INTEGRATED DATA MANAGEMENT AND ANALYSIS FOR THE FIELD LINGUIST

The Linguist’s Shoebox: Tutorial and User’s Guide

SIL International
Waxhaw, North Carolina

Produced and distributed by:
JAARS, Inc.
Computer and Communications Services
PO Box 248
Waxhaw, NC 28173-0248
USA

Telephone: (704) 843-6085
(800) 421-8024
Fax: (704) 843-6500

Contents

Acknowledgments

The Linguist's Shoebox reflects the contributions of many field workers actively engaged in linguistic and anthropological research. The usability and flexibility of Shoebox are directly related to the efforts of those colleagues in SIL International and other organizations who use this software and offered their suggestions for improving it.

Shoebox was originally developed for MS-DOS by John Wimbish, who worked closely with David Coward to meet their needs for data management as they studied languages in the province of Maluku in eastern Indonesia.

During development of Shoebox for Windows and Macintosh, the technical team has included programmers Tom Bogle, Alan Buseman, Rod Early, Mark Pedrotti, Brian Yoder, and Bryan Wussow. Significant contributions also came from Doug Case and Doug Rintoul (the Consistent Changes component), Tim Erickson (context-sensitive rendering using Script Definition Files), and Michael Lastufka (increased number of characters possible in sort orders). Doug Eberman, Josh Kelley, and Wolfgang Witteveen contributed their skills on a short-term basis.

Barbara Altork, Karen Buseman, and the programmers developed Shoebox Help.

Alan Buseman, Karen Buseman, David Coward, Dean Jordan, and the programmers wrote this *Tutorial and User's Guide*. Tom Bogle compiled its index. Jim Albright, Kim Blewett, and Don Horneman provided helpful editorial suggestions.

Phil Baer, Jr., John C. Bailey, Michael Buchanan, Chet Matheson, and Nathan Smith developed the Self-Paced Computer-Based Training (CBT) for Shoebox.

Many others in Computer and Communications Services at JAARS helped develop Shoebox, especially Phil Baer, Jr. and Neal Brinneman (computer training) and Jon Coblentz and Bob Wright (software support).

We are grateful to God for His leading and provision throughout the development and release of Shoebox over the years.

Serving together,
The Shoebox team

Introduction

The Linguist's Shoebox helps field researchers integrate various kinds of text data: lexical, cultural, grammatical, etc. It has flexible options for selecting, sorting, and displaying data. It is especially useful for helping researchers build a dictionary as they use it to analyze and interlinearize text. 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.

Shoebox runs on the following computer systems:

- Windows 3.1 (file name: Shw.exe)
- Windows 95 and later (file name: Shw32.exe)
- Macintosh
- Power Macintosh

The Shoebox Setup program installs the appropriate version of Shoebox on your computer. See *Installing Shoebox*, p. 7.

Shoebox 5 automatically converts older "settings" files from versions 3 and 4. For information about converting files from version 2 (for MS-DOS), see *Appendix A: Importing from MS-DOS*, p. 265.

Book overview

This book consists of three parts:

- *Tutorial*
- *User's Guide*
- *Appendixes*

The *Tutorial* is a guided tour of the program that demonstrates typical ways it can be used. Working through the step-by-step procedures in this section on a computer is much more effective than merely reading it.

- *Getting started* covers installing, starting, and exiting Shoebox.
- *Basic features* introduces the screen display, Shoebox Help, field markers, and commands for selecting, sorting, and displaying data.
- *Shoebox settings files* describes projects, language encodings, and database types.
- *Interlinear text* introduces morphological parsing and building a lexical database from text.

The *User's Guide* contains selected reference information. For more complete and detailed information, access Shoebox Help when you are running Shoebox.

- *Basic information* covers field markers, variables, and printing databases using the Print and Export commands.
- *Frequently asked questions* gives some helpful tips.
- *Notes on interlinearization* gives more details about parsing.

The *Appendixes* contain information that some users might need when first getting started with this version of Shoebox.

- *Appendix A: Importing from MS-DOS* gives step-by-step procedures for converting and importing existing files from Shoebox for MS-DOS.

- *Appendix B: Comparison to MS-DOS* lists some of the improvements to Shoebox and summarizes changes to filters and the menu bar.
- *Appendix C: Uninstalling Shoebox* tells how to remove Shoebox from your computer should that become desirable.
- *Appendix D: Reinstalling the samples* gives step-by-step procedures for installing the sample files again if you want to repeat parts of the tutorial.

Tutorial milestones

The *Tutorial* is organized into a series of milestones. Each milestone begins with a fresh set of sample files. If you have trouble with any section of the tutorial, you can start fresh at the next milestone.

Macintosh and Windows specifics

This book describes the Macintosh and Windows versions of the program. In data management and analysis, the two versions are very similar. Here are two basic differences:

- Mouse buttons. The Macintosh mouse has only one button. On a computer running Windows, a mouse has a right button. To simulate right-clicking on the Macintosh, hold down the OPTION button as you click the mouse button (OPTION+click). *Note:* Shoebox uses the right mouse button sparingly.
- Keyboard shortcuts. The Macintosh uses a COMMAND key, but Windows uses an ALT key. Therefore, some of the shortcuts use different keys. This book rarely mentions them. If you prefer to use keyboard shortcuts, you can find them on the corresponding menu items and ToolTips.

Windows 3.1 terminology is not identical with the other systems that Shoebox runs on. Here are two minor differences:

- A *directory* on Windows 3.1 (and MS-DOS) corresponds to a *folder*. This book usually refers to *folders*.
- A *program group* corresponds to a *folder window*.

Notation conventions

In this book and in Shoebox Help, a + between two or more keys indicates that they should be pressed together. That is, hold down the first key and then press the second (e.g., CTRL+N). This notation is also used for the OPTION+click key-mouse combination on the Macintosh (which corresponds to clicking the right mouse button on Windows).

This book uses the following special typographic conventions:

Font Style	Meaning
Italic	Book titles, references to section headings, emphasis, and data (in captions, footnotes, and the body text; but not in indented examples).
Bold	Data field markers (in captions, footnotes, and the body text; but not in indented examples).
Sans-serif	Menu items, list items, and buttons that you are to choose, select, or click. Data field markers in indented examples.
`Monospace`	Text items that you must type exactly as shown.
SMALL CAPS	Names of keys on the computer keyboard that you are to press.

PART 1

TUTORIAL

Getting started

Installing Shoebox

This section contains step-by-step procedures for installing The Linguist's Shoebox. The programs that install Shoebox on your computer are called Shoebox Setup on Windows and the Shoebox Installer on the Macintosh. If you want to stop the installation, click Cancel in any of the dialog boxes displayed by the Setup/Installer program.

Minimum requirements:

- Windows 3.1 or Macintosh OS 7.5
- At least 10 megabytes (MB) free space on the hard disk
- The required amount of memory (RAM) depends on the version of Windows or Macintosh, the requirements of other programs that you use, and the amount of data you will store in Shoebox. Installing additional memory in your computer can increase the speed and capacity of Shoebox.

Macintosh installation

To install Shoebox on the Macintosh do the following:

1. Insert the Shoebox CD-ROM or floppy disk #1 into the appropriate drive.
2. In the Finder, look at the Shoebox CD-ROM or disk and double-click the Shoebox Installer program.

The Installer will display a window with release information (the Read Me) for Shoebox. The release notes give installation information and additional insights into the program that will help you get the most from Shoebox.

3. Review the release notes, then click Continue.

The Installer will display a dialog that enables you to choose the drive or location where Shoebox will be installed. Usually the default drive is the best.

4. Read the information displayed and click Install when you are ready to continue.

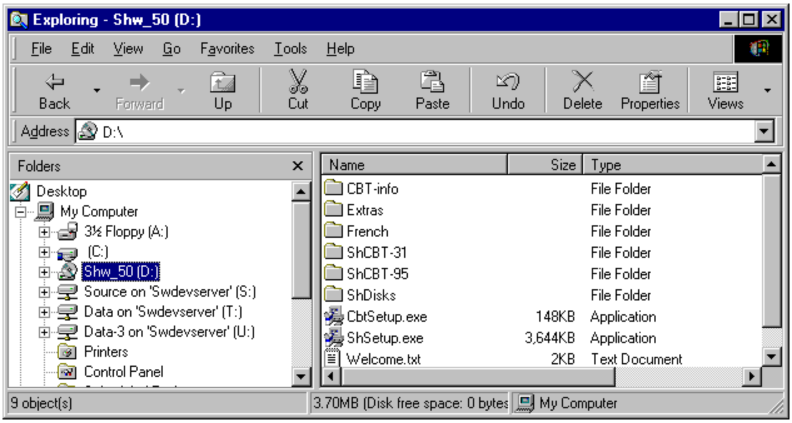
The installation will begin. When the software is completely installed, the Installer will notify you.

5. Click Quit to finish the installation.

Windows installation

The following steps show the details of installing the CD-ROM edition of Shoebox on Windows 98:

1. Insert the Shoebox CD-ROM into the CD-ROM drive.
2. In Windows Explorer, click the CD-ROM drive.



Address: D:\

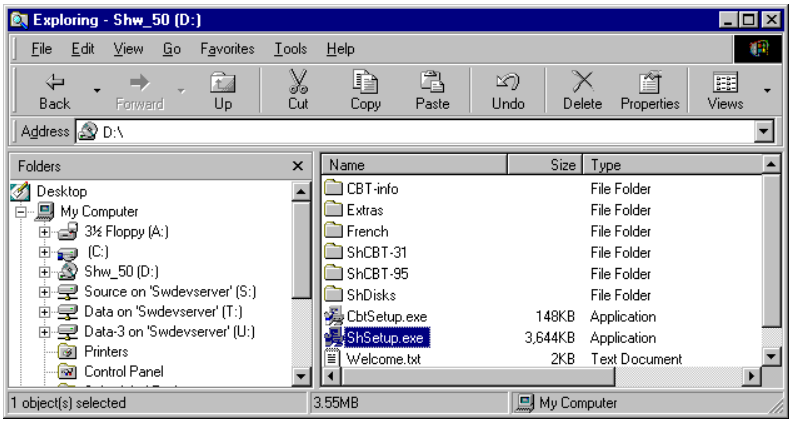
Name	Size	Type
CBT-info		File Folder
Extras		File Folder
French		File Folder
ShCBT-31		File Folder
ShCBT-95		File Folder
ShDisks		File Folder
CbtSetup.exe	148KB	Application
ShSetup.exe	3,644KB	Application
Welcome.txt	2KB	Text Document

9 object(s) 3.70MB (Disk free space: 0 bytes My Computer

Fig. 1. Shw_5.0, Shoebox for Windows 5.0, selected

Note: To see a similar view, select View, and then Details.

3. Double-click ShSetup.exe, the Shoebox Setup program.



Address: D:\

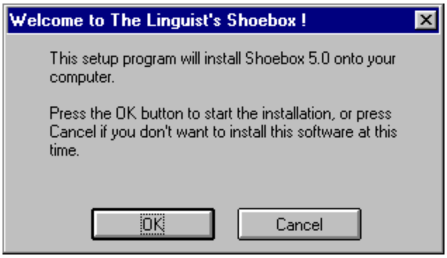
Desktop
My Computer
3½ Floppy (A:)
(C:)
Shw_50 (D:)
Source on 'Swdevserver' (S:)
Data on 'Swdevserver' (T:)
Data-3 on 'Swdevserver' (U:)
Printers
Control Panel

Name	Size	Type
CBT-info		File Folder
Extras		File Folder
French		File Folder
ShCBT-31		File Folder
ShCBT-95		File Folder
ShDisks		File Folder
CbtSetup.exe	148KB	Application
ShSetup.exe	3,644KB	Application
Welcome.txt	2KB	Text Document

1 object(s) selected | 3.55MB | My Computer

Fig. 2. Shoebox Setup program selected

Setup will display a welcome message.


Welcome to The Linguist's Shoebox !

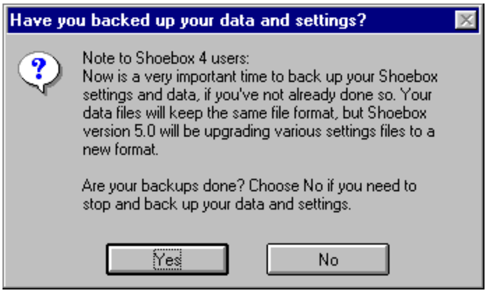
This setup program will install Shoebox 5.0 onto your computer.

Press the OK button to start the installation, or press Cancel if you don't want to install this software at this time.

Fig. 3. Welcome dialog box

4. Click OK.

If you are upgrading from an earlier version of Shoebox, the Setup program will remind you to back up your settings and data.



Have you backed up your data and settings?

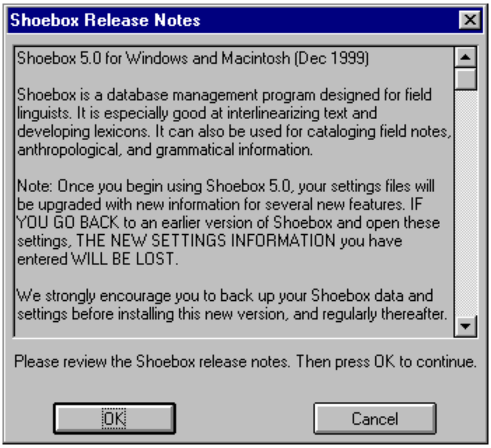
Note to Shoebox 4 users:
Now is a very important time to back up your Shoebox settings and data, if you've not already done so. Your data files will keep the same file format, but Shoebox version 5.0 will be upgrading various settings files to a new format.

Are your backups done? Choose No if you need to stop and back up your data and settings.

Fig. 4. Shoebox Setup backup reminder

5. If your backups are complete, click Yes to continue.

Setup will display a window showing release information (the Read Me) for Shoebox.



Shoebox Release Notes

Shoebox 5.0 for Windows and Macintosh (Dec 1999)

Shoebox is a database management program designed for field linguists. It is especially good at interlinearizing text and developing lexicons. It can also be used for cataloging field notes, anthropological, and grammatical information.

Note: Once you begin using Shoebox 5.0, your settings files will be upgraded with new information for several new features. IF YOU GO BACK to an earlier version of Shoebox and open these settings, THE NEW SETTINGS INFORMATION you have entered WILL BE LOST.

We strongly encourage you to back up your Shoebox data and settings before installing this new version, and regularly thereafter.

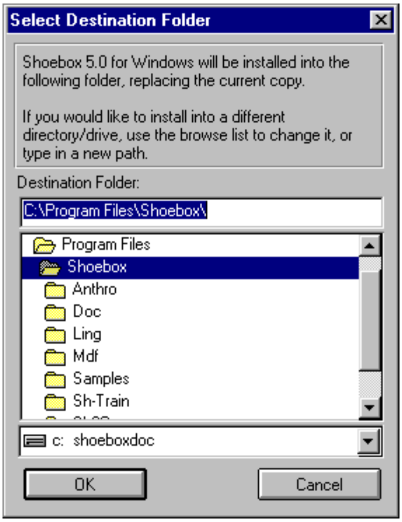
Please review the Shoebox release notes. Then press OK to continue.

Fig. 5. Shoebox Release Notes

The Release Notes give installation information and additional insights into the program that will help you get the most from Shoebox.

6. Review the release notes, then click OK.

Setup will then ask where you want to install Shoebox. Unless you wish otherwise, use the default destination folder.



Shoebox 5.0 for Windows will be installed into the following folder, replacing the current copy.

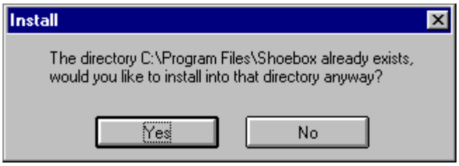
If you would like to install into a different directory/drive, use the browse list to change it, or type in a new path.

Destination Folder: C:\Program Files\Shoebox\

Fig. 6. Select Destination Folder dialog box

7. Click OK to continue.

If you are upgrading from a prior version, you will usually be overwriting the old version in the Shoebox folder. If this is the case, Setup will prompt you to confirm that it will install into an existing directory. (If installing to a new folder, Setup will skip this step.)

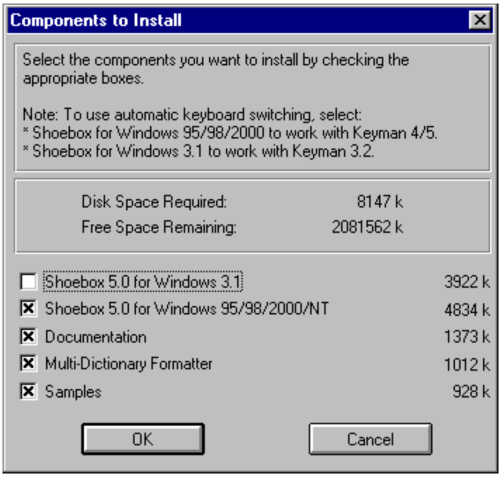


The directory C:\Program Files\Shoebox already exists, would you like to install into that directory anyway?

Fig. 7. Install dialog box

8. Unless you wish otherwise, click Yes here to continue the installation.

Setup will next ask you which parts of the Shoebox package you want to install.



Components to Install

Select the components you want to install by checking the appropriate boxes.

Note: To use automatic keyboard switching, select:
* Shoebox for Windows 95/98/2000 to work with Keyman 4/5.
* Shoebox for Windows 3.1 to work with Keyman 3.2.

Disk Space Required:	8147 k
Free Space Remaining:	2081562 k

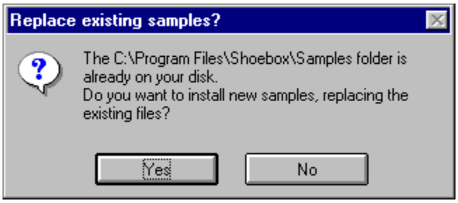
☐ Shoebox 5.0 for Windows 3.1	3922 k
☒ Shoebox 5.0 for Windows 95/98/2000/NT	4834 k
☒ Documentation	1373 k
☒ Multi-Dictionary Formatter	1012 k
☒ Samples	928 k

Fig. 8. Components to Install dialog box

When the dialog opens, most of the components are selected. You may change the selections if you have reason to, but the components selected by default are recommended.

9. After you have made your selections, click OK to continue.

Setup will begin copying files to your destination folders. If you are replacing a prior installation, Setup will prompt you before it overwrites the Samples folder.



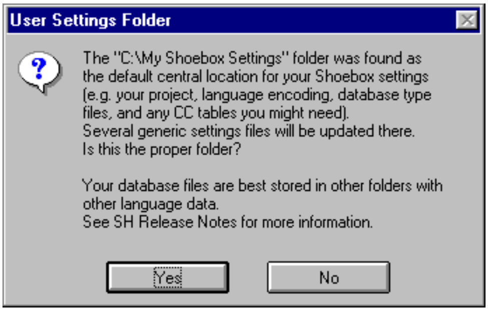
Replace existing samples?

The C:\Program Files\Shoebox\Samples folder is already on your disk.
Do you want to install new samples, replacing the existing files?

Fig. 9. Replace existing samples dialog box

10. If you wish to clear out everything in the present Samples folder and replace the samples, click Yes.

Setup asks or confirms where you want to keep your Shoebox settings.


User Settings Folder

The "C:\My Shoebox Settings" folder was found as the default central location for your Shoebox settings (e.g. your project, language encoding, database type files, and any CC tables you might need).
Several generic settings files will be updated there.
Is this the proper folder?

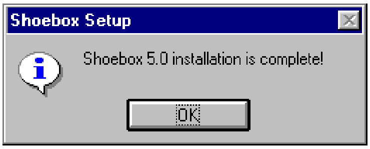
Your database files are best stored in other folders with other language data.
See SH Release Notes for more information.

Fig. 10. User Settings Folder dialog box

The user settings folder is the default central location for the various settings files that you will create and use in Shoebox. Unless you have reasons to do otherwise, use the default name: My Shoebox Settings.

11. Click Yes to continue.

When finished, Setup will report to you that the installation is complete.


Shoebox Setup

Shoebox 5.0 installation is complete!

Fig. 11. Shoebox Setup dialog box

12. Click OK to close the Setup program.

This concludes the installation instructions.

Finding the sample files

The Tutorial uses sample files that are installed with Shoebox. Here are the default locations of the folders that contain these files:

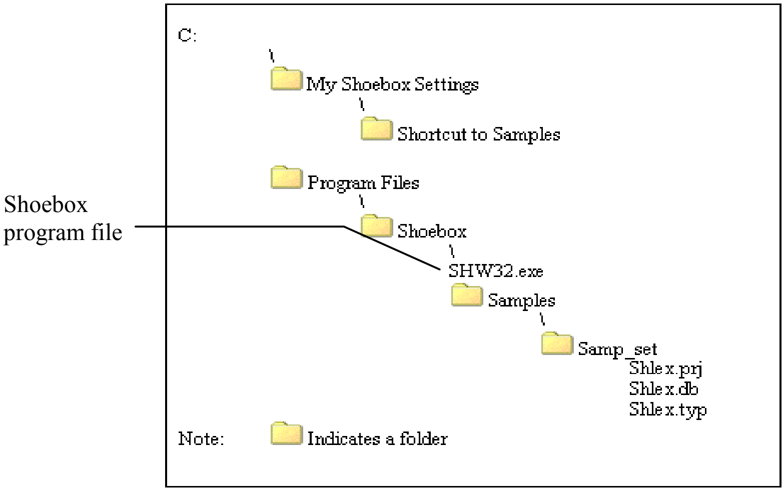


Fig. 12. Default locations of sample folders in Windows 95 and later

In the following section, you will open the Shlex.prj project file. If you have trouble finding it using the procedures given, you can still continue the tutorial by finding Shlex.prj in Windows Explorer (or the Macintosh Finder) and double-clicking it. The screen shown in Fig. 18 should then appear.

Opening a sample project

For the first part of the tutorial, you will use some sample data from a lexicon of the Selaru language of Indonesia.[1] This sample lexicon uses the standard set of field markers defined by the Multi-Dictionary Formatter (MDF) software that is included with this version of Shoebox.[2]

The following are step-by-step instructions for opening the project on Windows 95 and later. If you are upgrading from an older

[1] This is unpublished data collected by David and Naomi Coward.

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

version of Shoebox, you should back up your Shoebox settings and data files before continuing. If Shoebox is already running on your computer, skip to step 6 below.

If Shoebox is not already running,

1. Click the Windows Start button.
2. Click Programs.
3. Click Shoebox 5.0.
4. Click Shoebox.

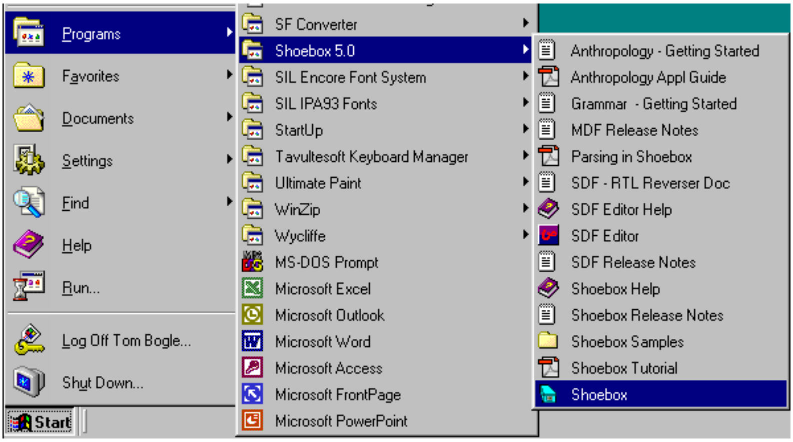


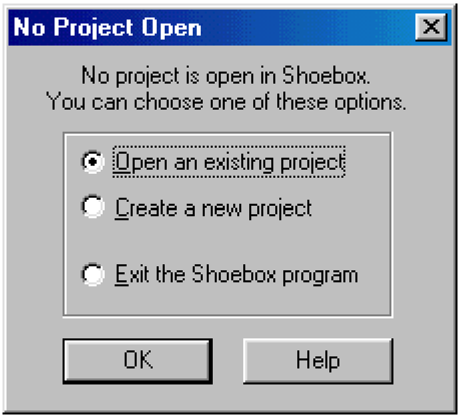
Fig. 13. Shoebox 5.0 Start menu

5. One of three things will happen:

   a. Windows will display the Shoebox No Project Open dialog box unless you have used Shoebox before. If this happens skip to step 6.

   b. If you have just upgraded from an earlier version of Shoebox, Shoebox will display the Convert Older Settings Files dialog box. If you have not yet backed up your Shoebox settings files, click No and, after backing them up, start over at step 1. To continue with this tutorial, click Yes.

c. If you have used Shoebox before (and exited with a project still open), Shoebox will open whatever you had open the last time. You will see the project name in the lower right corner of the Shoebox window. Choose Project, Close. The No Project Open dialog box (Fig. 14) will be displayed. Now go on to step 6.



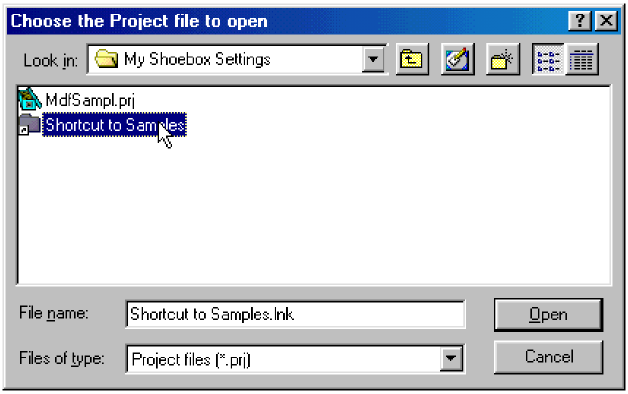
No project is open in Shoebox.
You can choose one of these options.

● Open an existing project
○ Create a new project
○ Exit the Shoebox program

Fig. 14. No Project Open dialog box

6. Select Open an existing project, then click OK.

Shoebox will display the Choose the Project file to open dialog box.



Look in: My Shoebox Settings

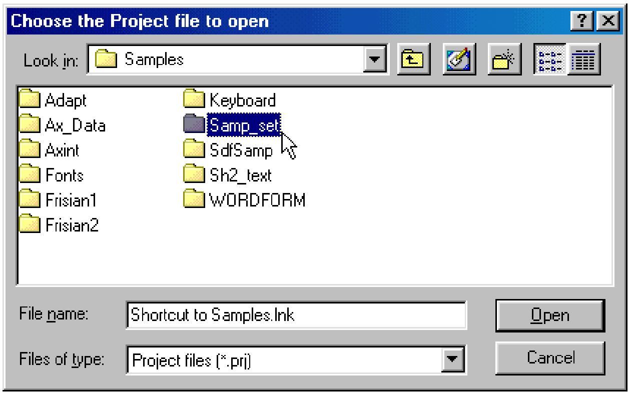
MdfSampl.prj
Shortcut to Samples

File name: Shortcut to Samples.lnk
Files of type: Project files (*.prj)

Fig. 15. Choose the Project file to open

7. To begin the tutorial, double-click Shortcut to Samples.

Shoebox will display the contents of the Samples folder.



Choose the Project file to open

Look in: Samples

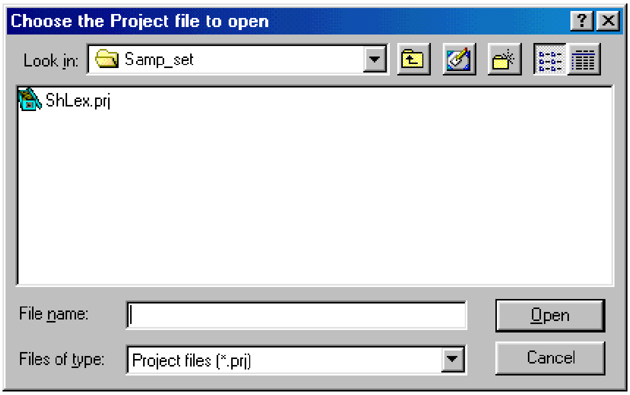
File name: Shortcut to Samples.lnk

Files of type: Project files (*.prj)

Fig. 16. Samples folder

8. Double-click Samp_set.

Shoebox will display the contents of the Samp_set folder.



Choose the Project file to open

Look in: Samp_set

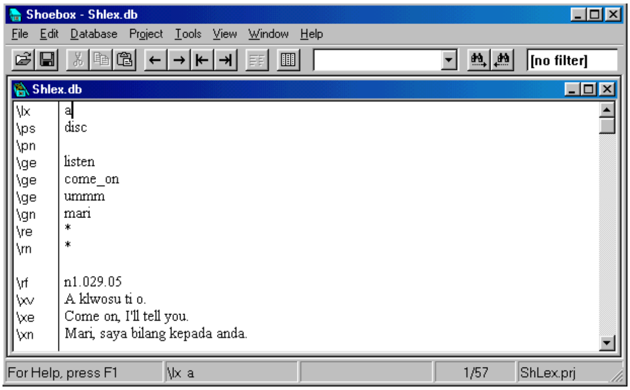
File name:

Files of type: Project files (*.prj)

Fig. 17. Samp_set folder

9. Double-click Shlex.prj.

Shoebox will open the project file and all of its associated files.



```
\lx   a
\ps   disc
\pn
\ge   listen
\ge   come_on
\ge   ummm
\gn   mari
\re   *
\rn   *

\rf   n1.029.05
\xv   A klwosu ti o.
\xe   Come on, I'll tell you.
\xn   Mari, saya bilang kepada anda.
```

Fig. 18. Shoebox - Shlex.db

A project file stores the names and paths of all the files that were open when you last used Shoebox.

If you open Shlex.prj and it does not look like Fig. 18, you should reinstall the sample files from your Shoebox CD-ROM, then start over at *Opening a sample project*, p. 14.

Exiting Shoebox

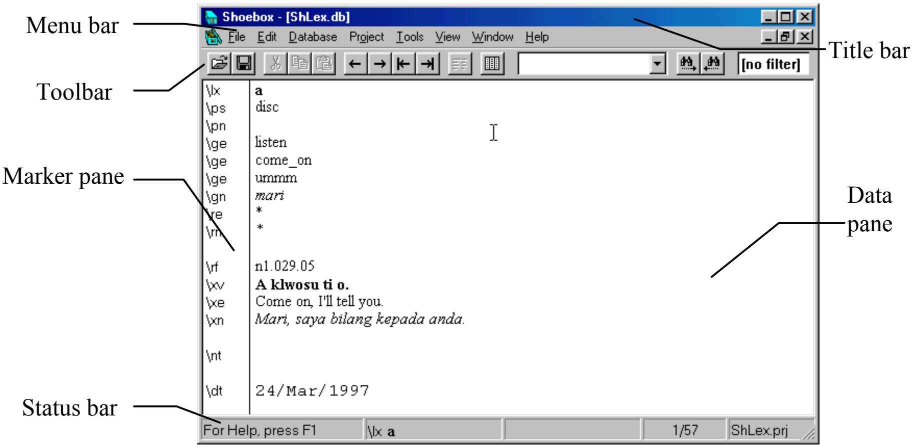
Important: If at any time during this tutorial you need to exit Shoebox, choose File, Exit with all databases and windows still open. This will save everything exactly as you had it. When you run the program again, it will automatically open all your files and position them exactly as you had them so you can continue from where you exited.

Basic features

This chapter is a guided tour that introduces basic features of Shoebox: screen display, Shoebox Help, field markers, and commands for selecting, sorting, and displaying data. These features are demonstrated using a sample project. Working through the step-by-step procedures on a computer is much more effective than merely reading the information.

Screen display

Here is the screen display for the Shlex.prj sample project:



```
\lx     a
\ps     disc
\pn
\ge     listen
\ge     come_on
\ge     ummm
\gn     mari
\re     *
\rn     *

\rf     n1.029.05
\xv     A klwosu ti o.
\xe     Come on, I'll tell you.
\xn     Mari, saya bilang kepada anda.

\nt

\dt     24/Mar/1997
```

For Help, press F1 | \lx a | 1/57 | ShLex.prj

Fig. 19. Shoebox Screen

The window displays the first record of the Shlex.db database file.[3]

[3] If any data is bold or italic or colored, the tutorial has already been done by someone. You should reinstall the samples (see *Appendix D: Reinstalling the samples*, p. 299). Then rerun Shoebox, and if necessary reopen the Shlex project as above.

A vertical line separates the marker pane (on the left) and the data pane (on the right). The marker pane contains the field markers for your database. The data pane contains the data fields for each record.

In the data pane, you can move the text insertion point with the arrow keys. You can also place the insertion point by clicking with the mouse. You can edit text by typing characters to insert or by deleting in the usual manner. Observe that when you move the mouse pointer into the marker pane it turns into an arrow. You cannot edit the Standard Format markers from this window as changes to the markers affect every record in your database. In the data pane you can edit the data entries and it affects only the record you are currently editing.

Toolbar Below the menu bar is the toolbar. Each button on the toolbar is a shortcut to a different command.[4]

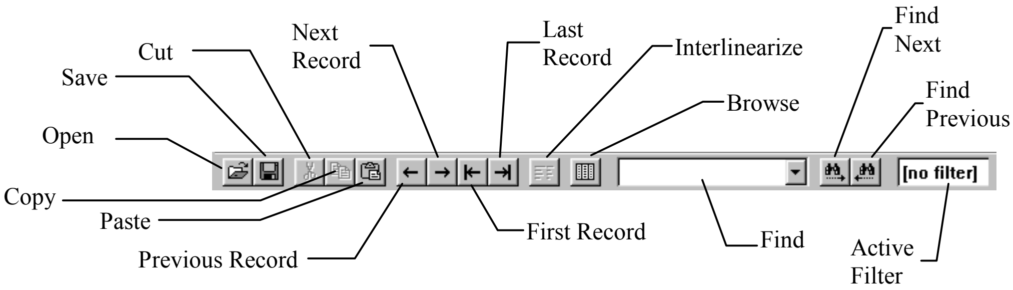
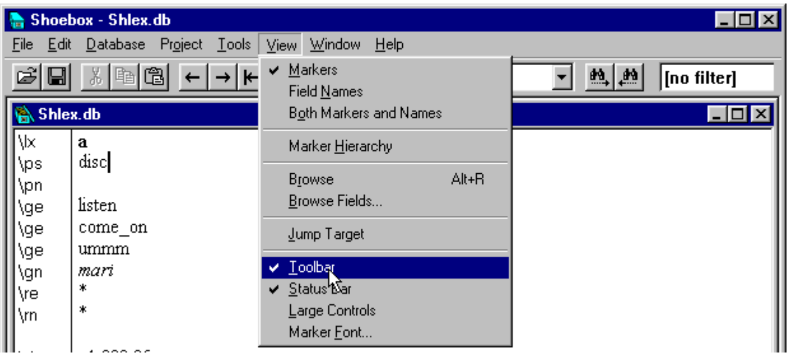


Fig. 20. Shoebox Toolbar

The commands associated with the toolbar buttons are discussed in detail further on.

You will use the toolbar during the tutorial. You can hide it later if you prefer not to use it in your own work. To turn the toolbar on or off, choose View, Toolbar.

[4] If you are running Windows, you may have to maximize the Shoebox window to see all the tools.



\lx	a
\ps	disc
\pn	
\ge	listen
\ge	come_on
\ge	ummm
\gn	*mari*
\re	*
\rn	*

Fig. 21. Turning off the Toolbar

Status bar

The status bar at the bottom of the screen shows a variety of information.

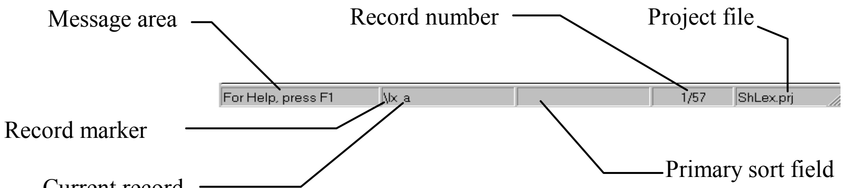


Fig. 22. Status bar

Starting on the left is a Message area. The message area gives a brief description of any toolbar button you have selected. This is usually not much help as the command has already been executed and it is too late to be looking for information. Also, while selecting menu items, the message area displays a brief description of the command being selected. Usually the message area says: “For Help, press F1”.

Note: ToolTips are a more useful source of information on toolbar items. When you move the pointer and pause on an item, a text window drops down for several seconds describing its function.

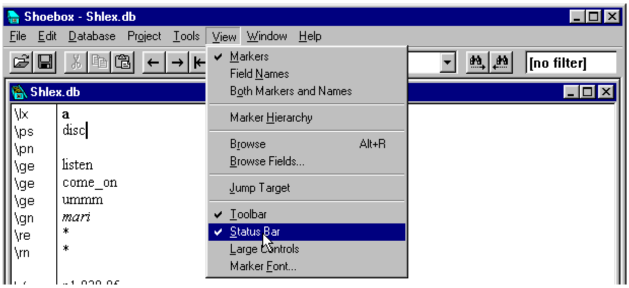
Next is the record marker in the active database followed by the contents of the record marker field. This is the most important field in a record as it identifies the record. It is like the family name, which identifies all of the members of the family; for database use it identifies all the entries of a record.

Next, the marker of the primary sort field and its contents are displayed. (At this point in the tutorial, this part of the status bar should be blank because Shoebox displays the primary sort marker and field content only if it is different from the record marker field. At this point, Shlex.db is sorted by the record marker, so to display both would duplicate what is already showing.) Together these two status fields help clarify where you are in the record sequence.

The status bar also shows the number of the current record and the total number of records in the database. It should read 1/57 (for Shlex.db) if you are at the top of the database. (Actually, it only shows the total number of records in a database when the database sorts by the record marker and no filter is active. This is described more fully under *Multiple views*, p. 56.)

In the final pane of the status bar, the name of the project file is displayed. It should currently say Shlex.prj.

If you are pressed for screen space or prefer not to see or use the status bar, you can hide it. Chose View and click Status Bar. (It can be restored again by the same method.)



```
\lx   a
\ps   disc
\pn
\ge   listen
\ge   come_on
\ge   ummm
\gn   mari
\re   *
\rn   *
```

Fig. 23. Status Bar turn off menu

Shoebox Help

The Help system is an important part of Shoebox. It contains more complete and detailed information than this book. Use Shoebox Help whenever you have a question. There are four main ways to access Help: context-sensitive help, help contents, help links, and help search.

Context-sensitive help

Context-sensitive help is used to explain the meaning or purpose of the various buttons on the toolbar and the numerous choices in the menus.

1. Choose Help, Context Sensitive. (On the Macintosh, the Help menu is a question mark near the upper right corner of the screen.)

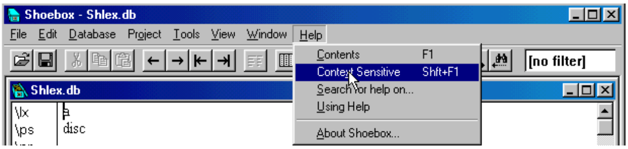


Fig. 24. Accessing context-sensitive help

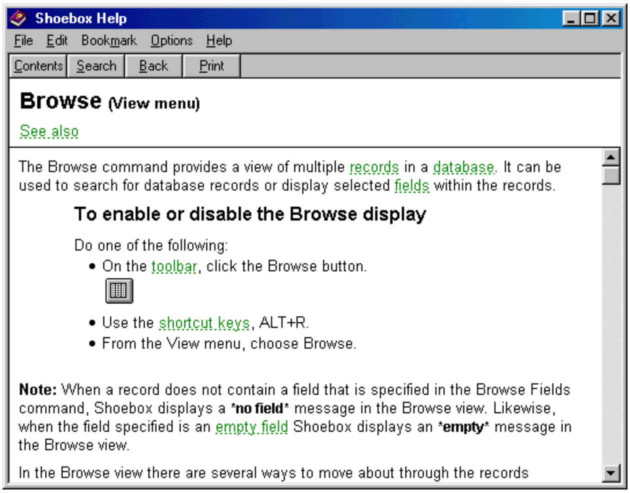
You will see that the mouse pointer now has a question mark beside it, indicating that it is ready to give context-sensitive help.

Fig. 25. Context-sensitive mouse pointer

2. Move the mouse pointer to ▦ on the toolbar and click it.

3. Shoebox Help will activate and you will see a discussion of Browse. (On the Macintosh, if you see a message saying Failed to launch Help, see footnote below.[5])

[5] On the Macintosh, Shoebox Help runs a separate program, which requires about 500 KB of memory. If the system does not have that much available, Help cannot start. There are various ways to deal with this. If you are running other programs, you may be able to close them to make room. A sure way to make extra memory is to turn on virtual memory or to enlarge it if already on and you have the free space available on your hard disk. Another thing to check is the Shoebox Memory Requirements numbers. You can see these by selecting the Shoebox program icon and choosing File, Get Info. If the preferred size is larger than the available memory when Shoebox is run, then it takes all available memory. Reducing the preferred size may help.



Browse (View menu)

See also

The Browse command provides a view of multiple records in a database. It can be used to search for database records or display selected fields within the records.

To enable or disable the Browse display

Do one of the following:

- On the toolbar, click the Browse button.
- Use the shortcut keys, ALT+R.
- From the View menu, choose Browse.

Note: When a record does not contain a field that is specified in the Browse Fields command, Shoebox displays a ***no field*** message in the Browse view. Likewise, when the field specified is an empty field Shoebox displays an ***empty*** message in the Browse view.

In the Browse view there are several ways to move about through the records

Fig. 26. Shoebox Helps Menu showing Browse

4. Close the Shoebox Help window.

Context-sensitive help on a menu choice

You can also use context-sensitive help on a menu choice.

1. Choose Help, Context Sensitive.
2. Choose Database, Sorting from the menu bar using your context-sensitive mouse pointer.

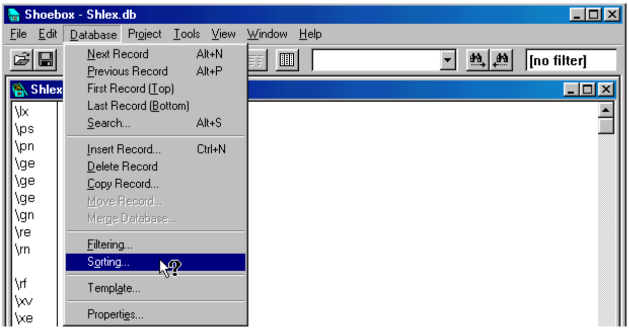
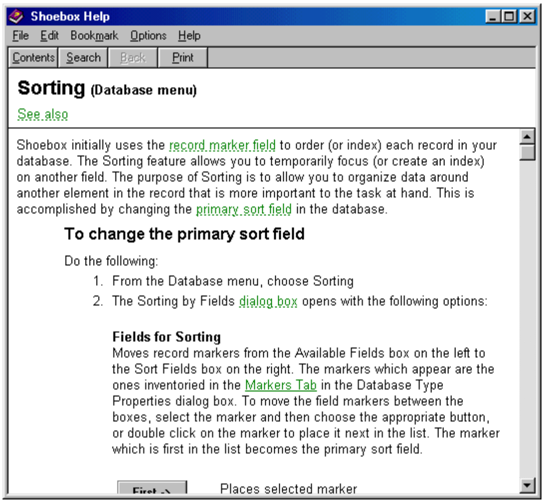


Fig. 27. Choose Database, Sorting

You will see a discussion of sorting.



Sorting (Database menu)

See also

Shoebox initially uses the record marker field to order (or index) each record in your database. The Sorting feature allows you to temporarily focus (or create an index) on another field. The purpose of Sorting is to allow you to organize data around another element in the record that is more important to the task at hand. This is accomplished by changing the primary sort field in the database.

To change the primary sort field

Do the following:

1. From the Database menu, choose Sorting
2. The Sorting by Fields dialog box opens with the following options:

Fields for Sorting
Moves record markers from the Available Fields box on the left to the Sort Fields box on the right. The markers which appear are the ones inventoried in the Markers Tab in the Database Type Properties dialog box. To move the field markers between the boxes, select the marker and then choose the appropriate button, or double click on the marker to place it next in the list. The marker which is first in the list becomes the primary sort field.

First -> Places selected marker

Fig. 28. Shoebox Helps showing Sorting

3. Close the Shoebox Help window.

Context-sensitive help on any dialog box

You can also get context-sensitive help on any dialog box.

1. Choose Database, Sorting.

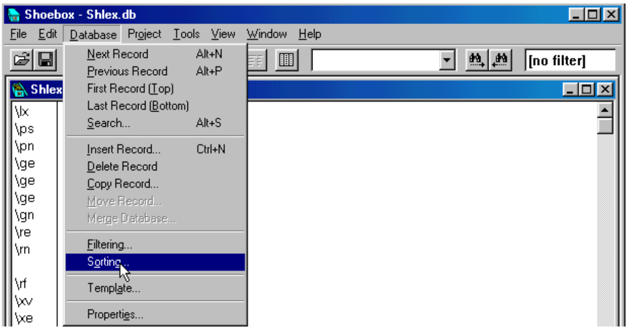
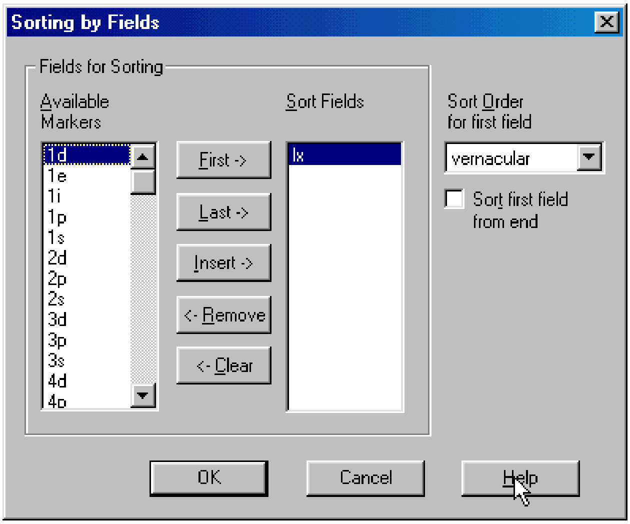


Fig. 29. Choose Database, Sorting

You will see a dialog box titled Sorting by Fields. One of the buttons on the box is labeled Help.



Sorting by Fields

Fields for Sorting

Available Markers: 1d, 1e, 1i, 1p, 1s, 2d, 2p, 2s, 3d, 3p, 3s, 4d, 4p

Sort Fields: lx

Sort Order for first field: vernacular

□ Sort first field from end

Fig. 30. Sorting by Fields dialog box

2. Click Help.

You will see the same discussion of sorting as you saw above in Fig. 28.

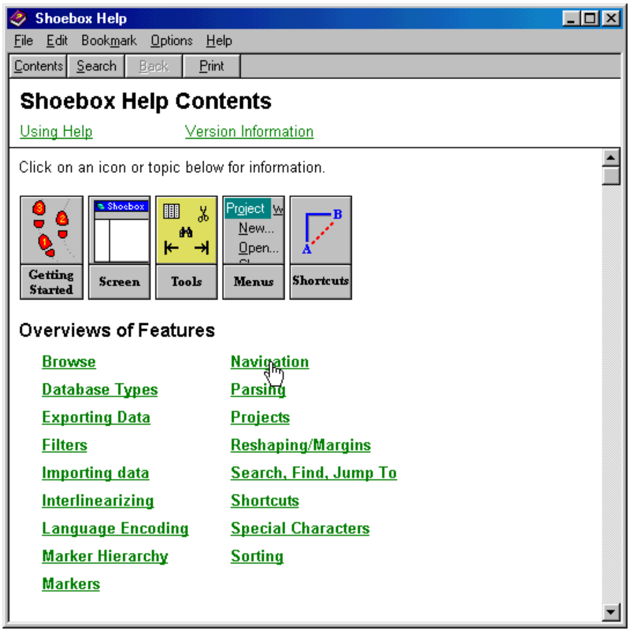
3. Close the Shoebox Help window.

4. Click Cancel on the Sorting by Fields dialog box to close the Sorting by Fields dialog box.

Help contents

Another way to access help information is through the table of contents.

1. Choose Help, Contents.



Shoebox Help Contents

Using Help Version Information

Click on an icon or topic below for information.

Overviews of Features

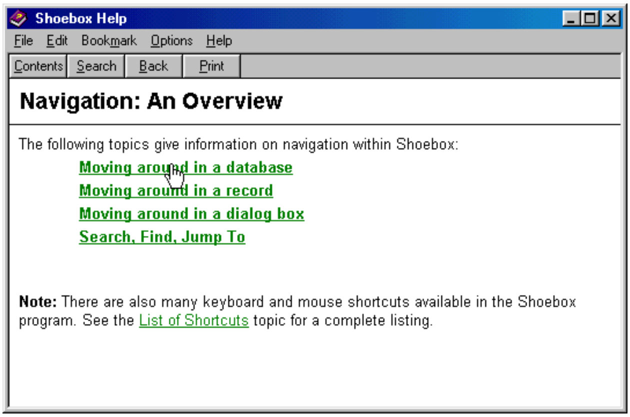
Browse	Navigation
Database Types	Parsing
Exporting Data	Projects
Filters	Reshaping/Margins
Importing data	Search, Find, Jump To
Interlinearizing	Shortcuts
Language Encoding	Special Characters
Marker Hierarchy	Sorting
Markers	

Fig. 31. Shoebox Help Contents

You will see a table of contents with a variety of topics that give you focused discussions on specific areas of the program. For example, under Overviews of Features you will see a topic named Navigation.

2. Click Navigation.

You will see more topics, including one named Moving around in a database.



Navigation: An Overview

The following topics give information on navigation within Shoebox:

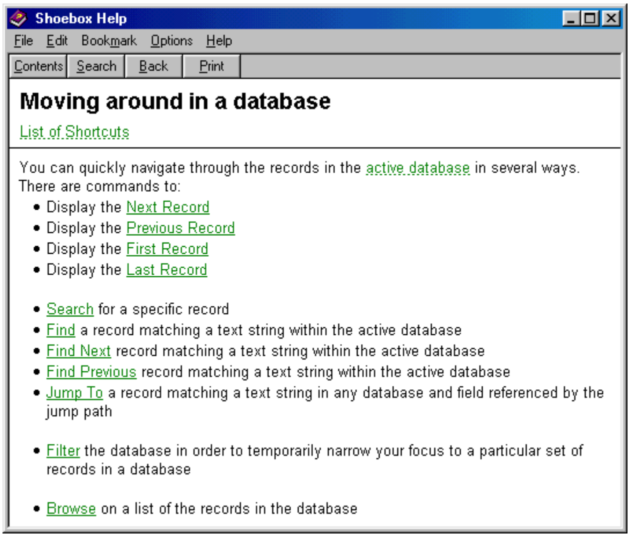
- Moving around in a database
- Moving around in a record
- Moving around in a dialog box
- Search, Find, Jump To

Note: There are also many keyboard and mouse shortcuts available in the Shoebox program. See the List of Shortcuts topic for a complete listing.

Fig. 32. Navigation: an Overview

3. Click **Moving around in a database** .

You will see a summary of ways to move in a database.



Moving around in a database

List of Shortcuts

You can quickly navigate through the records in the active database in several ways. There are commands to:

- Display the Next Record
- Display the Previous Record
- Display the First Record
- Display the Last Record

- Search for a specific record
- Find a record matching a text string within the active database
- Find Next record matching a text string within the active database
- Find Previous record matching a text string within the active database
- Jump To a record matching a text string in any database and field referenced by the jump path

- Filter the database in order to temporarily narrow your focus to a particular set of records in a database

- Browse on a list of the records in the database

Fig. 33. Moving around in a database

4. Do not close the Shoebox Help window yet.

Help links Help links are ways of getting from one help topic to other related topics. Any underlined word or phrase in a help window is a link to more information. For example, in the help window which is currently open, there are many words (like Next Record, Search, and Find) that are underlined.

1. Click Browse at the bottom of the help text.

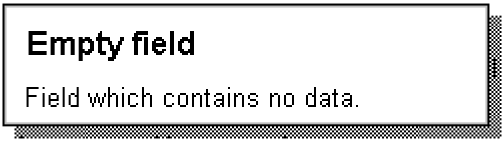
You will see the same discussion of Browse that you saw earlier, as well as links to other topics.

If a word or phrase has a solid underline, it will take you to a different topic.

Clicking Back will return you to the previous help topic.

If a word or phrase has a dotted underline, clicking this phrase will pop up a small window with a brief definition. The previous window remains on the screen. A second mouse-click will cancel the small window.

2. In the Browse help topic, click empty. You will see a small box with the definition of an empty field.



Empty field

Field which contains no data.

Fig. 34. Definition Box for an Empty field

3. Click anywhere to close it.

Another useful button at the top of the help window is Contents. Clicking this button takes you to the table of contents of the Help file.

4. Do not close the Shoebox Help window yet.

Help search

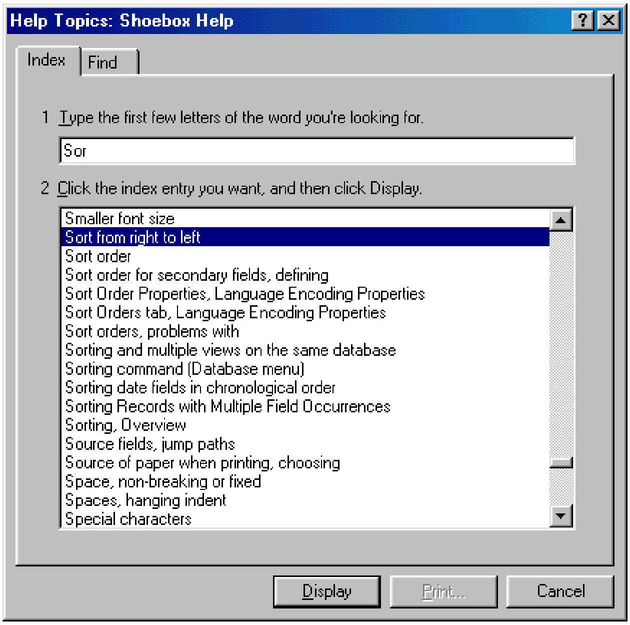
One of the most powerful ways of accessing help is with the help search, which is like the index of a book. Among the buttons at the top of the help window is one labeled Search.

To use the search feature to get help on sorting Shoebox databases:

1. Click Search at the top of the Shoebox Help window, which should still be open.

You will see a dialog box with a place to fill in a search topic and below it, an alphabetical list of topics.

2. Type `Sor` and observe that as you type a letter the list of topics jumps to the first one starting with the letters you have typed so far.



