makes the codes more useful when the time comes to compare related topics.

  In Shoebox, add the chunks as separate records:

\date	**1997-02-26a**
\date	**1997-02-26b**
\date	**1997-02-26c**

- Notice that the *date entered* is two days after the date of the observation. This indicates that the observation was written while fairly fresh. However, entering the date as **1997-02-28** requires the user to compare the **\date** and **\de** fields. As mentioned in the description of the **\de** field, another approach would be to enter the *time difference* rather than the date itself. For example, **d02** (two days later), **d15**, **w03** (w = weeks), **y03** (y = years), etc. Sorting by the **\de** field would show the most reliable data at the top and the least reliable at the bottom.[5]

- The **\cf** field refers to another entry (by its date, e.g., **1996-10-05**). This indicates it has closely-related data and it is worth reading along with this entry.

- Several of the marker descriptions refer to data consistency. For more information about data properties, range sets, and data links, access Shoebox Help for the Consistency Check command on the Tools menu. Data links also let you jump easily to related information. For more information about jumping, access Shoebox Help for the Jump Path tab on the Database Type Properties dialog box (from the Database menu, choose Properties).

Defining a database template

When you add new entries, Shoebox can automatically include the data fields that you normally use. Whenever you add a new entry to a database, Shoebox adds the fields in its database template. This saves typing and ensures consistent content in your records. To define a database template:

1. From the Database menu, choose Insert Record.

2. Add the field markers that you want to use from the preceding list (and any user-defined field markers).

---

[5] Using **m** for months would cause incorrect sorting relative to the other time units. You could use weeks instead (or possibly use **x** for months).

3. Fill in the contents of any field that you want inserted with default values (this is for fields that don't change very often, e.g., **\rscr bjm**, **\loc Tehini**).

4. Add any blank lines you want between fields.

5. From the Database menu, choose Template.

6. Click to select the Use field contents check box.

7. Click OK.

Cataloging the data

Well-organized categories are an essential part of data management. For cultural anthropology, several systems are available. One of the more popular is the Human Relations Area Files publication *Outline of Cultural Materials* (OCM).[6] It provides numbered categories arranged from broad general topics to specific subtopics.

Two keys to retrieving the relevant data when compiling the analysis are:

• cataloging (filing) the entries by pertinent categories

• cross-referencing closely related entries

Outline of Cultural Materials (OCM)

Here is a section of the outline:

76	DEATH
761	Life and Death
762	Suicide
763	Dying
764	Funeral
765	Mourning
766	Deviant Mortuary Practices
767	Mortuary Specialists
768	Social Readjustments to Death
769	Cult of the Dead

---

[6] Murdock, George, et al. 1987. *Outline of cultural materials*. 5th edition with revisions. New Haven, Conn.: Human Relations Area Files, Inc. (The OCM manual is available on the LinguaLinks Library CD-ROM in Folio View format.)

For each topic, the OCM manual provides an in-depth description and a listing of related topics. Here is the topic on *mourning*:

765   MOURNING—duration of the mourning period; behavior of spouse and other relatives after the funeral (e.g., seclusion, mourning garb, observance of taboos, sacrifices); treatment of relics (e.g., preservation of skull or head, wearing of bones, use of hair for artifacts); visits to the grave; mortuary feasts; exhumation, second funeral, and reburial; ceremonies terminating mourning, etc.

        See also:

Headhunting	721
Mortuary cannibalism	266
Status of widows	589

The OCM system is an excellent scheme for organizing data. Its descriptions of each topic can be a useful guide to your research.

As demonstrated in the example records on pages 7 and 12, you can enter the appropriate OCM topic codes as the categories in **\anth** fields to catalog your anthropology notes.