1 Type the first few letters of the word you're looking for.

Sor

2 Click the index entry you want, and then click Display.

Smaller font size
Sort from right to left
Sort order
Sort order for secondary fields, defining
Sort Order Properties, Language Encoding Properties
Sort Orders tab, Language Encoding Properties
Sort orders, problems with
Sorting and multiple views on the same database
Sorting command (Database menu)
Sorting date fields in chronological order
Sorting Records with Multiple Field Occurrences
Sorting, Overview
Source fields, jump paths
Source of paper when printing, choosing
Space, non-breaking or fixed
Spaces, hanging indent
Special characters

Fig. 35. Searching for a help topic

3. Look down this list to find and click Sorting command (Database menu).

The highlight will move from Sort from right to left to the item you selected, Sorting command (Database menu), and it will appear in the box where you previously typed `Sor`.

4. On Windows 95 and later, click Display. On Windows 3.1 or Macintosh, click Show Topics, and then click Go To.

You see the same Sorting help window you have seen before.

5. Close the Help window.

There are usually several ways to get to the windows you want. Another way to get to the Help Topics window:

6. Choose Help, Search for help on.

You will see the same Help Topics dialog box as before.

7. Click Cancel to return to Shoebox.

Field markers

Shoebox reads and writes database files in a text format called SIL Standard Format. This format has been used by SIL field workers for storing language data on computers since the late 1970s. Many other SIL programs use Standard Format markers to identify data in a similar way. This allows a smooth flow of information between programs.

In Standard Format, you use *field markers* to tag the different *kinds of information* in your database. Field markers are analogous to *styles* in word processors and *tags* in the Hypertext Markup Language (HTML) and Extensible Markup Language (XML) formats used on the World Wide Web.

Standard Format markers

Standard Format markers consist of a backslash character (\) at the beginning of a line, followed by one or more characters that are an abbreviation of the meaning of the marker. The marker is separated from its following data by a space. If you have trouble remembering the marker names, Shoebox allows you to display the names of the markers on your computer screen.

The following example shows a single record of a simple French dictionary entry using Standard Format markers. **\lx** is for a lexeme, **\ps** is for a part of speech, and **\ge** is for the English gloss:

```
\lx  garçon
\ps  n
\ge  boy
```

Garçon is the French word for boy and follows the lexeme field marker. The *n* signifies the part of speech is a noun.

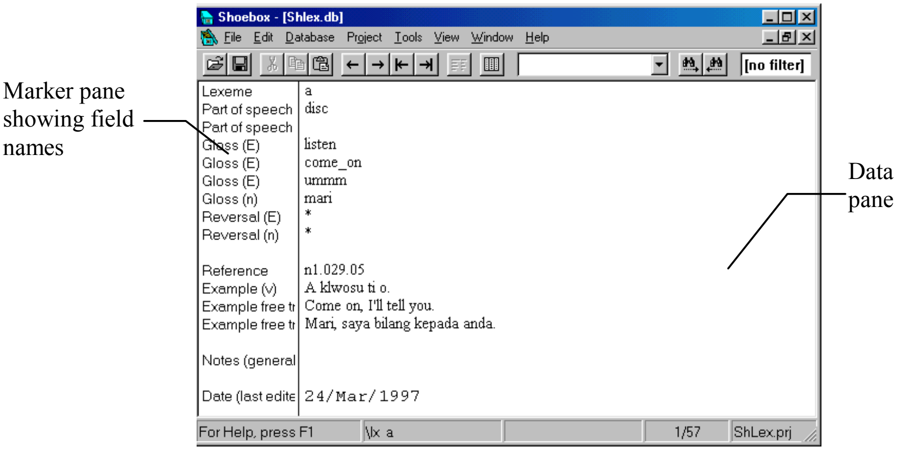
Records and fields

Shoebox organizes databases into records and fields. A data field consists of a Standard Format marker and the data that it marks (called the field contents). A record is a collection of related data fields. The record marker is the specially selected Standard Format marker that starts each new record. For example, most dictionaries have the **\lx** or lexeme field as their record marker.

A database is a collection of records. For example, a lexical database contains many records, usually one record for each word or morpheme in a language. In turn, each record contains various bits of information about its word or morpheme; these bits of information are the fields.

Field markers are necessary for identifying data types. Because they are repeated over and over again in every record, database designers abbreviate them to save disk space. As a result, field markers become too terse to be easily remembered. Shoebox gives the option of seeing the full name of the field along with or in place of the field marker, when you are willing to lose some viewing space for your data. For example,

1. Choose View, Field Names.


Marker pane showing field names
Data pane
Shoebox - [Shlex.db]
File Edit Database Project Tools View Window Help
[no filter]
For Help, press F1
\lx a
1/57
ShLex.prj

Lexeme	a
Part of speech	disc
Part of speech	
Gloss (E)	listen
Gloss (E)	come_on
Gloss (E)	ummm
Gloss (n)	mari
Reversal (E)	*
Reversal (n)	*
Reference	n1.029.05
Example (v)	A klwosu ti o.
Example free tr	Come on, I'll tell you.
Example free tr	Mari, saya bilang kepada anda.
Notes (general	
Date (last edite	24/Mar/1997

Fig. 36. Shlex Database file using the View, Field Names option

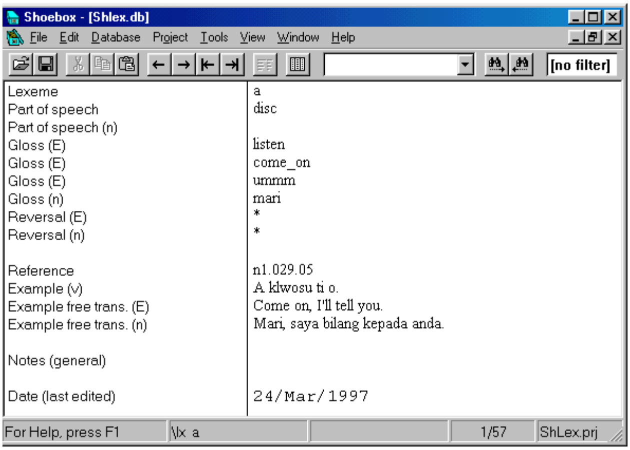
You see the vertical line separating the field markers from their data moved over to the right. The field names replaced the field markers in the marker pane.

The line dividing the marker pane and the data pane can be moved. To move the line,

2. Place the point of the mouse arrow on the vertical dividing line. The mouse pointer should look like this ↔.

3. While the mouse pointer looks like a double-headed arrow, hold down the left mouse button and drag the line to the right an inch or so.

You will notice that some field names are very long.



Lexeme	a
Part of speech	disc
Part of speech (n)	
Gloss (E)	listen
Gloss (E)	come_on
Gloss (E)	ummm
Gloss (n)	mari
Reversal (E)	*
Reversal (n)	*
Reference	n1.029.05
Example (v)	A klwosu ti o.
Example free trans. (E)	Come on, I'll tell you.
Example free trans. (n)	Mari, saya bilang kepada anda.
Notes (general)	
Date (last edited)	24/Mar/1997

Fig. 37. Shlex database with the dividing line moved right

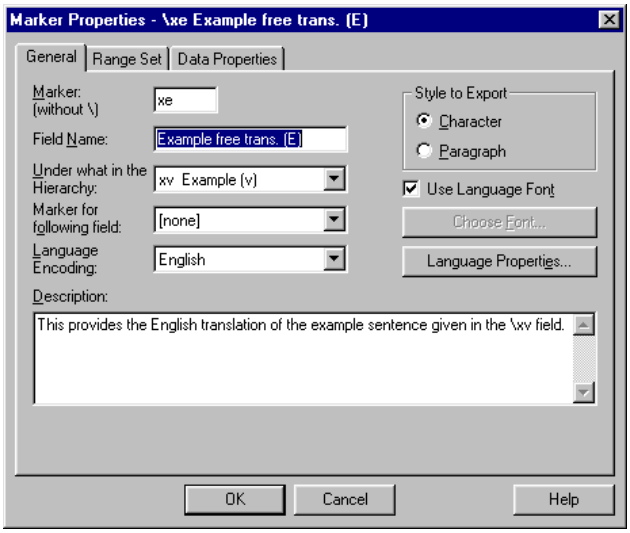
Notice that the field name **Example free trans. (E)** is very long. This is the 12th field down. If you cannot see it, use the vertical scroll bar to scroll down.

Changing a field name

Sometimes you will want to change the field name. Do this:

1. Right-click the field name **Example free trans. (E)**. (Macintosh: OPTION+click).

You will see a dialog box containing various properties of the **\xe** marker.



Marker Properties - \xe Example free trans. (E)

Marker: (without \) xe

Field Name: Example free trans. (E)

Under what in the Hierarchy: xv Example (v)

Marker for following field: [none]

Language Encoding: English

Style to Export
● Character
○ Paragraph

☑ Use Language Font

Description:
This provides the English translation of the example sentence given in the \xv field.

Fig. 38. Marker Properties window

2. In the Field Name box, type `Ex (E) free trans`.

3. Click OK.

You are now viewing the Shlex.db window again.

Viewing both markers and names

To change the view of the marker pane to show both the field marker and its corresponding name,

1. Choose View, Both Markers and Names.

To change the view back to show only the markers,

2. Choose View, Markers.

Notice the vertical line dividing the marker pane has readjusted itself back to the left.

Review

To review the procedure for changing a field name, try the following:

1. Change the field name for the **\xe** marker back to **Example free trans. (E)**. (Hint: Start by right-clicking the **\xe** field marker.)

2. Click OK when done to close the Marker Properties dialog box and return to the Shlex.db window.

Your Shlex database window should look as it did when you started.

Inserting data fields

You can insert a field using the same keystrokes you would use to insert it in any text file. For example, if you wanted to insert another **\ge** field in this record,

1. Move the insertion point to the right end of the line in the first **\ge** field. (This is the word *listen* if you are still in the first record.)

Note: If you are in a record other than the first, you can move to the first record of your database by clicking ⇤ on the toolbar.

2. With the insertion point at the end of the word *listen*, press ENTER (Macintosh: press RETURN).

3. Type a backslash (\).

Notice that a box opens up in the marker pane.

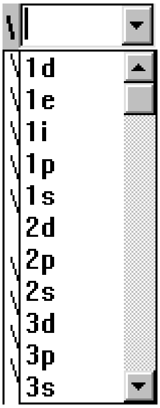
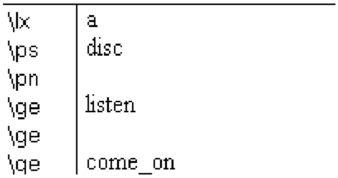


Fig. 39. Marker Entry Box

4. Type `ge` followed by a space.

Notice that when you type the space, the box closes and the insertion point moves out of the marker pane into the data pane. Now you are ready to type a data entry into your second **\ge** field.



```
\lx     a
\ps     disc
\pn
\ge     listen
\ge
\ge     come_on
```

Fig. 40. New marker insertion

Deleting data fields

Suppose you now decide you did not want the **\ge** field you just inserted.

1. Click the **\ge** marker of the field you want to delete.

Notice that the marker **\ge** is highlighted.

2. Press DELETE.

The marker vanishes and the insertion point is placed at the end of the previous field.

Inserting a field with a new marker

Shoebox keeps a list of all the field markers. The drop-down list used during marker entry has a scroll bar that allows you to move down through a list of the field markers. You can select a marker from this list by clicking it.

The marker list is also used for checking new field markers as they are inserted. For example,

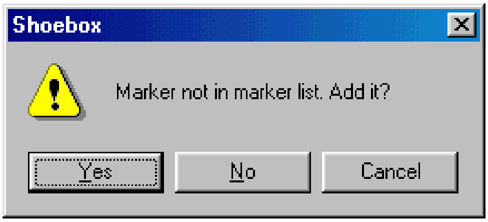
1. At the end of the word *listen* press ENTER (Macintosh: press RETURN) and type a backslash (\).

The marker box will drop down to display the marker list.

2. Type the marker xx in the box.

3. Press SPACEBAR.

When you type the xx and a space, Shoebox will ask you if you want to add a new field marker.



Shoebox

Marker not in marker list. Add it?

Fig. 41. New marker confirmation box

4. Click No.

Observe that the insertion point stays in the marker box and allows you to change the **xx** to something else.

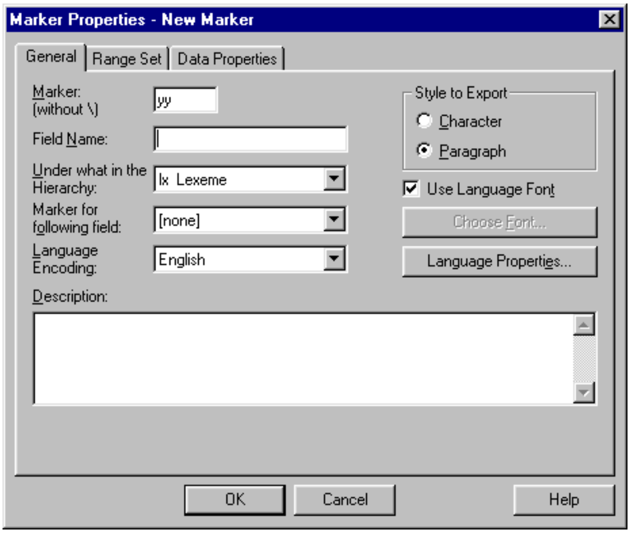
5. Change the **xx** marker to **yy**.

6. Press SPACEBAR.

Shoebox will again ask you if you want to add a new field marker.

7. Click Yes to add the marker to the list.

You will see the dialog box for the properties of the **yy** marker.



Marker Properties - New Marker

Marker: (without \) yy

Field Name:

Under what in the Hierarchy: lx Lexeme

Marker for following field: [none]

Language Encoding: English

Style to Export: ○ Character ● Paragraph

☑ Use Language Font

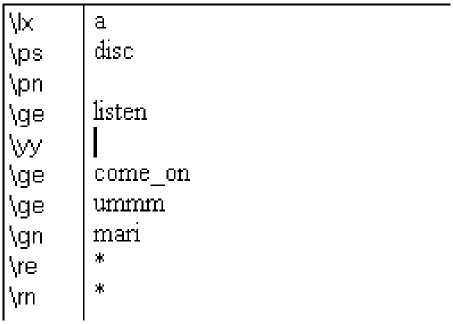
Description:

Fig. 42. Marker Properties dialog box

8. Type `temporary` as the Field Name.

9. Click OK.

You now have a **\yy** field in the record at the location where you inserted it.



\lx	a
\ps	disc
\pn	
\ge	listen
\yy	
\ge	come_on
\ge	ummm
\gn	mari
\re	*
\rn	*

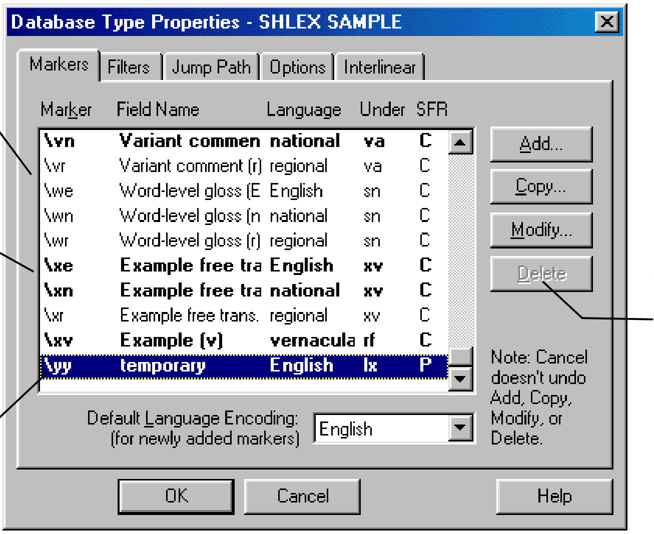
Fig. 43. Shlex record with new **\yy** field

Field marker list

The main field marker list, with more information than is available in the drop-down list box, is shown in the Markers dialog box.

1. Choose Database, Properties.

Shoebox will display the Database Type Properties dialog box showing a complete list of all field markers used in the file.



Marker	Field Name	Language	Under	SFR
\vn	**Variant commen**	**national**	**va**	**C**
\vr	Variant comment (r)	regional	va	C
\we	Word-level gloss (E	English	sn	C
\wn	Word-level gloss (n	national	sn	C
\wr	Word-level gloss (r)	regional	sn	C
\xe	**Example free tra**	**English**	**xv**	**C**
\xn	**Example free tra**	**national**	**xv**	**C**
\xr	Example free trans.	regional	xv	C
\xv	**Example (v)**	**vernacula**	**rf**	**C**
\yy	**temporary**	**English**	**lx**	**P**

Default Language Encoding: (for newly added markers) English

Note: Cancel doesn't undo Add, Copy, Modify, or Delete.

Unused markers in Shlex database

Markers used in Shlex database

Active field when you opened this window

Unable to delete an active marker

Fig. 44. Database Type Properties window

Notice that the **\yy** marker, which is highlighted in the list, is the marker for the field the insertion point was in when you brought up the Markers dialog box.

Notice also that some of the marker lines are bold and some are in normal (lighter) type. Each marker in bold occurs in at least one record of the current database; the markers in normal type are not currently being used.[6]

Notice the buttons on the right side of the marker list. The Delete button is grayed out. The **\yy** marker is used by at least one record in your database. You cannot delete from the list a marker that is in use.

2. Click OK to return to the database.

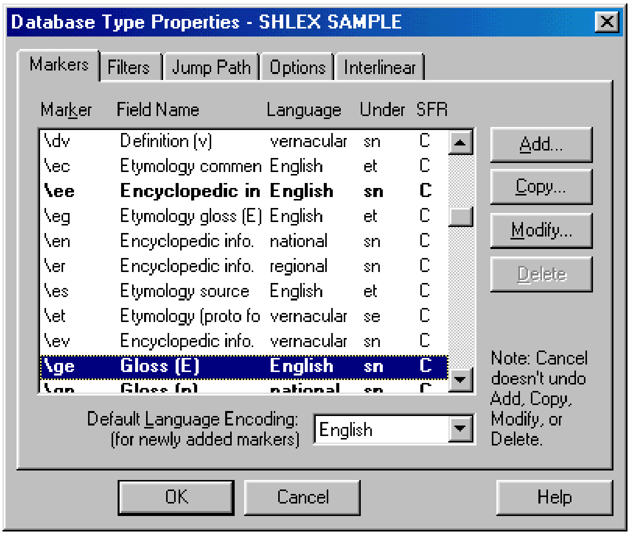
[6] The list includes all the field markers in the Multi-Dictionary Formatter (MDF) package. They are recommended if you want to use other functions of Shoebox like the dictionary printout.

3. Delete the **\yy** field from the record. (Click the field marker once to select it and press DELETE).

The **\yy** field marker still remains in your list of possible markers. You can now delete it also.

4. Click Database, Properties.

Since the insertion point is now on a different marker, that marker will be highlighted.



Marker	Field Name	Language	Under	SFR
\dv	Definition (v)	vernacular	sn	C
\ec	Etymology commen	English	et	C
\ee	**Encyclopedic in**	**English**	**sn**	**C**
\eg	Etymology gloss (E)	English	et	C
\en	Encyclopedic info.	national	sn	C
\er	Encyclopedic info.	regional	sn	C
\es	Etymology source	English	et	C
\et	Etymology (proto fo	vernacular	se	C
\ev	Encyclopedic info.	vernacular	sn	C
\ge	**Gloss (E)**	**English**	**sn**	**C**
\gn	Gloss (n)	national	sn	C

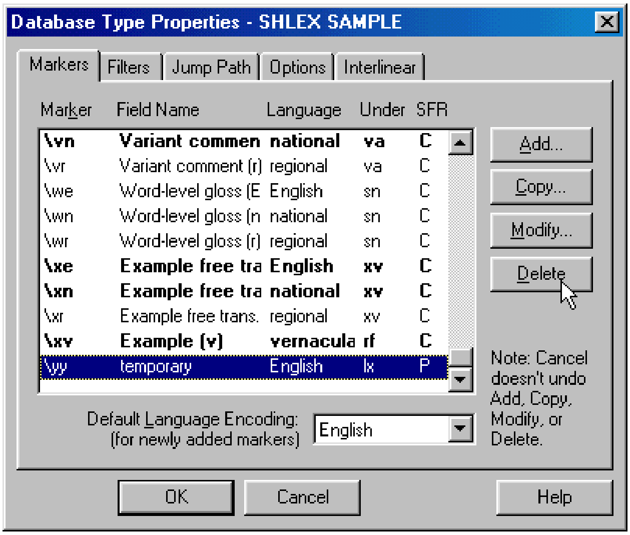
Default Language Encoding: (for newly added markers) English

Note: Cancel doesn't undo Add, Copy, Modify, or Delete.

Fig. 45. Database Type Properties, **\ge** selected

5. Scroll down to the **\yy** marker.

Notice that the marker line is not bold. No record is currently using that marker. When you click the line with the **\yy** marker, Delete will no longer be grayed out.



Database Type Properties - SHLEX SAMPLE

Marker	Field Name	Language	Under	SFR
\vn	**Variant commen**	**national**	**va**	**C**
\vr	Variant comment (r)	regional	va	C
\we	Word-level gloss (E	English	sn	C
\wn	Word-level gloss (n	national	sn	C
\wr	Word-level gloss (r)	regional	sn	C
\xe	**Example free tra**	**English**	**xv**	**C**
\xn	**Example free tra**	**national**	**xv**	**C**
\xr	Example free trans.	regional	xv	C
\xv	**Example (v)**	**vernacula**	**rf**	**C**
\yy	temporary	English	lx	P

Default Language Encoding: English
(for newly added markers)

Note: Cancel doesn't undo Add, Copy, Modify, or Delete.

Fig. 46. Database Type Properties, **\yy** selected

6. Click Delete.

The **\yy** marker disappears from the list of available markers.

7. Click OK.

You should now be back in your starting Shlex.db window.

Changing a field marker

Here is the procedure for changing a data field's marker:

1. Place the insertion point in front of the word *listen*.
2. Press BACKSPACE.

The marker box opens with **ge** highlighted.

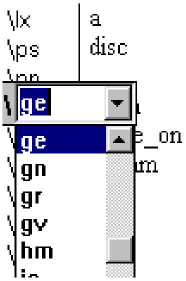


Fig. 47. Marker Entry List Box

3. Type gn and press SPACEBAR.

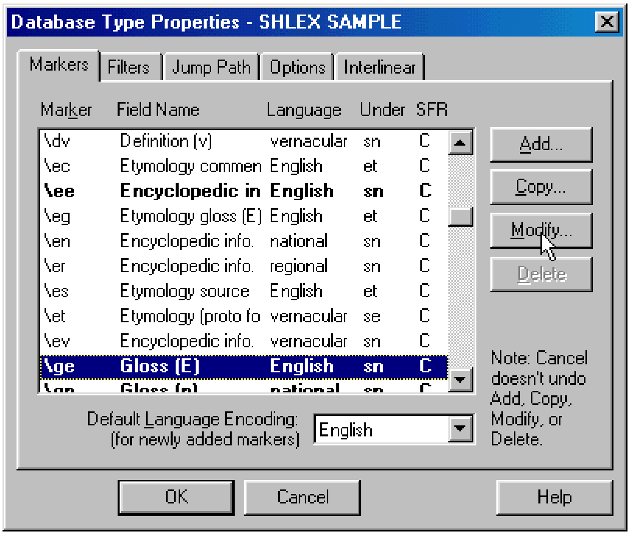
The marker is changed.

4. Change the marker back to **ge** using the procedure just described.

Changing a field marker everywhere

A marker can be changed in all the open databases of the same type. This should be done with caution. Unopened databases of the same type will not be changed. For example, you might want to change **\ge** to **\eng** (for English). To do that,

1. Choose Database, Properties.



Database Type Properties - SHLEX SAMPLE

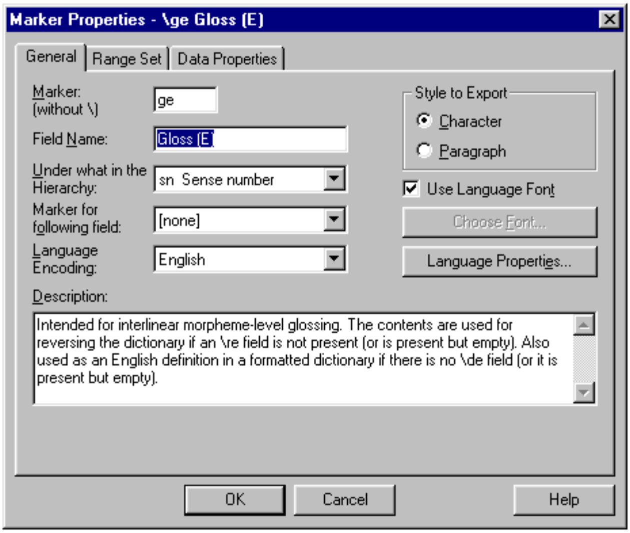
Marker	Field Name	Language	Under	SFR
\dv	Definition (v)	vernacular	sn	C
\ec	Etymology commen	English	et	C
\ee	**Encyclopedic in**	**English**	**sn**	**C**
\eg	Etymology gloss (E)	English	et	C
\en	Encyclopedic info.	national	sn	C
\er	Encyclopedic info.	regional	sn	C
\es	Etymology source	English	et	C
\et	Etymology (proto fo	vernacular	se	C
\ev	Encyclopedic info.	vernacular	sn	C
\ge	Gloss (E)	English	sn	C
\gn	Gloss (n)	national	sn	C

Note: Cancel doesn't undo Add, Copy, Modify, or Delete.

Default Language Encoding: (for newly added markers) English

Fig. 48. Database Type Properties Window

2. Select the **ge** marker and click Modify... (or double-click the **ge** line).



Marker Properties - \ge Gloss (E)

Marker: (without \) ge

Field Name: Gloss (E)

Under what in the Hierarchy: sn Sense number

Marker for following field: [none]

Language Encoding: English

Style to Export

● Character

○ Paragraph

☑ Use Language Font

Description:

Intended for interlinear morpheme-level glossing. The contents are used for reversing the dictionary if an \re field is not present (or is present but empty). Also used as an English definition in a formatted dictionary if there is no \de field (or it is present but empty).

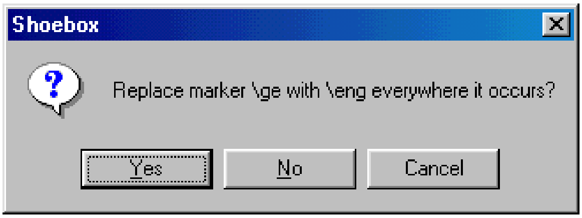
Fig. 49. Marker Properties dialog box

This is the same dialog box you reached previously by right-clicking the marker in the text window (Macintosh: OPTION+click).

3. Click in the Marker box and change the marker from **ge** to **eng**.

4. Click OK.

You will see a dialog box asking if you want to replace marker **\ge** with **\eng** everywhere it occurs.



Replace marker \ge with \eng everywhere it occurs?

Fig. 50. Replace Marker Confirmation Box

5. Click Yes.

6. Click OK in the Database Type Properties dialog box.

You will not want to leave this marker changed permanently, so you should change it back.

7. Change the **\eng** marker back to **\ge** everywhere. Hint: You can start by choosing Database, Properties.

Tutorial milestone

This is the first milestone. Up to this point you have undone every change you have made, so if at some later time you want to restart here, you can do it by reopening the project Shlex.prj.

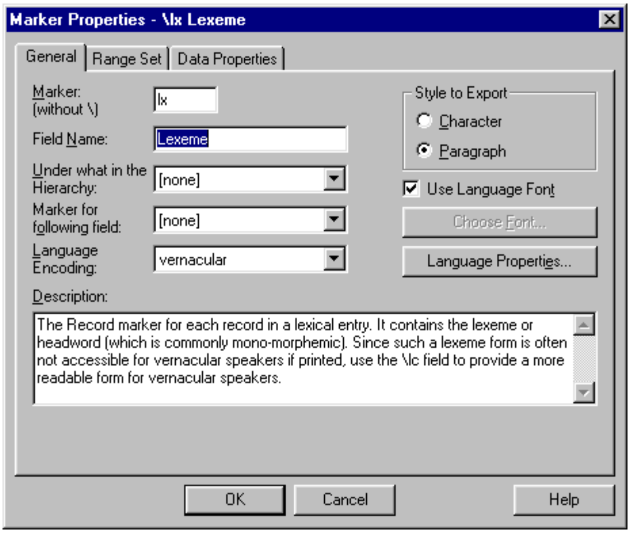
Note: If any data showing in this project is bold or italic or colored, then the next section has already been done. In this case, you should reinstall the samples. For more information, see *Appendix D: Reinstalling the samples*.

Field markers and language

All of your data is encoded in some language. It may be French, English, Swahili, etc. In a multilingual dictionary, different fields contain data from different languages.[7] Shoebox associates information about each field's language with its field marker. This language information includes things like: sorting order, the default font, and the keyboard. Shoebox calls this language information language encoding (see *Language encodings*, p. 113).

The Marker Properties dialog box allows you to review and change the language encoding associated with each marker.

1. Right-click the **\lx** marker (Macintosh: OPTION+click).



Marker Properties - \lx Lexeme

General | Range Set | Data Properties

Marker: (without \) lx

Field Name: Lexeme

Under what in the Hierarchy: [none]

Marker for following field: [none]

Language Encoding: vernacular

Description:

The Record marker for each record in a lexical entry. It contains the lexeme or headword (which is commonly mono-morphemic). Since such a lexeme form is often not accessible for vernacular speakers if printed, use the \lc field to provide a more readable form for vernacular speakers.

Style to Export: ○ Character ● Paragraph

☑ Use Language Font

Fig. 51. Marker Properties dialog box

Shoebox will display the Marker Properties dialog box. In the left side of the dialog box you will see the label Language Encoding. Next to it is a drop-down list containing the name of the language encoding associated with the **\lx** marker. It should say vernacular.

[7] It is necessary and desirable to change languages within some fields. That is a different issue which is discussed in answer to one of the questions on *Fonts*, p. 240, in the section *Frequently asked questions*.

2. Click the down arrow to see the other available language encodings.

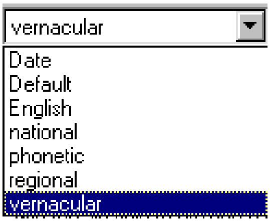
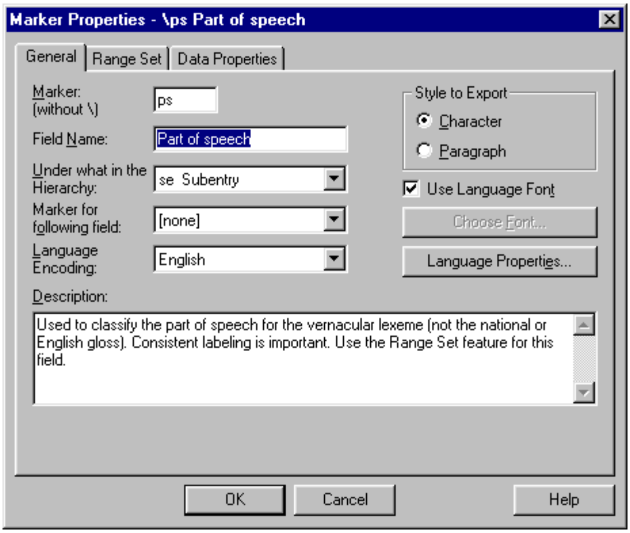


Fig. 52. Available language encodings for Shlex.db

3. Close the drop-down list without changing the language encoding from vernacular. Click Cancel again to close the Marker Properties dialog box for the **lx** marker.

4. Right-click the **ps** marker to bring up the Marker Properties dialog box (Macintosh: OPTION+click).



Marker: (without \) ps

Field Name: Part of speech

Under what in the Hierarchy: se Subentry

Marker for following field: [none]

Language Encoding: English

Style to Export: ● Character ○ Paragraph

☑ Use Language Font

Description:
Used to classify the part of speech for the vernacular lexeme (not the national or English gloss). Consistent labeling is important. Use the Range Set feature for this field.

Fig. 53. Marker Properties for Parts of Speech

This marker should have a language encoding of English.

5. Click Cancel to close the Marker Properties dialog box.

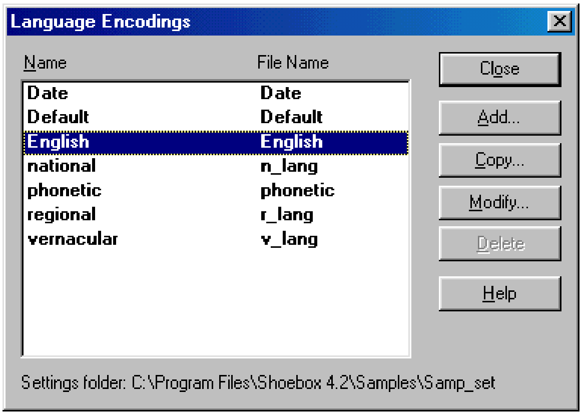
Language font

Two types of field-font relationships are available in Shoebox:

- A field can have specific font and font style.
- All the fields of a language can have the same font and font style. (This is the most common relationship.)

In one convention for making dictionaries, all words in the language being described are displayed using bold formatting. In the Multi-Dictionary Formatter, this language is called the vernacular. To change the font of the vernacular language to bold,

1. Choose Project, Language Encodings.



Name	File Name
Date	Date
Default	Default
English	English
national	n_lang
phonetic	phonetic
regional	r_lang
vernacular	v_lang

Settings folder: C:\Program Files\Shoebox 4.2\Samples\Samp_set

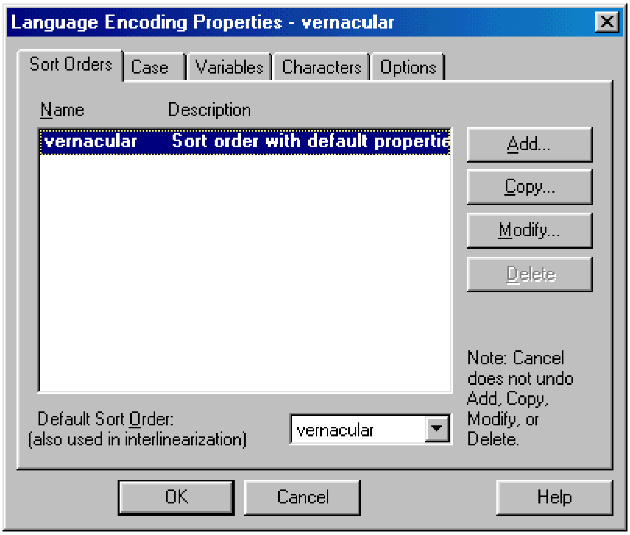
Fig. 54. Language Encodings dialog box

You will see a list of all the available language encoding settings. These are not built into the program, so you can make your own when you start on your own data.[8]

2. Select vernacular.
3. Click Modify....

[8] An advantage of keeping all of the project files together in the same folder is that they can share the various language encodings you develop—you do not have to keep recreating them for each project.

This brings you to a tabbed dialog box titled Language Encoding Properties - vernacular. The various tabs in this box show all the characteristics of the language encoding. The standard font for data in this language is on the Options tab.



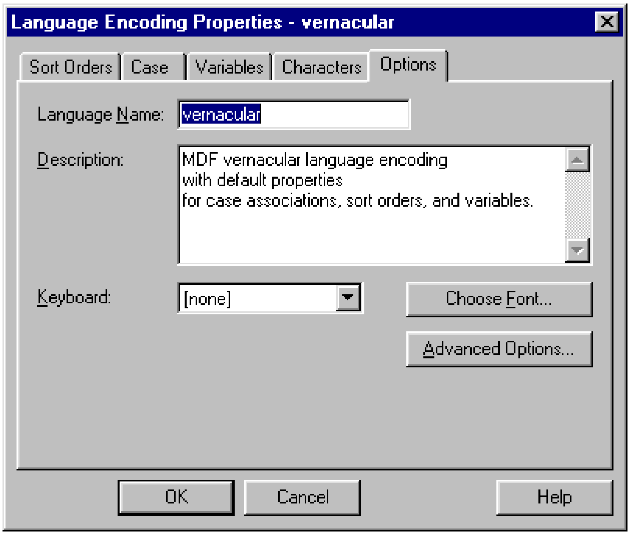
Name	Description
vernacular	Sort order with default propertie

Default Sort Order: (also used in interlinearization) vernacular

Note: Cancel does not undo Add, Copy, Modify, or Delete.

Fig. 55. Language Encoding Properties - vernacular

4. Click the Options tab.



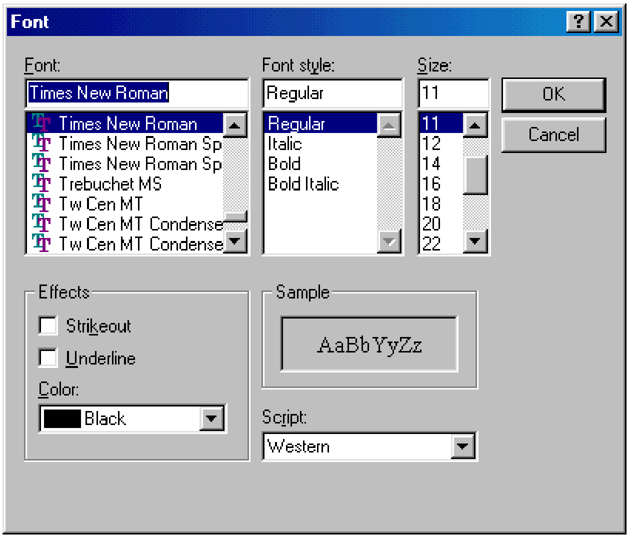
Language Name: vernacular

Description: MDF vernacular language encoding
with default properties
for case associations, sort orders, and variables.

Keyboard: [none]

Fig. 56. Language Encoding Properties - vernacular dialog box

5. Click Choose Font... .



Font: Times New Roman

Font style: Regular

Size: 11

□ Strikeout

□ Underline

Color: Black

Script: Western

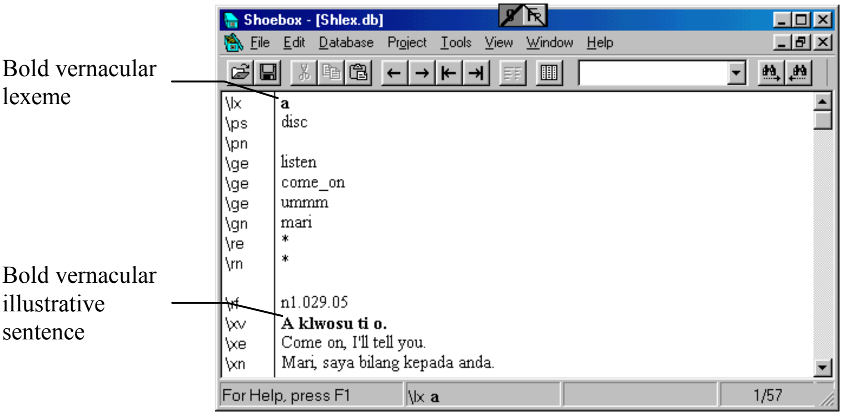
Fig. 57. Font box

6. Under font style, select Bold.

7. Click OK to close the Font dialog box.

8. Click OK again to close the Language Encoding Properties dialog box and to return to the Language Encodings dialog box.

9. Click Close to get back to the database.


Bold vernacular lexeme
Bold vernacular illustrative sentence
Shoebox - [Shlex.db]
File Edit Database Project Tools View Window Help
For Help, press F1
\lx a
1/57

\lx **a**
\ps disc
\pn
\ge listen
\ge come_on
\ge ummm
\gn mari
\re *
\rn *

\rf n1.029.05
\xv **A klwosu ti o.**
\xe Come on, I'll tell you.
\xn Mari, saya bilang kepada anda.

Fig. 58. Shlex Database with bold vernacular entries

Notice that now all vernacular language fields are bold. This will be more apparent if you scroll down the record so that you can see the illustrative sentence, marked **\xv**.

A convenient way to get to the language encoding of a language is through the properties of a marker in the language. You will use that approach to change all national language fields to italic.

10. Right-click the **gn** marker (Macintosh: OPTION+click). Shoebox displays the Marker Properties dialog box. The language encoding is national.

11. Click Language Properties... to bring up the properties of the national language encoding.

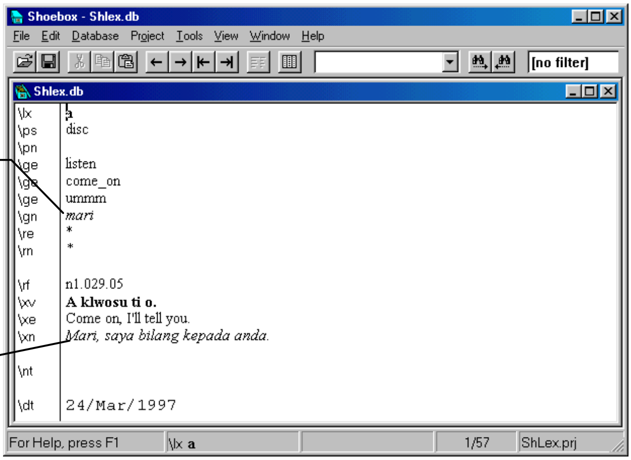
12. Click the Options tab.

13. Click Choose Font....

14. Under font style, select Italic.

15. Click OK until all dialog boxes are closed.

If you scroll down the record you will see that all national language fields are now italic.



```
\lx    a
\ps    disc
\pn
\ge    listen
\ge    come_on
\ge    ummm
\gn    mari
\re    *
\rn    *

\rf    n1.029.05
\xv    A klwosu ti o.
\xe    Come on, I'll tell you.
\xn    Mari, saya bilang kepada anda.

\nt

\dt    24/Mar/1997
```

For Help, press F1 | \lx a | 1/57 | ShLex.prj

Italic national language gloss

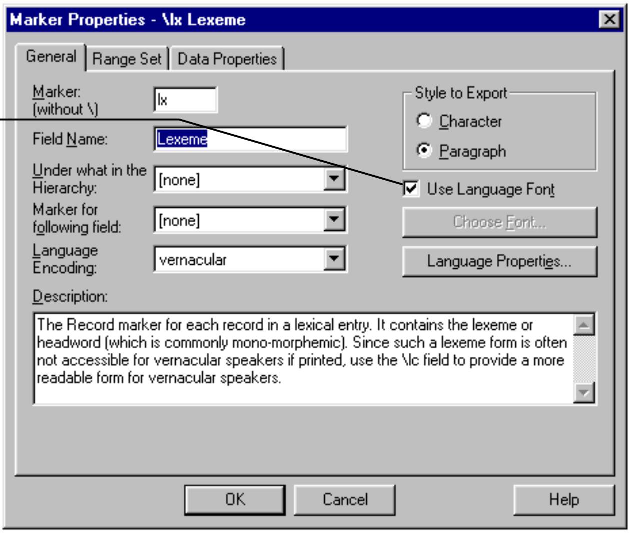
Italic national language translation

Fig. 59. Database Window with Italic National Field Fonts

Unique field font

It can be useful to associate an individual marker with a special font style to make its data stand out in some way. For example, you may want to have the lexeme larger than the other entries so you can read it at a quick glance.

1. Right-click the **lx** marker (Macintosh: OPTION+click). Shoebox displays the Marker Properties dialog box.



Marker: (without \) lx

Field Name: Lexeme

Under what in the Hierarchy: [none]

Marker for following field: [none]

Language Encoding: vernacular

Style to Export: ○ Character ● Paragraph

☑ Use Language Font

Description:

The Record marker for each record in a lexical entry. It contains the lexeme or headword (which is commonly mono-morphemic). Since such a lexeme form is often not accessible for vernacular speakers if printed, use the \lc field to provide a more readable form for vernacular speakers.

Use Language Font check box

Fig. 60. Marker Properties dialog box for the lexeme field

On the right side of the Marker Properties box is a check box labeled Use Language Font. This instructs Shoebox to use the default font specified for the vernacular language. If you want something different from the default selection for the lexeme field, you must turn off this option and then specify your choice. To change the lexeme field to a larger font,

2. Click ☑ Use Language Font. Notice the check disappears from the Use Language Font box and the Choose Font button is no longer grayed out.

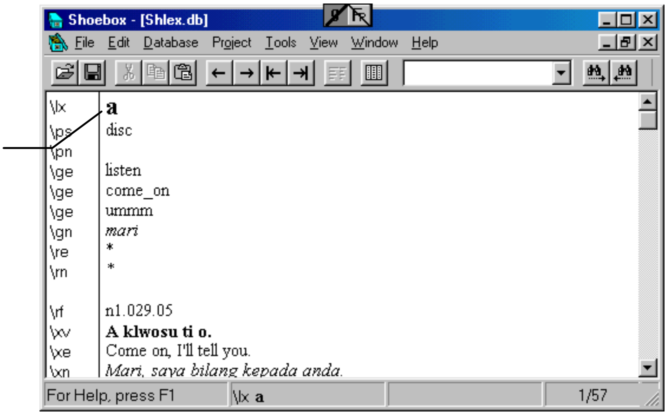
3. Click Choose Font....

You will see a Font dialog box.

4. Change the size to 16.

5. Click OK two times. You are back at the Shlex.db window.

Lexeme entry with a larger font



\lx a
\ps disc
\pn
\ge listen
\ge come_on
\ge ummm
\gn *mari*
\re *
\rn *

\rf n1.029.05
\xv **A klwosu ti o.**
\xe Come on, I'll tell you.
\xn *Mari, saya bilang kepada anda.*

Fig. 61. Shlex database with a large font for the lexeme entry

You will see the lexeme displayed much larger than the rest of the fields.

6. Change the lexeme font back to Use Language Font. Hint: To do this start by right-clicking **\lx** (Macintosh: OPTION+click).

Moving around in the database

You will be moving around in the database now. First notice the set of buttons with arrows in the toolbar.

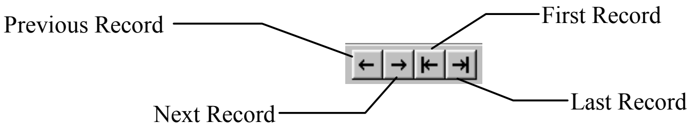


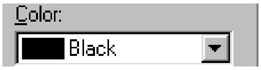
Fig. 62. Movement buttons from the toolbar

One points left, one right, with two more pointing at vertical lines. The left arrow moves the Previous record into your window. The right arrow moves the Next record into your window. The left arrow with the vertical bar moves the First record of your database into your window. The right arrow with the bar moves the Last record of your database into your window. Shoebox normally shows only one record at a time in a window. This is a convenient means of moving from record to record. (These commands are also available on the Database menu.)

1. Click ⇤ to go to the first record of the database.
2. Click → two times to locate the third record (the second occurrence of the *-a* lexeme). When you are at the third record 3/57 shows as the record number on the status bar.

Locate the national part of speech (**\pn** field). If you are having trouble finding it, look at the fourth line down.

3. Right-click the **pn** marker (Macintosh: OPTION+click).
4. Click ☑ Use Language Font check box to clear it.
5. Click Choose Font....
6. Locate the box labeled Color.



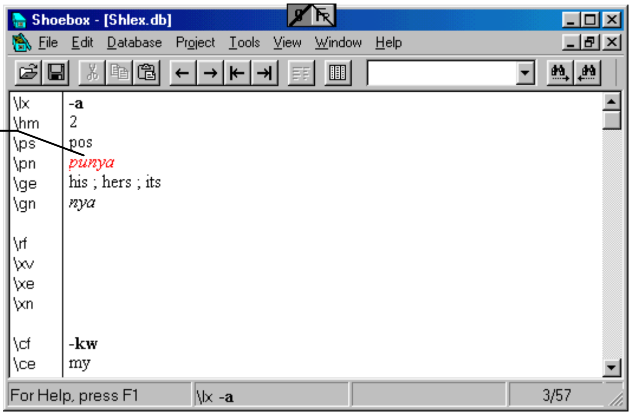
Color:
■ Black ▼

Fig. 63. Color selection box

The current color is black. Change it to red.

7. Click the down arrow beside the Color.
8. Select Red.[9]
9. Click OK two times.

[9] If you have a monochrome monitor, choose a color which will provide the most contrast, or do something else to provide noticeable contrast such as underlining, increasing the point size, or changing to a different font altogether.



```
\lx    -a
\hm    2
\ps    pos
\pn    punya
\ge    his ; hers ; its
\gn    nya

\rf
\xv
\xe
\xn

\cf    -kw
\ce    my
```

\lx -a

National part of speech displayed with red font

Fig. 64. Shlex record with red national part of speech

You are now back in the third record of the Shlex database. The national part of speech now clearly stands out on a color display.

10. Click → several times while watching the screen. Notice how easy it is to recognize a national part of speech filled in for a particular record. (Some records have the marker but no contents.)

This feature is very effective, especially as a short-term way to locate a field you want to focus on. When that field is no longer the focus of your work, you can change it back to the language font. You can also choose to set apart certain fields permanently, e.g., marking subentry fields so they always stand out clearly—like a section heading in a book.

Multiple views

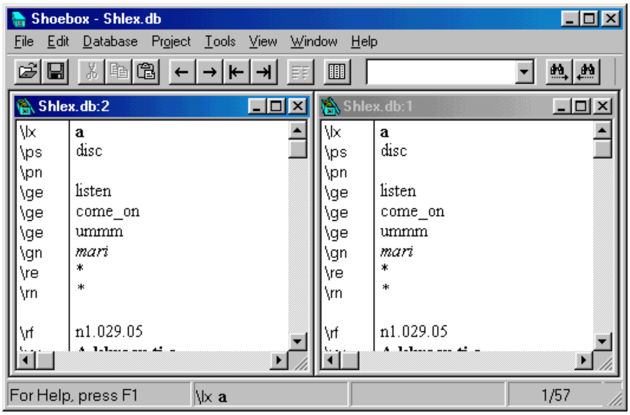
Shoebox can show two or more views of the same database in different windows. For example,

1. Click ⇤ to move to the top of the database.
2. Choose Window, Duplicate.

Shoebox displays the same record in a new window.[10] Since there are now two windows displaying the Shlex database, their title bars show Shlex.db:1 and Shlex.db:2.

Fig. 65. Shlex.db:2 title bar

3. Choose Window, Tile Side by Side.



```
\lx   a
\ps   disc
\pn
\ge   listen
\ge   come_on
\ge   ummm
\gn   mari
\re   *
\rn   *

\rf   n1.029.05
```

Fig. 66. Duplicate windows tiled side by side

This is a good layout for viewing multiple records.

Notice that the title bar of one of the windows is highlighted. This is the active window. The title bar of the other window appears dimmed. Many functions (such as moving between records, searching, etc.) work only in the active window. If a function does something unexpected, check to make sure that the window in which you are intending to work is the active window.

4. Change the word *listen* in the **\ge** field of the first database in the left window to *list*. Notice it changes in both windows.

[10] If you had the Shlex.db window maximized, the second window appears directly on top of it. It may not be obvious what happened, since the only apparent difference will be in the title bar.

5. Switch to the other window and change the word *list* after the **\ge** marker back to *listen*. Both windows show the word *listen* in the **\ge** marker field.

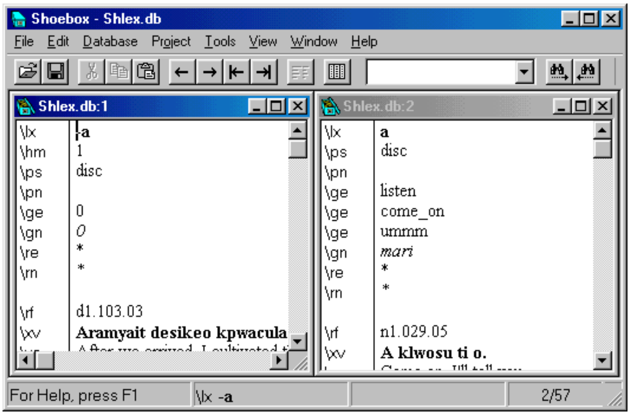
6. Click → on the toolbar.

Note: You moved to a different record in the *active* window, the last window you were working in. The record being displayed in the *inactive* window did not change. Remember that the *active* window has the highlighted background on the title bar, while the title bar of the *inactive* window appears dimmed.

For the rest of this section of the tutorial, the windows will be referred to as window 1 and window 2. You might find it is more natural to have window 1 on the left. To swap the positions of the two windows,

7. Click somewhere in the right window.

8. Click Window, Tile Side by Side.



```
\lx     -a
\hm     1
\ps     disc
\pn
\ge     0
\gn     0
\re     *
\rn     *

\rf     d1.103.03
\xv     Aramyait desikeo kpwacula
```

```
\lx     a
\ps     disc
\pn
\ge     listen
\ge     come_on
\ge     ummm
\gn     mari
\re     *
\rn     *

\rf     n1.029.05
\xv     A klwosu ti o.
```

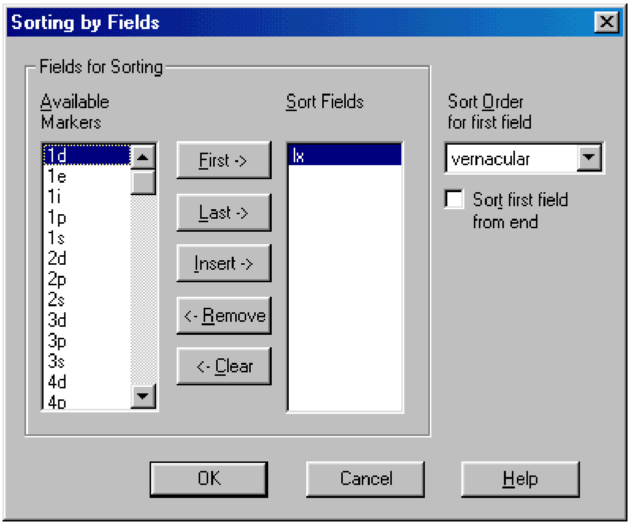
Fig. 67. Repositioned side-by-side windows

Sorting

To make window 2 sort by the gloss field while window 1 stays as it was,

1. Click somewhere in window 2 to make it the active window.

2. Choose Database, Sorting.



Sort Fields: lx

Sort Order for first field: vernacular

□ Sort first field from end

Fig. 68. Sorting by Fields window

Notice that sorting is by the **\lx** field (the record marker).