\date	**1992-07-10**
\de	d04
\wthr	*rain clouds, humid*
\rscr	jsw
\type	**observation**
\loc	Guarayu
\srce	self
\part	
\data	I became aware that someone had died as the brass band marched by the house, followed by the body, which was being carried on a wooden frame. The body was put into the hammock in which he had spent his nights sleeping… At the grave the body was gently lowered… following the service each of the bystanders reached down and grabbed a fistful of dirt from the pile and dropped it on the body. Perhaps this was done in order to share in helping to send the fellow on his way. Most of the crowd left, but a few stayed behind to complete the task of filling in the hole.
\anth	**763 Dying**
\anth	**764 Funeral**
\anth	**765 Mourning**
\anth	**292 Special Garments**
\mtrl	
\hypo	
\q	
\cf	
\nt	
\dt	14/Jul/1992

Including labels with numeric codes

In the examples, the categories consist of numeric codes and labels. Including labels is preferred to entering just the numbers for two reasons:

- On a later reading of this entry, we can tell at a glance whether all the relevant topics have been specified. (Since few of us have the entire OCM memorized, we would have to look up the meaning of many numeric codes.)

- We can add user-defined subtopics. For example, the OCM topic **136 Fauna** is not adequate to differentiate the various types of animals that are culturally significant. We could add labels like **136 Dog**, **136 Pig**, and **136 Jaguar** to get the distinctions we need without losing the common fauna grouping of code **136**.

Entering categories in a single field

Since this chapter is just a guide, we should note that there are other ways to enter the categories. Many researchers like to enter all the codes in one field (with or without labels):

\anth   **763 Dying, 764 Funeral, 765 Mourning,**
            292 Special Garments
\anth   **763, 764, 765, 292**

In these examples, the delimiter is a comma. Some researchers prefer semicolon or <> (angle brackets). Whatever punctuation you choose, be consistent.

Using this method of cataloging the data will make it impossible for Shoebox to index all of the topics. However, you can access them using filters and the Find command.

Embedding codes in the data

B. Grimes (1986: 9) suggests embedding the OCM codes in the **\data** field itself (using <> angle brackets to set them off from the text). Here is an example from her data:

\date	**1995-01-10**
\de	d02
\wthr	*sunny and clear*
\rscr	bdg
\type	**interview**
\loc	Buru
\srce	king
\part	Chuck
\data	<225> <321> <411> <626> <689> The king showed us the dried stingray tail he uses to punish serious offenders. He put a handle on it. It is about 90 cm long with 1/16-inch spikes along the length of it which will draw blood even through cloth. He told us a couple stories of when he's used it: -on a drunk in Wamkana who dared him to try it against the drunk's martial arts skills <579> -he also used it to break up a drunken lynch mob who were after two men who had fought. <579> <794> <578>
\mtrl	
\hypo	
\q	
\cf	
\nt	
\dt	12/Jan/1995

An advantage of embedded codes is that they are closer to the data they refer to. The Find command can move the insertion point to the relevant piece of data. A disadvantage of embedded codes is that they cannot be indexed by Shoebox.

Distinguishing primary categories

Some researchers want to distinguish primary and secondary categories in their data. Instead of the one **\anth** field, it is possible to use two category fields:

\anthm   *main* (primary) anthropology category
\antho   *other* (secondary) anthropology category

This method makes it possible to search for the records most pertinent to a particular topic. However, you cannot index the database for *all* topics, because Shoebox can sort by only one primary field marker, and the categories are stored in two different fields.

Here is another way to distinguish the primary category: enter the codes in a single **\anth** field with the primary code first and the secondary codes following it. Shoebox filters are capable of distinguishing the first code from any others.

Including categories in other databases

You can use the **\anth** field in your other databases. In this way, you can include relevant texts and lexical entries in the analysis of cultural topics. You can set up Shoebox data links from the **\anth** fields that you use to catalog sentences in the text database and words in the lexical database to the corresponding **\anth** record marker fields in the anthropology analysis database:

\anth   **765 Mourning**

Regardless of the method you use for entering the anthropology categories, you should enter as many as apply to the data. *The more thorough you are in cataloging the data, the more useful your data becomes.*

Anthropology analysis database

The *anthropology analysis* database is where you compile and organize your findings, conclusions, and analyses under specific cultural topics. Each record corresponds to an individual topic or subtopic from the OCM (or whatever system you are using). Here are the recommended data field markers:

\anth   *Anthropology topic*. The record marker contains a topic code, label, or both (depending on what system you are using). The examples in this chapter use OCM topic codes and labels, e.g., **\anth 765 Mourning**. It corresponds to the **\anth** category field in the notes database. It is a destination field in the data link in this database that has **\cf** as source fields, also in data links in anthropology notes, lexical, and text databases that have **\anth** as source fields. Data links ensure consistency when entering data and let you jump from the data to the topic record when compiling the analysis.