3. Click <- Clear.

4. Scroll down and select **ge** in the Available Markers list.

5. Click First -> to move it into the Sort Fields list.

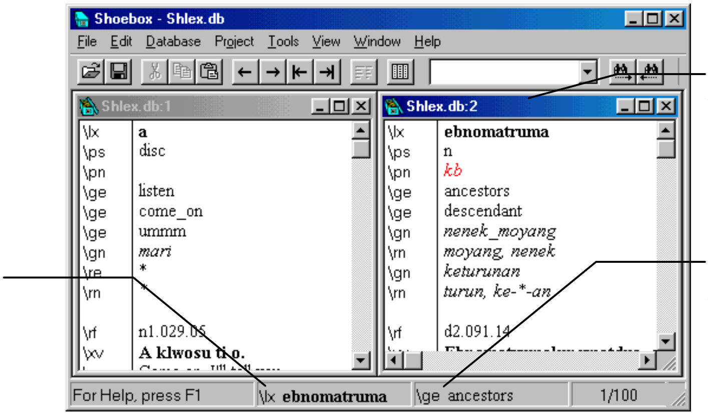
The result should be that only **ge** is in the Sort Fields list. (You can also move items from one list to the other by double-clicking them.)

6. Click OK.

Window 2 displays an index to the lexical data based on the **\ge** field. You can use it to search for a particular lexeme by its glosses or to search for lexemes that have a gloss in common.

7. Click ⇤, ←, and → a few times to observe that window 2 is now sorted by the contents of the English gloss field. The first record in window 2 is now *ebnomatruma* with the English

gloss of *ancestors* instead of *a* with an English gloss of *listen* shown in window 1.



```
\lx  a
\ps  disc
\pn
\ge  listen
\ge  come_on
\ge  ummm
\gn  mari
\re  *
\rn  *

\rf  n1.029.05
\xv  A klwosu ti o.
```

```
\lx  ebnomatruma
\ps  n
\pn  kb
\ge  ancestors
\ge  descendant
\gn  nenek_moyang
\rn  moyang, nenek
\gn  keturunan
\rn  turun, ke-*-an

\rf  d2.091.14
```

For Help, press F1 | \lx ebnomatruma | \ge ancestors | 1/100

Current active window

Record marker for active window

Field used for sorting of active window

Fig. 69. Shlex with window 2 sorted to the first record

Notice also that when you move to a new record the insertion point now appears at the beginning of the English gloss field after the **\ge** field marker, not after the **\lx** record marker in window 1. This is because the insertion point always goes to the beginning of the primary sort field when moving to a new record. Notice also that the status bar now shows both the record marker and the field used for sorting.

Multiple indexing of one record

In this example, Shoebox sorts the database in window 2 by the gloss field, and there may be multiple gloss fields in one record.

What does Shoebox do where there are several occurrences of the primary sort field? Shoebox displays the record for each occurrence of the primary sort field. This is slightly confusing to describe without an illustration. First you must position the database at the right record.

1. Click somewhere in window 2 (the one sorted by the **\ge** field) to make it the active window.

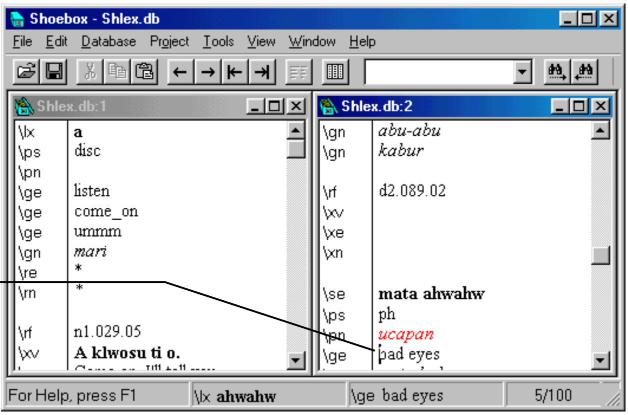
2. Choose Database, First Record or click ⇤ on the toolbar.

The window will display the record whose record marker field contains *ebnomatruma*. However, the insertion point will be sitting next to the word *ancestors* in the **\ge** field. This is because Shoebox now sorts the database alphabetically by the **\ge** field and the word *ancestors* comes first. Observe that the insertion point is always placed at the beginning of the primary sort field when moving to a different record. Also notice that the status bar shows the record marker **\lx** *ebnomatruma* and the current primary sort field **\ge** *ancestors*.

The status bar also claims this is the first of 100 records. You may recall that there were only 57 records when you started. Rather than being a count of the records, the total number is actually a count of the number of occurrences of the field by which the data is sorted. Since there are 100 occurrences of the **\ge** field in this database, Shoebox displays 100 in the status bar.

If the primary sort field marker occurs more than once in a record, the same record will appear at different places in the sequence. To show this, move forward to the record *ahwahw*, which contains four English glosses.

3. Still in window 2, click → four times.



```
\lx    a
\ps    disc
\pn
\ge    listen
\ge    come_on
\ge    ummm
\gn    mari
\re    *
\rn    *

\rf    n1.029.05
\xv    A klwosu ti o.
```

```
\gn    abu-abu
\gn    kabur

\rf    d2.089.02
\xv
\xe
\xn

\se    mata ahwahw
\ps    ph
\pn    ucapan
\ge    bad eyes
```

```
For Help, press F1 | \lx ahwahw | \ge bad eyes | 5/100
```

bad eyes is the gloss of the 5th record

Fig. 70. Fifth Record of Shlex sorted by the **\ge** field

The insertion point will move to *bad eyes*, which is one of the glosses in the record *ahwahw* (this gloss is for the subentry *mata ahwahw*). The insertion point may be too far down in the record for you to see the record marker, but you can still see which record it is by looking at the status bar (that is, **\lx** *ahwahw* and record 5/100).

4. Move to the top of the current record, if you cannot see it (use the scroll bar or press CTRL+HOME).

Notice that the same record also contains glosses *dusty*, *dirty*, and *cloudy*.

5. Click → nine more records to the gloss *cloudy* (which is in the same Selaru word *ahwahw*).
6. Click → six more records to the gloss *dirty* and another seven will take you to *dusty*.

These are all the *same record*. You are just seeing it repeated because the **\ge** field occurs multiple times.

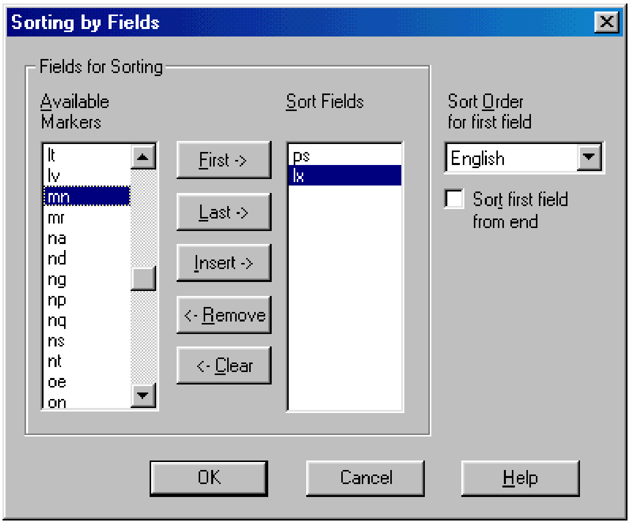
The two windows display different views of the *same database*. Sorting by a different field does *not* result in multiple copies of the records. Therefore, you can make changes to the database by editing in either window.

Sorting by multiple fields

It is also possible to sort by multiple fields. For example, you can sort by part of speech and then by lexeme. To do that,

1. If window 2 is not your active window, click anywhere in window 2.
2. Choose Database, Sorting.
3. To clear the Sort Fields, click <- Clear.
4. To move **ps** over to Sort Fields, select it and click First ->.
5. To move **lx** to be the second sort field, select it and click Last ->.

The result is **ps** and then **lx** as the Sort Fields.



Available Markers: lt, lv, mn, mr, na, nd, ng, np, nq, ns, nt, oe, on

Sort Fields: ps, lx

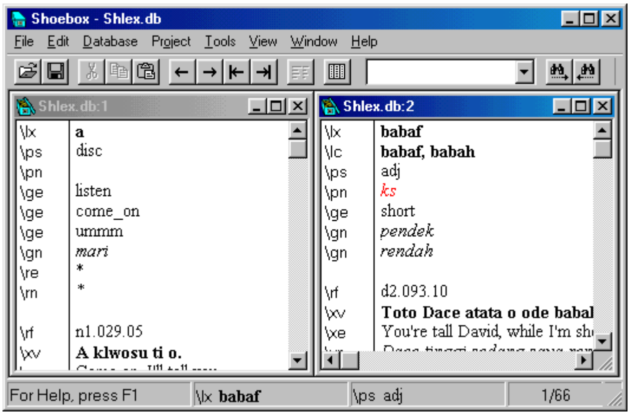
Sort Order for first field: English

☐ Sort first field from end

Fig. 71. Sorting by Fields box for Shlex.db:2

6. Click OK.

7. Go to the first record of the database (click ⇤).



```
Shlex.db:1
\lx   a
\ps   disc
\pn
\ge   listen
\ge   come_on
\ge   ummm
\gn   mari
\re   *
\rn   *

\rf   n1.029.05
\xv   A klwosu ti o.
```

```
Shlex.db:2
\lx   babaf
\lc   babaf, babah
\ps   adj
\pn   ks
\ge   short
\gn   pendek
\gn   rendah

\rf   d2.093.10
\xv   Toto Dace atata o ode babal
\xe   You're tall David, while I'm sh
```

For Help, press F1 | \lx babaf | \ps adj | 1/66

Fig. 72. Shlex.db:2 sorted by **\ps** and **\lx**

Notice that this places you in the *babaf* record. Adjectives sort first and *babaf* is the first lexical item to sort among the adjectives.

8. Click → seven times to move down to record 8 to the start of the nouns.

Observe that the database sorts first by part of speech (**ps**) and then by lexeme (**lx**) within the same part of speech. For example, the adjectives are listed with *babaf* first, then *babah*. The first few nouns you will see are *abat, abu, adar*.

Review

Sort window 2 by the **ge** field instead of by **ps** and **lx**.

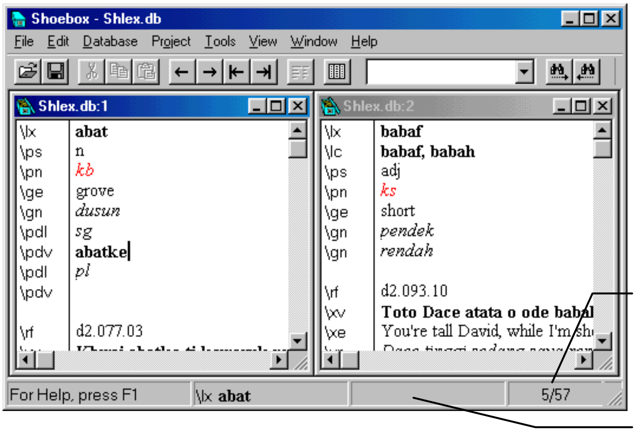
1. Choose Database, Sorting.
2. Remove the **ps** and **lx** field markers from the list of Sort Fields. Move the **ge** marker back in to make the database sorted only by the **ge** field again.
3. Click OK.

Sorting from the right

Examining suffixes and word structure can be aided by sorting words from the right end and seeing how they pattern.

1. Click somewhere in window 1 (which is sorted by the **lx** field).
2. Click ⇤ to ensure you are at the first record of the database.
3. Click → four times to move forward to the record *abat*.

Notice that the status bar numbers are 5/57. You are on the 5th record of 57 records. Also notice that the primary sort field on the status bar is empty. This means you are sorting on the **\lx** field shown in the current record marker on the status bar.



Shlex.db:1

```
\lx    abat
\ps    n
\pn    kb
\ge    grove
\gn    dusun
\pdl   sg
\pdv   abatke
\pdl   pl
\pdv

\rf    d2.077.03
```

Shlex.db:2

```
\lx    babaf
\lc    babaf, babah
\ps    adj
\pn    ks
\ge    short
\gn    pendek
\gn    rendah

\rf    d2.093.10
\xv    Toto Dace atata o ode babal
\xe    You're tall David, while I'm sh
```

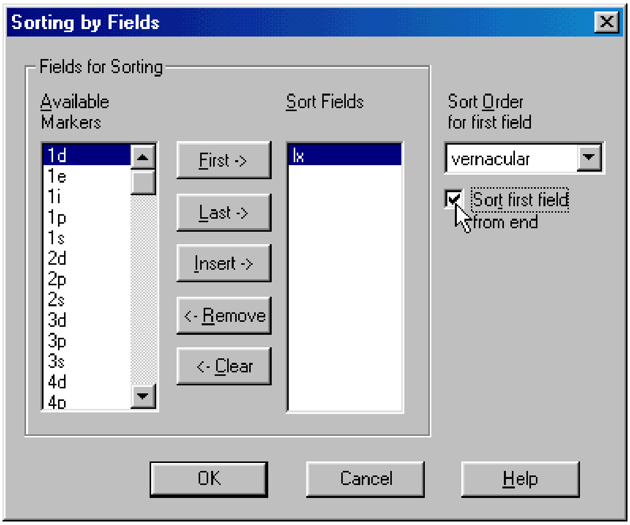
For Help, press F1 | \lx abat | | 5/57

Record 5 of 57 records

Primary sort field is empty

Fig. 73. Shlex.db set up with the left window active

4. Choose Database, Sorting.



Fields for Sorting

Available Markers: 1d, 1e, 1i, 1p, 1s, 2d, 2p, 2s, 3d, 3p, 3s, 4d, 4p

Sort Fields: lx

Sort Order for first field: vernacular

☑ Sort first field from end

Fig. 74. Sorting by Fields dialog box

5. Click to select the check box to the right that says Sort first field from end.

6. Click OK.

You will notice that now the number reads 39/57. This is because *abat* is now sorted by the *t* at the end rather than by the *a* at the beginning. This feature can be helpful to you when studying word final constructions.[11]

7. Click → several times to move forward to see that you are moving through lexemes that end with *t*.

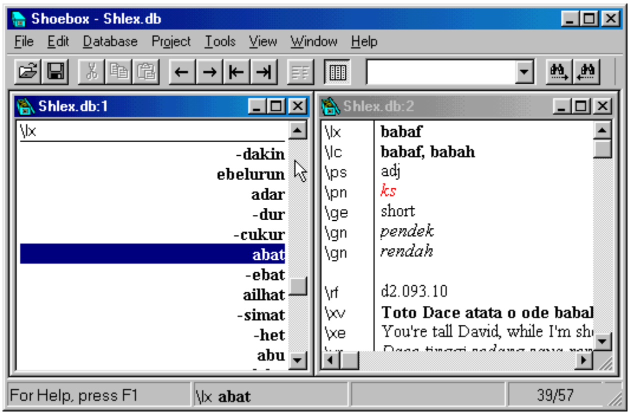
8. Click ← several times to move back to *abat*.

Browse view

It is difficult to get a good overview of the word patterns by stepping singly through the records. The Browse feature gives an overview of all the records in your database.

To see a list of all the records,

1. Choose View, Browse from the menu.



\lx

-dakin
ebelurun
adar
-dur
-cukur
abat
-ebat
ailhat
-simat
-het
abu

\lx	**babaf**
\lc	**babaf, babah**
\ps	adj
\pn	*ks*
\ge	short
\gn	*pendek*
\gn	*rendah*
\rf	d2.093.10
\xv	**Toto Dace atata o ode babal**
\xe	You're tall David, while I'm sh

Fig. 75. Browse view sorted from the right

This puts you into the browse view, which shows one record per line. You can now see more clearly that the ends of the words order the records. Notice also that the words are aligned flush

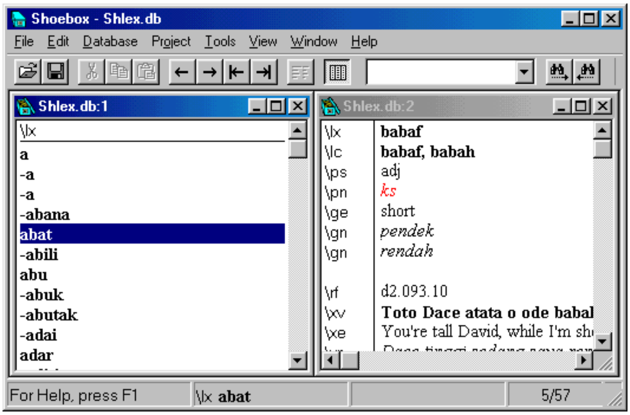
[11] Note: Sorting right to left is not the way to sort a script written right to left like Hebrew or Arabic. Shoebox stores all scripts in pronunciation order. Scripts written right to left are inverted on the display screen only, so use normal sorting for them.

against the right margin. This is because these words are sorted from the end, and this facilitates seeing any patterns that might be there.

2. Choose Database, Sorting.

3. Click to clear the Sort first field from end check box.

4. Click OK.



\lx
a
-a
-a
-abana
abat
-abili
abu
-abuk
-abutak
-adai
adar

\lx	babaf
\lc	babaf, babah
\ps	adj
\pn	ks
\ge	short
\gn	pendek
\gn	rendah
\rf	d2.093.10
\xv	Toto Dace atata o ode babal
\xe	You're tall David, while I'm sh

For Help, press F1 | \lx abat | 5/57

Fig. 76. Browse view with normal sorting

The view should change so that now you see *abat* in its normal alphabetical order. Notice the words are now aligned normally for this language.[12]

5. Choose View, Browse Fields.

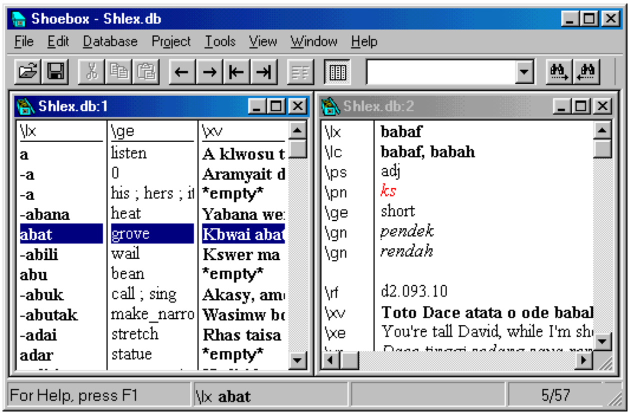
This brings up the dialog box for the fields that are shown on Browse. This works like the Sort Fields dialog box.

6. Remove any other markers in the browse view so only the **\lx** field is shown.

7. Insert the fields **\ge** and then **\xv** after the **\lx** field.

8. Click OK.

[12] The words starting with a hyphen are marked this way in the database to show that they are suffixes.



\lx	\ge	\xv
a	listen	A klwosu t
-a	0	Aramyait d
-a	his ; hers ; it	*empty*
-abana	heat	Yabana we:
abat	grove	Kbwai abat
-abili	wail	Kswer ma
abu	bean	*empty*
-abuk	call ; sing	Akasy, am
-abutak	make_narro	Wasimw bo
-adai	stretch	Rhas taisa
adar	statue	*empty*

\lx	babaf
\lc	babaf, babah
\ps	adj
\pn	ks
\ge	short
\gn	pendek
\gn	rendah
\rf	d2.093.10
\xv	Toto Dace atata o ode babal
\xe	You're tall David, while I'm sh

\lx abat

5/57

Fig. 77. Browse view showing the **\lx**, **\ge**, and **\xv** fields

You should now see three fields: the lexeme, the English gloss, and the vernacular example.

9. Drag the dividing line between the first two fields in the browse view to the left, until it is near the ends of the entry words.

Similarly, you can drag the line that divides the English gloss and the vernacular example to the left or right.

You can switch between browse view and record view by choosing View, Browse or clicking ▦ on the toolbar. In addition, from within the browse view, you can double-click a record or press ENTER (Macintosh: press RETURN) to return to record view.

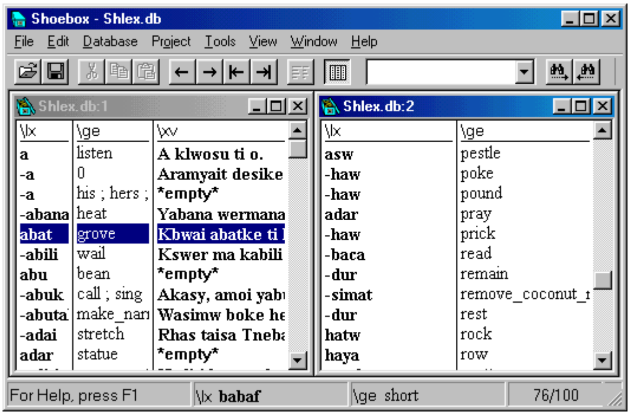
10. Select window 2.

11. From the menu, choose View, Browse Fields.

12. Remove any markers other than the **lx** marker in the Browse view.

13. Insert the **ge** field at the end of the list of fields to be browsed.

14. Click [OK].



Shlex.db:1

\lx	\ge	\xv
a	listen	A klwosu ti o.
-a	0	Aramyait desike
-a	his ; hers ;	*empty*
-abana	heat	Yabana wermana
abat	grove	Kbwai abatke ti
-abili	wail	Kswer ma kabili
abu	bean	*empty*
-abuk	call ; sing	Akasy, amoi yab
-abuta	make_nan	Wasimw boke he
-adai	stretch	Rhas taisa Tneba
adar	statue	*empty*

Shlex.db:2

\lx	\ge
asw	pestle
-haw	poke
-haw	pound
adar	pray
-haw	prick
-baca	read
-dur	remain
-simat	remove_coconut_
-dur	rest
hatw	rock
haya	row

For Help, press F1 | \lx babaf | \ge short | 76/100

Fig. 78. Browse view on both windows

Now it should be easier to see that window 2 is indeed sorted by the English gloss field. Scrolling up and down the list will show where some lexemes occur multiple times (because they have multiple glosses).

15. Click ⇤ for each window, so the first few records are showing in each view.

Notice that the window sorted by gloss shows the *ebnomatruma* record with the first gloss being *ancestors*, whereas the window sorted on the lexeme shows the first *a* record (glossed *listen*).

It can be useful to keep a window open in browse view with the records in your lexicon sorted by gloss.

16. Click ▦ with each window active to return both windows to their normal record view.

Marker hierarchy

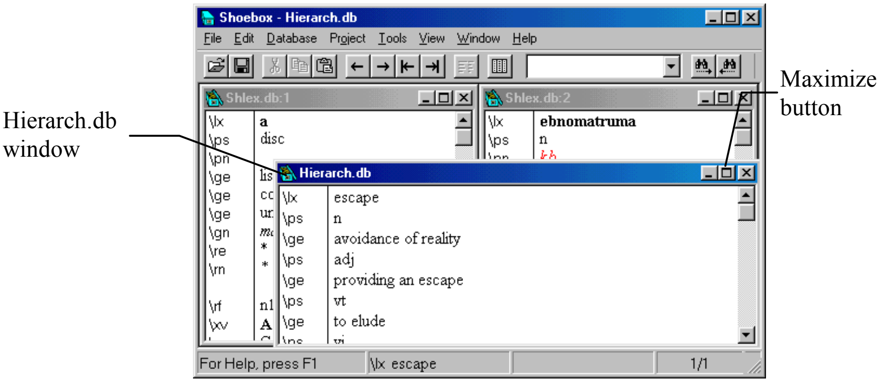
Marker hierarchy is a way to establish and indicate the structural relationship between different elements of a record. Hierarchy is useful for the browse view and for interlinearization.

The Shlex.db file already has the marker hierarchy established. It is impractical to remove the hierarchy, so a different database will be used to illustrate its significance. Leaving both windows of the Shlex.db file open,

1. Choose File, Open.

2. Move to and open the Samp_set folder. (If you have trouble finding this folder, see *Finding the sample files*, p. 14.)

3. Open the database Hierarch.db.



```
\lx   escape
\ps   n
\ge   avoidance of reality
\ps   adj
\ge   providing an escape
\ps   vt
\ge   to elude
```

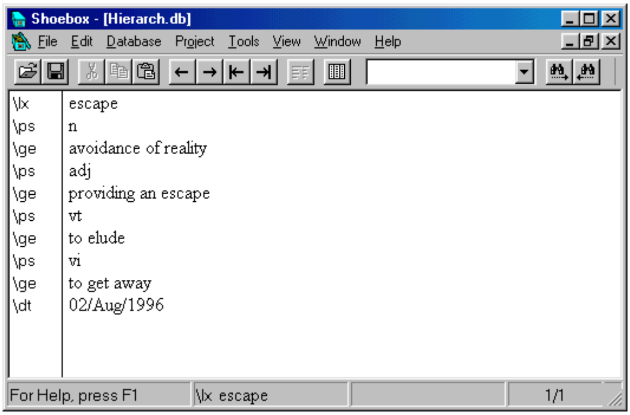
For Help, press F1 | \lx escape | 1/1

Hierarch.db window

Maximize button

Fig. 79. Hierarch.db database

4. Click the maximize button ☐ at the top of the Hierarch.db window to maximize the window and show more of the record.



\lx escape
\ps n
\ge avoidance of reality
\ps adj
\ge providing an escape
\ps vt
\ge to elude
\ps vi
\ge to get away
\dt 02/Aug/1996

Fig. 80. Hierarch.db window maximized

Note: Although Hierarch.db contains field markers that look the same as the markers in Shlex.db, the database is actually based on a different database type file (that has no hierarchy established).

You will look at the glosses in this data using three different settings.

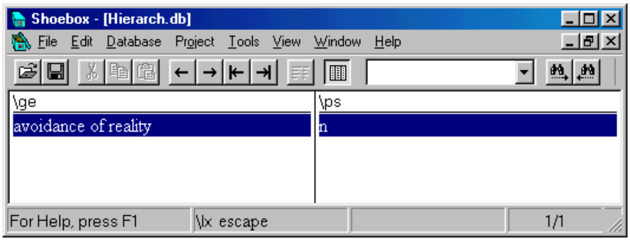
Set up the browse view to display only the **ge** and **ps** fields.

5. Choose View, Browse Fields.

6. Double-click the **ge** and **ps** fields to move them into the right column.

7. Double-click the **lx** field in the right column to move it out of the right column.

8. Click OK.



\ge	\ps
avoidance of reality	n

Fig. 81. Browse view of Hierarch.db showing the **\ge** and **\ps** fields

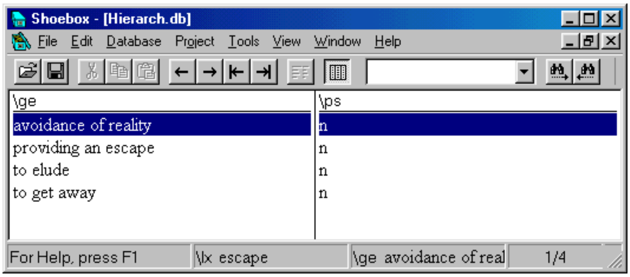
You will see only a single line displayed. This is because the database has only one record, and Shoebox sorts by the **\lx** field in this view (even though this field is not displayed). Notice that Browse View displays only the first occurrence of the **\ge** and **\ps** fields in this record.

Specify the **ge** field as the only field to sort by.

9. Choose Database, Sorting.

10. Double-click the **ge** field to move it into the right column; double-click the **lx** field to move it out.

11. Click OK.



\ge	\ps
avoidance of reality	n
providing an escape	n
to elude	n
to get away	n

For Help, press F1 | \lx escape | \ge avoidance of real | 1/4

Fig. 82. Browse view of the Hierarch database based on the **\ge** field

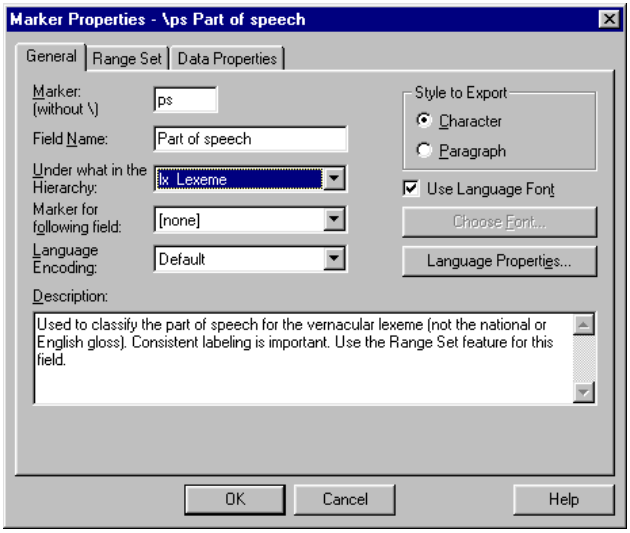
The view now shows all four English glosses, but it displays the same part of speech for each of them. This is because Shoebox has not been told about the relationship between the parts of speech and English gloss fields. Shoebox focuses on the gloss fields, which is the primary sort field, but includes only the first part of speech field that it encounters in the record. The same results would be obtained even if you sorted by both the **\ge** *and* the **\ps** fields. (If you were to sort *only* by the **\ps** field, you would see all the parts of speech, but would get the same gloss for all of them.)

Now you will set up the hierarchy. This will build a closer link between each gloss and its corresponding part of speech.

12. Click ▦ to return to viewing the record.

By default, Shoebox automatically lists all field markers in a record hierarchically under the record marker. So in this example, both the **\ps** and **\ge** fields have already been put under the **\lx** field. To verify this,

13. Right-click **\ps** in the marker pane (Macintosh: OPTION+click).



Marker Properties - \ps Part of speech

General | Range Set | Data Properties

Marker: (without \): ps

Field Name: Part of speech

Under what in the Hierarchy: lx Lexeme

Marker for following field: [none]

Language Encoding: Default

Style to Export: ● Character ○ Paragraph

☑ Use Language Font

Description:

Used to classify the part of speech for the vernacular lexeme (not the national or English gloss). Consistent labeling is important. Use the Range Set feature for this field.

Fig. 83. Marker Properties of the part of speech

This displays the Marker Properties dialog box for the **ps** marker.

14. In the box labeled Under what in the Hierarchy notice that it shows **lx Lexeme**.

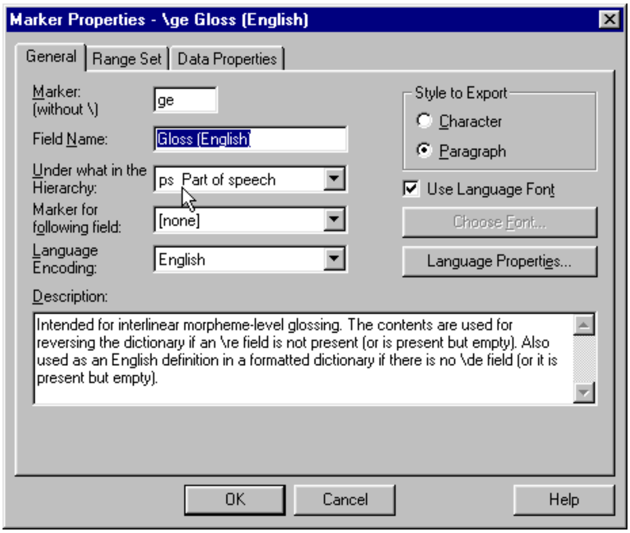
15. Click OK.

Now change the default hierarchy and specify that the **ge** marker is actually under the **ps** field.

16. Right-click **\ge** in the marker pane to display the Marker Properties dialog box for the **ge** marker (Macintosh: OPTION+click).

17. In the box labeled Under what in the hierarchy, change the **lx** field choice by selecting **ps** from the list.

18. Click OK.



Marker: (without \) ge

Field Name: Gloss (English)

Under what in the Hierarchy: ps Part of speech

Marker for following field: [none]

Language Encoding: English

Style to Export: ○ Character ● Paragraph

☑ Use Language Font

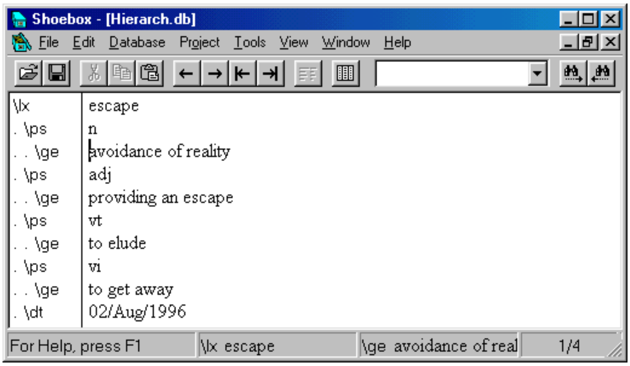
Description:

Intended for interlinear morpheme-level glossing. The contents are used for reversing the dictionary if an \re field is not present (or is present but empty). Also used as an English definition in a formatted dictionary if there is no \de field (or it is present but empty).

Fig. 84. Marker properties with the English gloss put under the part of speech

This establishes a simple hierarchy for this database. To view the record with the field markers showing the hierarchy,

19. Choose View, Marker Hierarchy from the menu.



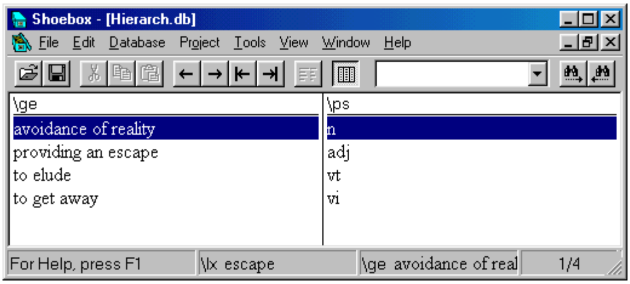
```
\lx       escape
. \ps     n
. . \ge   avoidance of reality
. \ps     adj
. . \ge   providing an escape
. \ps     vt
. . \ge   to elude
. \ps     vi
. . \ge   to get away
. \dt     02/Aug/1996
```

For Help, press F1 | \lx escape | \ge avoidance of real | 1/4

Fig. 85. Marker hierarchy shown in the marker pane

Notice that the field markers are now indented to various depths. This is similar to an outline display and shows graphically which marker is under which. This view allows you to see that the hierarchy has been established correctly. It is not necessary to view the hierarchy in order to reap the benefits of it.

20. Click ▦.



\ge	\ps
avoidance of reality	n
providing an escape	adj
to elude	vt
to get away	vi

Fig. 86. Browse view showing relationship between the **\ge** and **\ps** fields

Because of the hierarchy, the program can now tell that the different glosses belong to different parts of speech. That information is reflected in the browse view.

Since most databases have more than just three fields, it is generally best to establish the hierarchy for the whole database at once.

21. Choose File, Close to close the Hierarch.db database.

22. Click No to throw away the changes. You should now be back in the Shlex database.

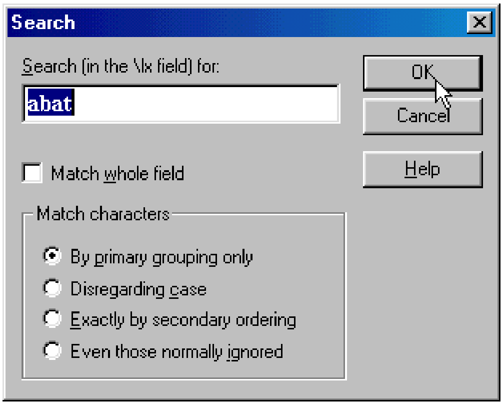
Search

Search is a fast way to get to a particular record. Search will change which record is visible in the current view. *You can use the Search command to look for data only in the primary sort field*. The string you specify must be the beginning or the entirety of the contents of the primary sort field.[13]

[13] If the primary sort field is sorted from the end, then the search string is also matched from the end, e.g., *ma* would match *ebnomatruma* (if the search is not set to match whole field).

1. Click somewhere in window 1 of the Shlex database, which is sorted by the **\lx** field (turn off browse view if on).[14]

2. Choose Database, Search.



Search (in the \lx field) for: abat

☐ Match whole field

Match characters
- ● By primary grouping only
- ○ Disregarding case
- ○ Exactly by secondary ordering
- ○ Even those normally ignored

Fig. 87. Search Dialog box searching the **\lx** field for *abat*.

3. Type the word `abat` and click OK.

Notice that the record is found immediately. Shoebox maintains an index of the primary sort fields. This enables Shoebox to find a match quickly in the primary sort field.

You can also search in a database that is sorted by something other than the record marker.

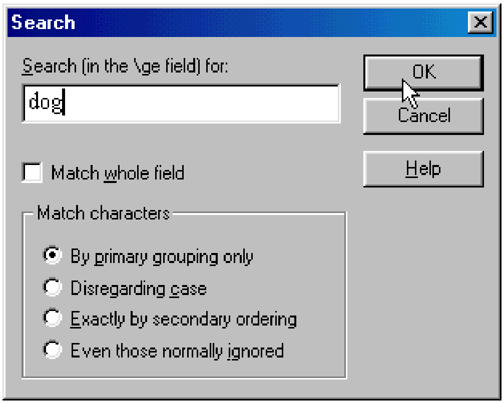
4. Click somewhere in window 2, which is sorted by the **\ge** field (turn off the browse view if still on).

5. Choose Database, Search.

Notice that the Search dialog box indicates it is searching in the **\ge** field.

[14] This is assuming the same window configuration as in the section on sorting. If you have skipped that section, please follow the steps at the beginning of *Multiple views,* p. 56, to establish two windows on the Shlex.db file, the first sorted by the record marker (**lx**) and the second sorted by the English gloss (**ge**) field.

6. Search for dog.



Search (in the \ge field) for:

dog

□ Match whole field

Match characters

● By primary grouping only

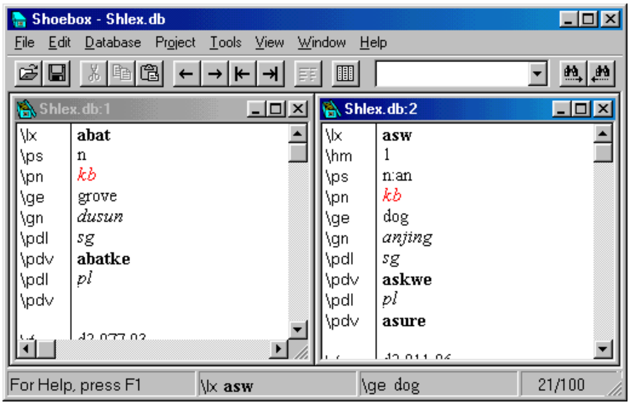
○ Disregarding case

○ Exactly by secondary ordering

○ Even those normally ignored

Fig. 88. Search Dialog box set to search the **\ge** field for *dog*.

The window will display the record *asw*, with the insertion point in front of *dog*.



Shlex.db:1

\lx	**abat**
\ps	n
\pn	*kb*
\ge	grove
\gn	*dusun*
\pdl	*sg*
\pdv	**abatke**
\pdl	*pl*
\pdv	

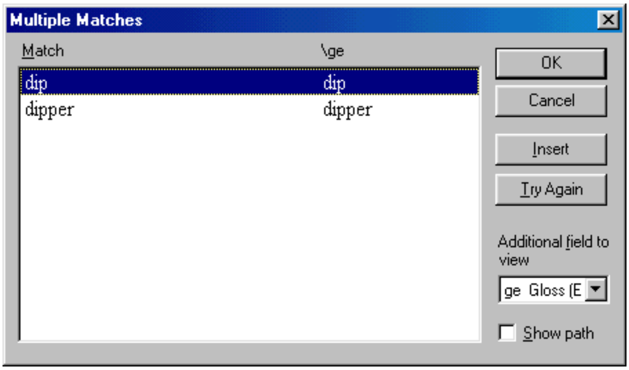
Shlex.db:2

\lx	**asw**
\hm	1
\ps	n:an
\pn	*kb*
\ge	dog
\gn	*anjing*
\pdl	*sg*
\pdv	**askwe**
\pdl	*pl*
\pdv	**asure**

For Help, press F1 | \lx **asw** | \ge dog | 21/100

Fig. 89. Shlex record for *dog*

7. Search for *dip*.



Multiple Matches

Match	\ge
dip	dip
dipper	dipper

Additional field to view: ge Gloss (E

☐ Show path

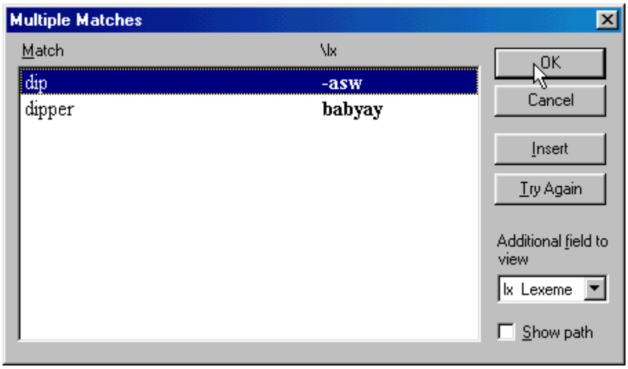
Fig. 90. Multiple Matches dialog box after searching for *dip*.

The Multiple Matches dialog box appears, indicating that both *dip* and *dipper* are possible matches. (If you had chosen the option ☐ Match whole field in the Search box, then only *dip* would have matched.)

Multiple matches

The Multiple Matches dialog box includes an option to show another field along with the field being searched. It currently defaults to showing the **\ge** field.

1. Click Additional field to view ge Gloss (E ▼ and select **lx**.



Multiple Matches

Match	\lx
dip	**-asw**
dipper	**babyay**

Additional field to view: lx Lexeme

☐ Show path

Fig. 91. Multiple Matches Dialog box showing the **\lx** field also.

The Multiple Matches dialog box now displays the **lx** field. This can help you choose whether you want *dip* or *dipper*.

2. Select *dip* and click OK.

3. You are now back at the Shlex record that contains *dip*.

Find

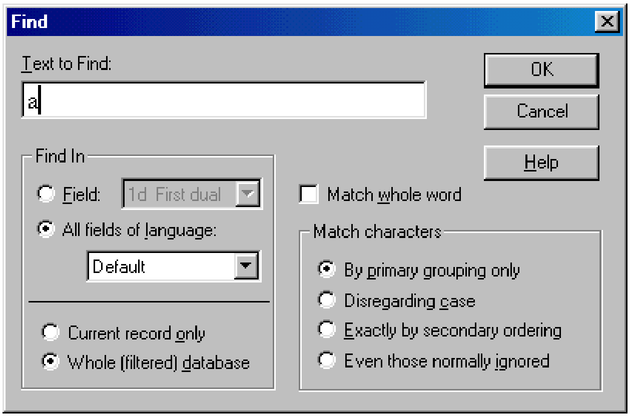
Whereas Search will find data only in the primary sort field, the Find command looks for data anywhere in a record.

1. Click in window 1 to make it the active window. Window 1 still sorts the Shlex database by **\lx** field.

2. Click ⇤ to move to the first record of the database.

3. Choose Edit, Find.

4. Type the single letter a in the Text to Find box (but do not start the Find yet).



Text to Find: a

Find In
○ Field: 1d First dual
● All fields of language: Default
○ Current record only
● Whole (filtered) database

□ Match whole word

Match characters
● By primary grouping only
○ Disregarding case
○ Exactly by secondary ordering
○ Even those normally ignored

Fig. 92. Find dialog box with *a* typed in the Text to Find box

A common letter like *a* will be found often. There are ways to reduce the number of unintended matches.

Notice that you have several options forming something like two columns. Immediately below the Text to Find box, there is a group of options labeled Find In. There you can choose to limit the Find by field or by language.

5. Choose All fields of language, and select vernacular from the drop-down list.

This limits your search to those fields where the vernacular language is used. Shlex uses the vernacular language in the lexeme field and the sample sentences.

Notice also that you can limit the Find by restricting it to the current record or you can let it examine the whole database.

6. Choose Whole (filtered) database.

Now notice that to the right of the Find In box, at the top of the second column of options, there is a check box specifying Match Whole Word.

7. Click the Match Whole Word check box to select it.

This will make the program skip over any *a* that is part of another word.

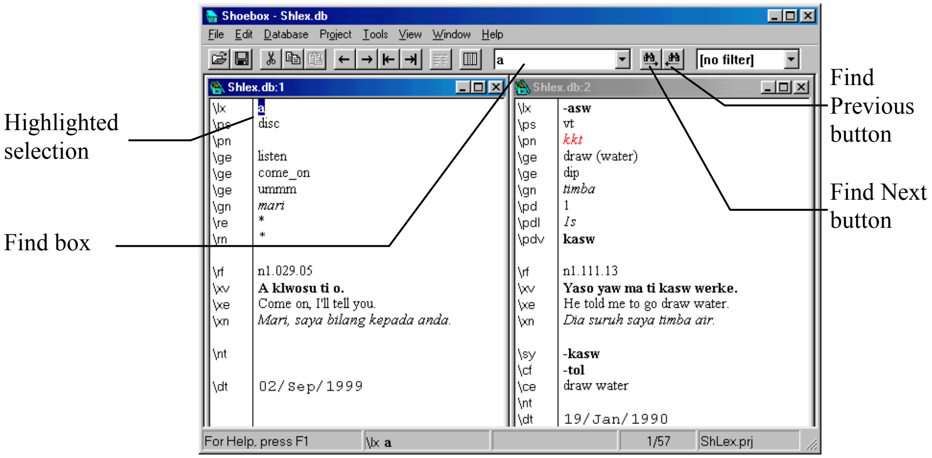
Below the check box is yet another set of options, titled Match Characters.

8. Choose Even Those Normally Ignored.

A hyphen is one of the characters that is normally ignored. Choosing this option will make the Find command “see” hyphens, thereby causing it to skip over the -*a* entries and find only the *a* entries.

9. Click OK.

The *a* at the top of the record will be highlighted to show it is found.



Highlighted selection

Find box

Find Previous button

Find Next button

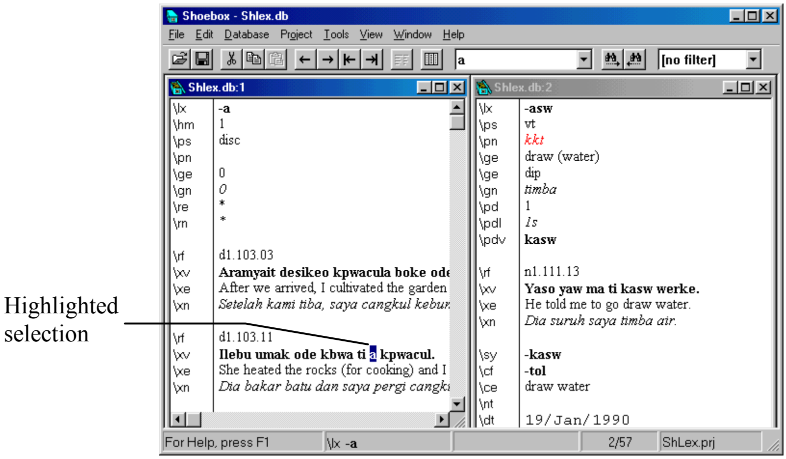
```
\lx     a
\ps     disc
\pn
\ge     listen
\ge     come_on
\ge     ummm
\gn     mari
\re     *
\rn     *

\rf     n1.029.05
\xv     A klwosu ti o.
\xe     Come on, I'll tell you.
\xn     Mari, saya bilang kepada anda.

\nt

\dt     02/Sep/1999
```

```
\lx     -asw
\ps     vt
\pn     kkt
\ge     draw (water)
\ge     dip
\gn     timba
\pd     1
\pdl    1s
\pdv    kasw

\rf     n1.111.13
\xv     Yaso yaw ma ti kasw werke.
\xe     He told me to go draw water.
\xn     Dia suruh saya timba air.

\sy     -kasw
\cf     -tol
\ce     draw water
\nt
\dt     19/Jan/1990
```

Fig. 93. Shlex record showing a Find on letter *a* in the vernacular

Now look on the toolbar, just to the right of the Browse View button. You will see a small text box, which now contains the letter *a*. It has a drop-down list of the last several things you have tried to find. At the moment, it only contains one item. To its right is a pair of binoculars with an arrow pointing right. This is the Find Next button.

10. Click 🔍.



Highlighted selection

```
\lx     -a
\hm     1
\ps     disc
\pn
\ge     0
\gn     0
\re     *
\rn     *

\rf     d1.103.03
\xv     Aramyait desikeo kpwacula boke ode
\xe     After we arrived, I cultivated the garden
\xn     Setelah kami tiba, saya cangkul kebun

\rf     d1.103.11
\xv     Ilebu umak ode kbwa ti a kpwacul.
\xe     She heated the rocks (for cooking) and I
\xn     Dia bakar batu dan saya pergi cangki
```

```
\lx     -asw
\ps     vt
\pn     kkt
\ge     draw (water)
\ge     dip
\gn     timba
\pd     1
\pdl    1s
\pdv    kasw

\rf     n1.111.13
\xv     Yaso yaw ma ti kasw werke.
\xe     He told me to go draw water.
\xn     Dia suruh saya timba air.

\sy     -kasw
\cf     -tol
\ce     draw water
\nt
\dt     19/Jan/1990
```

Fig. 94. Shlex showing the word *a* highlighted

Notice that rather than stopping at each *a* along the way, the program skips to the word *a* in one of the illustrative sentences. Observe that this is in the **\xv** field, which is a vernacular language field. You may need to scroll left to see the **\xv** marker.

11. Click .

There will be a pause while Find examines the data. Because Shoebox is looking through the whole database for the data, it can be slow if the database is large. It should stop in an **\xv** field in the **\lx** *baba* record.

Notice that the English translation in the next **\xe** field also contains the (English) word *a*. The next time you use the Find Next command, the program will *not* stop at that *a* because it is not in the vernacular language.

12. Click .

Again there will be a pause as Shoebox looks through the data until it finds the *-he tel* entry. The window may need to scroll horizontally to show the word.

Next to the Find Next button is another pair of binoculars with an arrow pointing to the left. This is the Find Previous button.

13. Click .

You will see the *baba* record again.

Jumping

The Jump To command is yet another way to find data. Think of jumping as a way to look up related data by searching one or more fields in specific databases. For example, a cross-reference or a lexical function value in your lexicon should refer to the headword of a lexical entry or a subentry. Similarly, words in the English free translation of an illustrative sentence often also appear as the English glosses for lexical entries. Shoebox uses jump paths to represent these types of relationships.

When you jump from a cross-reference (**\cf**) field or a lexical function value (**\lv**) field, you want Shoebox to search the lexeme (**\lx**) fields and subentry (**\se**) fields of your lexical database.

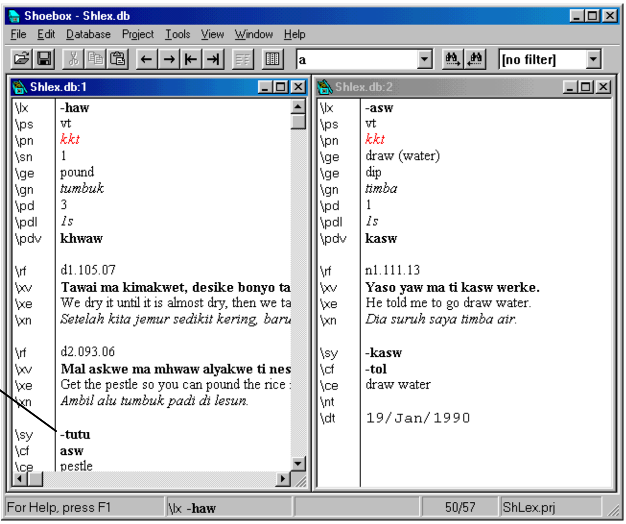
Likewise, when you jump from the English free translation (**\xe**) field, Shoebox should search the English gloss (**\ge**) field of your lexicon to find a match.

Jump To will bring up a new window showing the record that matches the criteria you have specified. Jump To is like Search in that it will only find data at the beginning of a field (or end if sorted from the end). It is unlike Search in that it is able to find data in fields other than the primary sort field.

1. In window 1, use the Search command to find the record *-haw*.

Notice that the first cross-reference is the word *asw*. You want Shoebox to jump to the entry for this word, whether it is a headword or a subentry in your lexicon.

2. Place the insertion point at the beginning of the word *asw* in the **\cf** field.



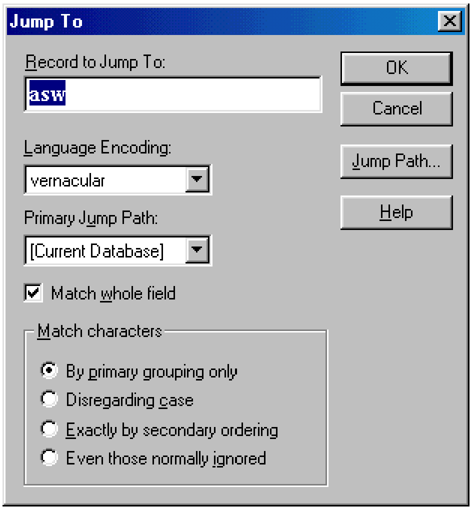
```
\lx     -haw
\ps     vt
\pn     kkt
\sn     1
\ge     pound
\gn     tumbuk
\pd     3
\pdl    1s
\pdv    khwaw

\rf     d1.105.07
\xv     Tawai ma kimakwet, desike bonyo ta
\xe     We dry it until it is almost dry, then we ta
\xn     Setelah kita jemur sedikit kering, baru

\rf     d2.093.06
\xv     Mal askwe ma mhwaw alyakwe ti nes
\xe     Get the pestle so you can pound the rice :
\xn     Ambil alu tumbuk padi di lesun.

\sy     -tutu
\cf     asw
\ce     pestle
```

```
\lx     -asw
\ps     vt
\pn     kkt
\ge     draw (water)
\ge     dip
\gn     timba
\pd     1
\pdl    1s
\pdv    kasw

\rf     n1.111.13
\xv     Yaso yaw ma ti kasw werke.
\xe     He told me to go draw water.
\xn     Dia suruh saya timba air.

\sy     -kasw
\cf     -tol
\ce     draw water
\nt
\dt     19/Jan/1990
```

Put your insertion point here

Fig. 95. Shlex database showing -*haw* record

3. Choose Edit, Jump To.



Jump To

Record to Jump To: asw

Language Encoding: vernacular

Primary Jump Path: [Current Database]

☑ Match whole field

Match characters

● By primary grouping only
○ Disregarding case
○ Exactly by secondary ordering
○ Even those normally ignored

Fig. 96. Jump To dialog box

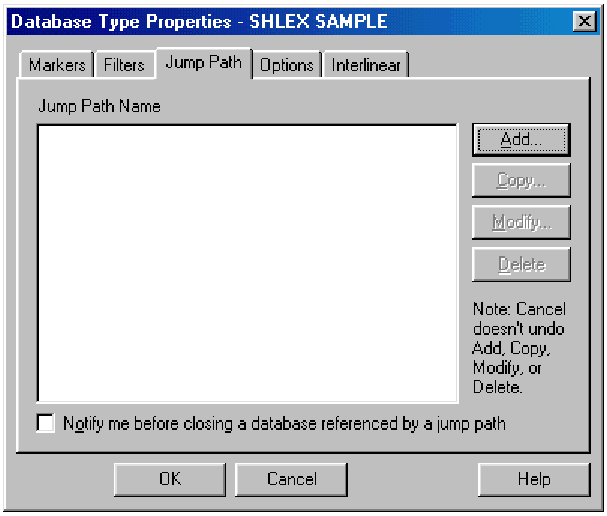
This will bring up a dialog box with the word *asw* already in the Jump To edit box. Below this box you will notice that Language Encoding is already specified. (If the word you are searching for is in a different language you can select it from the drop-down list.)

4. Select Match whole field (if it is not already checked).

Now the question arises: How will Shoebox know where to search for the word *asw* in order to complete the jump? It will start by searching the primary jump path. Below the Language Encoding box is the Primary Jump Path box. Shoebox automatically selects the correct jump path based on the field you are in, but if you have not yet set up any jump paths, this box says Current Database, meaning that Shoebox will simply do a normal Search operation. But you do not want to search just the lexemes, and when you later jump from a word in the English free translation, you do not want to search the lexemes at all. So you need to set up two jump paths to tell Shoebox how your data is related. To the right, under the OK and Cancel buttons, you will see another button called Jump Path....

To set up a new jump path,

5. Click Jump Path....



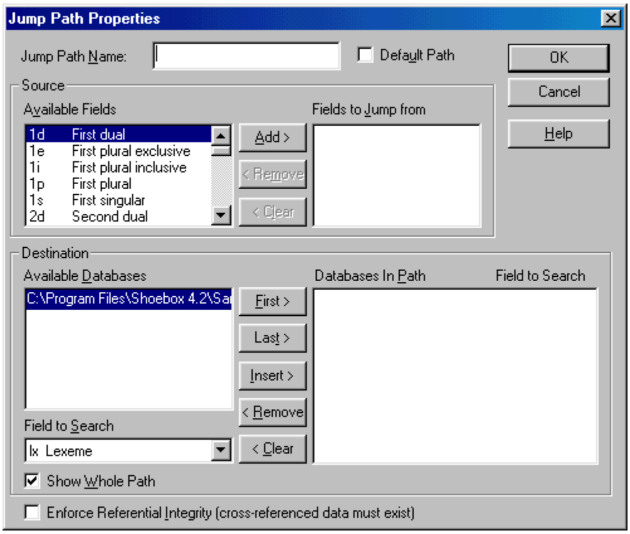
Note: Cancel doesn't undo Add, Copy, Modify, or Delete.

□ Notify me before closing a database referenced by a jump path

Fig. 97. Database Type Properties with Jump Path tab selected

This takes you to the Jump Path page of the Database Type Properties, where a list of jump paths is shown (it is empty at this point). Along the right-hand side of this list is a series of buttons: Add..., Copy..., Modify..., and Delete.

6. Click Add....



Jump Path Properties

Jump Path Name:

□ Default Path

Source

Available Fields

1d	First dual
1e	First plural exclusive
1i	First plural inclusive
1p	First plural
1s	First singular
2d	Second dual

Fields to Jump from

Destination

Available Databases

C:\Program Files\Shoebox 4.2\Sa

Databases In Path

Field to Search

Field to Search: lx Lexeme

☑ Show Whole Path

□ Enforce Referential Integrity (cross-referenced data must exist)

Fig. 98. Jump Path Properties dialog box

This takes you to a dialog box where you can define a new jump path. At the top of the dialog box is a place to enter the name of this jump path. The name you enter should describe the purpose or the function of this jump path. Since you want to create a jump path for looking up cross-referenced entries in the lexicon, you should enter a name that will help you remember this.

7. Type `Cross-reference` as the name.

The Default Path checkbox should not be selected.[15]

[15] Shoebox can use a default jump path to attempt to find a match if the primary jump path fails to find one. Here we will discuss only primary jump paths.

The next section of the dialog box is where you select the Source fields. Source fields are the fields you intend to jump from using this jump path. The fields listed in the box on the left are the available source fields for the jump path. The fields listed in the right box are the selected source fields (this box is empty). To tell Shoebox to use this jump path when you jump from a cross reference field or from a lexical function value field,

8. In the Available Fields box, scroll down to and select **cf Cross-reference**.
9. Click Add >.
10. In the Available Fields box, scroll down to and select **lv Lexical function lexeme**.
11. Click Add >.

The Destination portion of this dialog box is laid out much like the Sort Fields or Browse Fields dialog boxes. The left column has a list of currently open databases. Below it is a list of all the fields in the currently-selected database. The right column displays a list of all the paths that will be searched when you jump using this path (this box is empty).

The jump path destination defines a list of fields (and the databases they occur in) which is to be searched whenever this jump path is used. The string specified in the Jump To dialog box is compared to the contents of the fields in the jump path. (The contents of these fields are indexed, so the search is fast.) As mentioned before, unlike Search, which will only look in the primary sort field, Jump To will examine any field specified in the jump path.

12. In the Available Databases box, select Shlex.db.

Below this is the Field To Search box.

13. Display the Field to Search list by clicking the down arrow.

Notice that it lists all the fields in the database you have selected. It has defaulted to the record marker, **lx**, which is the first destination field you want.

14. Select the **lx** marker, then click First ->.

The destination database and field are added to the path displayed in the box on the right. You also want this jump path to search the subentry field.

15. Click Field to Search, and select the field **se**.

16. Click Last ->.

At the bottom of the Destination area of this dialog box is the Show Whole Path check box.

17. Click ☑ Show Whole Path to select it.

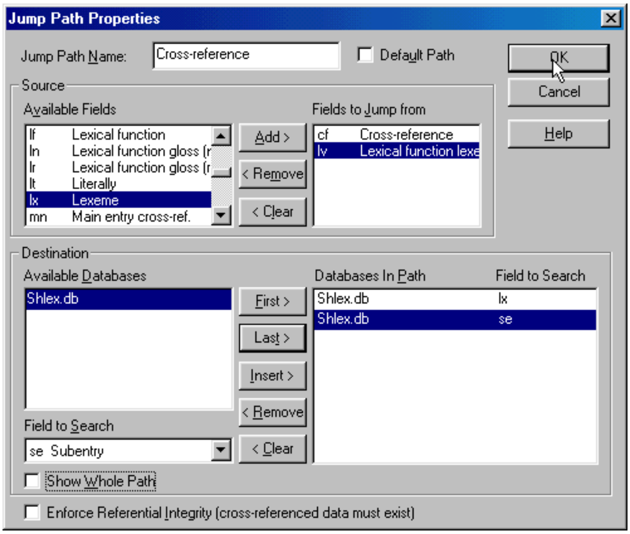
It will display the path on the disk of the databases in both the right and left boxes. Showing the path is most helpful for people working with a complex system for organizing data files.[16] If the paths are long, the names of the databases might be cut off.

The jump path retains all the destinations you have placed on the path, even if some of the destination databases are later closed. When you open them again, they are still in the jump path and available to be jumped to.

18. Click ☑ Show Whole Path again to clear it.

[16] For example, when you are experimenting with different test files, it is possible to add to the jump path a copy of a database in one folder. Later on, you might unload that database and load another one with the same name but in a different folder. The database name in the jump path looks right, but it is not (because it has a different path). When you first attempt to jump using this jump path, Shoebox will detect this situation and alert you to it. If you do not fix it at that time, it can be confusing later when the intended record cannot be found when you jump. Showing the full path can help clear up this confusion.

Your Jump Path Properties should look like the following:



Jump Path Name: Cross-reference

☐ Default Path

Source

Available Fields

lf	Lexical function
ln	Lexical function gloss (r
lr	Lexical function gloss (r
lt	Literally
lx	Lexeme
mn	Main entry cross-ref.

Fields to Jump from

cf	Cross-reference
lv	Lexical function lexe

Destination

Available Databases

Shlex.db

Databases In Path	Field to Search
Shlex.db	lx
Shlex.db	se

Field to Search: se Subentry

☐ Show Whole Path

☐ Enforce Referential Integrity (cross-referenced data must exist)

Fig. 99. Cross-reference jump path

19. Click OK to save this jump path and return to the Jump Path page of the Database Type Properties.

Notice that the list of jump paths now contains a jump path called Cross-reference.

Now add a second jump path to establish the relationship between words in the **\xe** and **\ge** fields. Note that this jump path will have only one source and one destination. Try to do this independently (name the jump path `English gloss lookup`). If you need help, follow these steps:

20. Click Add... to display the Jump Path Properties dialog box.

21. Type `English gloss lookup` as the name.

22. Scroll down in the Available Fields box, in the Source area, and select **xe Example free trans. (E)**.

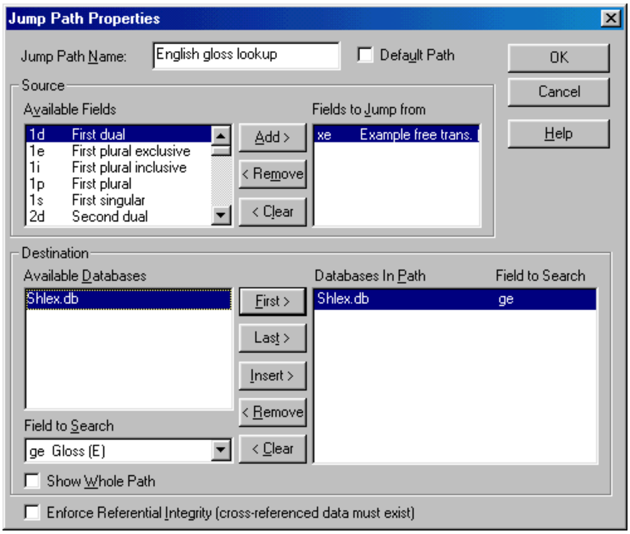
23. Click Add >.

24. In the Destination area, Available Databases box, select Shlex.db.

25. Click Field to Search [lx Lexeme ▼], and select the field **ge**.

26. Click First ->.

Now your Jump Path Properties dialog box should look like the following:



Jump Path Properties

Jump Path Name: English gloss lookup

□ Default Path

OK

Cancel

Help

Source

Available Fields	
1d	First dual
1e	First plural exclusive
1i	First plural inclusive
1p	First plural
1s	First singular
2d	Second dual

Add >

< Remove

< Clear

Fields to Jump from	
xe	Example free trans.

Destination

Available Databases: Shlex.db

First >

Last >

Insert >

< Remove

< Clear

Databases In Path	Field to Search
Shlex.db	ge

Field to Search: ge Gloss (E)

□ Show Whole Path

□ Enforce Referential Integrity (cross-referenced data must exist)

Fig. 100. English gloss lookup jump path

27. Click OK to save this jump path and return to the Jump Path page of the Database Type Properties dialog box.

Now the list contains two jump paths: Cross-reference and English gloss lookup. They establish the two necessary data relationships.

Observe that below the list of jump paths is a check box, which says: □ Notify me before closing a database referenced by a jump path.

It is all too easy to close extra windows and discover that you have closed the *last* window on a database. With this option selected, Shoebox warns you before you close the last window of a file that is in a jump path.

28. Click □ Notify me before closing a database referenced by a jump path.

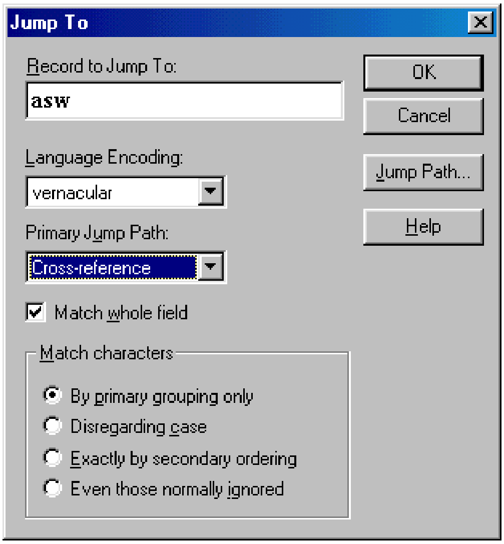
29. Click OK.

This brings you back to the Jump To dialog box.

If the jump paths had already been set up, Shoebox would have automatically chosen the correct primary jump path for jumping from the **\cf** field. This time you will have to select it manually.

30. Click Primary Jump Path: [Current Database] ▾, and select Cross-reference.

The Jump To dialog box should now look like the following:



Jump To

Record to Jump To: asw

Language Encoding: vernacular

Primary Jump Path: Cross-reference

☑ Match whole field

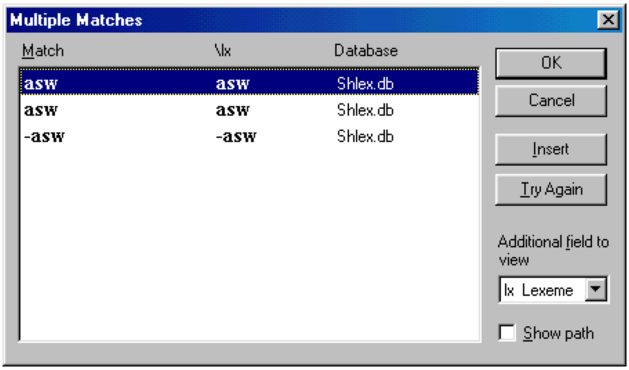
Match characters

● By primary grouping only
○ Disregarding case
○ Exactly by secondary ordering
○ Even those normally ignored

Fig. 101. Jump To dialog box ready for your first jump

31. Click OK.

The program will locate two *asw* entries and an *-asw* entry. This Multiple Matches dialog box is the same as when you use the Search command.[17]



Multiple Matches

Match	\lx	Database
asw	asw	Shlex.db
asw	asw	Shlex.db
-asw	-asw	Shlex.db

Additional field to view: lx Lexeme

□ Show path

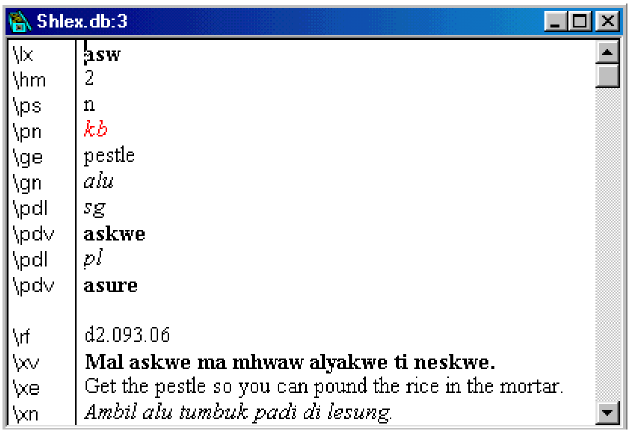
Fig. 102. Multiple Matches dialog box showing *asw* and *-asw*

Notice the Additional Field to View list in the lower right of the Multiple Matches dialog box.

32. If the **ge** field is not in view, change the Additional Field to View to the **ge** marker.

33. Select the entry with the meaning *pestle* and click OK.

[17] If there were more than one database on this Jump Path, the database paths would also be displayed. Sometimes this difference can be significant when choosing.



\lx	asw
\hm	2
\ps	n
\pn	*kb*
\ge	pestle
\gn	*alu*
\pdl	*sg*
\pdv	**askwe**
\pdl	*pl*
\pdv	**asure**
\rf	d2.093.06
\xv	**Mal askwe ma mhwaw alyakwe ti neskwe.**
\xe	Get the pestle so you can pound the rice in the mortar.
\xn	*Ambil alu tumbuk padi di lesung.*

Fig. 103. The *asw* record found using the Jump To command

Shoebox will now bring up a window titled Shlex.db:3 displaying the record.

34. Close the window Shlex.db:3

Testing the jump paths

Now you can test the jump paths you created and experiment with jumping.

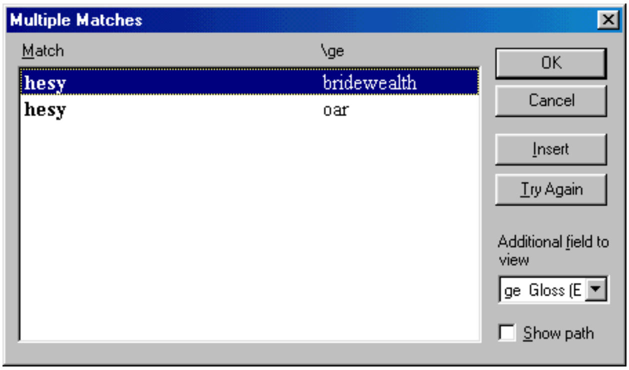
1. In the window Shlex.db:1, right-click *asw* (Macintosh: OPTION+click). Right-clicking a word is a convenient shortcut for the Jump To command.

Shoebox will display the Multiple Matches dialog box.

2. Click Cancel and return to the window you were in.

You will recall that the second jump path you created established a link from the English free translation to the English gloss field. Because of this you can jump from English words in the free translation as well.

3. Choose Database, Search for the word `hesy`.

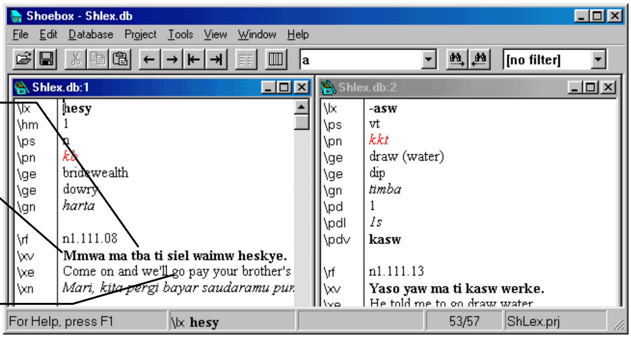


Match	\ge
hesy	bridewealth
hesy	oar

Fig. 104. Multiple Matches of the *hesy* search

4. Select the one that means *bridewealth*.

5. Click OK.



\lx	**hesy**
\hm	1
\ps	n
\pn	*kb*
\ge	bridewealth
\ge	dowry
\gn	*harta*
\rf	n1.111.08
\xv	**Mmwa ma tba ti siel waimw heskye.**
\xe	Come on and we'll go pay your brother's
\xn	*Mari, kita pergi bayar saudaramu pur*

\lx	**-asw**
\ps	vt
\pn	*kkt*
\ge	draw (water)
\ge	dip
\gn	*timba*
\pd	1
\pdl	*1s*
\pdv	**kasw**
\rf	n1.111.13
\xv	**Yaso yaw ma ti kasw werke.**
\xe	He told me to go draw water

Does *tba* mean *go*?

Example sentence

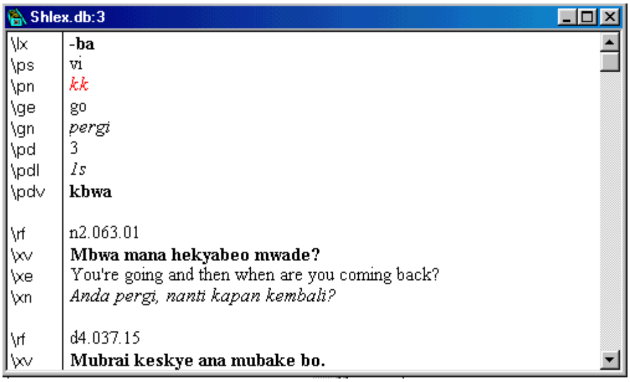
Put insertion point on *go* and jump to see.

Fig. 105. Entry for *hesy* meaning *bridewealth*

6. Scroll down to the example sentence.

Now assume that you need to determine whether the word *tba* is the vernacular word that is translated as *go*.

7. Jump from the word *go* in the English translation (Windows: right-click; Macintosh: OPTION+click).



```
\lx     -ba
\ps     vi
\pn     kk
\ge     go
\gn     pergi
\pd     3
\pdl    1s
\pdv    kbwa

\rf     n2.063.01
\xv     Mbwa mana hekyabeo mwade?
\xe     You're going and then when are you coming back?
\xn     Anda pergi, nanti kapan kembali?

\rf     d4.037.15
\xv     Mubrai keskye ana mubake bo.
```

Fig. 106. Shlex showing the record *-ba* meaning *go*

The program will quickly open a window displaying the lexeme *-ba* with the insertion point in the **\ge** *go* field.

8. Close this window.[18]

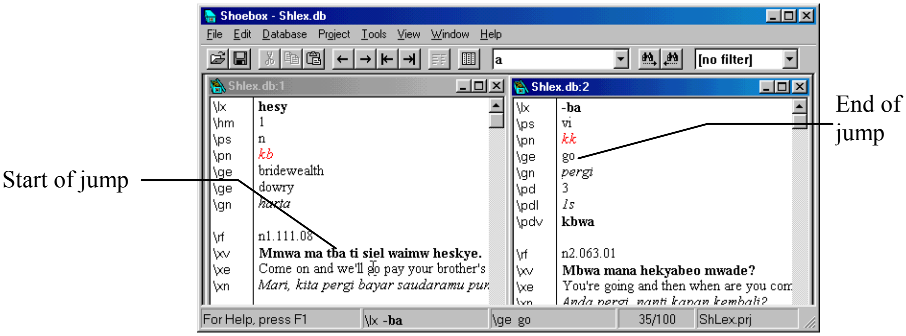
Another way to use jumping, which is very useful when interlinearizing, is to make an existing window a jump target. (You will use jumping and jump targets heavily later during interlinearization.)

9. Select the second window, which is sorted by English gloss.

10. Choose View, Jump Target.

11. Jump from the word *go,* which is still showing in the first window (Windows: right-click; Macintosh: OPTION+click).

[18] The Windows shortcut key for closing a window is CTRL+F4; on the Macintosh it is COMMAND+W.



```
\lx    hesy
\hm    1
\ps    n
\pn    kb
\ge    bridewealth
\ge    dowry
\gn    harta

\rf    n1.111.08
\xv    Mmwa ma tba ti siel waimw heskye.
\xe    Come on and we'll go pay your brother's
\xn    Mari, kita pergi bayar saudaramu pur
```

```
\lx    -ba
\ps    vi
\pn    kk
\ge    go
\gn    pergi
\pd    3
\pdl   1s
\pdv   kbwa

\rf    n2.063.01
\xv    Mbwa mana hekyabeo mwade?
\xe    You're going and then when are you com
\xn    Anda pergi, nanti kapan kembali?
```

```
For Help, press F1 | \lx -ba | \ge go | 35/100 | ShLex.prj
```

Fig. 107. Jump from the word *go* to the *-ba* record

Observe that the second window moves to show the word. This is very convenient because you do not have to keep closing jump windows.

Jump To can go only to a jump target window whose primary sort field is the same as one of the fields specified in the jump path. So a jump to a word that is in the gloss field must switch to a window that is sorted by the gloss field. (If no such window has been specified as a jump target, Shoebox opens a new window, as you saw earlier.)

If text is selected, the selection is used for the jump. If nothing is selected, the current word is used for the jump.

Search vs. Find vs. Jump

This table summarizes some important differences between Search, Find, and Jump:

Search searches only on the primary sort field.	*Find* can search any field, or any language.	*Jump* can jump to any field listed in the jump path.
Search searches the entire filtered database in the active window.	*Find* can be restricted to the current record or can search the entire filtered database.	*Jump* can search databases in other windows. It is *not* restricted to the active database, or even the same database type.
Search uses the index currently used for sorting. Sorting causes a delay initially, but once the index is built, searching is very fast.	*Find* always searches linearly through the entire record or the entire database, letter by letter, looking for the next occurrence of the string.	*Jump* builds an index for each database-field combination in its path. After building all the indexes, the jump is fast.
Search looks for the string as the first part of the primary sort field, or as the whole field if you so specify.	*Find* looks for the string anywhere it might occur, depending on whether it is finding by language or field. Find can be specified to match a whole word.	*Jump* (like Search) expects the string to be at the beginning of the field, but Jump is not restricted to the primary sort field; it can jump to any field in the jump path.

For example, if you use the Search command to find *the*—without specifying Match whole field—you will be given the choice among *the, the*m*, the*me*, the*n*, the*re*,* etc., i.e., all words that come at the beginning of that field and start with *the.* If you use the Find command to find *the*—without specifying Match whole word—it will stop at o*the*r, brea*the*, etc., in addition to all the words that Search would locate, i.e., anything containing the letters *the,* not just those words that begin with *the*. Jump behaves like Search when matching strings.[19]

Filters

Filters allow you to restrict your attention to a set of records that match a particular criterion. When you use a filter, the window will display only the matching records. Therefore, the Next Record and Previous Record commands will skip any records that do not match.

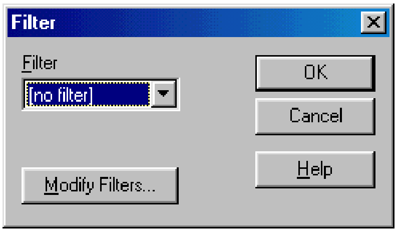
Like searching and jumping, filtering uses an index. When an index is built, there will be some delay depending on the speed of your machine and the size of your database. Once the index is built, finding data is quick. If you will be moving back and forth in the set of records that match some criteria, you might want the speed of an index. The Find command does not use an index. If you want to look up a data item only once, Find is probably a better choice.

Basic filters

For example, suppose you wish to examine lexemes in the Shlex database which contain the *ai* vowel combination.

1. Select window 1, which is sorted by lexeme.
2. Choose Database, Filtering.

[19] The Help file also has the chart information presented in a more terse form, organized by the three commands rather than by contrastive functions. Choose Help, Contents. Then, choose Search, Find, Jump To.



Filter: [no filter]

Fig. 108. Filter dialog box

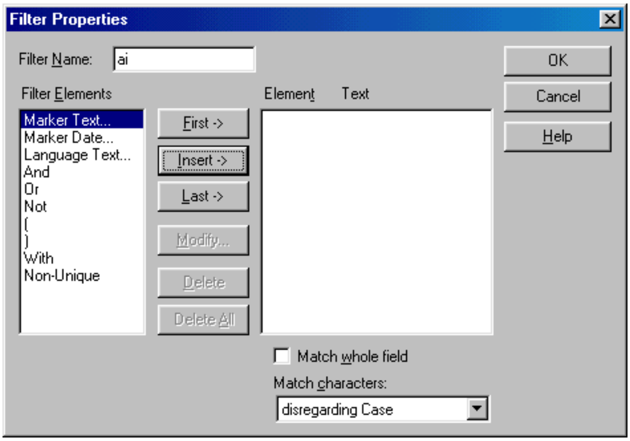
3. Click Modify Filters....

You see the Database Type Properties dialog box with the Filter tab selected.

4. Click Add....

5. Type `ai` as the Filter Name at the top of the dialog box.

Notice that there are two main columns, much as you have seen for Sorting, Jump Path, and elsewhere. The left column is labeled Filter Elements. It contains a list of all the types of filters available and the elements that connect them. The right column is empty since the filter has not yet been built.



Filter Name: ai

Filter Elements:
Marker Text...
Marker Date...
Language Text...
And
Or
Not
(
)
With
Non-Unique

□ Match whole field

Match characters: disregarding Case

Fig. 109. Filter properties dialog box

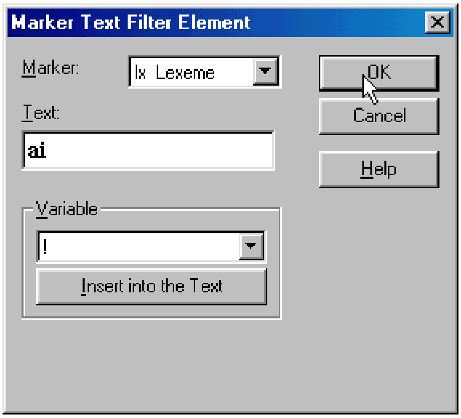
6. Select Marker Text in the left column.

7. Click Insert -> (or double-click Marker Text).

This will bring up another dialog box, which allows you to specify the marker and the text for the Marker Text filter element.

8. Select **lx** as the Marker.

9. Type `ai` as the Text.


Marker: lx Lexeme

Text: ai

Variable: !

Fig. 110. Marker Text Filter Element dialog box

10. Click OK several times until you get back to the Filter dialog box.

11. Select the filter ai in the drop-down list box.

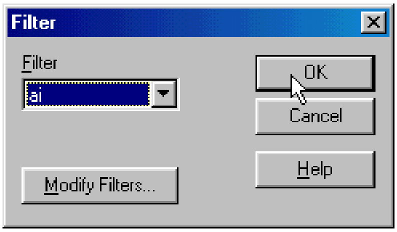

Filter: ai

Fig. 111. Filter selection dialog box

12. Click OK.

Shoebox will build the appropriate filter index. If the current record does not match the filter, you will be placed in the next record that does. Note that the status bar gives the count of the records matching the filter.

13. Choose View, Browse.

This is a good way to view the results of a filter. In this case, you can see the characteristics of the *ai* words in the **\lx** field.

Notice the rightmost box on the toolbar, the Active Filter box. It displays ai, the name of the active filter.

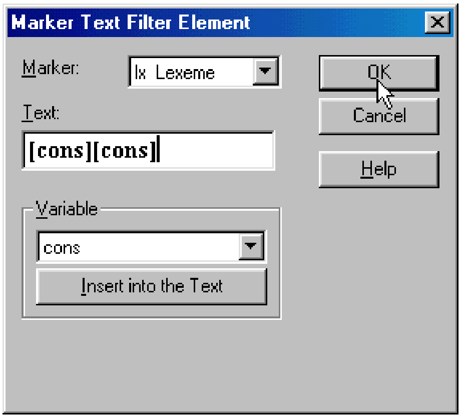
14. Select [no filter] in the Active Filter box (it is the first one in the list).

The window will again display all the records in the database.

Filters and variables

Here is a more complex example of using filters. Suppose you suspect that Selaru is an unusual Austronesian language because of its sound patterns. A simple syllable pattern of CVCV (consonant-vowel-consonant-vowel) is typical, but a quick glance at the data shows that Selaru has consonant clusters (CC patterns). You can use filters to examine this issue more carefully.

1. In window 1, choose Database, Filtering.
2. Click Modify Filters... . This displays the Filters tab on the Database Type Properties dialog box.
3. Click Add... .
4. Type `cc` as the new filter name.
5. Double-click Marker Text.
6. Select **lx** in the Marker window.
7. Drop down the list box under Variable, scroll down and click cons (the consonant variable).
8. Now click Insert into the Text two times. (This will put `[cons][cons]` into the Text box.)



Marker Text Filter Element

Marker: \lx Lexeme

Text: [cons][cons]

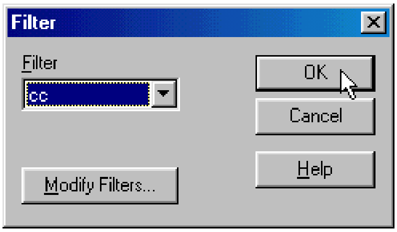
Variable: cons

Fig. 112. Marker Text Filter Element dialog box

9. Click OK three times.

You will see the Filter dialog box again.

10. Select the new cc filter from the drop-down list.



Filter

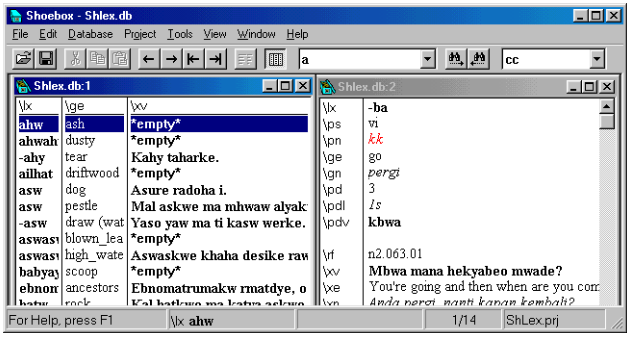
Filter: cc

Fig. 113. Filter selection dialog box

11. Click OK.

If the lexeme showing in window 1 did not contain a consonant cluster (a CC pattern), Shoebox will move to the next record in the database that does.

12. Click ▦.



\lx	\ge	\xv
ahw	ash	*empty*
ahwah	dusty	*empty*
-ahy	tear	Kahy taharke.
ailhat	driftwood	*empty*
asw	dog	Asure radoha i.
asw	pestle	Mal askwe ma mhwaw alyak
-asw	draw (wat	Yaso yaw ma ti kasw werke.
aswasw	blown_lea	*empty*
aswasw	high_wate	Aswaskwe khaha desike raw
babyay	scoop	*empty*
ebnon	ancestors	Ebnomatrumakw rmatdye, o

\lx	-ba
\ps	vi
\pn	*kk*
\ge	go
\gn	*pergi*
\pd	3
\pdl	*1s*
\pdv	kbwa
\rf	n2.063.01
\xv	Mbwa mana hekyabeo mwade?
\xe	You're going and then when are you com
\xn	*Anda pergi, nanti kapan kembali?*

For Help, press F1 | \lx ahw | 1/14 | ShLex.prj

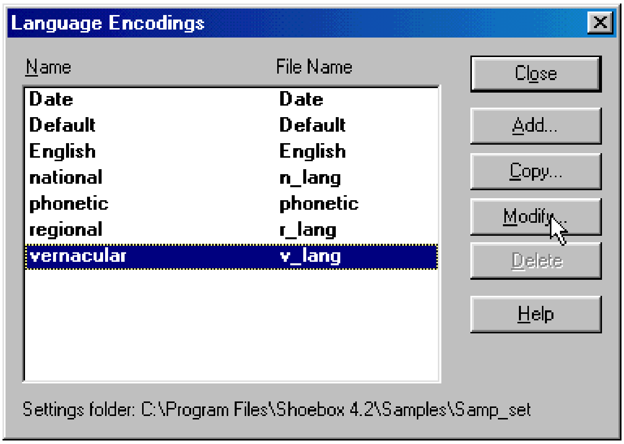
Fig. 114. Browse view showing consonant cluster filtering

You now have a listing of all the records that match the CC pattern you just specified for the filter. Looking down this list, you can see that most of the consonant clusters involve the glides *w* or *y*.

Adding a new variable

To help analyze Selaru glides, you can create a new variable.

1. Choose Project, Language Encodings.



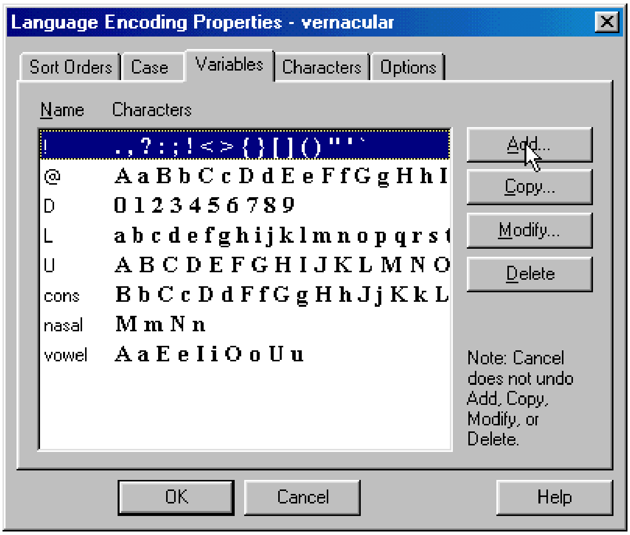
Name	File Name
Date	Date
Default	Default
English	English
national	n_lang
phonetic	phonetic
regional	r_lang
vernacular	v_lang

Settings folder: C:\Program Files\Shoebox 4.2\Samples\Samp_set

Fig. 115. Language Encodings dialog box

2. Select vernacular and click Modify... (or simply double-click vernacular).

3. Click the Variables tab.



Name	Characters
!	. , ? : ; ! < > { } [ ] ( ) " ' `
@	A a B b C c D d E e F f G g H h I
D	0 1 2 3 4 5 6 7 8 9
L	a b c d e f g h i j k l m n o p q r s t
U	A B C D E F G H I J K L M N O
cons	B b C c D d F f G g H h J j K k L
nasal	M m N n
vowel	A a E e I i O o U u

Note: Cancel does not undo Add, Copy, Modify, or Delete.

Fig. 116. Language Encoding Properties for the vernacular language

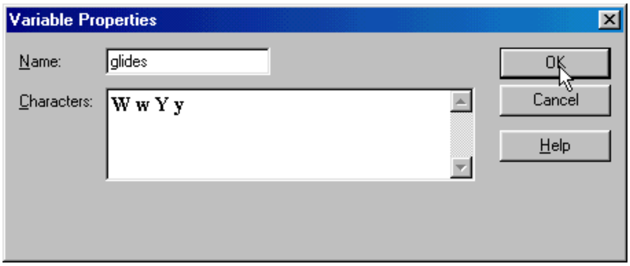
This shows all the variables that have been created for the vernacular language.[20] Variables are user definable, not fixed by Shoebox, so you may change any of these as you wish. Each language maintains its own set of variables, now you will add a variable for glides.

4. Click Add... .

5. In the Name box, type `glides`.

6. In the Characters box, type `W w Y y` (with one space between each letter).

[20] The concepts of projects and language encodings are described more fully in the section *Shoebox settings files*, p. 113.

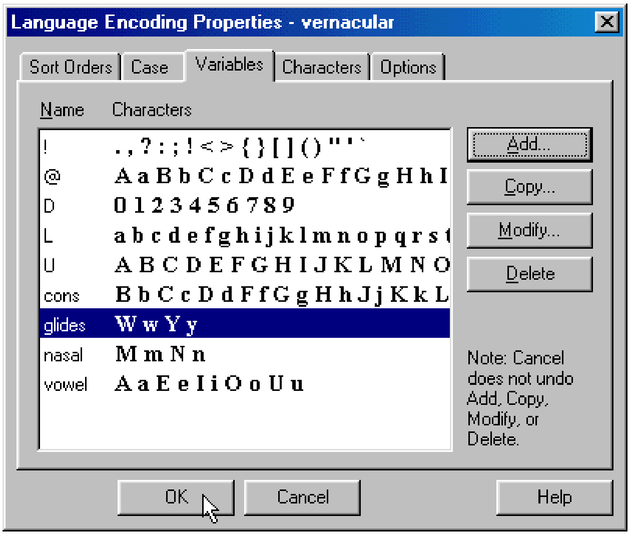


Name: glides

Characters: W w Y y

Fig. 117. Variable Properties with the glides *W*, *w*, *Y*, and *y* defined

7. Click OK.



Name	Characters
!	. , ? : ; ! < > { } [ ] ( ) " ' `
@	A a B b C c D d E e F f G g H h I
D	0 1 2 3 4 5 6 7 8 9
L	a b c d e f g h i j k l m n o p q r s t
U	A B C D E F G H I J K L M N O
cons	B b C c D d F f G g H h J j K k L
glides	W w Y y
nasal	M m N n
vowel	A a E e I i O o U u

Note: Cancel does not undo Add, Copy, Modify, or Delete.

Fig. 118. Language Encoding Properties - vernacular with glides defined

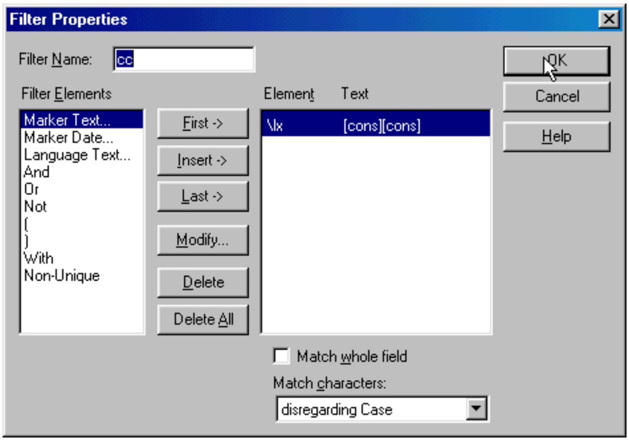
Here you can see that Shoebox has placed the new glide variable into the list of variables for vernacular.

Note that *W*, *w*, *Y*, and *y* are also still defined as consonants. This is okay since glides often act like consonants.

Complex filters

Now back to filtering. To see what impact those glides have on consonant cluster patterns in Selaru, you will weed out any lexeme with a consonant cluster that has a glide in it. This will enable you to see just how many lexemes have CC patterns *without* glides.

1. Click OK and then Close to return to the database.
2. Now choose Database, Filtering.
3. Click Modify Filters... .
4. Select the cc filter and click Copy... .



Filter Name: cc

Marker Text...
Marker Date...
Language Text...
And
Or
Not
(
)
With
Non-Unique

Element	Text
\lx	[cons][cons]

□ Match whole field

Match characters: disregarding Case

Fig. 119. Filter Properties for renaming cc filter

5. Give this copy of the cc filter a new name: `cc-no glide`.

Notice that this filter (since it is a copy of the cc filter) is already set up to constrain the database to show only lexemes with complex consonant patterns. You now need to add the further constraint of *not* showing those with glides.

6. Select And, then click Insert -> (or double-click And).
7. Select Not, then click Insert -> (or double-click Not).

This sets up the logic for the next constraint.

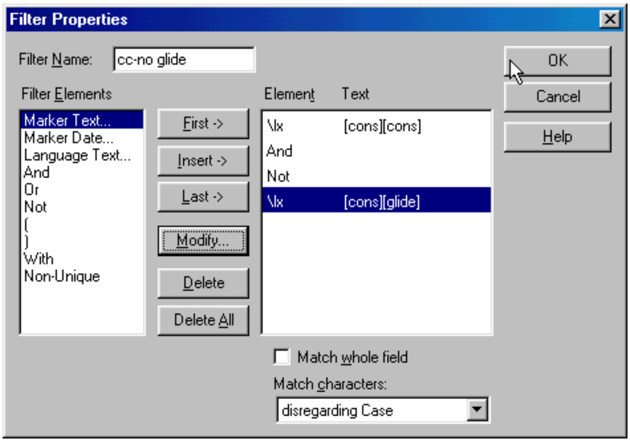
8. Select Marker Text, then click Insert -> (or double-click Marker Text).
9. Select **lx** in the Marker window.

10. Drop down the list box under Variable, scroll down and click the variable cons (as you did before).

11. Click Insert into the Text to put [cons] into the Text box.

12. Drop down the Variable list box again and click glides.

13. Click Insert into the Text to make the Text box read `[cons][glides]`.



Filter Properties

Filter Name: cc-no glide

Filter Elements: Marker Text..., Marker Date..., Language Text..., And, Or, Not, (,), With, Non-Unique

Element	Text
\lx	[cons][cons]
And	
Not	
\lx	[cons][glide]

□ Match whole field

Match characters: disregarding Case

Fig. 120. Filter Properties for consonant cluster, no glide filter

14. Click OK.

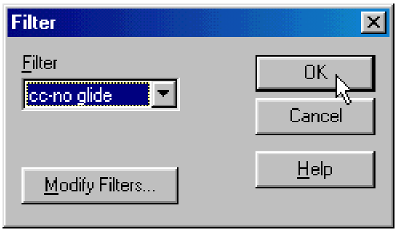
You are now back at the main cc-no glide filter dialog box. Notice how the complete filter can be understood as, “Show only records with lexemes that match these two constraints: the lexeme must contain a CC pattern *and* it must *not* contain a pattern with a glide as the second feature.”

To activate the filter,

15. Click OK two times.

This should put you back into the small Filter dialog box.

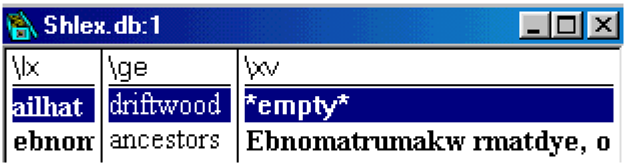
16. Select cc-no glide filter from the drop-down list.



Filter: cc-no glide

Fig. 121. Selecting the cc-no glide filter

17. Click OK.



\lx	\ge	\xv
ailhat	driftwood	*empty*
ebnorr	ancestors	Ebnomatrumakw rmatdye, o

Fig. 122. Browse view showing consonant clusters without glides

If window 1 is still in Browse mode, you should see the list of records drop down to only two. This shows you how rare non-glide complex consonant words are in Selaru. This example serves to demonstrate how filters can be used in analysis to help limit the clutter and focus on the pertinent data.[21]

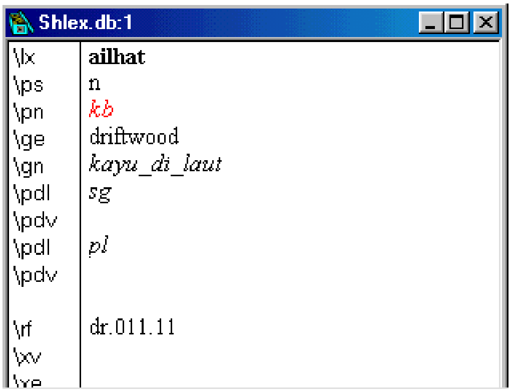
Undo

Shoebox keeps track of all the editing changes you make to a record. If you change your mind or make a mistake, you can usually reverse the actions you have made since starting to edit the record.

1. Go back to record view in database 1 (that is, not browse view).

2. Go to the first record of the database.

[21] A final comment on filters: If you plan to change the primary sort field and then use a filter, change the sorting first. If you select the filter first, it will build the filter index according to the current sorting. Then when you change the sorting, it will have to rebuild the filter index as well as build the sorting index for the whole (unfiltered) database.



\lx	**ailhat**
\ps	n
\pn	*kb*
\ge	driftwood
\gn	*kayu_di_laut*
\pdl	*sg*
\pdv	
\pdl	*pl*
\pdv	
\rf	dr.011.11
\xv	
\xe	

Fig. 123. First record of filtered database

3. Make several changes to the record.

4. Choose Edit, Undo.

This will undo the last change you made.

5. Keep choosing Edit, Undo to watch the record revert back to its initial state.

The Undo All command restores the original contents of a record. Use it only if you make a serious mistake and decide that you do not want to keep any of the changes you have made to a record.

6. Make several changes to the record again.

7. Choose Edit, Undo All.

Shoebox will display a message that asks: Are you sure you want to undo *all* the changes to the record?

8. Click Yes.

All the changes you made will be undone. The *ailhat* record will look exactly as it did before you made this group of changes.

Important: Shoebox keeps track of changes to a record in a window as long as that record is displayed. If you have multiple windows showing data, and you click a different window (and even edit there), then return to the first window (which still shows the original record), the changes in the first window can still be undone. But once you switch to a different record in that window (using Search, etc.), you *cannot* use Undo or Undo All to undo changes to the first record that were made before you switched records.

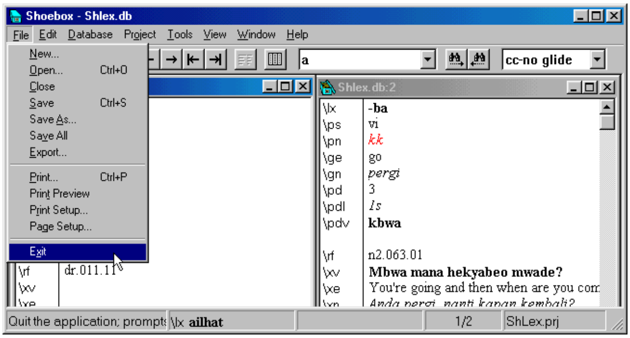
Undo and Undo All have some limitations:

- Undo does not undo movement—neither within a record nor from one to the next.
- If you save the database, you cannot undo changes made before the save.
- If you make a change to a marker's properties—for example, if you change its field name—that is not a change to the record. That change *cannot* be undone, but editing changes made before and after it can be undone.

Stopping and starting Shoebox

When you exit Shoebox and restart it again, it remembers all the windows you had open and their settings.

1. Choose File, Exit (Macintosh: choose File, Quit).



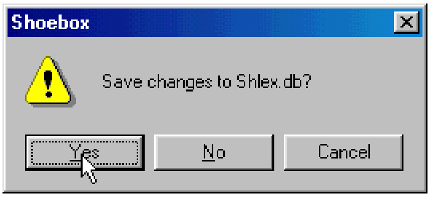
```
\rf   dr.011.11
\xv
\xe
```

```
\lx   -ba
\ps   vi
\pn   kk
\ge   go
\gn   pergi
\pd   3
\pdl  1s
\pdv  kbwa

\rf   n2.063.01
\xv   Mbwa mana hekyabeo mwade?
\xe   You're going and then when are you com
\xn   Anda pergi, nanti kapan kembali?
```

Quit the application; prompt | \lx ailhat | 1/2 | ShLex.prj

Fig. 124. Exiting Shoebox with current databases and settings

2. To save changes, click Yes (Macintosh: click Save).



Save changes to Shlex.db?

Fig. 125. Saving your changes

3. Run Shoebox again.

Observe that you get the same two windows sorted the same way and showing the same records as when you quit the program.

4. Close Shoebox before continuing on to the next section.

Tutorial milestone

This is a milestone in the *Tutorial*. Nothing further builds on changes you have made in this section.

Shoebox settings files

Overview

In Shoebox, there are three kinds of settings files: projects, language encodings, and database types. Understanding the kinds of "settings" that Shoebox stores and how they work together will enable you to create a new Shoebox project and database.

Projects

A project is primarily the configuration of your Shoebox workspace. A project settings file remembers all the database files that were open in Shoebox, the placement of the windows you had open, the data that each window was showing, and how each window was sorted and filtered.

When you exit Shoebox and later restart it, it automatically opens the last project you used and restores your workspace.

In Shoebox, you can have more than one project. Projects are an excellent way to organize your Shoebox work by tasks or academic domains. Each task or domain can have its own project file. For example, you can set up separate projects for:

- dictionary work
- anthropology research
- grammatical analysis
- text analysis
- translation work

This makes switching between tasks as simple as opening the appropriate project.

Language encodings

A language has many features. One is the script normally used when writing the language. In Shoebox, the set of script-related features is called a language encoding. (If a language can be written in two different scripts, each script is a language encoding.)

A language encoding consists of:

- The appropriate sorting order for that script.
- Upper and lower case forms of the characters (if any).
- Special groupings (called Variables) of the characters that are useful in examining or searching the data.
- A font to represent the character shapes.
- Often a special keyboard to facilitate entering the characters.

If two languages can be written using the same script and sorting order, it is possible to represent both with a single language encoding. While this might save a few minutes in setup, there are some significant advantages to representing each language with a unique language encoding:

- The Find command can be used to look for data in all fields of a specific language. (The example in the *Find* section of the *Tutorial* illustrates this.)
- If you want to distinguish between two languages by font style, it is much easier to do so with separate language encodings. Markers assigned to a particular language encoding will automatically inherit the font style of the language (unless you override this).
- When a marker is assigned to the appropriate language encoding, this can help document the purpose of the marker.

Database types

In Shoebox, a database type defines the structure of a particular kind of database, such as:

- lexicons
- anthropology notes
- interlinear texts

A database type file includes:

- All the field markers found in a database file of this type.
- A name for each field marker.
- Each field marker's position in a hierarchy.
- Each field marker's association with a language encoding.
- Any restricted contents (a range set) for any marker.
- A marker's style type for exporting to Microsoft Word.
- A template of field markers for new records (i.e., to insert a specified set of field markers for the new record).
- All of the filters for this database type (which often specify particular field markers).
- Any jump paths that have been set up.
- Any interlinear or adaptation information associated with this database type (to be described later in the *Tutorial*).
- All export processes for databases of this type.

How the three work together

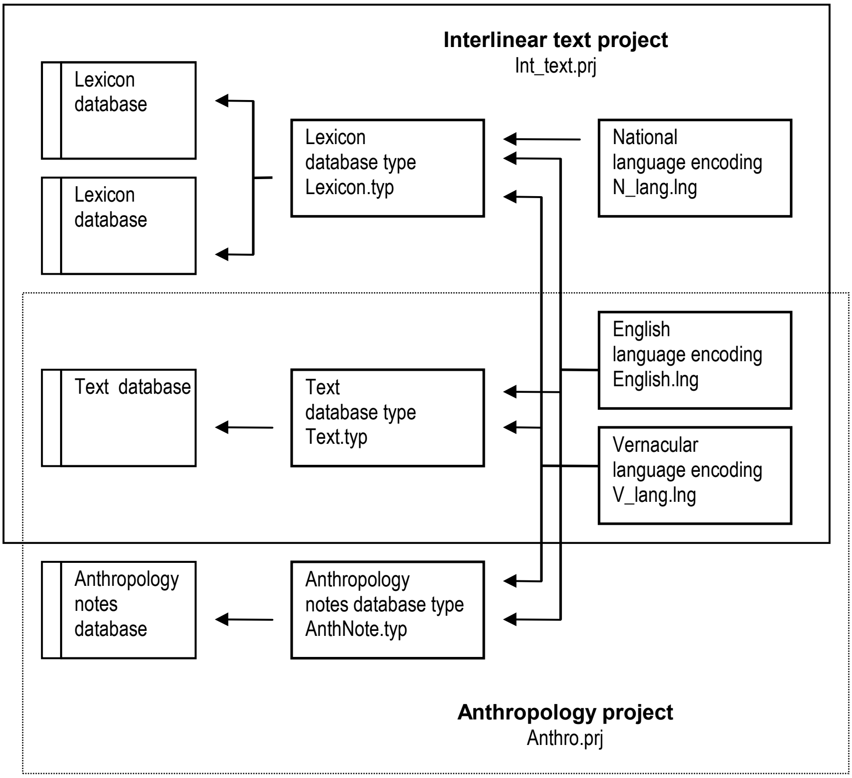
All currently open files are part of a project. However, a file does not belong to a particular project since several projects might share the same file. For example, the lexicon might be used by the grammar analysis project or the text analysis project, as well as by the dictionary project.

Each database file has a database type. Several files may have the same database types. For example, multiple text files might have the same interlinear text database type. A project may have files of various database types open at once.

Each marker in a database type has a language encoding. However, the language encoding does not belong to the database type. Field markers from different database types may have the same language encoding. For example, all field markers that contain vernacular data, regardless of the database type, would have the vernacular language.

Even when a file is part of more than one project, the database type of the file and the language encoding of its markers remain the same from one project to the next.

The following diagram shows two projects. The first is an interlinear text project which includes the lexicon (split into two files) and a file of texts to be interlinearized. The second is an anthropology project which includes a database of culture notes and the same text file used by the interlinear project. The text file is included in the anthropology project because it contains folk tales and other culturally relevant texts.