\s1 *Section level 1*. Use this field to enter section headings and give structure to the discussion of the topic. You can define more levels as needed (three is usually adequate, e.g., **\s1**, **\s2**, **\s3**).

\pp *Paragraph*. This field represents a generic paragraph of text. Use it to enter any type of text. For convenience of data entry, you might find it useful to select **\pp** as the default marker for the following field (in the Marker Properties dialog box).

\cf *Confer, reference*. Use this field to cross-reference related topics in the analysis database and also specific entries in the notes database (e.g., a particularly insightful observation). Enter the topic or date of the cross-referenced record. Set up a data link from this field to the **\anth** and **\date** fields. This will ensure consistency when entering data and let you jump to related data when compiling the analysis.

\data *Data field*. This is where relevant example data, quotations, stories, etc., from your anthropology notes database can be copied for use in confirming the analysis.

\bib *Bibliography*. Use this field to enter bibliographic references to specific articles relevant to the discussion.

\mtrl *Related Materials*. Use this field to catalog non-text information, such as the location of photographs, drawings, audio and video recordings, artifacts, published and unpublished material on the current topic.

\q *Questions for further study*. Use this field to list questions to help guide future elicitation sessions.

\nt *Notes*. Use this field for general notes or comments (to yourself).

\dt *Date last edited*. This field indicates how recent the generalization is, reflecting the potential need for revision. Select it as the date stamp, so that Shoebox will update it automatically.

Here is an example of a record in the analysis database:

\anth	**765 Mourning**
\pp	Mourning begins with a person's death, and includes an all-night wailing and chicha drinking session. During this wake, candles are burned at each corner of the bed. Therefore, guests bring candles, coffee, or chicha, as these are used throughout the night.
\pp	The wake progresses through several periods of first showing grief through loud wailing, then less frenzied activities. People talk about the dead person, the violin plays to quiet things down, people actually joke, and drinks are brought in. Thus, people get grief out of their systems, but they get to be calm some, also.
\pp	Wailing is distinctive from all other sounds; it is a loud wail, which means something terrible has happened! The word is 'oyase?o-iteanga', a combination of 'cry' and 'very emphatic' morphemes.
\q	How long do they really mourn their dead?
\q	How long must a man wait, after the death of his wife, before it is proper for him to consider remarriage?
\nt	
\dt	05/May/1998

This example is rather simplistic—a first attempt at summarizing *mourning*. As you do more research, you might include additional information in the analysis database, for example:

- references to literature (in the **\bib** field)

- specific examples—stories and relevant quotations from **\data** fields in the anthropology notes database—that you might use when you are ready to formalize a topic into a paper

Everything that you have entered into Shoebox concerning *mourning* is easily accessible to you for making and verifying hypotheses. The process of data retrieval and study that makes it possible to write the summary entries is the same for both preliminary and advanced research.

Compiling the analysis

You can use Shoebox to help you analyze the data as follows:

1. Choose a topic for study, e.g., *mourning*. In addition to its code **765**, think of key words, e.g., **death**, **mourn**, **wail**, etc.

2. Search for the topic and key words in your data files: anthropology notes, lexical, and text databases. Use sorting, filtering, and the Find command to select all the data you have collected that is relevant to the topic.

3. In the topic record in your anthropology analysis database, enter your hypotheses, generalizations, and conclusions.

4. As you notice "holes" in the data, you can plan for further data collection.

Notice the relationship between the anthropology notes and analysis databases:

- The notes database is cataloged by *categories* that correspond to *topics* in the analysis database.

- You select records (in the *notes* database) to gather the information that you use to form conclusions. You enter the conclusions in a topic record (in the *analysis* database).