Interlinear text project
Int_text.prj
Lexicon database
Lexicon database
Lexicon database type Lexicon.typ
National language encoding N_lang.lng
English language encoding English.lng
Text database
Text database type Text.typ
Vernacular language encoding V_lang.lng
Anthropology notes database
Anthropology notes database type AnthNote.typ
Anthropology project
Anthro.prj

- The lexicon databases use the Lexicon database type. These lexicon databases include data in three languages: the national language, the vernacular language, and English.

- The text database uses the Text database type. It includes data both in English and in the vernacular language.
- The anthropology notes database uses the Anthropology notes database type. It includes data both in English and in the vernacular language.

In Shoebox, you need to keep related project (.prj) files, database type (.typ) files, and language encoding (.lng) files in the same folder.

You will want to keep your database files (the cultural notes, lexical data, and texts) separate in their respective folders.[22] Arranging your files in this way allows Shoebox to use the same settings files between projects. In this example, even though two projects and three database types have data in the vernacular language, the information about sorting and font selection for the vernacular language is specified only once (in the V_lang.lng file).

Once you have set up the language encoding and database type files, you can create a new database by choosing File, New and selecting its type. For example, if the linguist using the above setup wanted to create another interlinear text database, he need only choose File, New and then select the Text database type from the list of choices. Shoebox knows what to do from there on.

Overview: starting a new project

In this section of the tutorial, you will build a very small sample lexical database, using Axininca Campa language data.[23] The following is an overview of the steps involved in starting a new Shoebox project:

1. Install any needed fonts.

[22] It is recommended that you *not* keep data files mixed in with your settings files in the projects folder. Mixing data files and settings files should generally be avoided unless you need to share data with others, you are testing other people's data, or you are developing sample data for training purposes, etc.

[23] Axininca Campa data used with permission from David L. Payne, 1981. *The Phonology and Morphology of Axininca Campa* (SIL International and The University of Texas at Arlington). Axininca is pronounced *aSHINinca.*

2. Set up Keyman, if needed.
3. Run Shoebox.
4. Create a new project.
5. Create a new database type, which includes setting up:
   a. field markers
   b. a database template
6. Create a new language encoding, which includes setting up:
   a. a keyboard
   b. a default font
   c. a sort order
   d. case associations
   e. variables

Special characters

Many languages have “special characters” (either modified Roman or a different script). Therefore, they need special sort orders, fonts, keyboards, etc. Steps 1, 2, and most of 6 deal with these issues. If you work with the plain Roman alphabet and do not need to learn about handling special characters, you can skip ahead to the *Creating a new project* section, p. 130. In step 6, an asterisk (*) marks the sections and steps you can skip as well.

Installing a font

Typical of most languages, Axininca has some special characters. The Shoebox samples include a font containing these special characters. The font files are in the Fonts folder. (In the folder in which Shoebox was installed you will find a folder called Samples. The Fonts folder is in the Samples folder.)

To install the font on your computer, follow the appropriate instructions. (The font will also be used for the *Interlinear text* section of the *Tutorial*. After you finish that section, you can delete them.)

Macintosh

These are the steps for installing a font on the Macintosh:

1. Go to the Fonts subfolder in the Samples folder.

2. Select Axininca Ft SILDoulos.
3. Drag it into the Fonts subfolder in the System folder.

Windows 3.1

These are the steps for installing a font on Windows 3.1:

1. In the Program Manager, in the Main Group, bring up the Control Panel.
2. Double-click Fonts (the ABC icon). You will see a dialog box with numerous fonts already listed.
3. Click Add.

This dialog box has an area for listing fonts (this is probably blank), a directory box, and a check box to the lower right of the directory box, plus the usual buttons.

4. Clear □ Copy fonts to Fonts folder if it is checked.
5. Skip to the *Windows 3.1 and 95* section.

Windows 95

These are the steps for installing a font on Windows 95 and later:

1. Choose Start, Settings, Control Panel.

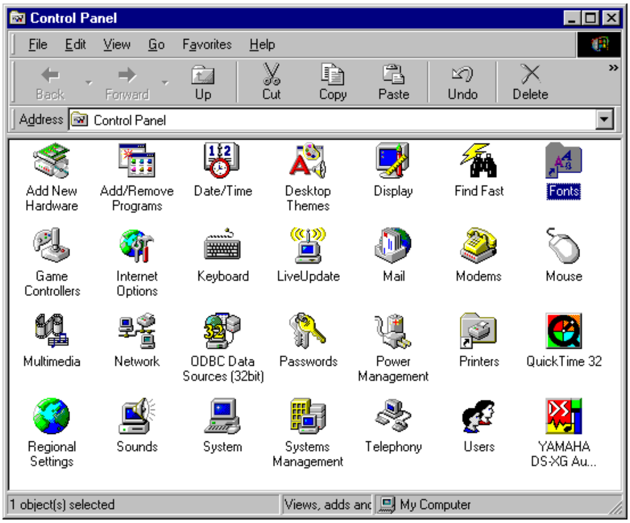
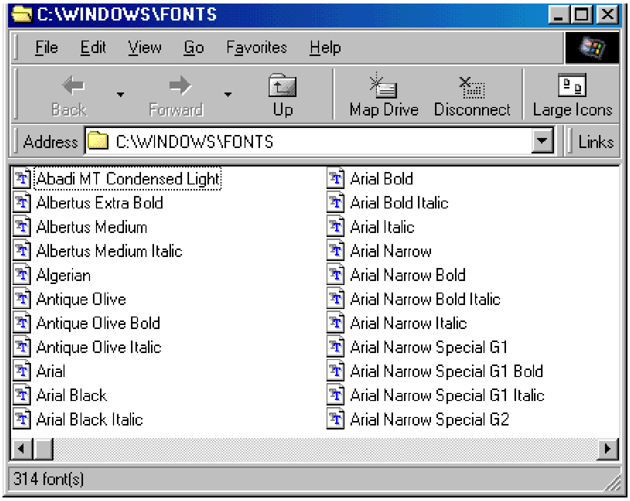


Fig. 126. Control Panel

You will see a number of icons displayed.

2. Double-click Fonts.



Address C:\WINDOWS\FONTS

Abadi MT Condensed Light
Albertus Extra Bold
Albertus Medium
Albertus Medium Italic
Algerian
Antique Olive
Antique Olive Bold
Antique Olive Italic
Arial
Arial Black
Arial Black Italic
Arial Bold
Arial Bold Italic
Arial Italic
Arial Narrow
Arial Narrow Bold
Arial Narrow Bold Italic
Arial Narrow Italic
Arial Narrow Special G1
Arial Narrow Special G1 Bold
Arial Narrow Special G1 Italic
Arial Narrow Special G2

314 font(s)

Fig. 127. C:\WINDOWS\FONTS folder

You open a window with a list of available fonts.

3. Choose File, Install New Font from the menu bar.

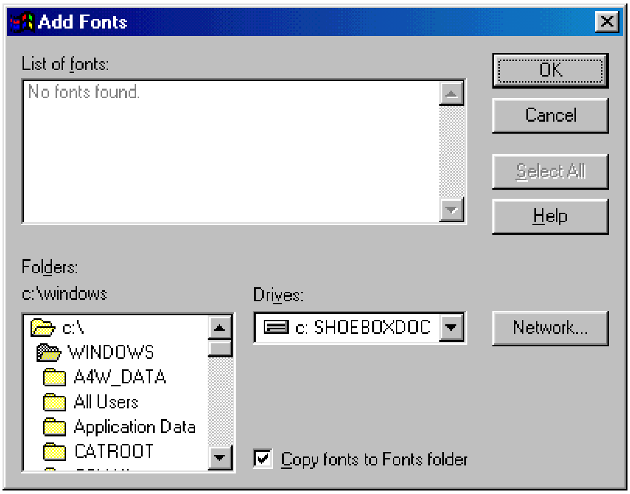


Fig. 128. Add Fonts dialog box

The Add Fonts dialog box has an area for listing fonts (this is probably blank), a folders box, and a check box to the lower right of the folders box, plus the usual buttons.

4. Clear ☐ Copy fonts to Fonts folder if it is checked.

Windows 3.1 and 95

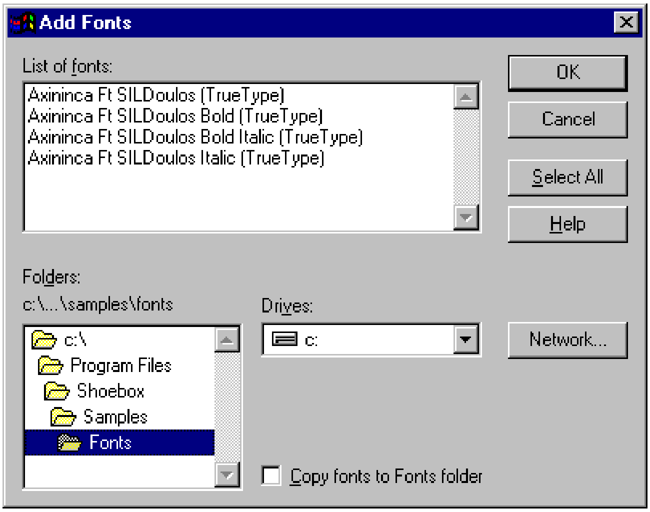
The following steps are common to both Windows platforms:

1. Navigate into the Shoebox program folder.

You can refer to Fig. 12 on p. 14 for information on the default directory structure of the Shoebox installation.

2. Navigate down through the Samples folder into the folder titled Fonts.

You will see four lines appear in the List of fonts.



List of fonts:
Axininca Ft SILDoulos (TrueType)
Axininca Ft SILDoulos Bold (TrueType)
Axininca Ft SILDoulos Bold Italic (TrueType)
Axininca Ft SILDoulos Italic (TrueType)

Folders:
c:\...\samples\fonts
c:\
Program Files
Shoebox
Samples
Fonts

Drives:
c:

□ Copy fonts to Fonts folder

Fig. 129. Add Fonts dialog box showing the folders and the Axininca fonts

3. Click Select All.

4. Click OK.

5. Close the Fonts and Control Panel windows to get back to your desktop.

Setting up Keyman (Windows only)

If you do not intend to use the Tavultesoft Keyboard Manager (Keyman), you can skip this section. Keyman runs only on Windows, not on Macintosh. In the tutorial, typing special characters is optional. However, if you do use Keyman, you will be interested to know how it works with Shoebox. The Shoebox CD-ROM includes two versions of Keyman.

- On Windows 3.1, use version 3.2 of Keyman.
- On Windows 95/98/2000/NT, you can use version 4 (or later) of Keyman. However, if you already have a Keyman 3.2 setup for Windows 95/98, you can continue to use it.

For more detailed information about using Keyman with Shoebox, search for Keyman in Shoebox Help.

Keyman 3.2

Here are the instructions for using Keyman 3.2:

1. If Shoebox is already running, choose File, Exit.

Shoebox only checks which keyboards are loaded when it starts.

2. On the Windows taskbar, click the Start button and follow the menus to find Keyman 3.2.

You probably have Keyman in your Startup folder or in the Tavultesoft Keyboard Manager folder.

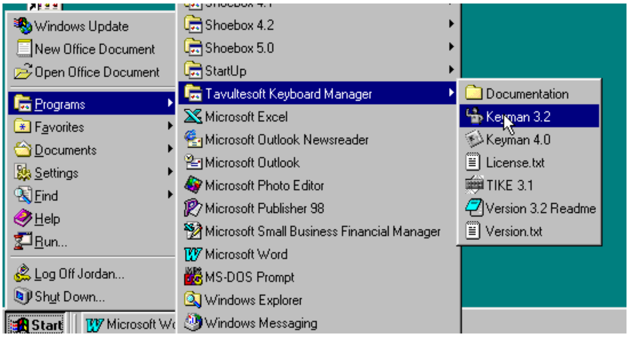


Fig. 130. Start menu showing Keyman 3.2

3. Click Keyman 3.2.

You will see the Tavultesoft Keyboard Manager dialog box (with the five large buttons—Hide, About, Options, Exit, and Help).

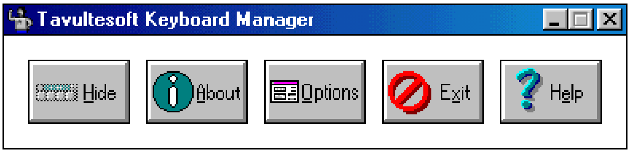
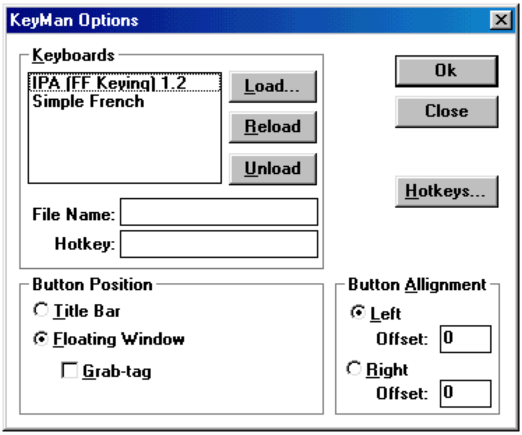


Fig. 131. Keyman Options buttons

4. Click Options.

You will see the Keyman Options dialog box with a number of options available.



Keyboards

IPA [FF Keying] 1.2
Simple French

File Name:

Hotkey:

Button Position

○ Title Bar
● Floating Window
□ Grab-tag

Button Allignment

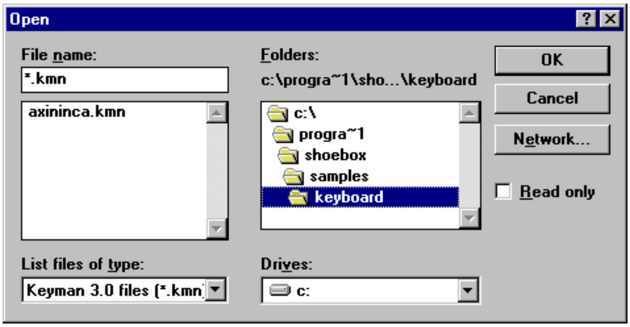
● Left
Offset: 0
○ Right
Offset: 0

Fig. 132. Keyman Options dialog box

5. Click Load....

This brings up a typical Open dialog box.

6. Navigate into the Keyboard subfolder that is in the Shoebox samples folder.



File name: *.kmn

axininca.kmn

Folders: c:\progra~1\sho...\keyboard

c:\
progra~1
shoebox
samples
keyboard

□ Read only

List files of type: Keyman 3.0 files (*.kmn

Drives: c:

Fig. 133. Open dialog box for Axininca keyboard selection

7. Double-click axininca.kmn.

This should return you to the five-button dialog box.

8. Click Hide.

Note: When you start Shoebox, it checks to see which keyboards Keyman has loaded. Therefore, you should always start Keyman before you start Shoebox.

Keyman 4

Here are the instructions for using Keyman 4:

1. On the Windows taskbar, click the Start button and follow the menus to find Keyman 4.0.

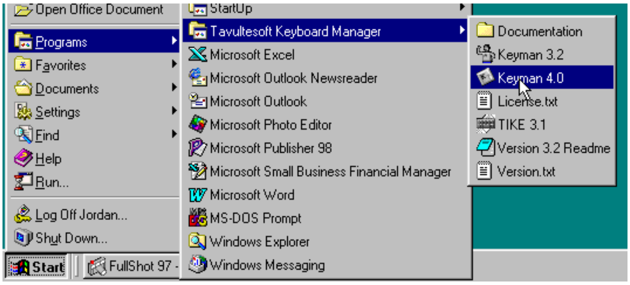
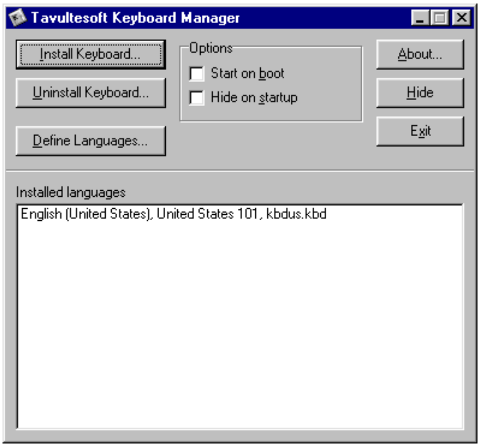


Fig. 134. Start menu showing Keyman 4.0

2. Click Keyman 4.0.

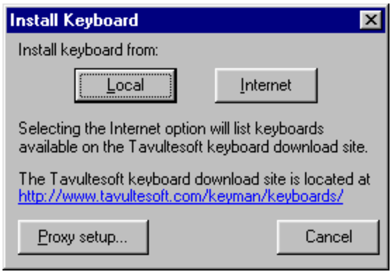


Options
□ Start on boot
□ Hide on startup

Installed languages
English (United States), United States 101, kbdus.kbd

Fig. 135. Tavulesoft Keyboard Manager window

3. Click Install Keyboard....



Install Keyboard

Install keyboard from:

Local | Internet

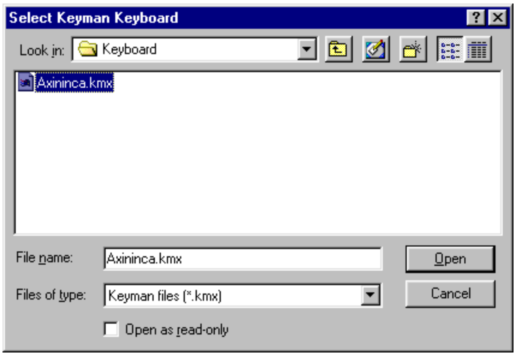
Selecting the Internet option will list keyboards available on the Tavultesoft keyboard download site.

The Tavultesoft keyboard download site is located at http://www.tavultesoft.com/keyman/keyboards/

Proxy setup... | Cancel

Fig. 136. Install Keyboard dialog box

4. Click Local.



Select Keyman Keyboard

Look in: Keyboard

Axininca.kmx

File name: Axininca.kmx

Files of type: Keyman files (*.kmx)

Open | Cancel

□ Open as read-only

Fig. 137. Select Keyman Keyboard dialog box

5. Navigate into the Keyboard subfolder that is in the Shoebox samples folder.

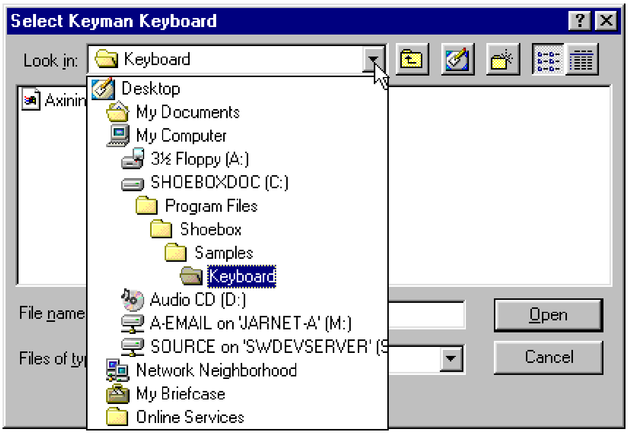
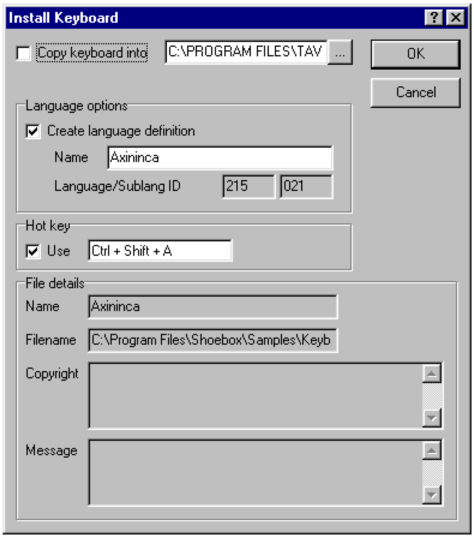


Fig. 138. Directory showing the location of the Keyboard folder

6. Double-click Axininca.kmx.



☐ Copy keyboard into C:\PROGRAM FILES\TAV

☑ Create language definition

Name: Axininca

Language/Sublang ID: 215 021

☑ Use Ctrl + Shift + A

Name: Axininca

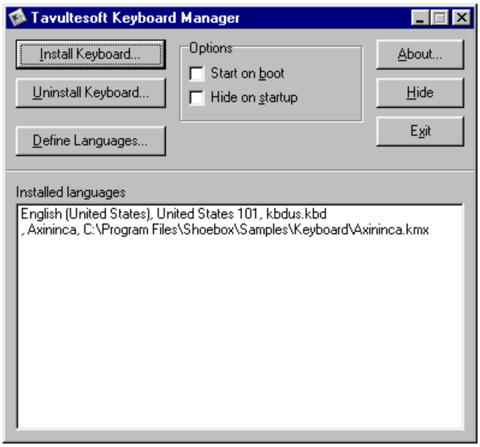
Filename: C:\Program Files\Shoebox\Samples\Keyb

Copyright:

Message:

Fig. 139. Install Keyboard dialog box for Axininca

7. Click OK.



Installed languages

English (United States), United States 101, kbdus.kbd
, Axininca, C:\Program Files\Shoebox\Samples\Keyboard\Axininca.kmx

Fig. 140. Tavultesoft Keyboard Manager showing Axininca installed

8. Click Hide.

You can use Keyman keyboards in the same way as standard Windows keyboard layouts.

9. On the Windows taskbar, click the Start button and follow the menus to find the Control Panel.

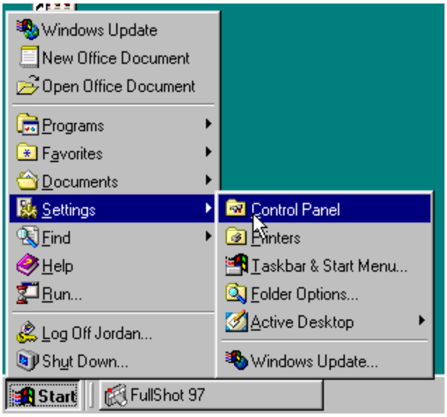
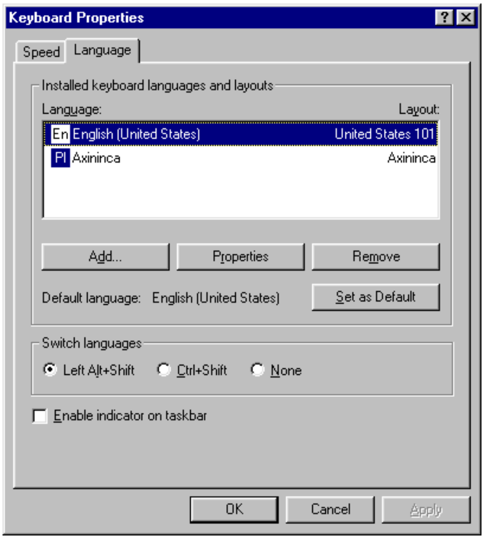


Fig. 141. Start menu showing path to the Control Panel

10. Click Control Panel.

11. Double-click Keyboard in the Control Panel window.

12. Click the Language tab.



Installed keyboard languages and layouts

Language:	Layout:
En English (United States)	United States 101
Pl Axininca	Axininca

Default language: English (United States)

Switch languages

● Left Alt+Shift ○ Ctrl+Shift ○ None

□ Enable indicator on taskbar

Fig. 142. Keyboard Properties showing Axininca installed

13. Clear □ Enable indicator on taskbar, if it is checked.

Keyman has its own indicator on the Windows taskbar, therefore the Windows indicator is not needed.

14. Click OK.

Note: After you have finished the tutorial, open this window again, select Axininca, and click Remove. You will probably not use this keyboard for your own language work.

Creating a new project

Shoebox establishes a project by creating a settings file containing the relevant information.

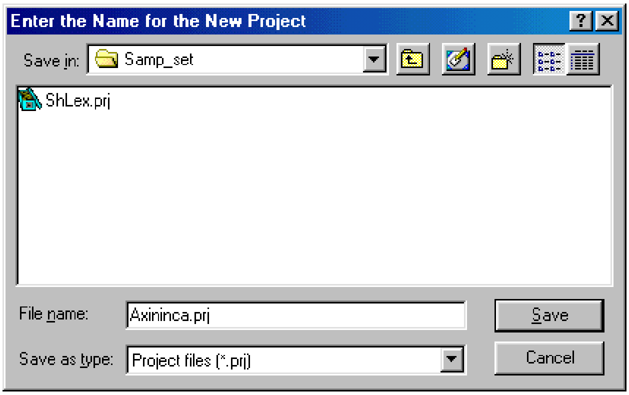
1. On the Windows taskbar choose Start, Programs, Shoebox 5.0, Shoebox 5.0 to run the Shoebox program.

2. Choose Project, New on the menu bar.[24]

This brings up a dialog box asking for a file name for the new project.

3. Navigate into the Samp_set folder. (In the folder in which Shoebox was installed you will find a folder called Samples. Samp_set is in the Samples folder.)

4. Type `Axininca.prj` as the project file name.[25]



Save in: Samp_set

ShLex.prj

File name: Axininca.prj

Save as type: Project files (*.prj)

Fig. 143. Enter Name for the New Project dialog box

5. Click Save.

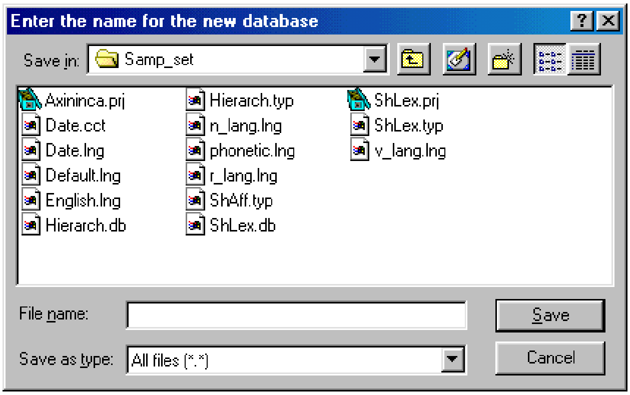
Shoebox creates a new project named Axininca.prj in the Samp_set folder.

[24] You do not need to close the current project before you create a new one. Shoebox can have only one project open at a time, so choosing Project, New will cause Shoebox to save and close the current project before it creates the new one. The same is true if you choose Project, Open while a project is still loaded.

[25] If you leave off the .prj extension, Shoebox will add it automatically.

There are no database windows open in this new project. Now you will create a small Axininca Campa–English lexicon in your new project.

6. Choose File, New.



Enter the name for the new database

Save in: Samp_set

Axininca.prj
Date.cct
Date.lng
Default.lng
English.lng
Hierarch.db
Hierarch.typ
n_lang.lng
phonetic.lng
r_lang.lng
ShAff.typ
ShLex.db
ShLex.prj
ShLex.typ
v_lang.lng

File name:

Save as type: All files (*.*)

Fig. 144. Enter the name for the new database dialog box

Shoebox displays a dialog box with options for specifying the location and file name of the new database.

7. Navigate to the Ax_Data folder (In the folder in which Shoebox was installed you will find a folder called Samples. Ax_Data is in the Samples folder.)

8. Type `Axininca.Lex` as the file name.

9. Click Save.

Creating a new database type

Shoebox displays the Select Database Type dialog box. Although a new database often belongs to an already-existing database type, you will create a new type since this is an example of how to start from scratch.

1. Click Add New Type.

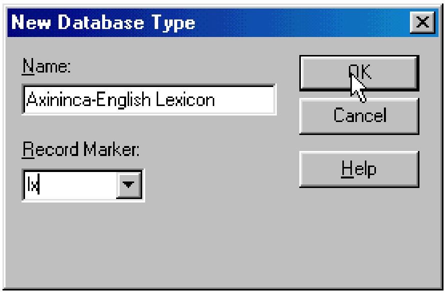
This brings you to the Database Types dialog box which lists all the available types.

2. Click Add....

This brings you to the New Database Type dialog box It is asking for the name of the new database type and for the record marker. The name does not have to be a single word.

3. Type `Axininca-English Lexicon` as the Name.

4. Type `lx` as the Record Marker.



Name: Axininca-English Lexicon

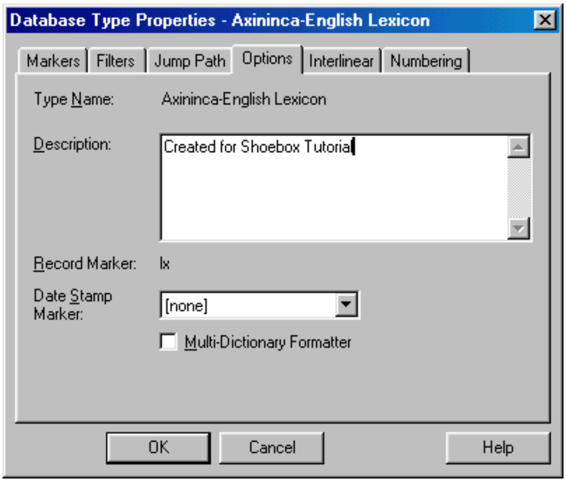
Record Marker: lx

Fig. 145. New Database Type dialog box

5. Click OK.

This brings you to a large, tabbed dialog box titled Database Type Properties. You are at the tab labeled Options. The program is now inviting you to enter a description of the database type.[26]

[26] This is useful when you are returning from a significant break in your work or when someone else is working with your data.


Database Type Properties - Axininca-English Lexicon

Type Name: Axininca-English Lexicon

Description: Created for Shoebox Tutorial

Record Marker: lx

Date Stamp Marker: [none]

□ Multi-Dictionary Formatter

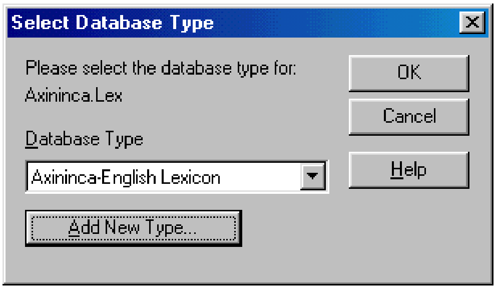
Fig. 146. Database Type Properties dialog box for Axininca

6. Type `Created for Shoebox Tutorial` as the Description.

7. Click OK from the Database Type Properties dialog box.

8. Click Close from the Database Types dialog box.

Now you are at the small Select Database Type dialog box. You should see Axininca-English Lexicon showing as the choice for the database type.


Select Database Type

Please select the database type for: Axininca.Lex

Database Type: Axininca-English Lexicon

Fig. 147. Select Database Type dialog box for the Axininca Lexicon

9. Click OK.

You will see a window titled Axininca.Lex. It shows an empty record with one **\lx** marker.

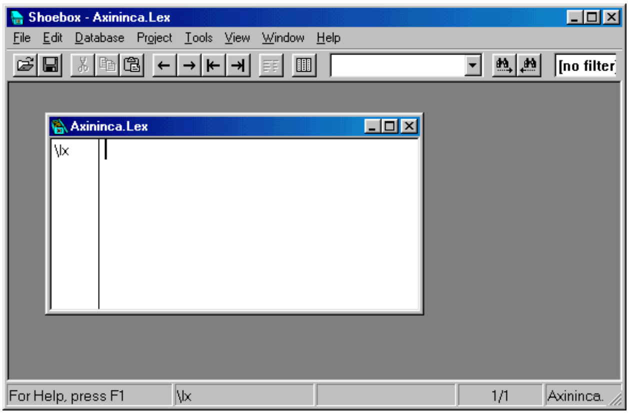


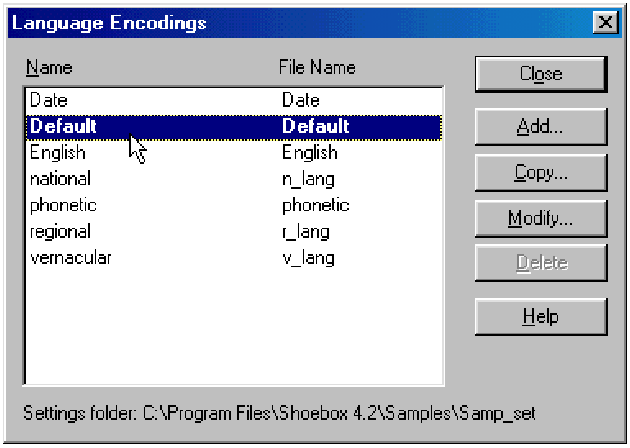
Fig. 148. Shoebox Axininca.Lex Empty record in the Axininca Lexicon

Creating a new language encoding

Axininca is written with a number of special characters. You will want to specify an appropriate font and other language characteristics for use in Axininca fields. Language encoding information is handled separately from the database type information.

1. Choose Project, Language Encodings.

Shoebox will display the Language Encodings dialog box. It contains a list of available language encodings.



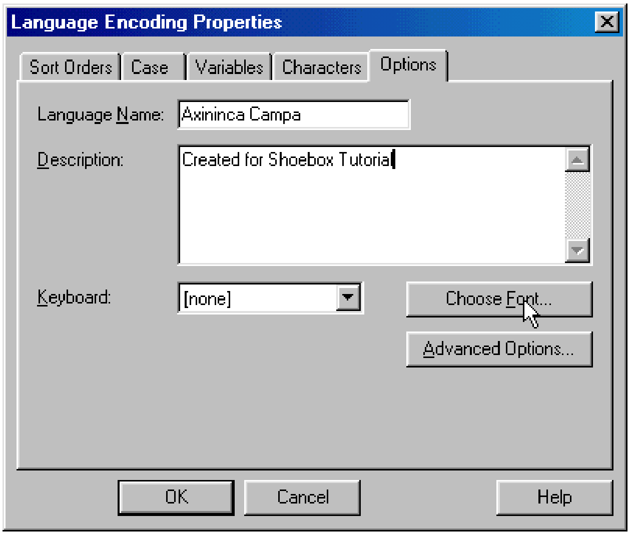
Name	File Name
Date	Date
Default	Default
English	English
national	n_lang
phonetic	phonetic
regional	r_lang
vernacular	v_lang

Settings folder: C:\Program Files\Shoebox 4.2\Samples\Samp_set

Fig. 149. Language Encodings dialog box

2. Click Add....

This will bring you to the Language Encoding Properties dialog box. You will be looking at the Options tab.



Language Name: Axininca Campa

Description: Created for Shoebox Tutorial

Keyboard: [none]

Fig. 150. Language Encoding Properties, showing the Options page

3. Type `Axininca Campa` as the Language Name.

4. Type `Created for Shoebox Tutorial` as the Description.

*Choosing a keyboard

If you are not using Keyman skip this step. Below the Description box, you will see a small drop-down box labeled Keyboard.

1. In the Keyboard list, select Axininca.

*Choosing a font

If you did not install the Axininca font, skip to the *Modifying the sort order* section, p. 137.

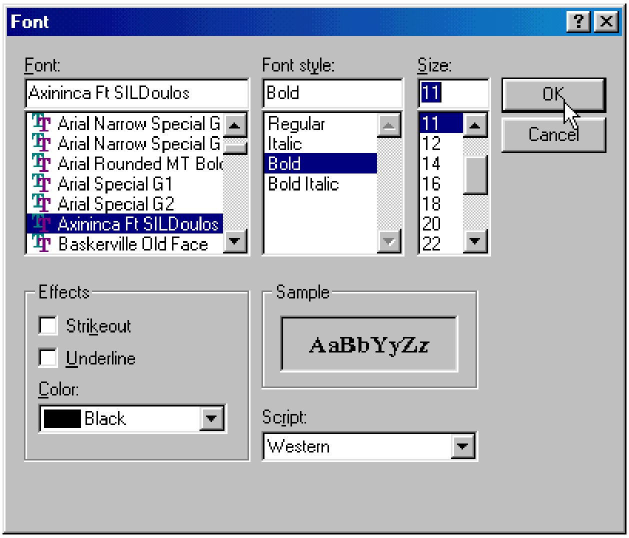
1. Click Choose Font....

You will see a large dialog box with the standard font options. This font option establishes the default font for the language. That is, if you assign a field to the Axininca language, this is the font that will be used for that field unless you specifically override it using the Font feature of the Field Marker Properties dialog box.

2. Select Axininca Ft SILDoulos.

3. Select Bold.

4. Enter 11 point as the Size (Macintosh: 12 point reads better).



Font

Font: Axininca Ft SILDoulos

Font style: Bold

Size: 11

□ Strikeout

□ Underline

Color: Black

Script: Western

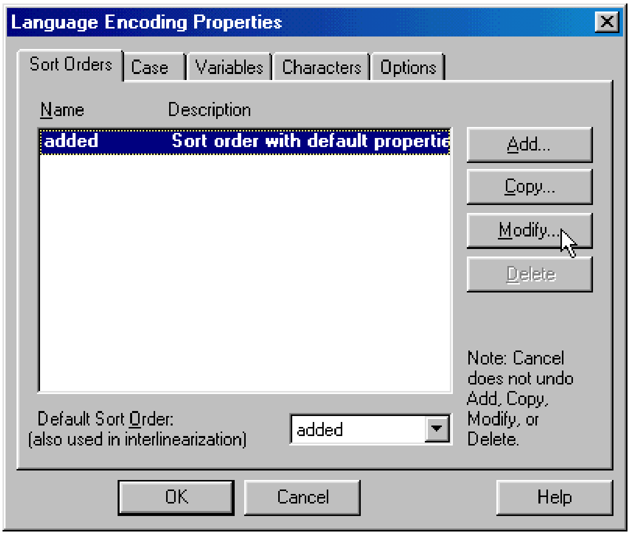
Fig. 151. Font dialog box choosing the Axininca Font

5. Click OK to return to the Language Encoding Properties dialog box.

Modifying the sort order

To see a list of the available sort orders for this particular language encoding,

1. Click the Sort Orders tab.



added Sort order with default propertie

Note: Cancel does not undo Add, Copy, Modify, or Delete.

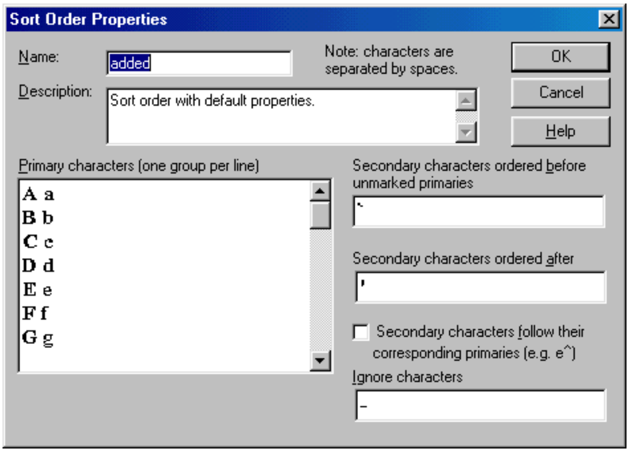
Default Sort Order: (also used in interlinearization) added

Fig. 152. Language Encoding Properties, Sort Orders tab selected

Notice that Shoebox has generated a default sort order and named it added.

2. Click Modify... .

You will see the Sort Order Properties dialog box.



Name: added

Description: Sort order with default properties.

Primary characters (one group per line)

A a
B b
C c
D d
E e
F f
G g

Secondary characters ordered before unmarked primaries: `

Secondary characters ordered after: '

□ Secondary characters follow their corresponding primaries (e.g. e^)

Ignore characters: -

Fig. 153. Sort Order Properties dialog box

3. Change the Name to `Axininca Campa`.

4. Change the Description to `Created for Shoebox Tutorial`.

Note: If you are unfamiliar with the terminology being used in this dialog box, please click Help. This gives examples of the different types of characters (primary, secondary, etc.).

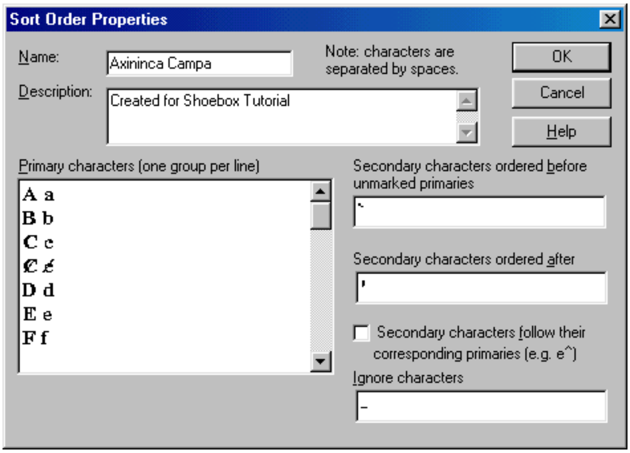
Shoebox sorting defaults to the standard order for English. This is almost right for Axininca, but there are some extra characters which should be added. If you did not install the Axininca font, you can skip these steps.

5. *Click in the Primary characters box to set the insertion point at the end of the *C c* line.

6. *Press ENTER (Macintosh: press RETURN).

7. *If you are using Keyman, type the following keystrokes: `/C /c` (forward slash followed by uppercase *C*, space,

forward slash followed by lowercase *c*). These will become an uppercase and a lowercase ¢.[27]



Sort Order Properties

Name: Axininca Campa

Description: Created for Shoebox Tutorial

Note: characters are separated by spaces.

Primary characters (one group per line)

A a
B b
C c
Ȼ ȼ
D d
E e
F f

Secondary characters ordered before unmarked primaries: ˋ

Secondary characters ordered after: '

□ Secondary characters follow their corresponding primaries (e.g. e^)

Ignore characters: -

Fig. 154. Sort Order Properties for Axininca

8. *If you are not using Keyman, press ALT+0201,[28] SPACEBAR, and ALT+0193 (Macintosh: press OPTION+SEMICOLON, SPACEBAR, and OPTION+1). An uppercase and a lowercase ¢ will be displayed.

Note: The uppercase ¢ and lowercase ¢ must be separated by a space. Characters which are not separated are considered part of a multigraph.

Axininca has a number of special characters. Since the point of this exercise is not to teach you to type Axininca, no further characters will be added.

A note about the Axininca character set. The transcription of Axininca that was used by Payne is described as "a low-level

[27] If you have not installed the Axininca font, the uppercase and the lowercase ¢ look like *É* and *Á* in most Windows fonts, and like ... and ¡ on the Macintosh.

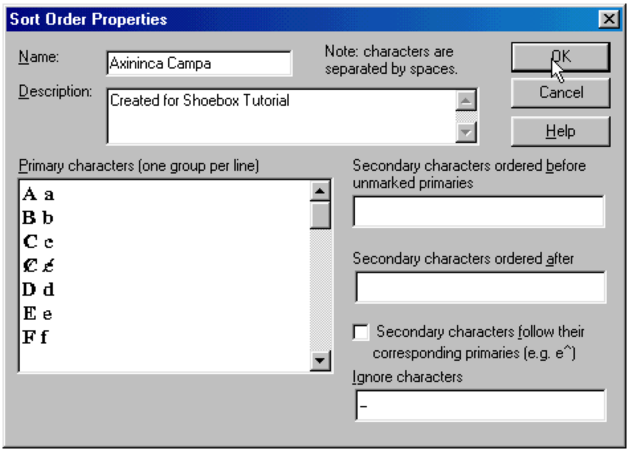
[28] Use the numeric keypad to enter the numbers while holding down the ALT key. *Note:* The keypad must be in the Num Lock position. Laptop users will need to activate the built-in keypad (usually under the right-hand letters) according to the manufacturer's instructions.

phonological representation"[29] rather than a practical orthography. It had no capitalization in the texts, except for *N*—which was not the same as *n*. However, you will leave the upper case in the sort order to make it more like a modified Roman alphabet.

Now, finish modifying the sort order.

9. From the Secondary characters ordered before unmarked primaries box, delete all characters.
10. From the Secondary characters ordered after box, delete all characters.

Your dialog box should now look like the following example.



Name: Axininca Campa

Note: characters are separated by spaces.

Description: Created for Shoebox Tutorial

Primary characters (one group per line)

A a
B b
C c
Ȼ ȼ
D d
E e
F f

Secondary characters ordered before unmarked primaries:

Secondary characters ordered after:

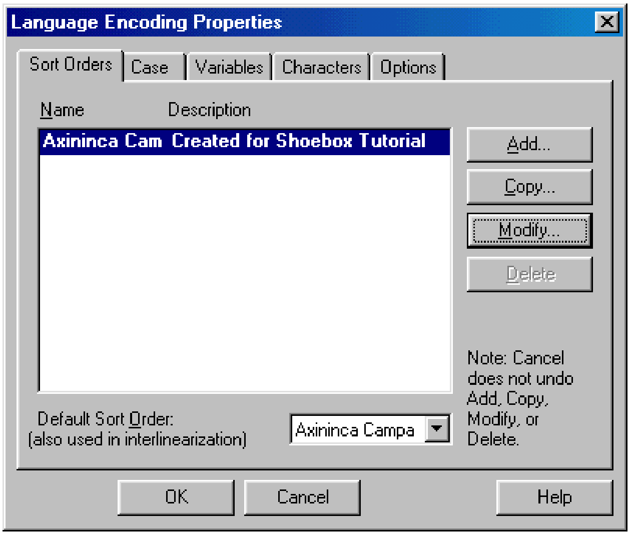
☐ Secondary characters follow their corresponding primaries (e.g. e^)

Ignore characters: -

Fig. 155. Sort Order Properties dialog box complete for Axininca

11. Click OK to return to the Language Encoding Properties dialog box.

[29] Payne 1981:6, footnote.



Name Description
Axininca Cam Created for Shoebox Tutorial

Default Sort Order: (also used in interlinearization) Axininca Campa

Note: Cancel does not undo Add, Copy, Modify, or Delete.

Fig. 156. Language Encoding Properties Sort Orders

Below the list of available sort orders is a small box labeled Default Sort Order. It is now showing Axininca Campa.

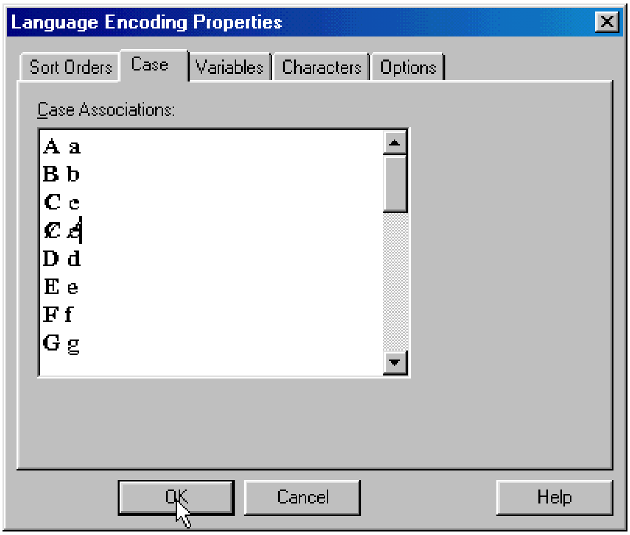
*Modifying case associations

You will next want to set up the case associations. You must include all characters which have an upper-lower case relationship in the Case description. You can skip this section if you did not install the Axininca font.

1. Click the Case tab.

2. Type an uppercase and a lowercase ¢ as you did for the sort order.

The case associations should now look like the following:



Case Associations:

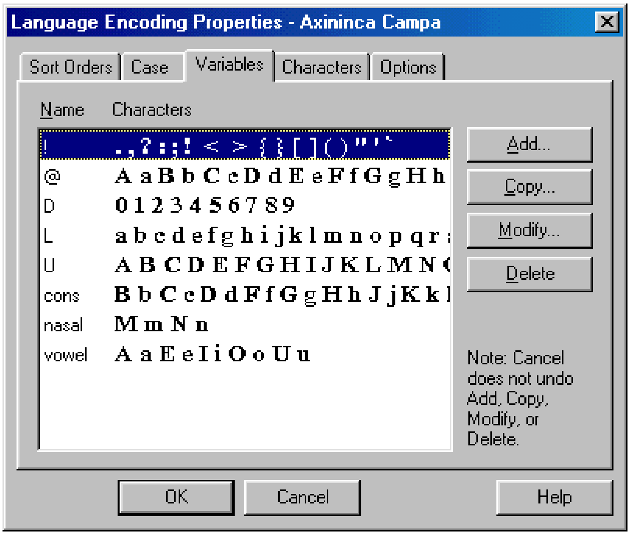
A a
B b
C c
Ȼ ȼ
D d
E e
F f
G g

Fig. 157. Language Encoding Properties dialog box showing Case setup

Modifying variables

You will now need to set up the variables for your language.

1. Click the Variables tab.



Name	Characters
!	. , ? : ; ! < > { } [ ] ( ) " ' `
@	A a B b C c D d E e F f G g H h
D	0 1 2 3 4 5 6 7 8 9
L	a b c d e f g h i j k l m n o p q r
U	A B C D E F G H I J K L M N
cons	B b C c D d F f G g H h J j K k
nasal	M m N n
vowel	A a E e I i O o U u

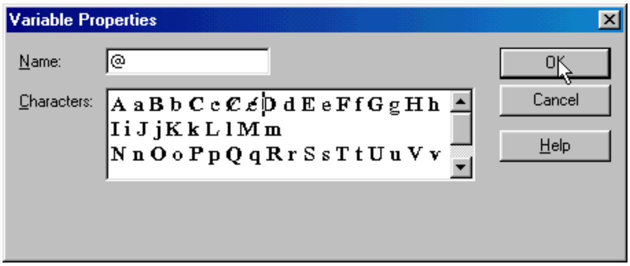
Note: Cancel does not undo Add, Copy, Modify, or Delete.

Fig. 158. Language Encoding Properties, Axininca Campa Variables

These are the user variables. As demonstrated earlier in the tutorial, variables are a user-defined collection of characters which are helpful in Filters and in the Find command. Typical variables for a language include consonants, nasals, and vowels. Other linguistically useful groupings could be specified.[30]

2. Select @. This represents "any letter."

3. Click Modify... .



Name: @

Characters: A a B b C c Ȼ ȼ D d E e F f G g H h I i J j K k L l M m N n O o P p Q q R r S s T t U u V v

Fig. 159. Variable Properties with Axininca characters added

[30] In the Variable Properties dialog box, click Help for the list of default and built in variables.

4. *If you installed the Axininca font, type an uppercase and a lowercase *¢* as you did for the sort order. Type them between the *c* and *D*.

5. Click OK.

6. Click OK in the Language Encoding Properties dialog box.

7. Click Close in the Language Encodings dialog box.

Setting up markers

All dialog boxes should, by now, have been closed. You should see only an empty data window labeled Axininca.Lex. There should be an **\lx** in the marker pane.

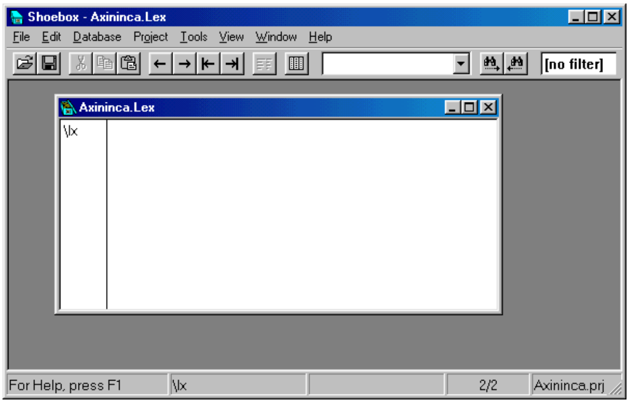


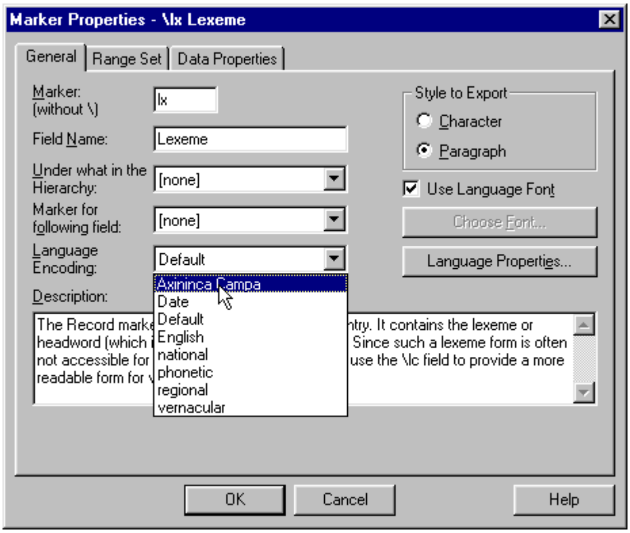
Fig. 160. New Axininca.Lex record

1. Right-click the **\lx** marker (Macintosh: OPTION+click).

This brings up the Marker Properties dialog box for the **lx** marker.

2. Type `Lexeme` as the Field Name.

3. Select Axininca Campa as the Language Encoding.



Marker Properties - \lx Lexeme

General | Range Set | Data Properties

Marker: (without \) lx

Field Name: Lexeme

Under what in the Hierarchy: [none]

Marker for following field: [none]

Language Encoding: Default

Axininca Campa
Date
Default
English
national
phonetic
regional
vernacular

Style to Export
○ Character
● Paragraph

☑ Use Language Font

Choose Font...

Language Properties...

Description:

The Record marke ... ntry. It contains the lexeme or
headword (which i ... Since such a lexeme form is often
not accessible for ... use the \lc field to provide a more
readable form for v

OK | Cancel | Help

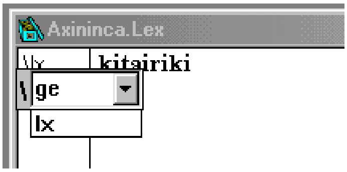
Fig. 161. Defining marker properties for the **lx** marker

4. Click OK.

5. Type `kitairiki` as the first lexical entry.

6. Press ENTER (Macintosh: press RETURN).

7. Type `\ge`.



Axininca.Lex

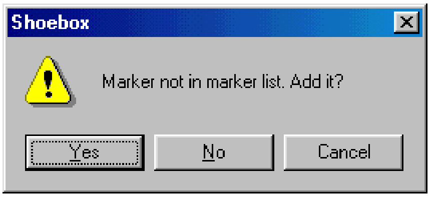
\lx kitairiki

\ge

lx

Fig. 162. Entering a new Field Marker

8. Press SPACE.



Shoebox

Marker not in marker list. Add it?

Fig. 163. Add another marker dialog box.

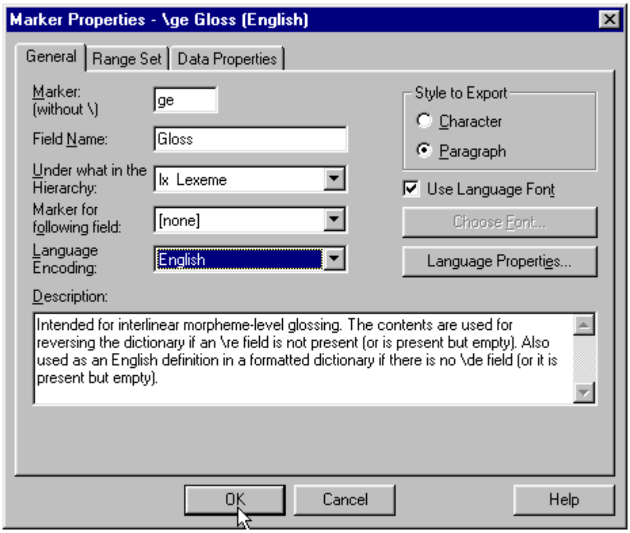
You will be asked whether or not to add the new marker to the list.

9. Click Yes.

This will again put you in the Marker Properties dialog box (with **ge** shown as the marker).

10. Type `Gloss` as the Field Name.

11. Select English as the Language Encoding.



Marker Properties - \ge Gloss (English)

Marker: (without \) ge

Field Name: Gloss

Under what in the Hierarchy: lx Lexeme

Marker for following field: [none]

Language Encoding: English

Style to Export: ○ Character ● Paragraph

☑ Use Language Font

Description:

Intended for interlinear morpheme-level glossing. The contents are used for reversing the dictionary if an \re field is not present (or is present but empty). Also used as an English definition in a formatted dictionary if there is no \de field (or it is present but empty).

Fig. 164. Marker Properties for the Gloss field

12. Click OK.

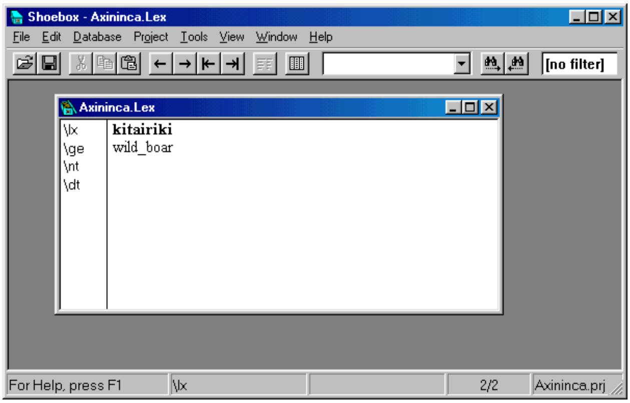
This will return you to the Axininca lexicon window.

13. Type `wild_boar` as the gloss.

Add two more data fields to this lexical database record. Go through the same process just described to add a field marker, a field name, and specify the appropriate language for each of them.

14. Marker: **\nt** Field Name: `notes` Language: English

15. Marker: **\dt** Field Name: `date` Language: Default



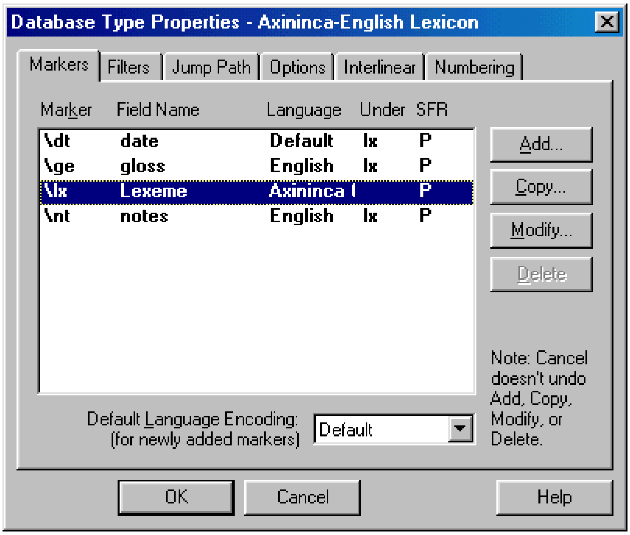
```
\lx   kitairiki
\ge   wild_boar
\nt
\dt
```

Fig. 165. Data fields in the Axininca Lexicon

Now make **\dt** a date stamp field (to remember when a record was last edited).

16. Choose Database, Properties.

Notice the new field markers in the list.



Marker	Field Name	Language	Under	SFR
\dt	date	Default	lx	P
\ge	gloss	English	lx	P
\lx	Lexeme	Axininca I		P
\nt	notes	English	lx	P

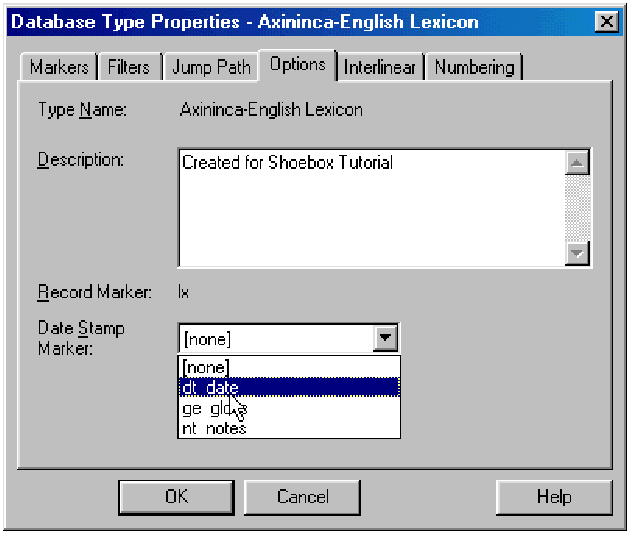
Default Language Encoding: (for newly added markers) Default

Note: Cancel doesn't undo Add, Copy, Modify, or Delete.

Fig. 166. Marker properties for the Axininca lexicon

17. Click the Options tab.

18. Select dt date from the Date Stamp Marker list.



Type Name: Axininca-English Lexicon

Description: Created for Shoebox Tutorial

Record Marker: lx

Date Stamp Marker: [none]

Fig. 167. Setting up the date stamp marker

19. Click OK.

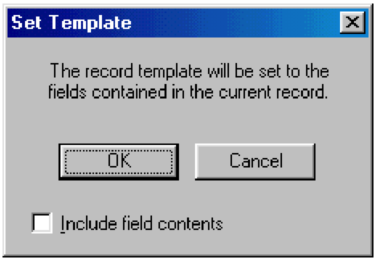
You have now set up some basic lexical data field markers for this tutorial.

Creating the new record template

Next, create the new record template—the set of field markers to be inserted automatically whenever a new record is created.

1. Choose Database, Template.

Shoebox will display the Set Template dialog box. Notice the check box which gives you the option of including the field contents. Sometimes it is a convenient shortcut to have some fields already filled in, but most of the time it is not useful. Do not select this check box now.



The record template will be set to the fields contained in the current record.

□ Include field contents

Fig. 168. Set Template dialog box

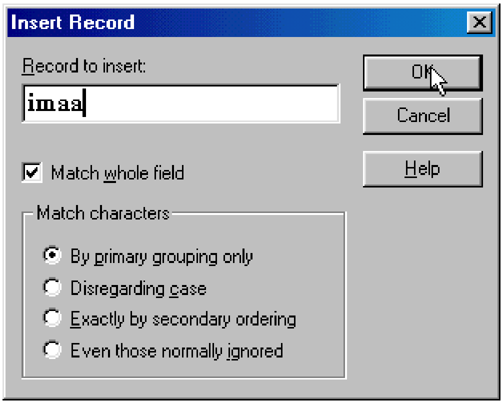
2. Click OK and the dialog box will close.

To insert a second record to test the template,

3. Choose Database, Insert Record.

This will bring up the Insert Record dialog box.

4. Type `imaa` as the Record to insert.



Record to insert: imaa

☑ Match whole field

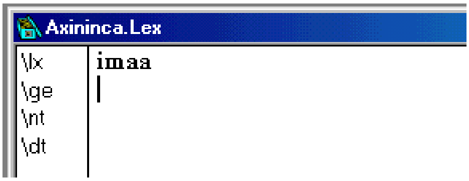
Match characters

- ● By primary grouping only
- ○ Disregarding case
- ○ Exactly by secondary ordering
- ○ Even those normally ignored

Fig. 169. Insert Record dialog box

5. Click OK.

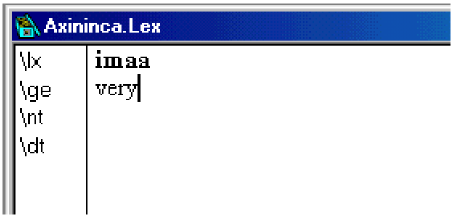
The program will return you to the lexicon window, with *imaa* already inserted as the contents of the record marker field. The insertion point will be in the **\ge** field. Notice that the template has inserted the **\ge**, **\nt**, and **\dt** field markers for you.



```
\lx   imaa
\ge
\nt
\dt
```

Fig. 170. Template with first record

6. Type `very` as the gloss for the **\ge** field.

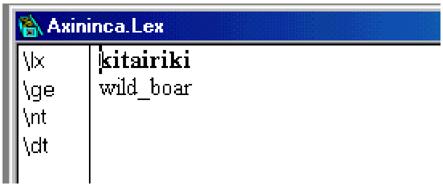


```
\lx   imaa
\ge   very
\nt
\dt
```

Fig. 171. Axininca.Lex *imaa* record

Notice that the date stamp is not filled in the **\dt** field yet. It is not filled in (or updated) until you leave a record.

7. Click → to go to the *kitairiki* record.



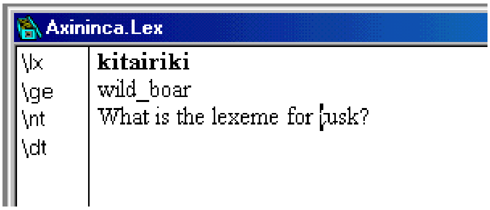
Axininca.Lex

```
\lx    kitairiki
\ge    wild_boar
\nt
\dt
```

Fig. 172. Axininca.Lex *kitairiki* record

Notice that its date stamp is still empty too. This is because you defined the date stamp marker *after* editing the record. If you edit this record now, then Shoebox will fill in the date stamp.

8. Type a note in the **\nt** field, such as: `What is the lexeme for tusk?`

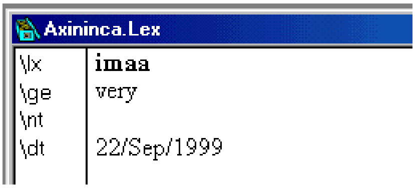


Axininca.Lex

```
\lx    kitairiki
\ge    wild_boar
\nt    What is the lexeme for tusk?
\dt
```

Fig. 173. Note field

9. Now click ← to go back to *imaa* to see the date stamp field filled in with today's date.

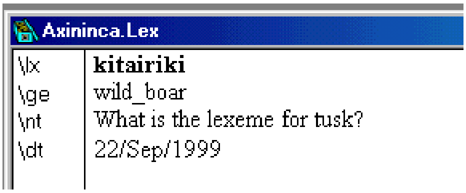


Axininca.Lex

```
\lx    imaa
\ge    very
\nt
\dt    22/Sep/1999
```

Fig. 174. Date stamp field

10. Click → again to see the *kitairiki* record. Now its date stamp is filled in too.



```
\lx   kitairiki
\ge   wild_boar
\nt   What is the lexeme for tusk?
\dt   22/Sep/1999
```

Fig. 175. Date stamp field filled in by Shoebox

With this exercise, you should have gained a better understanding of the type of information stored in the Shoebox settings files. You have started from scratch to create a new project, database type, and language encoding. In addition, you have learned how to install a font, set up Keyman, create a database template and use the date stamp feature.

11. Exit Shoebox. Save changes.

Tutorial milestone

This is a milestone in the *Tutorial*. Nothing further builds on changes you have made in this section.

Interlinear text

Interlinear text is text with additional information under each word. This usually includes a breakdown of the word into morphemes (called a parse), a gloss of each morpheme, and sometimes a grammatical category for each morpheme (usually a part of speech). Interlinearizing text is a very effective way of building a lexicon and studying the morphology of a language.

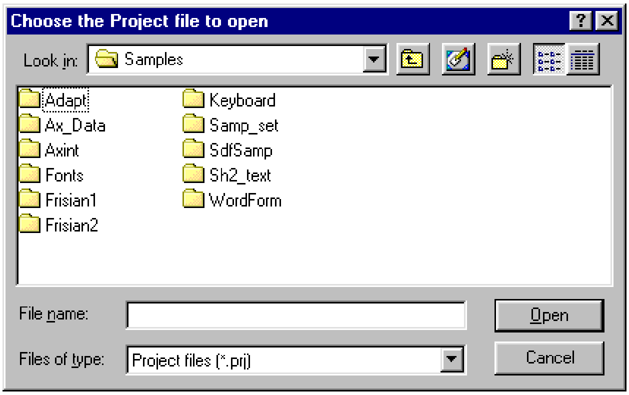
Shoebox can break words into morphemes and generate interlinear text. It does this by looking up stems, roots, and affixes in a lexicon.

Basic interlinearization

The first sample is from the Frisian language.[31]

1. Start Shoebox (if not already running).
2. Choose Project, Open.
3. Navigate into the Samples folder.

[31] Frisian is spoken in some parts of the Netherlands and Germany. The text is from the book *Frisian Reference Grammar* by Pieter Meijes Tiersma.



Look in: Samples

Adapt
Ax_Data
Axint
Fonts
Frisian1
Frisian2
Keyboard
Samp_set
SdfSamp
Sh2_text
WordForm

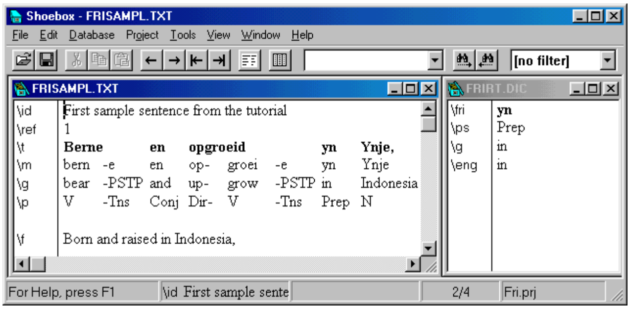
File name:

Files of type: Project files (*.prj)

Fig. 176. The Samples folder showing the Frisian1 folder

4. Navigate into the Frisian1 folder.

5. Open the project Fri.prj.



```
\id    First sample sentence from the tutorial
\ref   1
\t     Berne      en   opgroeid          yn   Ynje,
\m     bern -e    en   op-  groei -e     yn   Ynje
\g     bear -PSTP and  up-  grow  -PSTP  in   Indonesia
\p     V    -Tns  Conj Dir- V     -Tns   Prep N

\f     Born and raised in Indonesia,
```

```
\fri   yn
\ps    Prep
\g     in
\eng   in
```

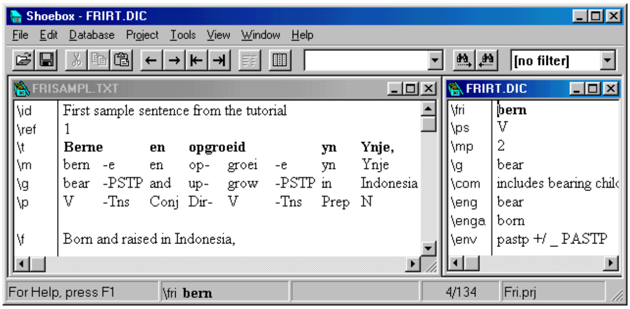
For Help, press F1 | \id First sample sente | 2/4 | Fri.prj

Fig. 177. Frisian1 project

Observe that the Frisian text in the left window has morpheme breaks, glosses, and parts of speech shown. For example, the first word *berne* is made up of the morphemes *bern* and *-e*. The morpheme *bern* has the gloss *bear*, and the part of speech *V* (for verb).

To see where the root *bern* came from,

6. Right-click (Macintosh: OPTION+click) *bern* to jump to its lexical entry.



```
\id   First sample sentence from the tutorial
\ref  1
\t    Berne       en    opgroeid         yn    Ynje,
\m    bern  -e    en    op-  groei -e    yn    Ynje
\g    bear  -PSTP and   up-  grow  -PSTP in    Indonesia
\p    V     -Tns  Conj  Dir- V     -Tns  Prep  N

\f    Born and raised in Indonesia,
```

```
\fri   bern
\ps    V
\mp    2
\g     bear
\com   includes bearing chil
\eng   bear
\enga  born
\env   pastp +/ _ PASTP
```

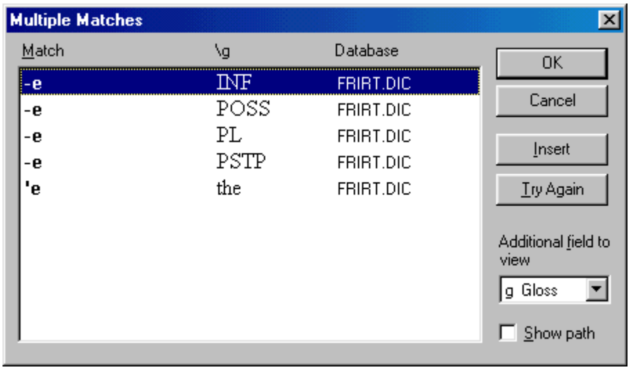
```
For Help, press F1   \fri bern      4/134   Fri.prj
```

Fig. 178. Frisian1 project showing the jump to the *bern* record

You will see the Frisian lexicon in the right window move to the entry for *bern*. In this entry you can see the part of speech *V* and the gloss *bear*. Notice that the part of speech is marked **\ps** in the lexicon, while it is marked with **\p** in the text file. The marker used in interlinear text does not have to be the same as the marker in the lexicon.

To see where the suffix *-e* came from,

7. Right-click (Macintosh: OPTION+click) *-e*.

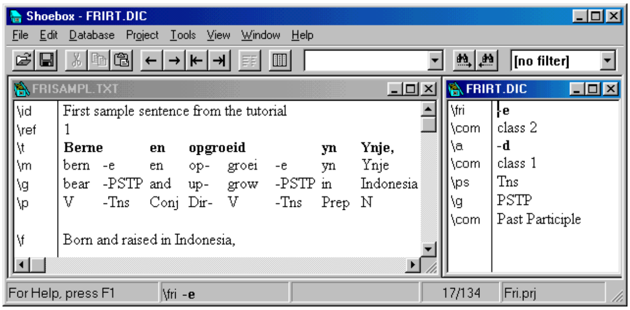


Match	\g	Database
-e	INF	FRIRT.DIC
-e	POSS	FRIRT.DIC
-e	PL	FRIRT.DIC
-e	PSTP	FRIRT.DIC
'e	the	FRIRT.DIC

Fig. 179. Multiple Matches dialog box

You will see a dialog box titled Multiple Matches containing a list of -*e* suffixes.[32]

8. Select the suffix that has a gloss of *PSTP* and click OK.



```
\id    First sample sentence from the tutorial
\ref   1
\t     Berne      en    opgroeid          yn    Ynje,
\m     bern  -e   en    op-  groei  -e    yn    Ynje
\g     bear  -PSTP and  up-  grow   -PSTP in    Indonesia
\p     V     -Tns Conj  Dir- V      -Tns  Prep  N

\f     Born and raised in Indonesia,
```

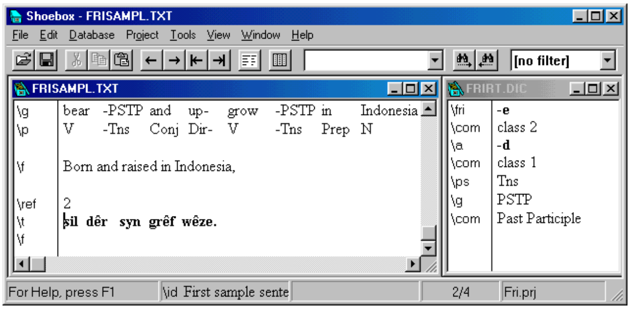
```
\fri   -e
\com   class 2
\a     -d
\com   class 1
\ps    Tns
\g     PSTP
\com   Past Participle
```

Fig. 180. Frisian1 project showing the -*e* suffix

You see the relevant -*e* lexicon entry. The hyphen shows that it is a suffix.

Next you will interlinearize a new sentence.

9. Scroll the text window down so that you can see the second clause of the sentence, which starts with *sil*.



```
\g     bear  -PSTP and  up-  grow   -PSTP in    Indonesia
\p     V     -Tns Conj  Dir- V      -Tns  Prep  N

\f     Born and raised in Indonesia,

\ref   2
\t     sil dêr  syn gref wêze.
\f
```

```
\fri   -e
\com   class 2
\a     -d
\com   class 1
\ps    Tns
\g     PSTP
\com   Past Participle
```

Fig. 181. Frisian project prepared for interlinearization

10. Place the insertion point in front of *sil*.

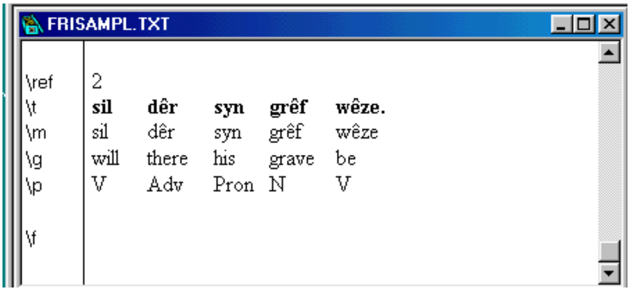
[32] It might also contain *'e* if the hyphen is being ignored. This could be fixed by changing a Jump setting, but it is not relevant at this point.

In Shoebox, there are three ways to choose the Interlinearize command:

- Menu: Choose Tools, Interlinearize
- Shortcut key: Press ALT+I (Macintosh: press COMMAND+I)
- Toolbar: Click ☰

In this tutorial, you will use the toolbar.

11. Click ☰.

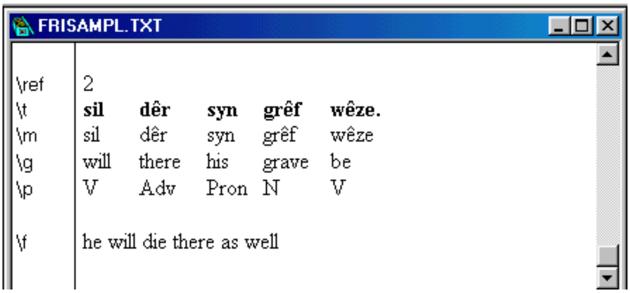


```
\ref   2
\t     sil    dêr     syn    grêf    wêze.
\m     sil    dêr     syn    grêf    wêze
\g     will   there   his    grave   be
\p     V      Adv     Pron   N       V

\f
```

Fig. 182. Frisian Sample text interlinearized

You will see Shoebox interlinearize the line automatically.

12. Move the insertion point into the **\f** (free translation) field and type: `he will die there as well.`



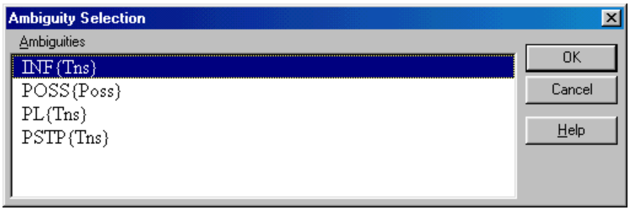
```
\ref   2
\t     sil    dêr     syn    grêf    wêze.
\m     sil    dêr     syn    grêf    wêze
\g     will   there   his    grave   be
\p     V      Adv     Pron   N       V

\f     he will die there as well
```

Fig. 183. Frisian Sample text with free translation

This particular text line was easy to interlinearize because all the words consist of one morpheme and each morpheme has exactly one gloss.

13. Move the insertion point up to the word *Berne* in the first clause of the sentence.

14. Click ▤.



Ambiguity Selection

Ambiguities

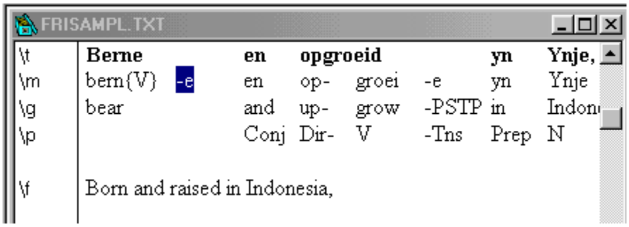
INF{Tns}
POSS{Poss}
PL{Tns}
PSTP{Tns}

Fig. 184. Ambiguity Selection for interlinearization

Shoebox will display the Ambiguity Selection dialog box. It is covering the word you interlinearizing, so you will want to move it. Shoebox will remember where you move it and will display it there the next time.

15. Drag the Ambiguity Selection dialog box to the right side of the screen.

You will see that *-e* is highlighted in the text. This shows that the four choices in the box are four possible glosses for *-e*. Notice that each gloss is followed by its part of speech in curly braces. This can sometimes help clarify the choices.



FRISAMPL.TXT

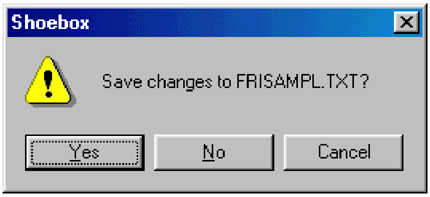
```
\t   Berne          en    opgroeid           yn    Ynje,
\m   bern{V}  -e    en    op-   groei  -e    yn    Ynje
\g   bear           and   up-   grow   -PSTP in    Indon
\p                  Conj  Dir-  V      -Tns  Prep  N

\f   Born and raised in Indonesia,
```

Fig. 185. Frisian Sample text showing highlighted *-e*

16. Select PSTP{Tns} and click OK (or double-click PSTP{Tns}).

Shoebox will finish interlinearizing *Berne*. Notice that the insertion point stops in front of the next word. When text has already been interlinearized, subsequent interlinearizing works only one word at a time.

17. Choose Project, Close.



Save changes to FRISAMPL.TXT?

Fig. 186. Save Changes dialog box

You will see a box asking if you want to save the changes to the text file.

18. Click Yes (Macintosh: click Save).

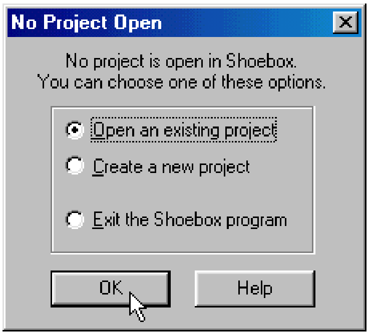
Tutorial milestone

This is a milestone in the *Tutorial*. Nothing further builds on changes you have made in this section.

Interlinear setup

To illustrate interlinear setup, you will perform the setup steps that were done to set up the Frisian example you just did. You will start with a project that has the lexicon and the uninterlinearized text but no interlinear setup.

If you have a project open in Shoebox, close it. Shoebox will display the No Project Open dialog box.



No Project Open

No project is open in Shoebox.
You can choose one of these options.

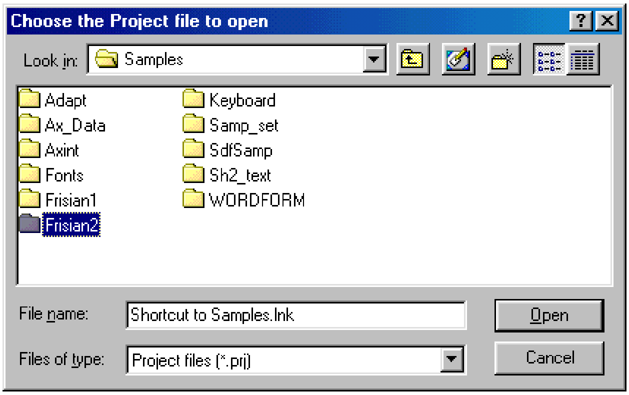
● Open an existing project

○ Create a new project

○ Exit the Shoebox program

Fig. 187. No Project Open dialog box waiting for instructions

1. Choose Open an existing project.
2. Click OK.



Choose the Project file to open

Look in: Samples

Adapt	Keyboard
Ax_Data	Samp_set
Axint	SdfSamp
Fonts	Sh2_text
Frisian1	WORDFORM
Frisian2	

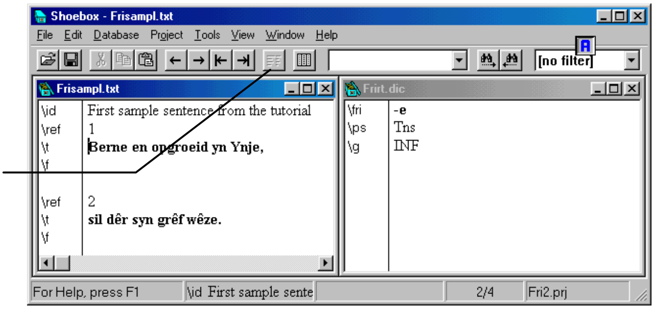
File name: Shortcut to Samples.lnk

Files of type: Project files (*.prj)

Fig. 188. Samples folder ready to choose a project

You are in the Choose the Project file to open dialog box.

3. Navigate to the Frisian2 folder in the Samples folder.
4. Open the project Fri2.prj.



```
\id    First sample sentence from the tutorial
\ref   1
\t     Berne en opgroeid yn Ynje,
\f

\ref   2
\t     sil dêr syn grêf wêze.
\f
```

```
\fri   -e
\ps    Tns
\g     INF
```

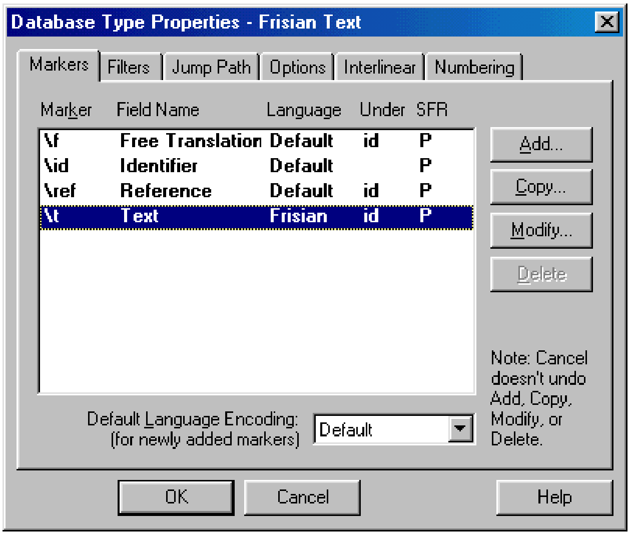
Interlinearize button

Fig. 189. Frisian2.prj open

You will see the same text as before, but without any interlinearization. Notice that the Interlinearize button is gray. This shows that interlinearization is not available because it has not been set up yet.

Important: Before you start setting up the interlinearization process, be sure the *Frisampl.txt* window is active, not the *Frirt.dic* (lexicon) window. Interlinearization is normally set up as part of the database properties of *text* files (not *lexical* files). If the title bar of the Frisampl.txt window appears dimmed, click anywhere in the Frisampl.txt window to make it your active window.

5. Choose Database, Properties.



Database Type Properties - Frisian Text

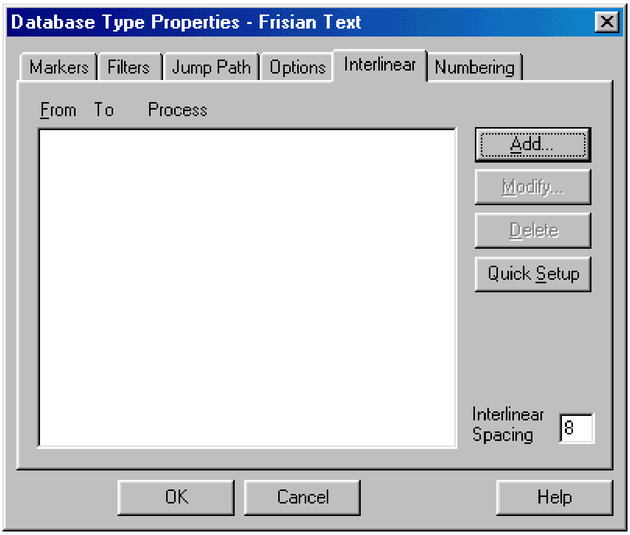
Marker	Field Name	Language	Under	SFR
\f	Free Translation	Default	id	P
\id	Identifier	Default		P
\ref	Reference	Default	id	P
\t	Text	Frisian	id	P

Default Language Encoding: (for newly added markers) Default

Note: Cancel doesn't undo Add, Copy, Modify, or Delete.

Fig. 190. Database Type Properties - Frisian Text dialog box

6. Click Interlinear to set up the interlinear processes.



Database Type Properties - Frisian Text

Interlinear Spacing 8

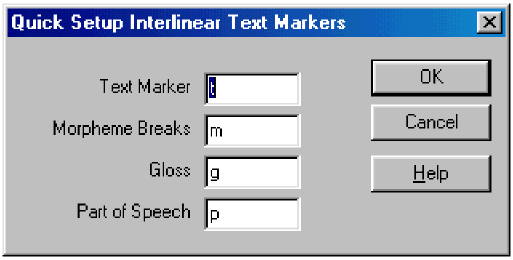
Fig. 191. Dialog box ready to interlinearization set up

Notice the From To Process list is empty. To set this one up like the Frisian interlinear text you saw before, you need to add three interlinear processes. The processes define the fields to be used in the actual interlinear text and the fields in the lexicon which contain the information it needs to fill in the interlinear fields.[33]

From (in interlinear text)	Source of data (in lexicon)	To (in interlinear text)
t (text)	**fri** (lexeme)	**m** (morpheme)
m (morpheme)	**g** (gloss)	**g** (gloss)
m (morpheme)	**ps** (part of speech)	**p** (part of speech)

You could add the processes one by one, but the best way to start an interlinear setup is with Quick Setup. Quick Setup asks for a minimum of information and adds three processes at once.

7. Click Quick Setup.



Quick Setup Interlinear Text Markers

Text Marker: t
Morpheme Breaks: m
Gloss: g
Part of Speech: p

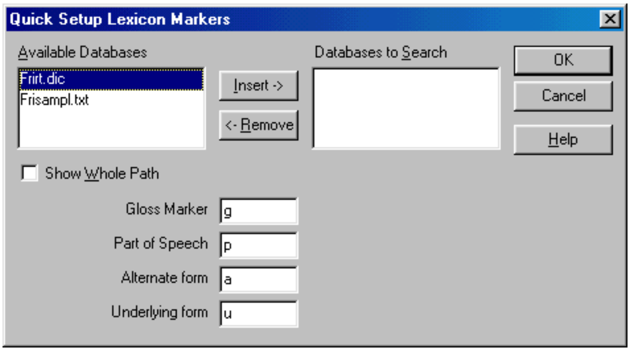
Fig. 192. Quick Setup dialog box for Frisian text markers

You will see a dialog box asking for Quick Setup Interlinear Text Markers. This dialog box is asking for the field markers to be used in the actual interlinear text. As described above, the original text is marked **t**. The morphemes are to be marked **m**, the glosses are

[33] A setup based on the Multi-Dictionary Formatter (MDF) would use **lx**, **ge**, and **ps**.

to be marked **g**, and the parts of speech are to be marked **p**. These are the default values, so you do not need to change them.

8. Click OK.


Quick Setup Lexicon Markers

Available Databases
Frirt.dic
Frisampl.txt

Databases to Search

☐ Show Whole Path

Gloss Marker: g
Part of Speech: p
Alternate form: a
Underlying form: u

Fig. 193. Quick Setup Lexicon Markers for interlinearization

You will see a dialog box asking for Quick Setup Lexicon Markers. This dialog box is asking for the field markers to be found in the lexicon. It also asks for the name of the lexicon or lexicons to use.

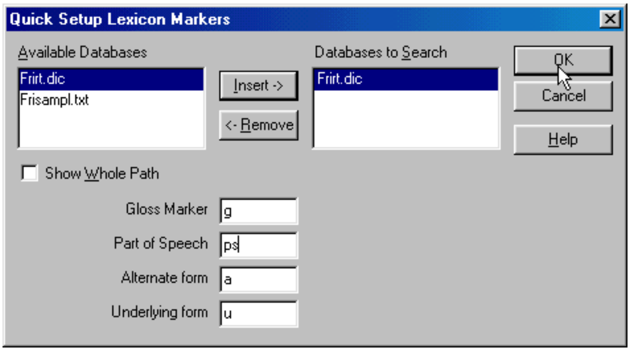
Use the lexicon Frirt.dic.

9. Select Frirt.dic, and click Insert -> to copy it to the list of Databases to Search.

The default gloss marker is **g**, which is what you need for this lexicon, but you need **ps** for the part of speech marker, not the default **p**.

10. Change the Part of Speech marker to **ps**.

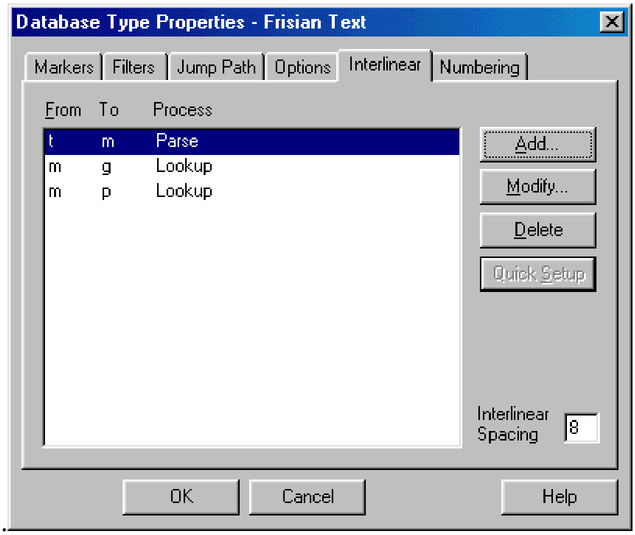
The field markers for Alternate form and Underlying form will be explained later. For now, they can be left as is. The Quick Setup Lexicon Markers dialog box should look like the following:



Available Databases: Frirt.dic, Frisampl.txt
Databases to Search: Frirt.dic
Gloss Marker: g
Part of Speech: ps
Alternate form: a
Underlying form: u

Fig. 194. Lexicon Markers set up for interlinearization

11. Click OK.



From	To	Process
t	m	Parse
m	g	Lookup
m	p	Lookup

Interlinear Spacing: 8

Fig. 195. Frisian Database Properties set up for interlinearization

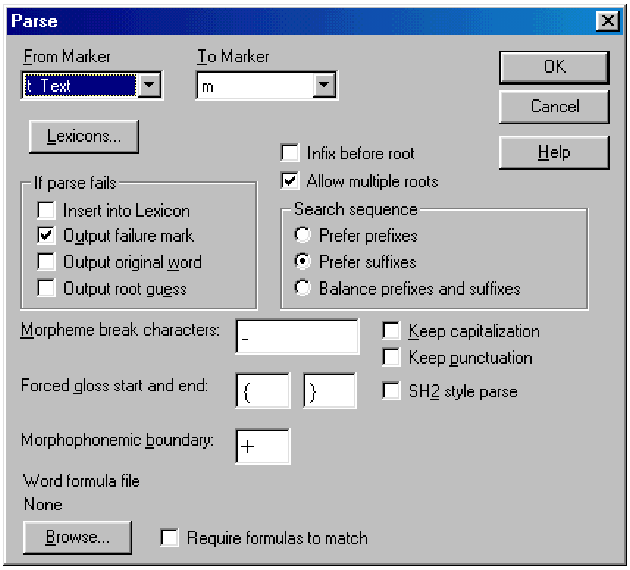
You have completed the Quick Setup. You will see that there are now three interlinear processes listed. The first is a Parse process, and the other two are Lookup processes. Each process has a From and a To column. These give the marker from which a process starts and the marker to which it annotates in the interlinear text. For example, the Parse process goes from **t** to **m**. This means that it parses the words on the **t** line and produces an **m** line containing morphemes. The other two processes are both Lookup processes starting from the **m** line. One process looks up the morphemes and produces a **g** line containing glosses. The other process looks up the morphemes and produces a **p** line containing parts of speech.

The list also shows the interlinear layout that will be produced in the text file. The From marker of the first process is the original text line. The To marker of each process is the interlinear line that will be added. You can see that in this example three interlinear lines will be added. They will be marked **m**, **g**, and **p**.

Notice that the Quick Setup button is no longer available. To redo a Quick Setup, you must first delete all the interlinear processes in the list.

While you are in this dialog box, take a few moments to explore these processes. You do not need to change anything, but you should be aware of where you can change options.

12. Select the Parse process and click Modify... .



From Marker: t Text

To Marker: m

☐ Infix before root

☑ Allow multiple roots

If parse fails
- ☐ Insert into Lexicon
- ☑ Output failure mark
- ☐ Output original word
- ☐ Output root guess

Search sequence
- ○ Prefer prefixes
- ● Prefer suffixes
- ○ Balance prefixes and suffixes

Morpheme break characters: -

☐ Keep capitalization

☐ Keep punctuation

Forced gloss start and end: { }

☐ SH2 style parse

Morphophonemic boundary: +

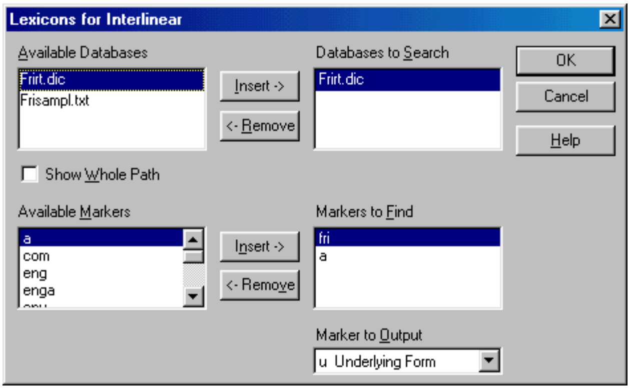
Word formula file
None

☐ Require formulas to match

Fig. 196. Parse dialog box set up for Frisian2

You will see a dialog box with the From marker and To marker at the top. It also contains options that show what to do if the parse fails, and places to change the morpheme break characters, forced gloss characters, and the morpheme boundary character.

13. Click Lexicons... to see more details of the process.



Available Databases: Frit.dic, Frisampl.txt

Databases to Search: Frit.dic

☐ Show Whole Path

Available Markers: a, com, eng, enga

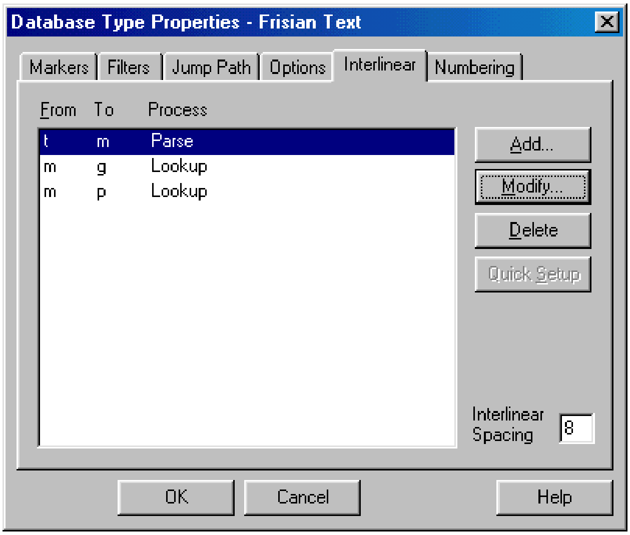
Markers to Find: fri, a

Marker to Output: u Underlying Form

Fig. 197. Lexicons for Interlinearization dialog box

You will see a dialog box showing the Databases to Search, Markers to Find, and the Marker to Output. It is set to search the Frirt.dic, find the **fri** and **a** field markers in that lexicon, and output the **u** marker. (The functions of the **\a** alternate form and the **\u** underlying form fields will be discussed later.)

14. Click OK two times to return to the list of processes.

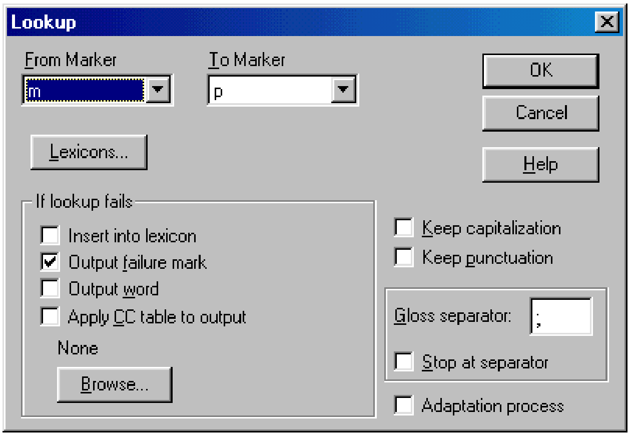


From	To	Process
t	m	Parse
m	g	Lookup
m	p	Lookup

Interlinear Spacing 8

Fig. 198. List of processes dialog box

15. Select the second Lookup process (i.e., from **m** to **p)** and click Modify... .



From Marker: m

To Marker: p

If lookup fails

☐ Insert into lexicon
☑ Output failure mark
☐ Output word
☐ Apply CC table to output

None

☐ Keep capitalization
☐ Keep punctuation

Gloss separator: ;

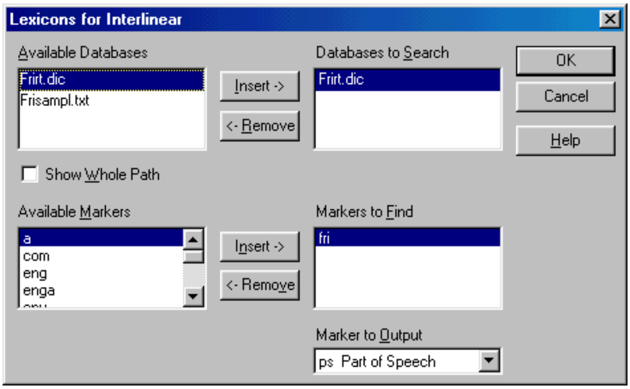
☐ Stop at separator

☐ Adaptation process

Fig. 199. Lookup dialog box

You will see a dialog box somewhat like the one for the Parse process. It shows a From Marker, a To Marker, and options for what to do if Lookup fails. It also allows changing the gloss separator character and has an option to stop at the first separator. Interlinear processes are very flexible so that they can be used in a variety of ways, including various forms of language adaptation.

16. Click Lexicons....



Available Databases: Frirt.dic, Frisampl.txt

Databases to Search: Frirt.dic

☐ Show Whole Path

Available Markers: a, com, eng, enga

Markers to Find: fri

Marker to Output: ps Part of Speech

Fig. 200. Lexicons for Interlinear dialog box

You will see the dialog box showing the Databases to Search, Markers to Find, and the Marker to Output. It is set to search the

Frirt.dic, find the **fri** marker in that lexicon, and output the **ps** marker.

Notice that this is where the difference between the text marker **\p** and the lexicon marker **\ps** appears. In the main dialog box, the To Marker is the **p** marker, indicating that the interlinear text will have a line marked **\p**. In the Lexicons box, the Marker to Output is the **ps** marker, meaning that the contents of the **\ps** field in the lexicon will be output into the interlinear text line (which is marked **p**).

17. Click OK two times to return to the list of processes.

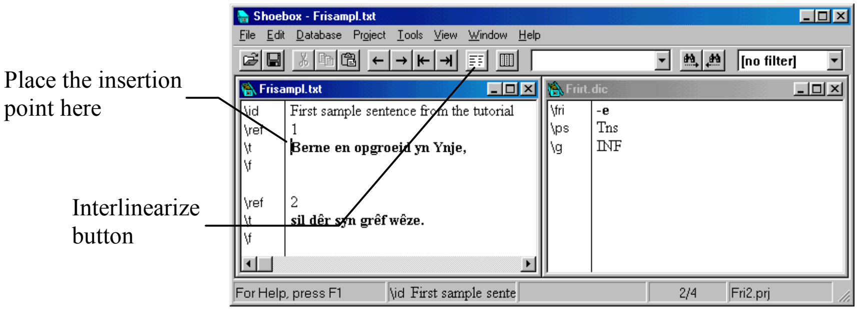
18. Click OK again to return to the text window.

That completes the interlinear setup. Notice that the interlinearize button, ☷, is now active.

Testing the interlinear setup

To demonstrate that the interlinear setup works, interlinearize the sample data.

1. Place the insertion point at the beginning of the first **\t** line, just before the word *Berne*.



Place the insertion point here

Interlinearize button

Shoebox - Frisampl.txt

Frisampl.txt

```
\id    First sample sentence from the tutorial
\ref   1
\t     Berne en opgroeid yn Ynje,
\f

\ref   2
\t     sil dêr syn grêf wêze.
\f
```

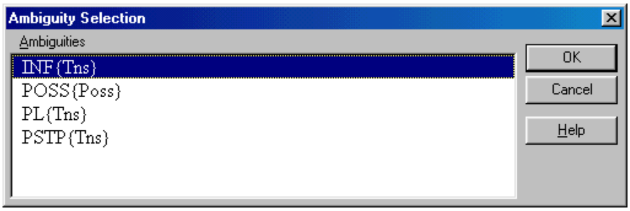
Frirt.dic

```
\fri   -e
\ps    Tns
\g     INF
```

For Help, press F1 | \id First sample sente | 2/4 | Fri2.prj

Fig. 201. Frisian2 project text window

2. Click ☷.



INF{Tns}
POSS{Poss}
PL{Tns}
PSTP{Tns}

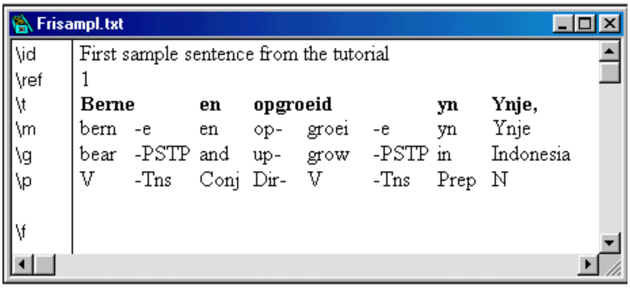
Fig. 202. Ambiguity Selection box for the *-e* morpheme

Shoebox will display the Ambiguity Selection dialog box.

3. Select PSTP{Tns} and click OK.

Shoebox will display the Ambiguity Selection dialog box again. You might have to drag the box lower to see which morpheme is highlighted in the text. Notice that it is the same morpheme, but Shoebox is parsing a different word. This time the *-e* occurs in the word *opgroeid*. The final *-d* of *opgroeid* is an allomorph of *-e*. In Shoebox parsing, it is called an alternate form.

4. Select PSTP{Tns} and click OK.

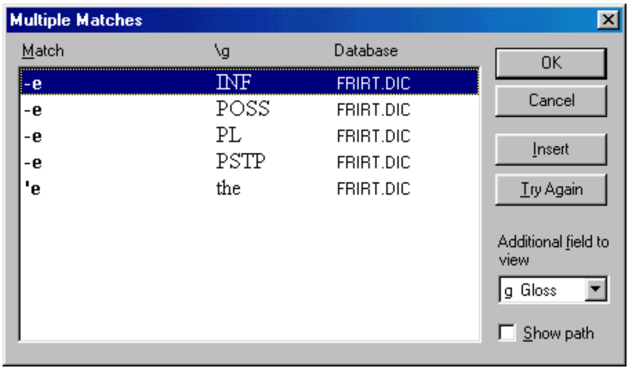


```
\id   First sample sentence from the tutorial
\ref  1
\t    Berne       en    opgroeid            yn    Ynje,
\m    bern  -e    en    op-   groei   -e    yn    Ynje
\g    bear  -PSTP and   up-   grow    -PSTP in    Indonesia
\p    V     -Tns  Conj  Dir-  V       -Tns  Prep  N

\f
```

Fig. 203. Frisian2 text interlinearized

To see the *-d* alternate form of *-e*, jump to that record entry in the lexicon.

5. Right-click (Macintosh: OPTION+click) *-e* in the text on the morpheme (**\m**) line.

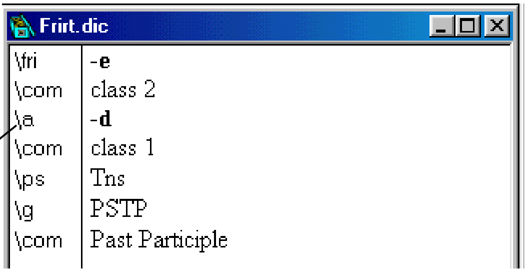


Match	\g	Database
-e	INF	FRIRT.DIC
-e	POSS	FRIRT.DIC
-e	PL	FRIRT.DIC
-e	PSTP	FRIRT.DIC
'e	the	FRIRT.DIC

Fig. 204. Multiple Matches dialog box for *-e*

6. From the Multiple Matches dialog box, select the *-e* with the gloss *PSTP* (the fourth item in the list).

7. Click OK.



```
\fri   -e
\com   class 2
\a     -d
\com   class 1
\ps    Tns
\g     PSTP
\com   Past Participle
```

Notice that *-d* is an alternate form for *-e*.

Fig. 205. Frisian2 dictionary entry

8. Notice that the *-e* record, which appears in the Frirt.dic window, has the field: **\a** *-d*. It is this field which tells Shoebox that *-d* is an alternate form of *-e*.

9. Choose Project, Close.

You are done with this project. You will see a box asking if you want to save the changes to the text file.

10. Click Yes (Macintosh: click Save).

Tutorial milestone This is a milestone in the *Tutorial*. Nothing further builds on changes you have made in this section.

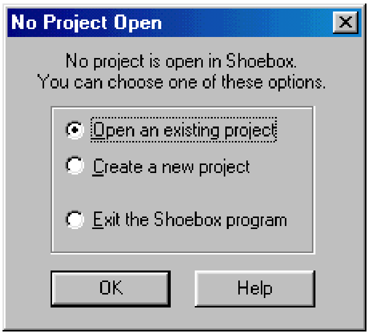
More on interlinearization

To further illustrate interlinearization, you will process a text from the Axininca Campa language of Peru. Axininca is an agglutinative language with significant morphophonemics, so it illustrates parsing well.[34]

The *Shoebox settings files* chapter of the *Tutorial* sets up the Axininca language and then starts an Axininca lexicon. It also explains how to install the Axininca fonts. If you did that section of the tutorial and installed the Axininca fonts, the data in this sample will be displayed correctly. If you did not install the Axininca fonts, some of the Axininca words will look different in Shoebox than described in this section (a minor issue which has no effect on the lesson).

This interlinear sample uses an Axininca lexicon and settings files from a different project folder than the one used earlier in the tutorial.