Sorting by topic

As explained earlier, Shoebox can sort by the categories in the anthropology notes database:

1. From the Window menu, choose Duplicate.

2. From the Database menu, choose Sorting.

3. Select **\anth** as the primary sorting field.

4. Select **\date** as a secondary sorting field (to arrange the entries in each category from earliest to most recent).

5. Click OK.

6. Search for **765** to move to the set of records that are indexed by the anthropology code for *mourning*.

7. Read through the **\data** fields to gather all the information you have entered on the topic.

Filtering topics

Alternatively, you could use a filter to display only the records that contain **\anth 765**. Filtering hides records that are irrelevant to a topic.

Note: One of the main reasons for dividing your **\data** fields into smaller, more manageable chunks (during data entry) is to make the selected records succinct and relevant to the topic of focus (during data analysis).

A Shoebox command for displaying data that works well with both sorting and filtering is Browse Fields on the View menu.

Finding key words

It is a good idea to use the Find command. Finding records that contain key words such as **death**, **mourn**, and **wail** double-checks for data in your anthropology notes database that you might not have cataloged adequately. This is how you can search for related information in your text and lexical databases, if you have not been entering anthropology categories in them.

To avoid overlooking potential key words for a topic, you can make a word list of the contents of the **\data** fields in the anthropology notes database. (You might find it helpful to include the **\hypo**, **\q** and **\nt** fields in the word list.) Access Shoebox Help for more information about the following commands:

- Text Corpora on the Project menu

- Word List on the Tools menu

- Concordance on the Tools menu

Printing the data

If you would prefer to study this information on paper, you can:

- print filtered databases directly (Print on the File menu)

- export them in Rich Text Format (RTF) to a word processor, like Microsoft Word, for formatting and printing (Export on the File menu)

Dividing topics   As a topic expands, you might want to divide it into subtopics. Derive the numeric codes and labels for these subtopics from the main topic, for example:

\anth   **765 Mourning**
\anth   **765a Mourning: People's Roles in a Wake**
\anth   **765b Mourning: Spiritual Implications**

Cut the relevant parts out of the main topic record in the analysis database and paste them into the appropriate subtopic record. Because the current Shoebox program is able to handle much larger records than earlier versions, you can keep all of a discussion in one record if you prefer. Dividing up a large topic is suggested here merely to make data access easier.

Goal for data analysis   A good goal for data analysis is to develop each topic record to the point that it could become a major section in an ethnography (or a paper on a single topic). By excluding the "working" fields (e.g., **\q**, **\nt**, **\dt**) and with a little additional polishing, the analysis database could be exported as a rough manuscript to be worked towards publication.

You might begin research by starting with topics related to social organization and worldview. Here are some good references:

- Spradley (1980) discusses social organization, covering social situations, procedures of analysis, and cultural scenes.

- Kearney (1984, especially Chapter 3) discusses worldview.

- Bernard (1994) discusses research techniques.

- Van Willigen (1986) discusses planned cultural change.

References

Bernard, H. Russell. 1994. *Research methods in cultural anthropology*. 2nd edition. Newbury Park, Calif.: Sage Publications.

Coward, David F. and Charles E. Grimes. 1995. *Making Dictionaries: A guide to lexicography and the Multi-Dictionary Formatter*. Waxhaw, NC: SIL International.

Grimes, Barbara D. 1986. Wanted: computer programs for anthropology. *Notes on Translation.* 115:7–19. Dallas: SIL International.

Kearney, Michael. 1984. *World view*. Novato, Calif.: Chandler and Sharp Publishers, Inc.

The Linguist's Shoebox: Tutorial and User's Guide. 2000. Waxhaw, NC: SIL International.

Murdock, George, et al. 1987. *Outline of cultural materials*. 5th edition with revisions. New Haven, Conn.: Human Relations Area Files, Inc.

Spradley, James P. 1980. *Participant observation*. New York: Holt, Rinehart, and Winston.

van Willigen, John. 1986. *Applied anthropology: an introduction*. South Hadley, Mass.: Bergin and Garven Publishers, Inc.

Index