1. If you see a No Project Open dialog box as shown below, choose Open an existing project and click OK.



No Project Open

No project is open in Shoebox.
You can choose one of these options.

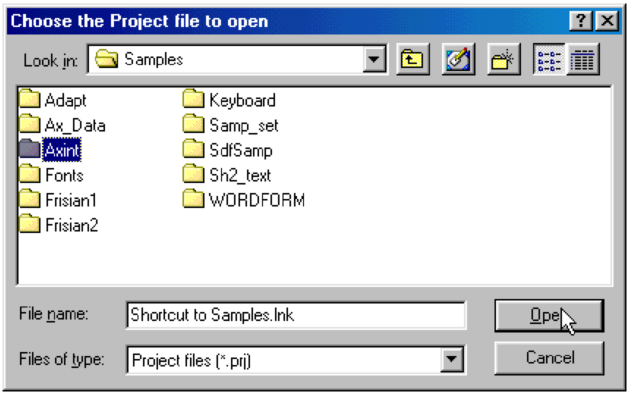
● Open an existing project
○ Create a new project
○ Exit the Shoebox program

Fig. 206. No Project Open dialog box

2. If Shoebox has a project open already, choose Project, Open from the menu.

[34] The text is from the appendix of *The Phonology and Morphology of Axininca Campa* by David Payne. We have tried to generate the same analyses as shown in the book, except that we have abbreviated some of the glosses.

3. Navigate into the Axint folder. (In the folder in which Shoebox was installed you will find a folder called Samples. Axint is in the Samples folder.)



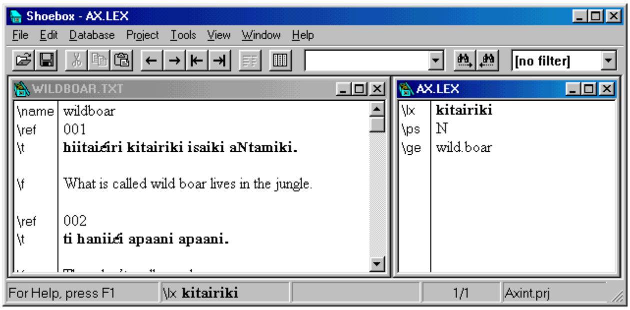
Look in: Samples

File name: Shortcut to Samples.lnk

Files of type: Project files (*.prj)

Fig. 207. Opening the Axint project

4. Open the project Axint.prj.



```
\name  wildboar
\ref   001
\t     hiitaiśiri kitairiki isaiki aNtamiki.

\f     What is called wild boar lives in the jungle.

\ref   002
\t     ti haniiśi apaani apaani.
```

```
\lx  kitairiki
\ps  N
\ge  wild.boar
```

For Help, press F1 | \lx kitairiki | 1/1 | Axint.prj

Fig. 208. Axint project open

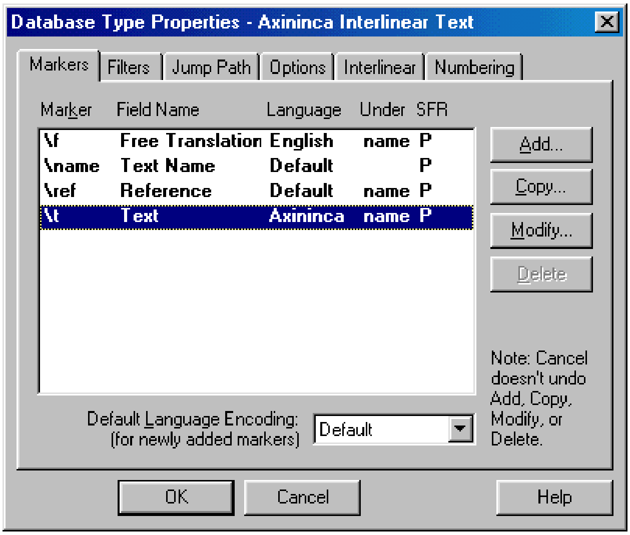
Shoebox will open a text file named Wildboar.txt and a lexicon named Ax.Lex. The text has a free translation after each sentence but is not interlinearized.

Interlinear setup

This example does not yet have an interlinear setup, so you need to set one up now. The interlinear setup is part of the database type.

1. Click somewhere in the Wildboar.txt window to make sure it is the active window.

2. Choose Database, Properties.



Marker	Field Name	Language	Under	SFR
\f	Free Translation	English	name	P
\name	Text Name	Default		P
\ref	Reference	Default	name	P
\t	Text	Axininca	name	P

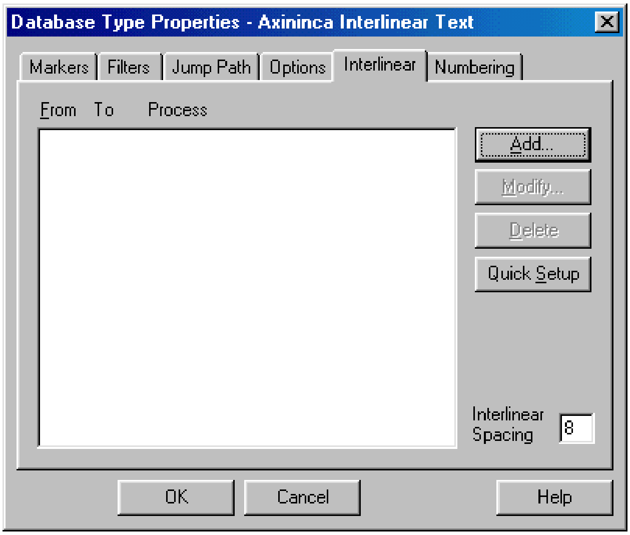
Default Language Encoding: (for newly added markers) Default

Note: Cancel doesn't undo Add, Copy, Modify, or Delete.

Fig. 209. Database Properties dialog box

You will see a dialog box titled Database Type Properties—Axininca Interlinear Text.

3. Click the Interlinear tab.

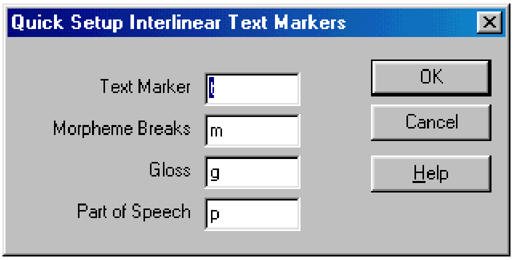


Interlinear Spacing 8

Fig. 210. Interlinear setup dialog box for Axininca

You will see an empty list titled From To Process. On the right is a button Quick Setup. Again, as shown in the last lesson, this is the best way to start an interlinear setup.

4. Click Quick Setup.

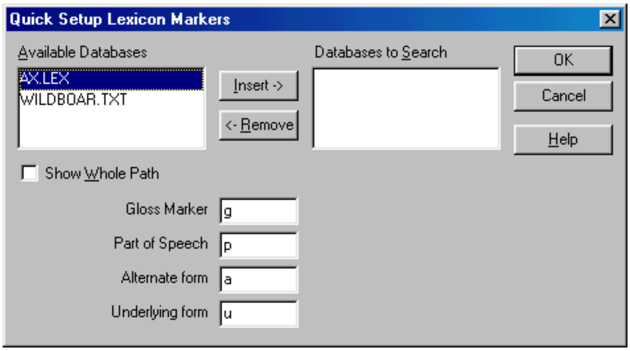


Text Marker t
Morpheme Breaks m
Gloss g
Part of Speech p

Fig. 211. Quick Setup Interlinear Text Markers for Axininca

You will see the dialog box titled Quick Setup Interlinear Text Markers. These are the field markers that will be used in the interlinear text. The dialog box shows four field markers: Text Marker, Morpheme Breaks, Gloss, and Part of Speech. The field markers default to **t**, **m**, **g**, and **p**. You will be using these values, so you do not need to change them.

5. Click OK.



Quick Setup Lexicon Markers

Available Databases

AX.LEX

WILDBOAR.TXT

Databases to Search

□ Show Whole Path

Gloss Marker g

Part of Speech p

Alternate form a

Underlying form u

Fig. 212. Quick Setup Lexicon Markers for Axininca

You will see the dialog box titled Quick Setup Lexicon Markers. This dialog box is asking for the field markers that will be searched for in the lexicon when interlinearizing.

The top of the dialog box has a list of Available Databases, showing the open databases and an empty list of Databases to Search. This is where you tell it the name(s) of the lexical database(s) Shoebox must search. In this example, there is only one.[35]

6. Select AX.LEX in the list of Available Databases.

7. Click Insert -> to copy it to the Databases to Search box.

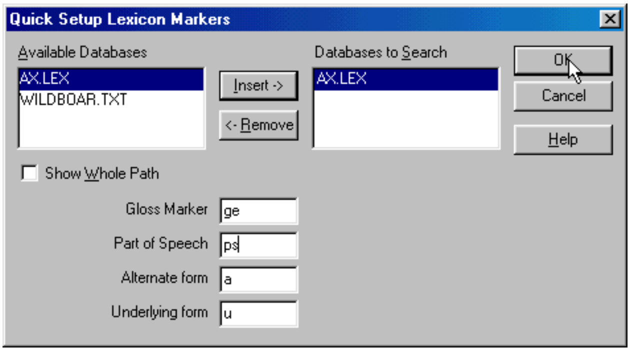
The lower part of this dialog box has four field markers: Gloss Marker, Part of Speech, Alternate form, and Underlying form. The field markers default to **g**, **p**, **a**, and **u**. These are not the field markers used for gloss and part of speech in the Axininca lexicon, so you will need to change them.

8. Change the Gloss Marker to **ge**.

9. Change the Part of Speech marker to **ps**.

[35] You would have more than one lexical database if, for example, you are keeping stems or proper names separate from your main lexical database.

The other field markers are correct and can be left alone. Their meanings will be explained later. The Quick Setup Lexicon Markers dialog box should look like the following:



Available Databases: AX.LEX, WILDBOAR.TXT

Databases to Search: AX.LEX

☐ Show Whole Path

Gloss Marker: ge

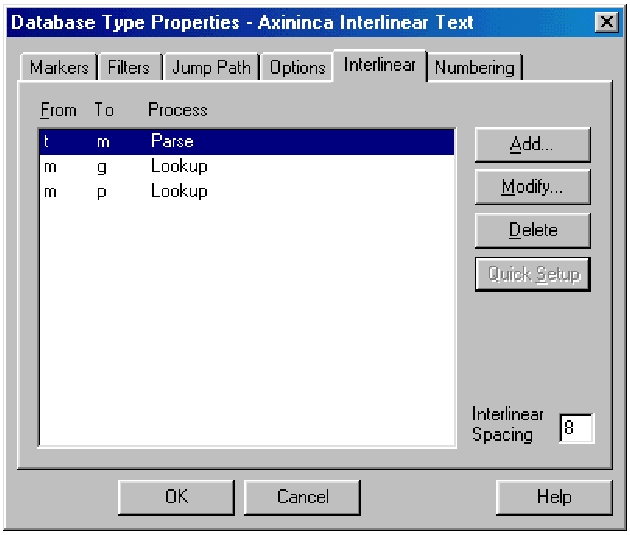
Part of Speech: ps

Alternate form: a

Underlying form: u

Fig. 213. Quick Setup Lexicon Markers for interlinearization

10. Click OK.



From	To	Process
t	m	Parse
m	g	Lookup
m	p	Lookup

Interlinear Spacing: 8

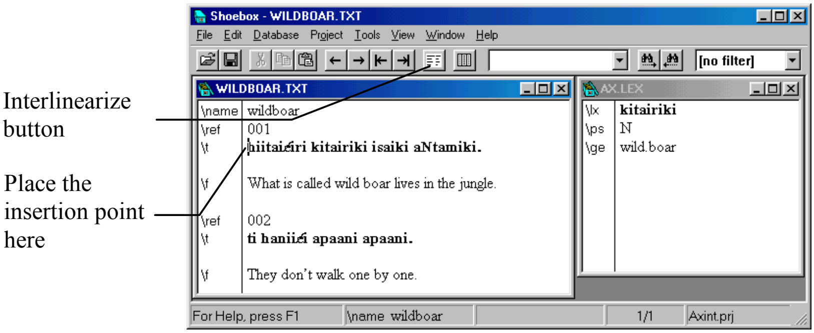
Fig. 214. Interlinear Process List for Axininca

As in the last exercise, you will see the Interlinear Process List now showing the three processes set up for you. The list of To markers shows the interlinear lines that will be generated by

Shoebox as it interlinearizes. The generated lines will be **m**, **g**, and **p**.

The names of the processes tell what they do. A Parse process parses words into morphemes. A Lookup process looks up glosses of words or morphemes in a lexicon. Notice that the two Lookup processes both look up morphemes from the **m** line. The first finds the gloss of the morpheme and the second finds the part of speech.

11. Click OK to return to the main window.


Interlinearize button
Place the insertion point here
Shoebox - WILDBOAR.TXT
File Edit Database Project Tools View Window Help
[no filter]
WILDBOAR.TXT
AX.LEX

```
\name  wildboar
\ref   001
\t     hiitaićiri kitairiki isaiki aNtamiki.

\f     What is called wild boar lives in the jungle.

\ref   002
\t     ti haniići apaani apaani.

\f     They don't walk one by one.
```

```
\lx   kitairiki
\ps   N
\ge   wild.boar
```

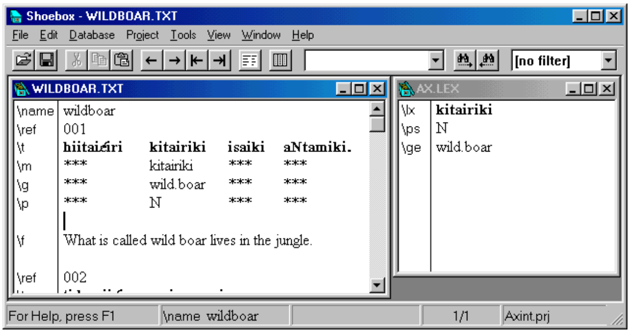
```
For Help, press F1   \name  wildboar   1/1   Axint.prj
```

Fig. 215. Axininca project ready for interlinearization

Interlinearizing

Put the text insertion point at the beginning of the first sentence of text (in the **\t** field).

1. Click ☰ to start interlinearization.



```
\name  wildboar
\ref   001
\t     hiitaićiri   kitairiki   isaiki   aNtamiki.
\m     ***          kitairiki   ***      ***
\g     ***          wild.boar   ***      ***
\p     ***          N           ***      ***

\f     What is called wild boar lives in the jungle.

\ref   002
```

```
\lx   kitairiki
\ps   N
\ge   wild.boar
```

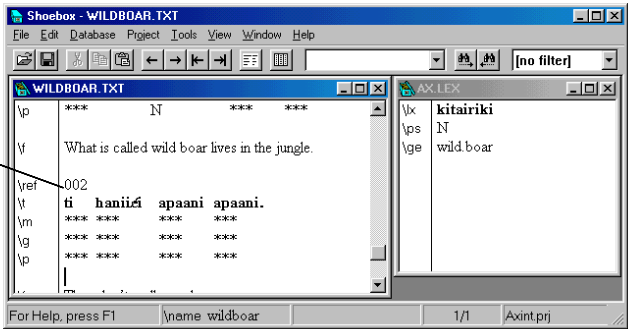
```
For Help, press F1   \name  wildboar   1/1   Axint.prj
```

Fig. 216. Wildboar text interlinearization

You will see Shoebox interlinearize the first sentence. The word *kitairiki* was the only entry in the lexicon. It interlinearizes with a gloss of *wild.boar* and a part of speech of *N*, as shown in the lexicon.[36] The other words have asterisks under them to show that Shoebox failed to parse them. (If *kitairiki* does not interlinearize successfully, go back to the start of interlinear setup, delete all of the processes, and run Quick Setup again. Return to the text, delete all of the wrong interlinear lines and try again.)

After interlinearizing a sentence, the insertion point is placed beneath the last interlinear line, where you can insert a free translation (this has already been done for you in this example text). To continue interlinearizing, you do *not* need to move the insertion point down to the next text line. The interlinearization process automatically moves on to the next sentence.

2. Click ☰ again to process the next text line.



```
\p     ***          N          ***     ***

\f     What is called wild boar lives in the jungle.

\ref   002
\t     ti    haniiɨ   apaani  apaani.
\m     ***   ***      ***     ***
\g     ***   ***      ***     ***
\p     ***   ***      ***     ***
```

```
\lx    kitairiki
\ps    N
\ge    wild.boar
```

For Help, press F1 | \name wildboar | 1/1 | Axint.prj

Point to the word *ti*

Fig. 217. Wildboar interlinear text - second line

You will see the second sentence interlinearized. The words have failure marks because they are not in the lexicon. The first word of the second sentence *ti* means *no*. To insert the word *ti* into the lexicon, you will use the Jump Insert command. Jump Insert works like the Jump To command, except that it saves you the

[36] It is common to use a period (full-stop) to tie words together in a multiple-word gloss, e.g., *wild.boar*. Though not required, it makes it easier to read multi-word glosses in interlinearized text.

step of clicking Insert in the No Matches dialog box if the morpheme is not found.

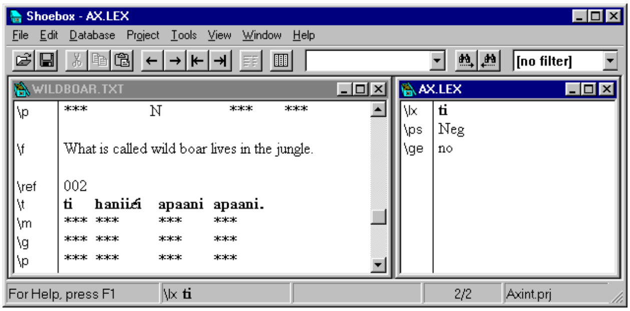
3. Point to the word *ti* and use Jump Insert (Windows: CTRL+right-click; Macintosh: CONTROL+OPTION+click).

The lexicon window will change to a new entry for *ti*. The part of speech and gloss fields are blank.

4. Type `Neg` as the part of speech.

5. Type `no` as the gloss.

The entry for *ti* should now look like the following:



```
\p     ***          N           ***      ***

\f     What is called wild boar lives in the jungle.

\ref   002
\t     ti   haniiŕi   apaani  apaani.
\m     ***  ***       ***     ***
\g     ***  ***       ***     ***
\p     ***  ***       ***     ***
```

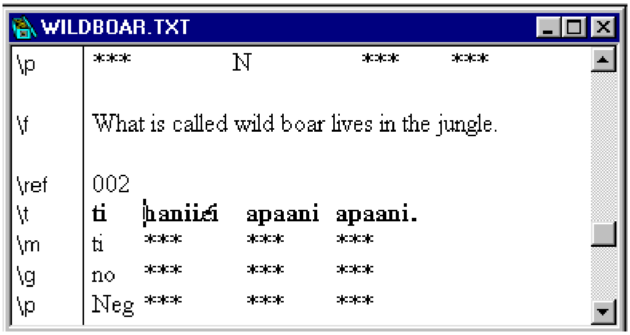
```
\lx   ti
\ps   Neg
\ge   no
```

Fig. 218. Axininca lexical entry for *ti*

Now you can interlinearize the word again and see it succeed.

6. Put the insertion point in front of the word *ti* in the Wildboar.txt window.

7. Click ☰.



```
\p     ***         N          ***      ***


\f     What is called wild boar lives in the jungle.


\ref   002
\t     ti   haniiri   apaani  apaani.
\m     ti   ***       ***     ***
\g     no   ***       ***     ***
\p     Neg  ***       ***     ***
```

Fig. 219. Axininca text with *ti* interlinearized

You will see the word *ti* interlinearize with a gloss of *no* and a part of speech of *Neg*.

Review You have completed the normal sequence of operations in interlinearizing. Here is a review of the steps:

1. Interlinearize a sentence.
2. For each failed word, select the word or its root and use Jump Insert to insert it into the lexicon.[37]
3. Fill in all the needed fields in the lexical entry.
4. Interlinearize the word again.

For further practice,

5. Use Jump Insert to insert the word *apaani* into the lexicon.
6. Type `Pron` as the part of speech.
7. Type `one` as the gloss.

[37] When Shoebox fails to parse a word, you can have it stop interlinearizing and automatically insert the word into your lexicon. To do this, select the Insert into Lexicon option in the Parse process. Also keep the option Output failure mark checked, so if you choose to skip a word and not insert it into the lexicon, Shoebox will mark that word for you. When Insert into Lexicon is checked, Shoebox will give a message, "Lookup failure. Insert into lexicon?" If you choose yes, the effect is the same as a Jump Insert.

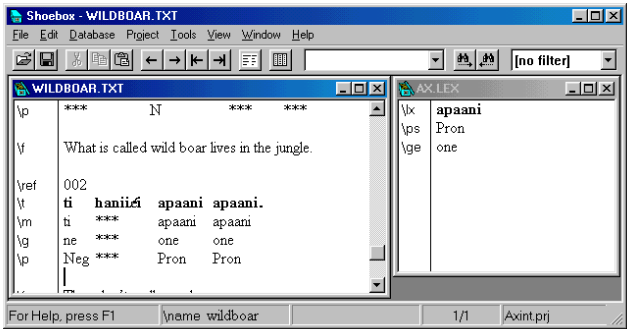
The result should be:

```
\lx   apaani
\ps   Pron
\ge   one
```

8. Place the insertion point somewhere in the first *apaani* in the text window and click ☰.

Notice that there are two occurrences of *apaani* in a row, so the second one can now succeed as well. (We will discuss reduplication in the section titled *Reduplication*, p. 252.)

9. Click ☰ again to interlinearize the second *apaani*.



```
\p     ***          N            ***      ***

\f     What is called wild boar lives in the jungle.

\ref   002
\t     ti   haniiʎi   apaani  apaani.
\m     ti   ***       apaani  apaani
\g     ne   ***       one     one
\p     Neg  ***       Pron    Pron
```

```
\lx   apaani
\ps   Pron
\ge   one
```

```
For Help, press F1 | \name  wildboar |  | 1/1 | Axint.prj
```

Fig. 220. Wildboar text with *apaani* glossed

Affixes

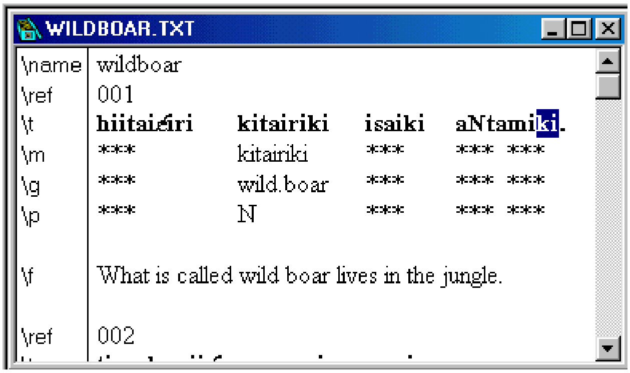
The Parse process breaks off affixes and handles morphophonemic changes to affixes and roots.

The simplest multiple morpheme word in the sample is the last word of the first sentence *aNtamiki* which means *in the jungle*. It has one suffix and is parsed *aNtami-ki*, glossed *jungle-LOC*. First you will need to enter -*ki* into the lexicon.

1. Scroll the text window to the first sentence.

To insert only part of a word, first select the part of the word you want and then use Jump Insert.

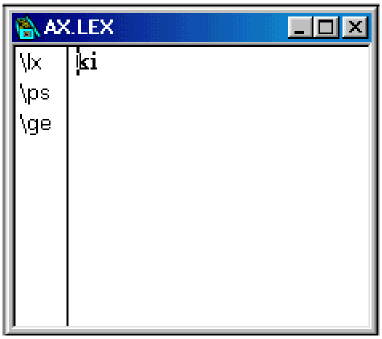
2. Select *ki*.



```
\name   wildboar
\ref    001
\t      hiitaiśri    kitairiki    isaiki    aNtamiki.
\m      ***          kitairiki    ***       ***  ***
\g      ***          wild.boar    ***       ***  ***
\p      ***          N            ***       ***  ***

\f      What is called wild boar lives in the jungle.

\ref    002
```

Fig. 221. Wildboar text with the *-ki* suffix selected

3. Insert *ki* into the lexicon using Jump Insert (Windows: CTRL+right-click; Macintosh: CONTROL+OPTION+click).



```
\lx   ki
\ps
\ge
```

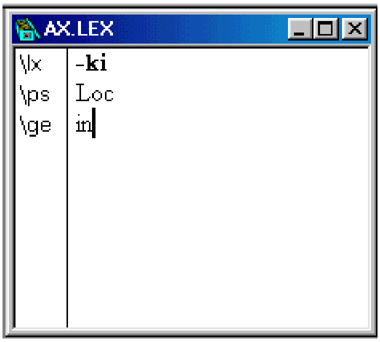
Fig. 222. AX.LEX with suffix *ki* inserted

4. Now put a hyphen on the front of *ki* in the lexical entry (i.e., *-ki*). This is important! Without a hyphen, *ki* is not a suffix and Shoebox will treat it like a root.

Note: The hyphen on an affix is *very* important and is very easy to miss. If ever an affix fails to interlinearize properly, make certain you did not forget the hyphen.

5. Add a part of speech of `Loc` and a gloss of `in`.[38]

The *-ki* suffix record in the lexicon should be as follows:



```
\lx  -ki
\ps  Loc
\ge  in
```

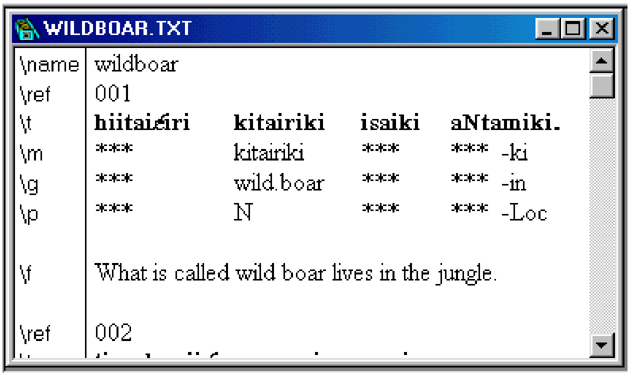
Fig. 223. AX.LEX with the suffix *-ki* glossed

To see what happens when the affix is in the lexicon but the root is not,

6. Interlinearize *aNtamiki*.[39]

[38] The term part of speech actually applies only to roots. Affixes are not classified by part of speech, and affix entries in a printed dictionary should not show a part of speech. But for interlinearizing, you can use the part of speech field to give helpful grammatical category information about the affix. In this example, the part of speech says that this affix is a locative affix but does not say what location it refers to. The gloss adds the information that this locative affix means *in*. Other common types of affixes include tense, aspect, person, and number. The category names chosen will vary depending on the morphological characteristics of the language.

[39] If you have forgotten how to interlinearize, click somewhere in the word *aNtamiki*, then click the interlinearize button on the toolbar.



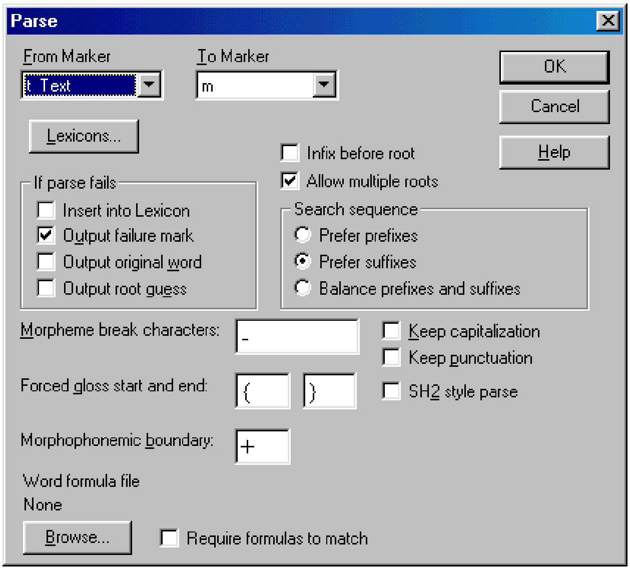
```
\name  wildboar
\ref   001
\t     hiitaiári    kitairiki   isaiki   aNtamiki.
\m     ***          kitairiki   ***      ***  -ki
\g     ***          wild.boar   ***      ***  -in
\p     ***          N           ***      ***  -Loc

\f     What is called wild boar lives in the jungle.

\ref   002
```

Fig. 224. Interlinearizing *aNtamiki*

Shoebox will parse and gloss the *-ki* suffix properly. It will display asterisks where the root would be if it were known.

There is an option in the interlinear setup that causes the Shoebox parser to guess that the remainder of a word, after possible affixes have been parsed, is a root. Shoebox will display a single asterisk in front of the root guess. When the root guess is correct, you can insert the root into the lexicon quickly using Jump Insert without selecting it first. Quick Setup does not turn on this helpful option. To set this option:

7. Choose Database, Properties.
8. Click the Interlinear tab.
9. Select the Parse process and click Modify... .



Parse

From Marker: t Text

To Marker: m

OK | Cancel | Help | Lexicons...

□ Infix before root

☑ Allow multiple roots

If parse fails
- □ Insert into Lexicon
- ☑ Output failure mark
- □ Output original word
- □ Output root guess

Search sequence
- ○ Prefer prefixes
- ● Prefer suffixes
- ○ Balance prefixes and suffixes

Morpheme break characters: -

□ Keep capitalization

□ Keep punctuation

Forced gloss start and end: { }

□ SH2 style parse

Morphophonemic boundary: +

Word formula file
None

Browse...

□ Require formulas to match

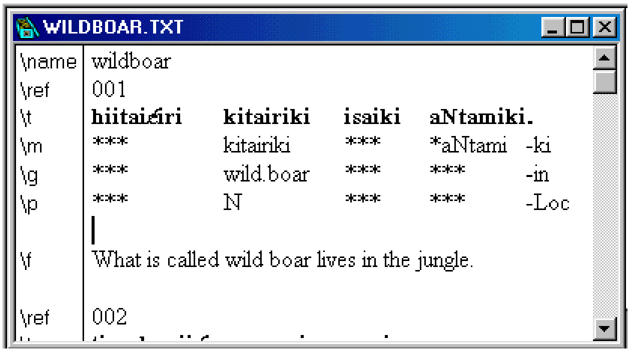
Fig. 225. Parse dialog box

You will see the Parse dialog box with a number of options.

10. Click □ Output root guess to select this option. (Leave the Output failure mark option checked also).

11. Click OK two times to return to the main window.

12. Now interlinearize *aNtamiki* again to see the difference this makes.



WILDBOAR.TXT

```
\name  wildboar
\ref   001
\t     hiitaićiri   kitairiki   isaiki   aNtamiki.
\m     ***          kitairiki   ***      *aNtami   -ki
\g     ***          wild.boar   ***      ***       -in
\p     ***          N           ***      ***       -Loc

\f     What is called wild boar lives in the jungle.

\ref   002
```

Fig. 226. *aNtamiki* interlinearized with the root guess

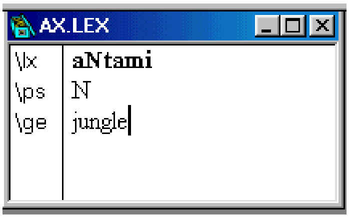
Shoebox will display **aNtami -ki* as the parse. The root guess is correct.

13. Use Jump Insert to insert *aNtami* in the lexicon.[40]

14. Type `N` as the part of speech.

15. Type `jungle` as the gloss.

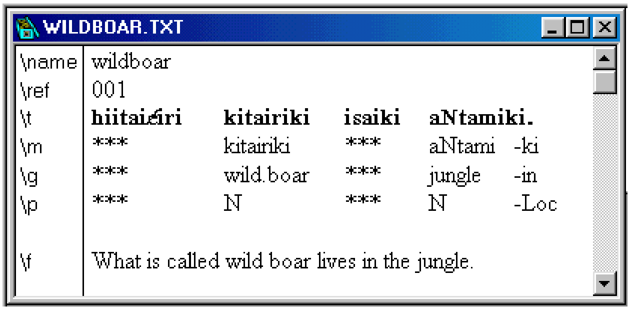
The result should be:



```
\lx   aNtami
\ps   N
\ge   jungle
```

Fig. 227. The *aNtami* record in the AX.LEX database

16. Interlinearize *aNtamiki* again to see that it parses correctly now.



```
\name  wildboar
\ref   001
\t     hiitaiŕiri   kitairiki   isaiki   aNtamiki.
\m     ***          kitairiki   ***      aNtami   -ki
\g     ***          wild.boar   ***      jungle   -in
\p     ***          N           ***      N        -Loc

\f     What is called wild boar lives in the jungle.
```

Fig. 228. *aNtamiki* with a correct parse

Morphophonemics

Morphophonemics refers to changes in the sounds of morphemes based on surrounding morphemes. This is often reflected in the spelling. Shoebox has a powerful morphophonemic system, and the Axininca language has lots of morphophonemics.

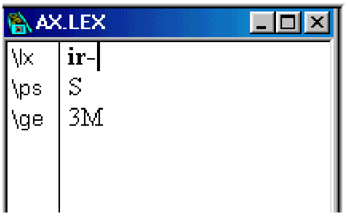
[40] When you use Jump Insert, the asterisk preceding *aNtami* will not be included because it is *not* a letter in the Axininca language.

Consider first the word *isaiki*. It is made up of three morphemes: *ir- saik -i*, glossed *3M- be,sit -NF*. (*3M* stands for third person masculine and *NF* means non-future.) The *ir-* prefix shows a morphophonemic process in which *r* is lost before *s*.

You cannot select *ir* and use Jump Insert since it is not visible in the word. So insert just the *i* into the lexicon and then add the *r-*:

1. Select the initial letter *i* and use Jump Insert to insert it.
2. Fix the new entry by adding *r* and a hyphen to make it *ir-*. Do not forget the hyphen!
3. Type `S` (for subject) as the part of speech.
4. Type `3M` as the gloss.

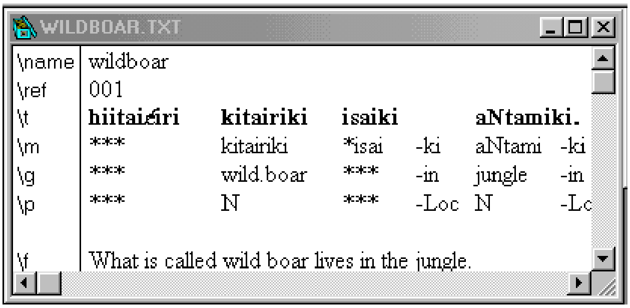
The *ir-* prefix entry should look like the following:



```
\lx   ir-
\ps   S
\ge   3M
```

Fig. 229. *ir-* prefix record

5. Interlinearize the word *isaiki* again to see what happens.



```
\name  wildboar
\ref   001
\t     hiitaiśiri   kitairiki   isaiki        aNtamiki.
\m     ***          kitairiki   *isai   -ki   aNtami  -ki
\g     ***          wild.boar   ***     -in   jungle  -in
\p     ***          N           ***     -Loc  N       -Lc

\f     What is called wild boar lives in the jungle.
```

Fig. 230. Parsing *isaiki* in the wildboar text

You will see that *-ki* has been cut off the back, but nothing has been cut off the front. The *-ki* is a false guess, which will go away when the root is in the lexicon.

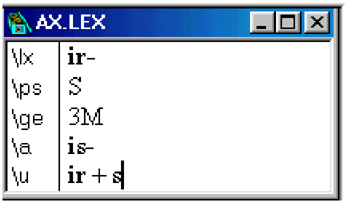
The prefix *ir-* was not found because you have not yet accounted for the morphophonemic change that makes it *i-* before *s*. To do that, use the Alternate form and Underlying form fields in the lexicon. (You may recall that the Quick Setup defaulted to **\u** and **\a** as the field markers for those fields.)

In Shoebox, to say that *r* goes away before *s*, you have to say that the morpheme *ir-* becomes *i-* before *s*. To express that, you say the underlying form *ir+s* becomes *is*. To put this information into the lexical database using the Alternate and Underlying form fields,

6. Enter the following two fields at the bottom of the *ir-* entry:[41]

```
\a    is-
\u    ir+s
```

Your *ir-* prefix record should look like the following:



\lx ir-
\ps S
\ge 3M
\a is-
\u ir + s

Fig. 231. The *ir-* prefix record with alternate and underlying forms

Again, this can be read as: *ir-* plus *s* becomes *is*, or, if you prefer, the surface string *is-* comes from an underlying form of *ir-* plus *s*.

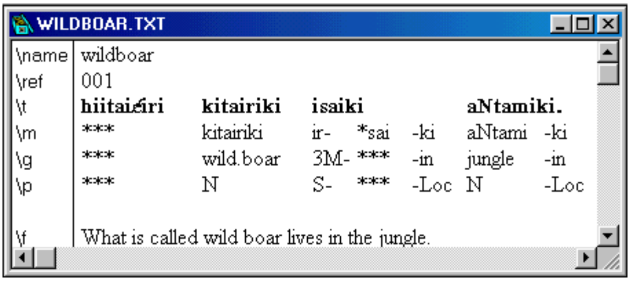
Because the main focus of this notation is parsing, the surface form (*is-*) is placed above the underlying form (*ir+s*). If you are thinking generatively, it may help you to regard the underlying form as primary, but be sure to order the fields as shown.

[41] The plus sign in the Axininca font might look like it has spaces before and after it, but it does not. The parse will not work if you put spaces before or after the plus sign.

Important: The alternate surface form *must* have a hyphen to show that it is an affix. If the underlying form has a morpheme boundary symbol, no hyphen is needed. If there is no morpheme boundary, the hyphen is required. The default morpheme boundary symbol is a plus sign (+).

Explanation: When Shoebox is trying to parse a word like *isaiki* and it finds that *is*- is a possible prefix in the lexicon, it looks at the underlying form (*ir*+*s*). Because *is*- is a prefix, Shoebox assumes that the part of the underlying form before the morpheme boundary (the *ir*) is also a prefix. Then it puts the part after the morpheme boundary (the *s*) back onto the remainder of the word (*aiki* to get *saiki*) and tries to parse the result. The parse of the result will succeed or fail depending on what else is in the lexicon (i.e., if *saiki* were in the lexicon as a root form, the parsing would be successful; if not, the parsing would fail).[42]

7. Interlinearize the word *isaiki* again.



```
\name  wildboar
\ref   001
\t     hiitairiri  kitairiki  isaiki              aNtamiki.
\m     ***         kitairiki  ir-  *sai  -ki      aNtami  -ki
\g     ***         wild.boar  3M-  ***   -in      jungle  -in
\p     ***         N          S-   ***   -Loc     N       -Loc

\f     What is called wild boar lives in the jungle.
```

Fig. 232. Parse of *isaiki*

You will now see the prefix *ir*- and the suffix -*ki* cut off, leaving a root guess of **sai*. This is not quite right since the root is really *saik*, but it is getting closer. It shows that the morphophonemic change on *ir*- is working.

[42] If it were a *suffix* with the underlying form *ir*+*s*, then Shoebox would assume that the part *after* the morpheme boundary (the *s*) is a suffix and would put the part before the morpheme boundary (the *ir*) back onto the remaining string for further parsing.

You are still missing two lexical entries needed to fully parse this word: the root and the correct suffix. First add the suffix *-i*.

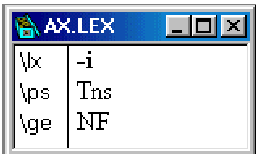
8. Select the final letter *i* and use Jump Insert to insert it.

9. Add a hyphen before the *i* in the entry. Remember that this is very important!

10. Type `Tns` as the part of speech.

11. Type `NF` as the gloss.

The result should be:



```
\lx  -i
\ps  Tns
\ge  NF
```

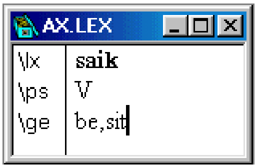
Fig. 233. The *-i* suffix record

12. Select *saik* in the **\t** line and use Jump Insert to insert it. (Since *saik* is a root, do not put a hyphen on this one.)

13. Type `V` as the part of speech.

14. Type `be,sit` (with no space after the comma) as the gloss.

The *saik* root record should be as follows:

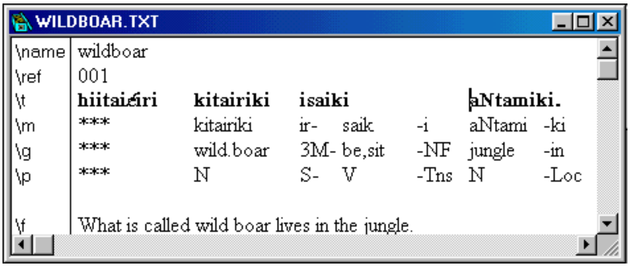


```
\lx  saik
\ps  V
\ge  be,sit
```

Fig. 234. The *saik* root record

The comma is often used to mean *or* in a multiple-word gloss. So *saik* means *be* or *sit*. It is usually impossible to find a gloss that is fully equivalent, so sometimes techniques like this are used to help clarify the meaning.[43]

15. Interlinearize *isaiki* again.



```
\name  wildboar
\ref   001
\t     hiitaiciri    kitairiki   isaiki                aNtamiki.
\m     ***           kitairiki   ir-   saik     -i     aNtami   -ki
\g     ***           wild.boar   3M-   be,sit   -NF    jungle   -in
\p     ***           N           S-    V        -Tns   N        -Loc

\f     What is called wild boar lives in the jungle.
```

Fig. 235. *Isaiki* parsed correctly

This time it will parse.[44] You now have three of the four words in the first sentence interlinearized.

Formatting the interlinear lines

One thing you have not yet done is set the language encodings and fonts for the interlinear lines. The second line showing morpheme breaks is in the Axininca language.

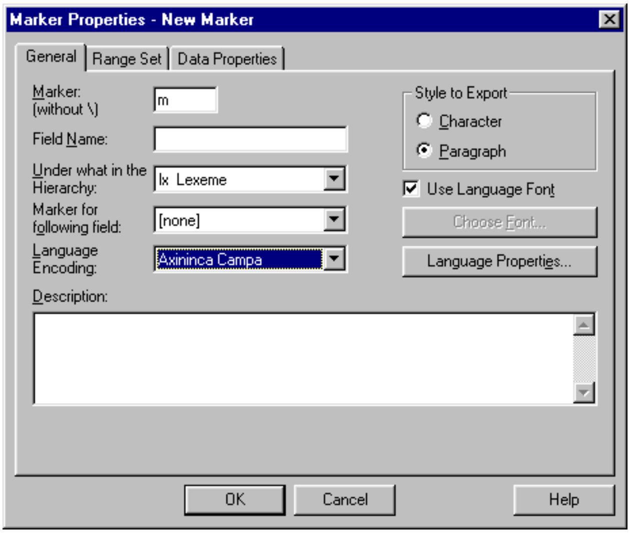
1. Right-click (Macintosh: OPTION+click) **\m** to see the Marker Properties dialog box.

2. Select Axininca as the Language Encoding.

[43] If a morpheme has a single meaning, it should not be given multiple senses even if it does not correspond to a single word in the glossing language. One way to deal with this is to use the most common equivalent as the gloss and explain the full range of meaning in the lexical entry. The other way is to join alternative glosses into one.

If a morpheme truly has multiple senses, then it should have multiple glosses. When interlinearizing a text, you should choose the sense which corresponds to the meaning in context.

[44] If it does not parse, verify that you have entered the hyphens correctly and that your morphophonemic fields are exactly as shown earlier.



Marker: (without \): m

Field Name:

Under what in the Hierarchy: lx Lexeme

Marker for following field: [none]

Language Encoding: Axininca Campa

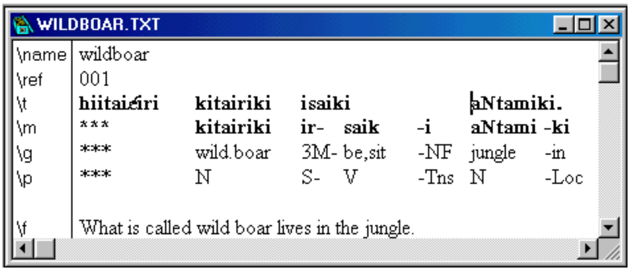
Description:

Style to Export: ○ Character ● Paragraph

☑ Use Language Font

Fig. 236. Morpheme encodings set to Axininca

3. Click OK.



```
\name  wildboar
\ref   001
\t     hiitaiɨ́ri   kitairiki   isaiki             aNtamiki.
\m     ***         kitairiki   ir-  saik    -i    aNtami -ki
\g     ***         wild.boar   3M-  be,sit  -NF   jungle -in
\p     ***         N           S-   V       -Tns  N      -Loc

\f     What is called wild boar lives in the jungle.
```

Fig. 237. All Axininca text using the Axininca font

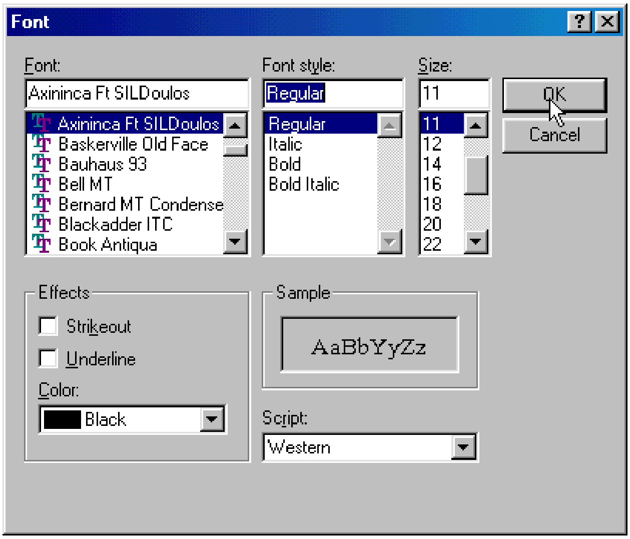
The second line is now in the Axininca font and is bold. The main text line will show up better if this line is not bold, so change the font style of the second line to regular.

4. Right-click (Macintosh: OPTION+click) **\m** to see the Marker Properties dialog box.

5. Click ☑ Use Language Font to clear it.

6. Click Choose Font... .

7. Change the font style from Bold to Regular.



Font

Font: Axininca Ft SILDoulos

Font style: Regular

Size: 11

Axininca Ft SILDoulos
Baskerville Old Face
Bauhaus 93
Bell MT
Bernard MT Condense
Blackadder ITC
Book Antiqua

Regular
Italic
Bold
Bold Italic

11
12
14
16
18
20
22

Effects

□ Strikeout

□ Underline

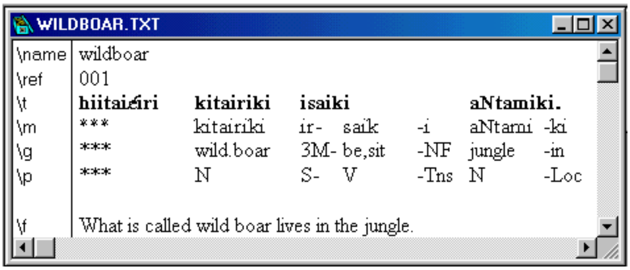
Color: Black

Sample: AaBbYyZz

Script: Western

Fig. 238. Selecting a special font for the Axininca morphemes

8. Click OK two times to return to the main window.



WILDBOAR.TXT

```
\name  wildboar
\ref   001
\t     hiitaiŕiri   kitairiki   isaiki               aNtamiki.
\m     ***          kitairiki   ir-  saik    -i      aNtami  -ki
\g     ***          wild.boar   3M-  be,sit  -NF     jungle  -in
\p     ***          N           S-   V       -Tns    N       -Loc

\f     What is called wild boar lives in the jungle.
```

Fig. 239. Wildboar text window with special fonts

The morpheme line is now in the regular font style.

Another thing that can help the readability of interlinear text is to use style or color to distinguish the lines. If you have a color display, make the gloss red and the part of speech blue. If you have a monochrome display, make the gloss bold italic and the part of speech italic.

9. Right-click (Macintosh: OPTION+click) **\g** to see the Marker Properties dialog box.

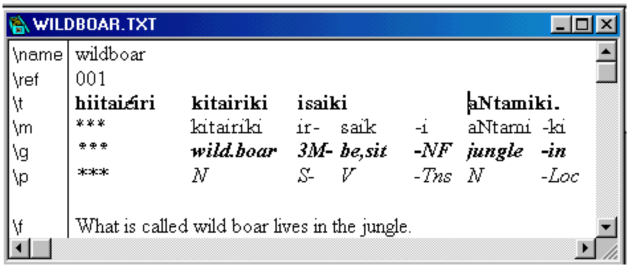
10. Click ☑ Use Language Font to turn it off.

11. Click Choose Font... .

12. Change the color to Red, or if you have a monochrome display, select Bold Italic.

13. Click OK two times to return to the main window.

14. Use the same process to change the **\p** line to Blue, or if you have a monochrome display, to Italic.



```
\name  wildboar
\ref   001
\t     hiitai¢iri   kitairiki   isaiki                 aNtamiki.
\m     ***          kitairiki   ir-  saik      -i      aNtami  -ki
\g     ***          wild.boar   3M-  be,sit    -NF     jungle  -in
\p     ***          N           S-   V         -Tns    N       -Loc

\f     What is called wild boar lives in the jungle.
```

Fig. 240. Wildboar interlinearization with font enhancements

Notice how much easier it is now to pick out the different lines.

Epenthesis

Now look at the first word in the first sentence. It has morphophonemics in the suffixes. The *¢* is epenthesized between the third and fourth *i*, and the *t* is epenthesized between the second *i* and the *a*. The word parses as *ir- ii -ai -i -ri*. It is glossed as *3M-name -PASS -NF -REL*. The affix glosses stand for third person masculine, passive, non-future, relative. It means *what is called*.

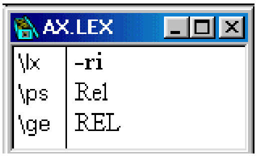
In the following steps, the hardest part is remembering to put hyphens on the affix entries in the lexicon. If at any time you do not get the expected parse result, check to make sure you did not forget a hyphen.

1. Enter the suffix `-ri` in the lexicon. Be sure to put a hyphen in front of *ri*.

2. Type `Rel` as the part of speech.

3. Type `REL` as the gloss.

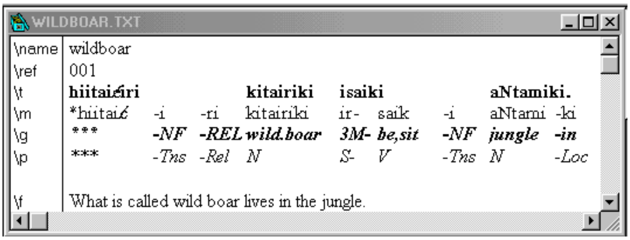
The result should look like the following:



```
\lx   -ri
\ps   Rel
\ge   REL
```

Fig. 241. The *-ri* suffix record in Axininca

4. Interlinearize *hiitai¢iri*.



```
\name  wildboar
\ref   001
\t     hiitai¢iri              kitairiki  isaiki              aNtamiki.
\m     *hiitai¢  -i    -ri    kitairiki  ir-  saik    -i    aNtami  -ki
\g     ***       -NF   -REL   wild.boar  3M-  be,sit  -NF   jungle  -in
\p     ***       -Tns  -Rel   N          S-   V       -Tns  N       -Loc

\f     What is called wild boar lives in the jungle.
```

Fig. 242. Interlinearization of *hiitai¢iri*

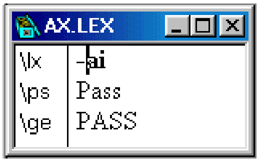
Shoebox will parse the *-ri* and *-i* suffixes correctly.

5. Enter the suffix `-ai` in the lexicon. (Do not forget the hyphen.)

6. Type `Pass` (for Passive) as the part of speech.

7. Type `PASS` as the gloss.

The result should look like the following:



```
\lx   -ai
\ps   Pass
\ge   PASS
```

Fig. 243. The *-ai* suffix in the Axininca lexicon

8. Interlinearize again and observe that nothing changes since the epenthesized *¢* has not been accounted for.

To deal with that, enter an alternate form under the *-i* morpheme. Here, do not put the epenthesis with *-ai,* but rather with *-i*, which is farther out in the word.

Note: A morphophonemic change between two morphemes is always placed in the lexical entry of the *outer* affix.

9. To jump to the *-i* lexical entry, right-click *-i* (Macintosh: OPTION+click).
10. At the bottom of the entry add the following:

```
\a   -i¢i
\u   i+i
```

To enter the *¢*, you can copy it from the text window or type it directly. On Windows, press ALT+0193 on the numeric keypad, or type `/c` if the Axininca keyboard is active. (Macintosh: OPTION+1)

These two fields say that when the suffix *-i* is after another *i*, *¢* is inserted between them. Do not forget to put the hyphen in front of *i¢i*. Notice that even though this looks like a phonological rule, it is not fully so, because the part to the right of the + must be a full suffix, not just a letter.

As a general rule, the underlying form must include a whole affix and may include part of the adjacent affix or root. Also, it cannot be expressed as a feature or class, such as *vowel*.

Morphophonemic notation

The same morphophonemic notation applies to both prefixes and suffixes but in opposite directions, which can be confusing.

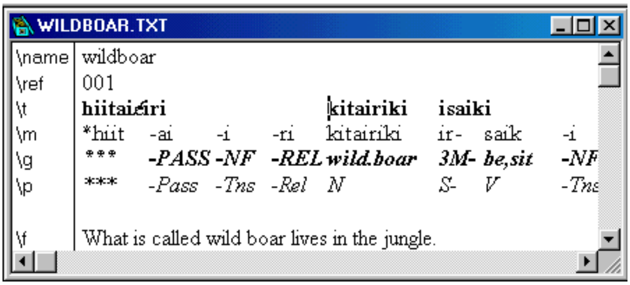
Two basic principles regarding morphophonemic notation are:

- Morphophonemic changes are expressed in Shoebox notation in the entry of the *outer morpheme*, not the inner one. Roots are always considered *inner*. So changes to roots are expressed in the entries of the adjacent affixes, not in the entries of the roots. This is good because affixes are a much smaller class than roots, so the number of morphemes that need alternate forms to account for morphophonemics is quite small.[45]

- The part of the underlying form on the same side of the morpheme boundary symbol (+) as the hyphen (the inner side) is kept for further parsing. The part on the other side of the morpheme boundary (the outer side) is cut off and matched with the underlying form of the affix.

To verify that what you have done thus far is actually working,

1. Interlinearize the word *hiitai¢iri* again.



```
\name  wildboar
\ref   001
\t     hiitai¢iri               kitairiki   isaiki
\m     *hiit  -ai   -i   -ri    kitairiki   ir-  saik   -i
\g     ***    -PASS -NF  -REL   wild.boar   3M-  be,sit -NF
\p     ***    -Pass -Tns -Rel   N           S-   V      -Tns

\f     What is called wild boar lives in the jungle.
```

Fig. 244. Interlinearization of *hiitai¢iri* again

Observe that *-ai* is now parsed correctly. (If it is not, check to be sure that you did not forget the hyphen.)

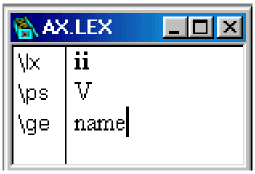
2. Enter the root `ii` in the lexicon.

3. Type `V` as the part of speech.

4. Type `name` as the gloss.

[45] For example, if voiceless stops become voiced after a nasal, *pat* would become *bat* after *in-*. This would be expressed in the *in-* entry with the alternate **\a** *inb* and the underlying form **\u** *in+p*. If the change were that *pat* became *pad* (voicing the final consonant) before a nasal, a suffix like *-na* would be where the root change is expressed with **\a** *dna* and **\u** *t+na*. In either case, if there were a change in the affix (e.g., assimilation to the point of articulation), that would also be expressed in the affix entry. So whether the change is in the root or the affix, it is expressed in the affix entry.

The *ii* lexeme should look like:



```
\lx   ii
\ps   V
\ge   name
```

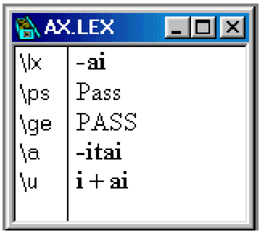
Fig. 245. The *ii* lexeme entry

There are still two morphophonemic changes to account for before this word will parse completely. The first is that a *t* is epenthesized between the final *-i* of *ii* and the *a* of *-ai*.

5. Jump to the entry for *-ai* and add the following:

```
\a    -itai
\u    i+ai
```

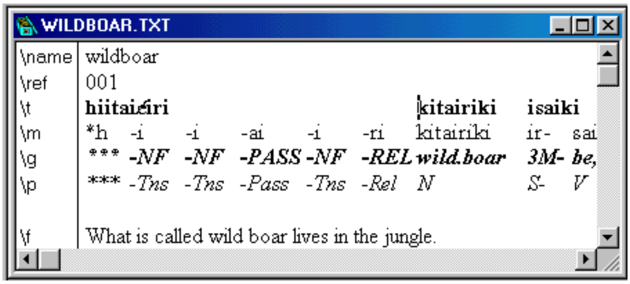
This says that when the suffix *-ai* comes after an *i*, a *t* is added. The *-ai* record now looks like the following:



```
\lx   -ai
\ps   Pass
\ge   PASS
\a    -itai
\u    i + ai
```

Fig. 246. The *-ai* lexeme with alternate and underlying forms

6. Interlinearize *hiitai¢iri* again.



```
WILDBOAR.TXT
\name  wildboar
\ref   001
\t     hiitai¢iri                         kitairiki   isaiki
\m     *h   -i   -i   -ai    -i   -ri    kitairiki   ir-  sai
\g     ***  -NF  -NF  -PASS  -NF  -REL   wild.boar   3M-  be,
\p     ***  -Tns -Tns -Pass  -Tns -Rel   N           S-   V

\f     What is called wild boar lives in the jungle.
```

Fig. 247. Parsing *hiitai¢iri* again

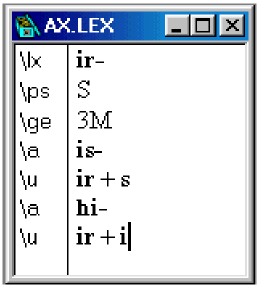
Observe that now the parse shows **h -i -i -ai -i -ri*. The first part of the word is parsed incorrectly. When morphemes are missing from the lexicon or morphophonemic changes have not been accounted for, Shoebox often parses incorrectly, resulting in more affixes than are really present.

The last morphophonemic change to account for is that the prefix *ir-* changes to *h* before an *i*.

7. Jump to the entry for *ir-* and add the following fields at the bottom of the record (underneath the other alternate form fields):

```
\a    hi-
\u    ir+i
```

The *ir-* record should look like the following:



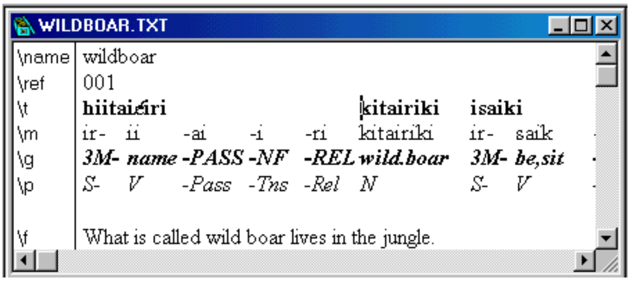
```
AX.LEX
\lx   ir-
\ps   S
\ge   3M
\a    is-
\u    ir + s
\a    hi-
\u    ir + i
```

Fig. 248. Record for *ir-* lexeme in Axininca

This says that the prefix *ir-* changes to *h* when it comes before an *i*.

8. Interlinearize *hiitai¢iri* again.

Observe that the word now parses correctly.



```
\name  wildboar
\ref   001
\t     hiitai¢iri                         kitairiki   isaiki
\m     ir-  ii    -ai   -i   -ri   kitairiki   ir-  saik   -
\g     3M-  name  -PASS -NF  -REL  wild.boar   3M-  be,sit -
\p     S-   V     -Pass -Tns -Rel  N           S-   V      -

\f     What is called wild boar lives in the jungle.
```

Fig. 249. The word *hiitai¢iri* correctly parsed

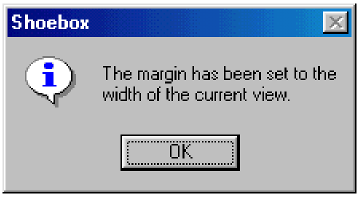
That may have seemed like a lot of work, but once the affixes and morphophonemic changes are entered, interlinearizing is faster and easier. All that is needed after that is to enter the root of each new word.

Reshaping interlinear text

The first sentence should be completely interlinearized now. You may have noticed that sentences lengthen as they are interlinearized. For example, the first sentence now stretches past the right edge of the window. If you want to keep a text segment visible without having to scroll horizontally, you can reshape the text. Set the margin by sizing the window to the width you want and setting the wrap margin to that width.

1. Choose Tools, Set Wrap Margin.

You will see a message confirming that the wrap margin has been set to the width of the current view.



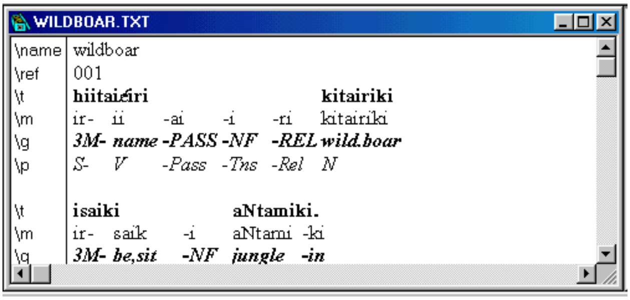
The margin has been set to the width of the current view.

Fig. 250. Margin set verification dialog box

2. Click OK.

3. Put the insertion point in the first sentence.

4. Choose Tools, Reshape.



```
WILDBOAR.TXT
\name  wildboar
\ref   001
\t     hiitai¢iri                    kitairiki
\m     ir-  ii    -ai   -i    -ri   kitairiki
\g     3M-  name -PASS -NF   -REL  wild.boar
\p     S-   V    -Pass -Tns  -Rel  N

\t     isaiki           aNtamiki.
\m     ir-  saik   -i   aNtami -ki
\g     3M-  be,sit -NF  jungle -in
```

Fig. 251. Reshaping interlinear text within new margins

Observe that the sentence and all of the interlinear fields are now broken so that they are all visible again within the width of the window. (You can also reshape an entire interlinear file to a new margin width by choosing Tools, Reshape Entire File.)

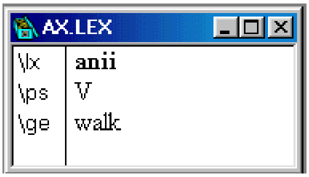
More practice

The second word in sentence two *hanii¢i* is the only failed word left. It needs to be parsed as *ir- anii -i*, glossed as *3M- walk -NF*, with the parts of speech of *S- V -Tns*. To get this to parse, enter the root *anii* and add another morphophonemic change to *ir-* to indicate that *ir-* becomes *h* before *a* (the suffix *-i* has already been dealt with).

This will give you more practice on what you have learned about morphophonemics.

1. Enter the root and morphophonemic change required to make *hanii¢i* parse properly.

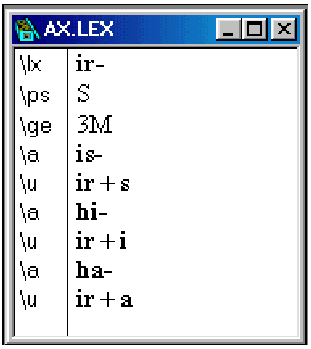
The root *anii* should look like the following:



```
AX.LEX
\lx   anii
\ps   V
\ge   walk
```

Fig. 252. Axininca record for the lexeme *anii*

The entry for the prefix *ir-* should become:

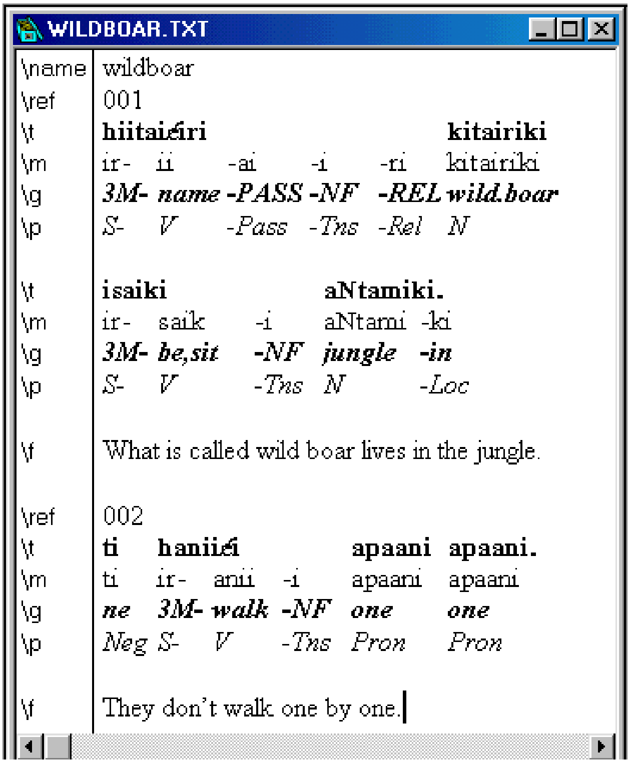


\lx	ir-
\ps	S
\ge	3M
\a	is-
\u	ir+s
\a	hi-
\u	ir+i
\a	ha-
\u	ir+a

Fig. 253. Axininca record for the *ir-* lexeme

2. Interlinearize the word *hanii¢i* again.

You should now see the text fully interlinearized, thus:



```
\name  wildboar
\ref   001
\t     hiitaiɨ́ri                      kitairiki
\m     ir-  ii    -ai   -i    -ri   kitairiki
\g     3M-  name -PASS -NF   -REL  wild.boar
\p     S-   V    -Pass -Tns  -Rel  N

\t     isaiki            aNtamiki.
\m     ir-  saik   -i    aNtami -ki
\g     3M-  be,sit -NF   jungle -in
\p     S-   V      -Tns  N      -Loc

\f     What is called wild boar lives in the jungle.

\ref   002
\t     ti   haniiɨ́            apaani  apaani.
\m     ti   ir-  anii -i     apaani  apaani
\g     ne   3M-  walk -NF    one     one
\p     Neg  S-   V    -Tns   Pron    Pron

\f     They don't walk one by one.
```

Fig. 254. Fully interlinearized wildboar text

If you do not get this result, you may have entered the alternate form incorrectly. The following is a common mistake when specifying morphophonemic processes:

* \a h-
* \u ir+a

This does not work because it removes the *h* and changes *it* to *ir* but also adds an *a* onto the root (yielding *aan*). You must include an *a* in the match string too.

\a ha-
\u ir+a

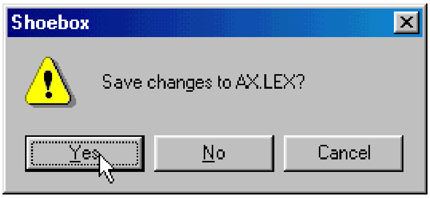
It might seem as if you could leave the *a* off altogether:

```
\a   h-
\u   ir-
```

This would result in a proper parsing of this example. However, this is too general, since it allows Shoebox to remove any *h-* and change it to *ir-*. This tends to produce ambiguities.

That completes the interlinearization of the first two lines of the Axininca wildboar text. It is interesting to note that in so doing, you have expanded the Axininca lexicon from only one entry to twelve. Admittedly, at first this can be a slow and somewhat painful process (especially for morphologically complex languages), but as the basic affixes become well defined, the process of interlinearizing text (and expanding your lexicon at the same time) quickly picks up speed.

3. Choose Project, Close.



Save changes to AX.LEX?

Fig. 255. Save Changes dialog box

For each of the open database files, Shoebox will ask if you want to save the changes.

4. Click Yes each time (Macintosh: click Save).

This concludes the tutorial on interlinearization. There are probably many issues and concepts that are not yet clear to you. For more information on the theory and details of interlinearization in Shoebox, see the chapter *Notes on interlinearization*, p. 247.

Tutorial milestone

This concludes the *Tutorial*. The remainder of this documentation describes several key features of Shoebox not covered in the *Tutorial*, gives tips on overcoming some of the limitations of Shoebox, and gives practical answers to the more common questions we have encountered.

PART 2

USER'S GUIDE

Basic information

Projects

In Shoebox, the current workspace is the current project. It consists of the number and position of open windows, which files are open, and how they are sorted and filtered.

The location of the project settings file determines the location of the language encoding files and the database type files which are used by that project, because a project file expects to find all of the other settings files with it in the same folder. (Data files should normally be kept elsewhere.)

Different project files may share the same folder and thus share the same language encoding and database type files.

Individual users will gain the most benefit by keeping all projects in a single folder.

Computer consultants will prefer to have multiple project folders, one for each person they are helping. This will prevent unintentional redefinition of a database type or language encoding.

To change to a different project while Shoebox is running, choose Project, Open and specify the project you want to work on.

To create a new project, choose Project, New, then configure the workspace as desired.

The current project is saved in any one of these ways:

- Opening another project (Choose Project, Open).
- Creating a new one (Choose Project, New).
- Saving it directly (Choose Project, Save).

- Saving it to a new name (Choose Project, Save As).
- Closing the current project (Choose Project, Close), which then gives you a chance to open another one, create a new one, or exit Shoebox.
- Exiting the program (Choose File, Exit).

Whenever you start Shoebox, it will normally open the project you were using in the previous session. However, assuming that Shoebox is not already running, you can start it and have it open a particular project. To do this,

- On Windows, double-click the project file in Windows Explorer (File Manager on Windows 3.1).[46]
- On the Macintosh, double-click the project file in the Finder.

As a matter of convenience, you may wish to create a shortcut icon for a project so that you can open it from the desktop or another convenient location. This instructions for doing this are included on p. 244 in the *Frequently asked questions* chapter.

Field markers

Field markers identify each piece of data. Each marker consists of a backslash character followed by any combination of alphabetic characters, numbers, and punctuation.[47]

The first marker in a record, called the record marker, is the dividing point between records. Thus the record marker can occur only once per record.

Any other marker can occur as often as is appropriate for the type of data and for the particular record.

Field marker list

A list of field markers is maintained for each database. The database type file holds this information from one session to another, so it is not lost when you exit the program.

[46] This works because the Shoebox setup program associates all .prj files with the Shoebox program.

[47] Some MS-DOS programs restrict field markers to a maximum of four characters (not including the backslash character).

The marker list is on the Markers tab of the Database Type Properties dialog box. To display this dialog box for the active database, choose Database, Properties. On the Markers tab:

- Field markers in use are bold; unused field markers are not bold.
- A field marker is unused if it is not contained in any record.
- Only unused field markers may be deleted from the marker list.

A field marker may be added to the list even though no record yet contains it. This allows the creation of standard field marker sets (such as the standard set which the Multi-Dictionary Formatter program recognizes). To add a new marker, choose Database, Properties, Markers tab, Add.

Marker properties

The database type file also maintains a set of information about each marker, including its field name, language encoding, and place in the hierarchy. This information is displayed on the Marker Properties dialog box. This dialog box can be accessed in either of two ways:

- Right-click (Macintosh: OPTION+click) the marker in the marker pane of an open window.
- Choose Database, Properties, Marker tab, select the marker, and click Modify....

Field name

A marker may be given a name, called a Field Name.

- To have Shoebox show the field names in the marker pane of the record window, choose View, Field Names.
- To see just the markers, choose View, Markers.
- To see both, choose View, Both Markers and Names.

If a field is pasted into a record from another application and its marker is not already included in the list of markers for that database, a warning is given. Shoebox automatically adds the new marker to the list. Be sure to go to the Marker Properties dialog box for that marker to type a meaningful field name and make sure the other properties (like language and hierarchy, etc.) are correct. The new marker is given a single asterisk (*) as its field name, making it easy to find when you look at the marker list.

Marker language encoding

The data in a field is displayed and interpreted according to the language encoding of the field marker. In Shoebox, the language encoding includes font, sorting information, and keyboard. For the data to be interpreted correctly, it is important that each marker have the appropriate language encoding and that each data field have the appropriate marker.

Shoebox automatically selects the default language encoding for new markers. This happens either when a database with unknown markers is imported the first time or when you add a new marker to an existing database. You can change which language encoding is the default. Go to the Markers tab in the Database Type Properties dialog box. Change the language specified in the Default Language Encoding list.

If the default language encoding is not appropriate for an individual marker, you must select the appropriate language encoding. Modify the Marker Properties for that marker (right-click the marker).[48]

Marker font

You can specify a different font or a special enhancement for an individual marker. In the Marker Properties dialog box, clear ☑ Use Language Font then click Choose Font.... Choose the font, point size, color, attribute, etc. Changes made for this marker will have no effect on any other marker. On Windows 3.1, special fonts for individual markers consume system resources faster than using the Language Font. In some cases, this could cause an error message from the system.

[48] There is no convenient way to change the database properties of a group of markers. For advanced users: Exit Shoebox, edit the database type file, and find and replace all occurrences of language encoding you want to change. for example, to change the Default language to National, look for **\lng** Default and replace it with **\lng** National. Note that these names are case-sensitive. Save as text and restart Shoebox.

Range sets

Fields can be restricted to a limited set of elements (e.g., the parts of speech). The elements of this set are listed on the Range Set tab in the Marker Properties dialog box. To create a range set for a field marker:

1. Go to the Marker Properties dialog box for that field and click the Range Set tab.
2. Click the Use Range Set check box (if it is not already checked).
3. In the Range Set Element box, type a word, abbreviation, code, or phrase that should be regarded as acceptable for this type of field (e.g., `v.t.`).
4. Click the Add button to add the element to the list.
5. Repeat steps 3 and 4 until all the acceptable values for this field have been added to the list (as your needs change, you can later add more elements or remove elements that are no longer needed).
6. Click OK.

To create a range set for a data field that is already in use, click the Build Range Set from the data button. Shoebox compiles a list of range set elements from the data fields in the active window for you automatically. You may need to edit the list Shoebox produces to delete or modify invalid elements or to add additional valid elements that did not occur in the data.

Shoebox checks the consistency of data in three ways:

- If you edit a record, its data fields are checked before any operation that could cause a different record to be displayed in the active window (e.g., Next Record on the Database menu).
- When you export a database, it is checked if you have selected that option in the export process properties.
- When you choose Check Consistency on the Tools menu, the records in the active window are checked. (If the window is

filtered, any records that do not match the filter are not checked.)

If an inconsistency is found in the data, it is selected in the active window and a dialog box is displayed.

Marker hierarchy

Most records have an internal structure. For example, everything in a dictionary entry is “under” the lexeme; i.e., you will not (should not) find information in an entry that is irrelevant to the lexeme. Similarly, illustrative sentences are under a particular sense—they illustrate *that* sense of meaning, not some other. In Shoebox, this relationship can be reflected using the Hierarchy feature. For a simple example which illustrates the purpose and use of hierarchy, see *Marker hierarchy* on p. 69.

Variables

In Shoebox, a variable matches a group of characters. By default, the variable vowel matches the letters *A a E e I i O o U u*. The section *Filters and variables*, p. 101, gives a step-by-step example of how to use variables in filters.

To use variables in a Filter or a Find, enclose them in square brackets, e.g., `[vowel]`.

In the Find dialog box, you must type the variable (enclosed in square brackets) by hand.

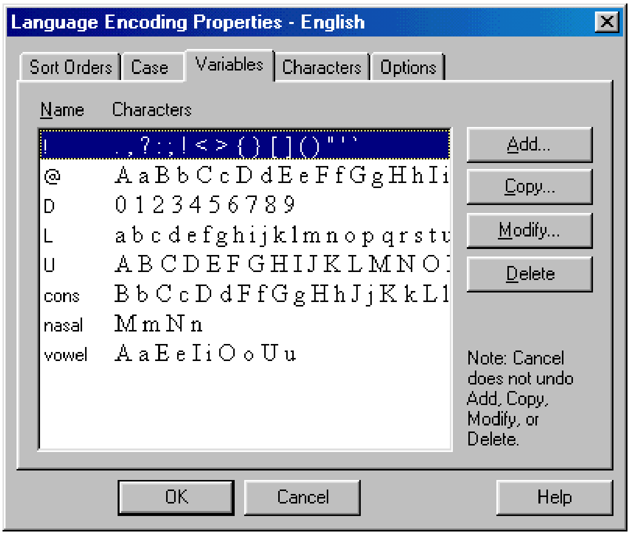
When building a filter in the Marker Text Filter Element and Language Text Filter Element dialog boxes, you can select a user-defined variable from the drop-down list and click Insert into the Text to insert it automatically. For an example of how to enter user-defined variables, see p. 101. However, to use built-in variables you must type them by hand.

The built-in variables are:

Variable	Description
[#]	**word boundary**—matches white space, any character from the [!] punctuation variable for the current field's language, or the beginning or end of the field.
[?]	**any character**—matches any single character, including space. If this language has multiple-character sorting units such as *gb*, they count as one unit when matching.
[…]	**zero or more characters**—matches any sequence of characters, including spaces.
	space—matches one or more of any white space (i.e., space, tab, or newline character). This does not require brackets.
[-]	**morpheme boundary**—matches hyphens and everything that the word boundary variable matches (if your morpheme break character is not a hyphen, this variable will not work as you would expect).

The match character options in the Find dialog box can significantly affect how these variables work in the Find command. Learn by experimenting.

All other variables are defined uniquely for each language encoding. This is because variables contain letters that must belong to a language of some kind. To view or modify the default variables Shoebox creates for each language, or to create new ones, go to the Language Encodings dialog box (on the Project menu), select the appropriate language, and click Modify.... This displays the Language Encoding Properties dialog box for that language. Click the Variables tab to view, modify, add, or delete the available variables as needed.



Name	Characters
!	. , ? : ; ! < > { } [ ] ( ) " ' `
@	A a B b C c D d E e F f G g H h I i
D	0 1 2 3 4 5 6 7 8 9
L	a b c d e f g h i j k l m n o p q r s t u
U	A B C D E F G H I J K L M N O
cons	B b C c D d F f G g H h J j K k L l
nasal	M m N n
vowel	A a E e I i O o U u

Note: Cancel does not undo Add, Copy, Modify, or Delete.

Fig. 256. Variables definition dialog box setup for English

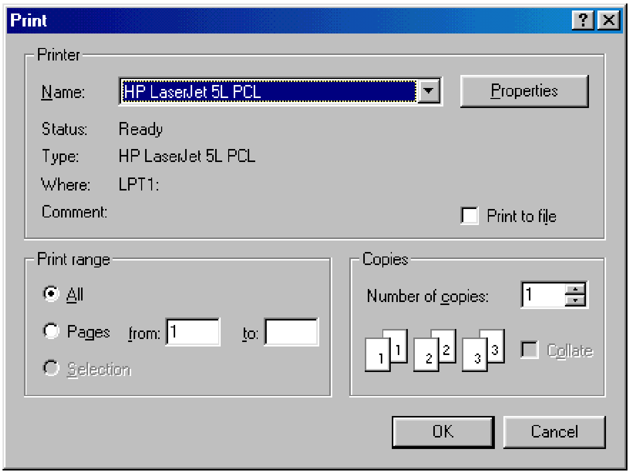
Printing files

To print your Shoebox data, begin by choosing between either of two approaches:

- Obtain a draft printout directly using the Print command.
- If you wish to have more control over formatting, export your data either as a Standard Format file or as a Rich Text Format (RTF) file and print it through a word processor like Microsoft Word. See *Exporting files*, p. 222, for more on this method.

Printing directly

The four commands related to printing directly from Shoebox are all on the File menu. These are: Print, Print Preview, Print Setup, and Page Setup.


Print

Printer

Name: HP LaserJet 5L PCL

Status: Ready

Type: HP LaserJet 5L PCL

Where: LPT1:

Comment:

□ Print to file

Print range

● All

○ Pages from: 1 to:

○ Selection

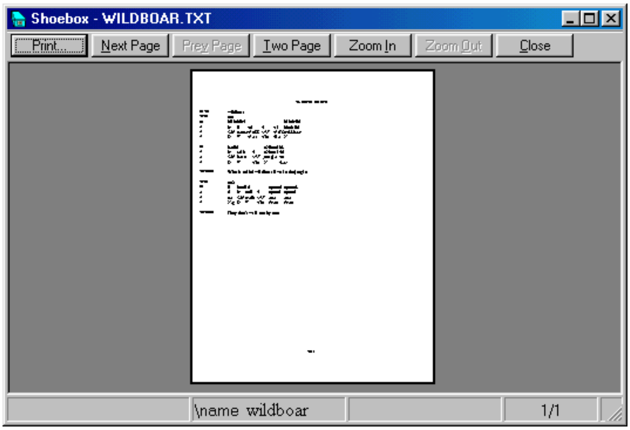
Copies

Number of copies: 1

□ Collate

Fig. 257. Print dialog box

The Print dialog box allows you to specify the number of copies and whether to print all pages or a selected range of pages. The concept of pages does not quite fit the Shoebox concept of records, but if you have short records and specify a page break before each record—via the Page Setup option—the page numbers and record numbers will line up fairly well.

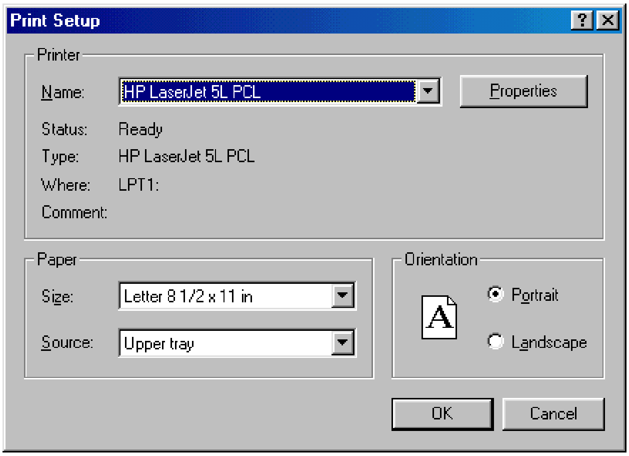

Shoebox - WILDBOAR.TXT

\name wildboar

1/1

Fig. 258. Print preview dialog box

The Print Preview dialog box is the standard print preview which shows one or two pages at a time and can be used to review page breaks.



Print Setup

Printer

Name: HP LaserJet 5L PCL

Status: Ready

Type: HP LaserJet 5L PCL

Where: LPT1:

Comment:

Paper

Size: Letter 8 1/2 x 11 in

Source: Upper tray

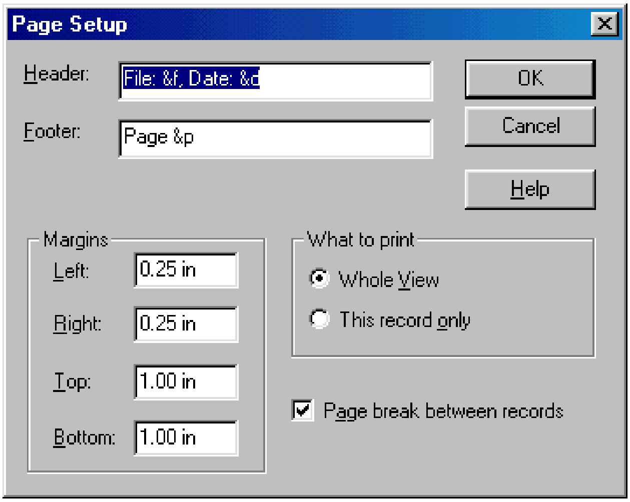
Orientation

● Portrait

○ Landscape

Fig. 259. Print Setup dialog box

The Print Setup dialog box allows specification of the printer as well as certain features that are printer specific such as page orientation, paper size, etc.



Header: File: &f, Date: &d

Footer: Page &p

Margins

Left: 0.25 in

Right: 0.25 in

Top: 1.00 in

Bottom: 1.00 in

What to print

● Whole View

○ This record only

☑ Page break between records

Fig. 260. Page Setup dialog box

The Page Setup dialog box allows specification of the format of the data on the page. You can also choose whether to print:

- all the records in the active window (If a filter is active, only the records that match it will be printed. If the data is sorted by a field which occurs multiple times in a record, that record will be printed once for each occurrence of that field.)
- the current record in the active window

Margins can be specified in either inches or centimeters.

If you turn off the option to start each record on a new page, records will be printed continuously with a gap between each.

Headers and footers can contain text, as well as a set of special printing codes which allow the inclusion of items such as the file name, date, record marker, etc. The printing codes are case-sensitive and must be preceded with an & (ampersand).

&F	file name (without the path)
&P	the number of the current page
&n	the current record number
&N	the number of records in the view being printed
&R	the record marker
&r	the record field contents
&K	the record key marker
&k	the record key field contents
&D	the current date in yy/mm/dd format
&T	the current time in 24-hour format, hh:mm
&L	the name of the current filter if any
&&	prints an ampersand (&)

As an example:

File &F Record &n/&N

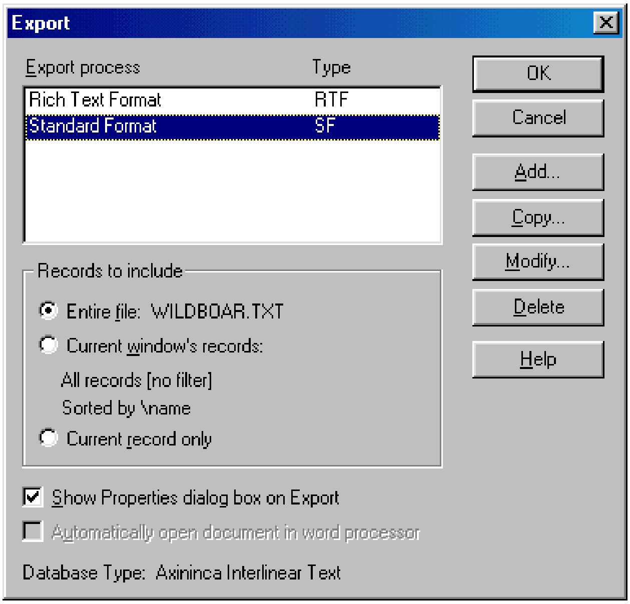
produces a header of this form:

File Shlex.db Record 8/57

Exporting files

The File Export feature of Shoebox enables you to format and print a database as a Rich Text Format (RTF) file using Microsoft Word (or any other RTF-aware program). But File Export is for more than just printing. It is also useful for exporting a Shoebox database as a pure Standard Format (SF) file.

Whether you export your database as RTF or SF, you can have it exported through a Consistent Changes Table in order to make systematic changes to the output file. You can also specify whether to export the entire database, only the records in the current window, or only the current record. You can select which fields to include in or exclude from the exported file. For RTF output, you can have Shoebox automatically attach a Microsoft Word Document Template to the file, and open the file in your word processor (Windows users only). You can even create special export configurations and save them for future use. To access these various features, select File, Export from the menu.



Export process	Type
Rich Text Format	RTF
Standard Format	SF

Records to include

● Entire file: WILDBOAR.TXT

○ Current window's records:

All records [no filter]
Sorted by \name

○ Current record only

☑ Show Properties dialog box on Export

□ Automatically open document in word processor

Database Type: Axininca Interlinear Text

Fig. 261. Export dialog box showing export processes

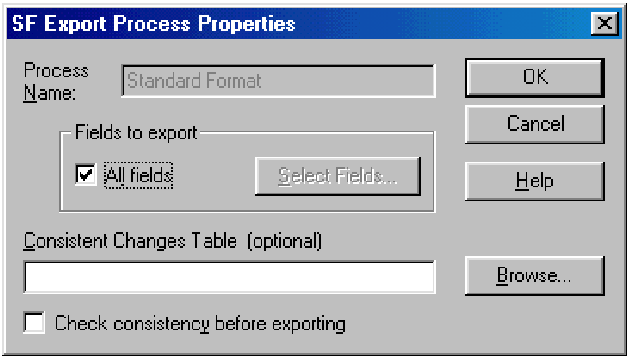
Basic File Export

There are two basic export processes available to every database file:[49]

- Standard Format (SF) export
- Rich Text Format (RTF) export

Regardless of which export option you choose, there are four basic steps to exporting a file:

1. Choose File, Export from the menu.
2. Select the export process you want (SF, RTF, or some other you have already set up) in the Export dialog box, and click OK.



SF Export Process Properties

Process Name: Standard Format

Fields to export

☑ All fields

Consistent Changes Table (optional)

☐ Check consistency before exporting

Fig. 262. SF Export Process Properties dialog box

3. Shoebox will then show you the appropriate Export Process Properties dialog box.[50] Here you can change these settings as needed for this export. Shoebox will remember them for the next time you use the export process. When ready, click OK.

[49] If a database is based on the MDF database type files, then an MDF export option should be available as well. Refer to the MDF documentation for a detailed description of how its export option works.

[50] If you want to skip this step and not see the properties dialog box, clear the Show Properties dialog box on Export before choosing OK in Step 2.

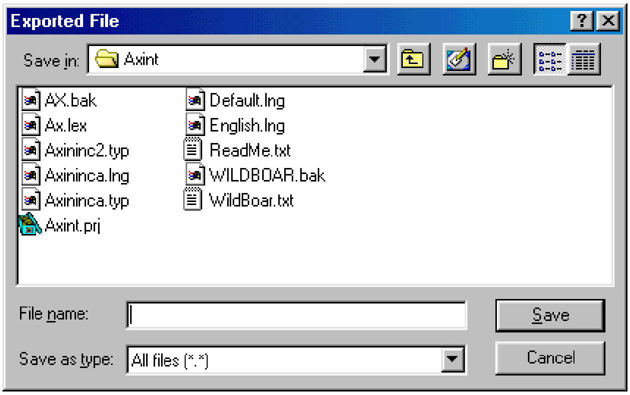


Fig. 263. Exported File dialog box open for typing in a file name

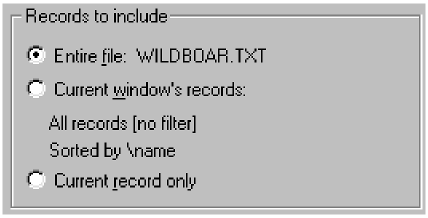
4. In the Exported File dialog box, enter a file name for the file that Shoebox will create, and click Save.

Shoebox then exports the file as requested. If it is to a Standard Format file, Shoebox writes the file to disk and returns to its main screen. If it is an RTF export process and you requested your word processor to be run automatically (Windows only), Shoebox will cause that word processor to open the new RTF file. Otherwise Shoebox simply exports the RTF file to disk and returns to its main screen.

If you did not select Automatically open document in word processor when you ran an RTF export, you will have to run your word processor and open the RTF file yourself in order to view it or print it.

Records to include

The Records to include options in the main Export dialog box gives you control over exactly how much of the database is exported and which records are selected. The three choices under this section are common to both SF and RTF export processes.



Records to include

● Entire file: WILDBOAR.TXT

○ Current window's records:

All records [no filter]

Sorted by \name

○ Current record only

Fig. 264. Records to include section of the Export dialog box

- Entire file ignores the current window's sorting and filtering properties and exports every record in the database, sorted in the basic record marker sort order.
- Current window's records exports all of the current window's records, sorted and filtered as you see them. If the primary sorting field occurs multiple times in a record, that record is exported multiple times.
- Current record outputs only the current record.

If you are exporting the current window's records and the current window is sorted by a marker other than the record marker, then before you are asked to give a file name for the output file, you will see an Export Current Window dialog box which gives you two more choices:

- Leave the fields in each record as they are in the current window—no reordering.
- Move the relevant sort field to the top of each record. Exclude other occurrences of the primary sort field in the record. This option is recommended if primary sort fields can occur multiple times in records.

Multiple export processes

File Export in Shoebox lets you name and define multiple export processes. This allows you to create export processes for specific tasks (e.g., a specific set of interlinear markers for a grammar publication). For each export process, you can specify what fields to include and a Consistent Changes table (if needed).

Shoebox has one SF export process and one RTF export process built-in. If needed, you can customize these or create your own.

To modify a process that is already defined:

1. Select the process and click Modify....
2. Make any changes you wish and click OK.

The changes you made will be saved for the next time you want to export something using that process.

To create a new process, either Copy one that is already there or Add a new one (in which you are given a choice between creating an SF or RTF export process).[51]

If you are sure you will not need one of your processes again, you can delete it.

You may modify the original SF and RTF processes as needed. However, if you delete them, Shoebox will create new ones when it reopens the current database type file. This ensures that at least the basic export options are always available.

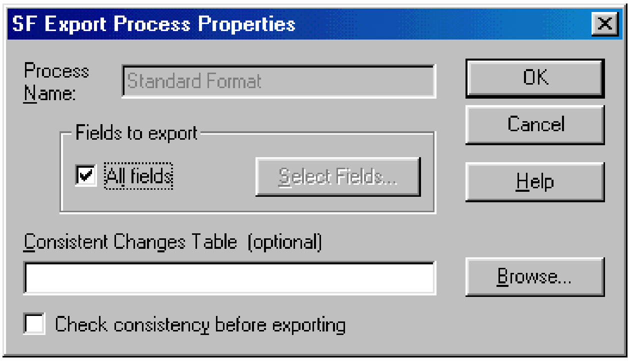
Whether you Add, Copy, or Modify a process, the resulting dialog box depends on whether you are editing an SF export process or an RTF export process.

Standard Format (SF) export

The Standard Format (SF) export process basically exports the file in plain Standard Format structure, without any special formatting or the Shoebox header field. This makes it easy to open the file using any program that recognizes SF.

[51] If the database is an MDF type, you will also have the choice of defining a new MDF export process.

The SF Export Process Properties dialog box gives the following options for the SF export process:


SF Export Process Properties

Process Name: Standard Format

Fields to export

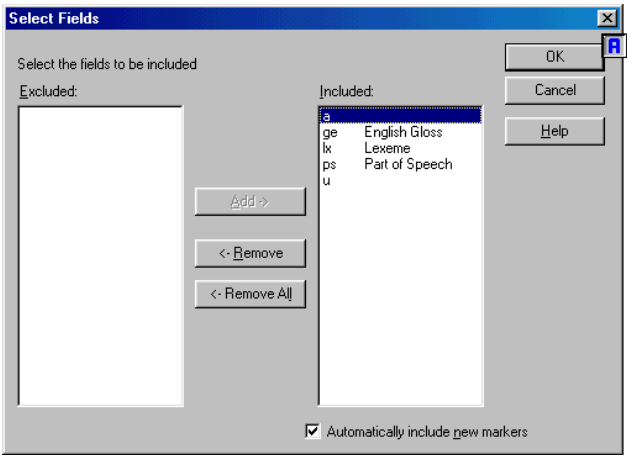
☑ All fields

Consistent Changes Table (optional)

☐ Check consistency before exporting

Fig. 265. SF Export Process Properties dialog box

Fields to export allows you to specify exactly which fields from the database are included and which are excluded from the output file. If you clear the All Fields check box, the Select Fields button will become available. If you click Select Fields..., the Select Fields dialog box will appear, allowing you to include and exclude specific fields in your export process.


Select Fields

Select the fields to be included

Excluded:

Included:

a	
ge	English Gloss
lx	Lexeme
ps	Part of Speech
u	

☑ Automatically include new markers

Fig. 266. Select Fields dialog box for selecting export fields

The Consistent Changes Table box allows you to specify a CC table to be used to modify the data as it is exported. For example, if you are exporting a file to open the data in a DOS program, you may want to modify the characters using the ANSI_IBM.cct table or some similar table which converts the Windows (ANSI) characters back to the DOS (ASCII) character set. (This table may need to be modified to account for your set of characters.)

Rich Text Format (RTF) export

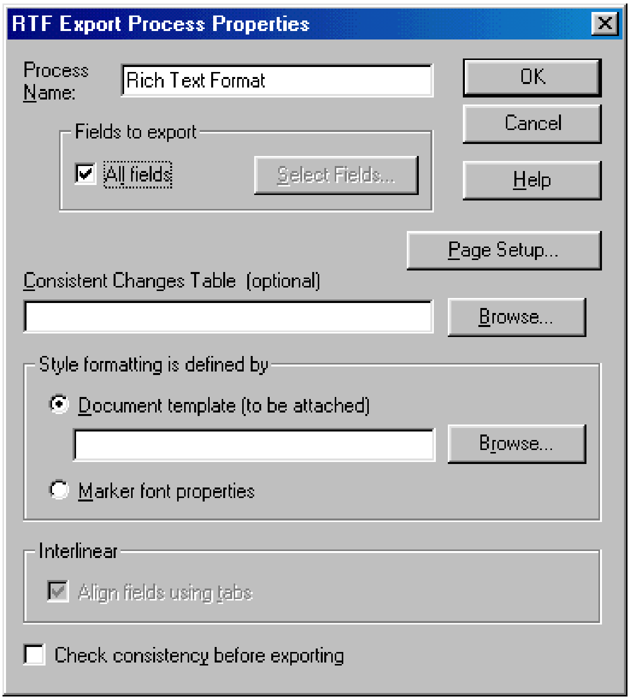
The Rich Text Format (RTF) export process uses the paragraph and character style specifications for each marker (defined in the Marker Properties dialog box) in your database to produce an RTF file which is readable by Microsoft Word and other RTF-aware programs.

To better understand the RTF process, notice the parallels between your Shoebox files and the resulting document files:

- The database file becomes a formatted document file.
- The database type file supplies the information needed to create a document template.
- The marker field names (defined in the Marker Properties dialog box) become the style names for every piece of text in the document.[52]

[52] Shoebox also automatically converts subfields to their appropriate character styles. Subfields are discussed briefly in the answer to a question about embedding fonts in another field, p. 240.

The RTF Export Process Properties dialog box gives the following options for the RTF export process:



RTF Export Process Properties

Process Name: Rich Text Format

Fields to export

☑ All fields

Consistent Changes Table (optional)

Style formatting is defined by

● Document template (to be attached)

○ Marker font properties

Interlinear

☑ Align fields using tabs

☐ Check consistency before exporting

Fig. 267. RTF Export Process Properties dialog box

Fields to export allows you to specify exactly which fields from the database are included and which are excluded from the output file.

The Consistent Changes Table box allows you to specify a CC table to be used to modify the data as it is exported. This CC process occurs before the file is converted to RTF, so it is useful if you need to combine fields, reorder fields, or make some other global change that you want reflected in the output.

The Style formatting is defined by box allows you to use the marker properties as defined in Shoebox or to specify a Microsoft Word document template (.dot) file.

If you are setting up an RTF export process for the first time, export the file using Marker font properties. The exported file contains styles whose names correspond to the field names of all your markers. To create a document template,

1. Open the RTF file in Microsoft Word.
2. Delete all the text in the document (so it is blank).
3. Modify the formatting of any of the styles as needed.
4. Save it as a document template.

Word will create a .dot file that you can then specify in the above export option. This allows the fonts and font styles used in an exported document to be different from those used in Shoebox.

The Page Setup button allows you to specify paper size, margins, and the number of columns for the exported document.

Other RTF export information

On the Database Properties Markers tab, the rightmost column in the marker list (with the heading SFR for Style-Font-Range set) contains one or more letters which indicate some of the attributes about each marker. A marker with

- P—will be assigned to a paragraph style when exported.
- C—will be assigned to a character style when exported.
- F—has a font chosen (i.e., *not* using the language font).
- R—has a range set defined.

These can be changed for each marker through the Marker Properties dialog box. The paragraph and character style choices are in the radio button group labeled Style to Export. The paragraph choice is the default for newly added markers.

Frequently asked questions

Memory requirements

Q: How much memory do I need to open my database files in Shoebox?

A: To determine approximately how much random access memory (RAM) your data and the Shoebox program will need,

1. Add the sizes of your database files in KB. (If sizes are expressed in bytes, divide by 1000. If they are expressed in MB, multiply by 1000.)
2. Multiply that size by
   - 2.75 on Windows
   - 3.5 on the Macintosh
3. Add the amount of memory the Shoebox program itself takes:
   - 1100 KB on Windows 3.1
   - 1600 KB on Windows 95 and later
   - 2200 KB on the Macintosh
   - 1000 KB on the Power Macintosh[53]

The actual amount of installed RAM you will need to have in your computer will depend on the memory requirements of the operating system and other applications you wish to run simultaneously with Shoebox. As a general rule, a computer that runs Microsoft Office comfortably will be able to run Shoebox with a typical set of databases.

[53] If your Power Macintosh does not have virtual memory turned on, Shoebox will take about 3900 KB.

If, based on the above information, you think that your computer has enough memory to open your data in Shoebox, but Shoebox is running very slowly, read the answer to the next question.

Q: Why is Shoebox running so slowly? Why does it take so long to open databases?

A: Most operations in Shoebox will be quite fast as long as the program has access to enough random access memory (RAM). When Shoebox cannot load all the data into physical RAM, it will begin to access virtual memory. When this happens, many operations will be noticeably slower. When Shoebox opens a database, it builds one or more indexes. Since this is a memory-intensive operation, opening databases is very slow if your computer does not have enough RAM available to Shoebox.

Following are some possible remedies if Shoebox does not have access to enough memory to keep all your data in physical RAM:

- Install additional RAM in your computer.
- Close other applications.[54]
- Reduce the amount of memory used for disk caching.
- Allocate additional memory to Shoebox (Macintosh only).
- Avoid unnecessary sorting and filtering.
- Open fewer databases.[55]

Macintosh

On the Macintosh, every application is assigned a minimum and recommended (maximum) memory allocation. Even if you have adequate RAM installed in your computer, a program cannot use more than the amount allocated to it. The Shoebox Installer initially sets up the Shoebox application to use up to 8 MB of RAM (10 MB on the Power Macintosh). If the memory allocated

[54] On the Macintosh, you might also consider turning off unneeded extensions.

[55] With some care, it may be possible to split an existing database and work on the parts separately. This is likely only a temporary solution, but you might consider it until you are able to upgrade.

to Shoebox is insufficient, Shoebox will begin to run slowly and might fail. To confirm that a performance problem is the result of inadequate memory allocation,

1. Start Shoebox and open your databases.
2. Switch to the Finder.
3. On the Apple menu, choose About this Macintosh (or About this Computer).

A bar graph shows how much memory is allocated to Shoebox, and the dark portion of the bar shows how much of that is currently used by the program and data. If the dark portion of the bar fills all or nearly all of the bar, you should consider increasing the memory allocated to Shoebox.

If you have enough memory for Shoebox and any other programs you need to run simultaneously with Shoebox, you can increase the memory allocated to Shoebox to prevent it from having to use virtual memory. To increase the RAM allocation,

4. Quit Shoebox if it is running.
5. Select the Shoebox application file in the Finder.
6. Choose File, Get Info (System 8.5 and later: File, Get Info, Memory).
7. In the Info dialog box, increase the number in the lower right corner. (It will be labeled Preferred size or Current size, depending on your operating system version.)

Note that if you allocate all your memory to Shoebox itself, Shoebox Help (and other programs) will not have enough memory to run.

Windows 95 On Windows 95 and later, to find out how much physical memory your computer has available,

1. Choose Start, Settings, Control Panel.
2. In the Control Panel, choose Help, About Windows 95.

Physical memory available to Windows is the amount of memory Windows can use. If you calculated (see answer to previous question) that the amount of memory needed to open your databases in Shoebox is anywhere near the physical memory available, you will not be able to open those databases simultaneously without accessing virtual memory. However, if it seems you have enough RAM but Shoebox still runs slowly, you might be able to keep it from accessing virtual memory by closing other applications or by reducing the size of the disk cache.

By default, Windows will often use 25% of your RAM as a disk cache. If you have 32 MB of RAM or more, try setting the cache size at a fixed number. We recommend 4 MB for systems with less than 48 MB of RAM, and 8 MB for those with 48 MB or more. Here is how to set it up:

3. Run SysEdit. (Choose Run from the Start menu, type `SysEdit` in the text box, and press ENTER.)
4. Click in the system.ini window.
5. Scroll down to the [vcache] section. If this section heading does not exist, type `[vcache]` as a new section heading.
6. If you have 48 MB of RAM or less, enter these two lines:

   ```
   MinFileCache=0
   MaxFileCache=4096
   ```

7. If you have more than 48 MB of RAM, enter these two lines:

   ```
   MinFileCache=0
   MaxFileCache=8192
   ```

8. Choose File, Save.
9. Choose File, Exit.
10. Restart Windows.

Windows 3.1 To find out how much physical memory your computer has available, in Program Manager, choose Help, About Program Manager. Near the bottom of the About box is a number labeled Memory.

Physical memory available to Windows is the amount of memory Windows can use. If you calculated (see answer to previous question) that the amount of memory needed to open your databases in Shoebox is anywhere near the physical memory available, you will not be able to open those databases simultaneously without accessing virtual memory. However, if it seems you have enough RAM but Shoebox still runs slowly, you might be able to keep it from accessing virtual memory by closing other applications or by reducing the size of the disk cache. To reduce the size of the disk cache:

1. Run Control Panel.
2. Double-click Enhanced.
3. Click Virtual Memory.
4. Click Change.
5. Decrease the size of the disk cache to the desired level. See the numbers in 6 and 7 on the previous page.
6. Click OK two times.
7. Restart your computer.

Getting started with Shoebox

Q: Isn’t there an easier way to start a new database or to import an existing one? Do I really have to deal with database types and language encodings?

A: There is an easier approach, but the results will lack some of the power that could be available to you as a user. However, the “quick-and-dirty” approach can be gradually cleaned up as you see the need for the different features.

The key points are:

- You cannot avoid the reality that files of the same database type must have the same record marker and different databases often have different record markers and fields, so you must create some different database types. However, you do not have to specify anything more than the record marker. If you are importing an existing database, you do not have to fill in the marker list (the program will do that automatically), type field names, nor specify the hierarchy.

- You can also leave the language for each marker as Default. This will give everything the same font, sorting order, etc.

If you need to specify a special sorting order and case relationships, you can modify the default language encoding. On the other hand, you could acknowledge that this *really is* a different language, define it as such, and so gain the benefits of it. You might have a font to consider as well—it is more efficient to assign the language a font than to assign a special font to each field.

Q: When I am creating a new database type for an existing file, must I fill in all the markers? What if I can't remember some of them?

A: No, you need not remember anything more than the record marker. The program will automatically generate the list of markers when it imports the file.

Field markers

Q: What is MDF?

A: The Multi-Dictionary Formatter (MDF) is a system for developing, exporting, and printing a multilingual dictionary and reversed glossaries. The capability of the original version has been enhanced and fully integrated into Shoebox and so is readily available to all Shoebox users.

But while MDF is a helpful lexical system for many languages, it has its limitations and is not the solution for all languages or all types of lexicons. Refer to the MDF documentation to see how it works and whether it will work for you.

Q: Do I have to use the MDF markers? I already have my own set.

A: The MDF database marker set is optional. If you already have a working set of lexical markers, certainly use your own set.

Q: I want to delete a marker in the marker list that is not used in my database, but Shoebox won't let me. How can I delete it?

A: Even if a marker does not occur in the data, it cannot be deleted if certain settings refer to it.

- Choose Database, Sorting to see if the marker is in the Sort Fields list.
- Choose Database, Properties, Interlinear tab. Click Modify for each line, then click Lexicons button. See if the marker is the Marker to Output or in the Markers to Find list.

For advanced users, perhaps the quickest way to find where the marker is referenced is to exit Shoebox and to examine the database type file for that database with a word processor or text editor. Search through the .typ file to find all references to the marker you want to delete. Do *not* modify the file directly; instead note where the marker is referenced, and remove the references using Shoebox.

Q: How can I locate records containing a particular marker?

A: Here are three possible approaches:

- Use the Find command. Under Find In, select the Field option and select the marker. Leave the Text to Find box empty. (Shoebox will automatically insert […] as the text to be found.)
- Use a Marker Text filter (Database, Properties, Filters tab). Leave the Text box empty. Shoebox will display all records containing that marker.
- Sort by the marker (Database, Sorting) and Browse (View, Browse Fields) on the field. Shoebox will display all occurrences of the marker. Records which do not contain the marker will be labeled **no field**. Records which contain the marker but have no data in the field will be labeled **empty**.

Q: How can I make the TextPrep.cct table insert the desired markers when I use it to import text files?

A: There are comments in the TextPrep.cct file itself which describe what should be changed. Use any word processor or text editor to edit the table.

Fonts

Q: Why do my overstriking diacritics not show up on the Language Encoding Characters page? Why do they overlap where I typed them as secondary characters in the sort order?

A: These issues might affect users who have to enter solitary overstriking diacritics into any of the edit boxes on the Sort Order Properties or Variables Properties dialog boxes or on the Case tab. Using some fonts:

- a diacritic at the beginning of a line is not visible.
- diacritics separated by a single space overlap in a confusing clump.

The quickest remedy for the first problem seems to be for the diacritic character to be preceded by spaces. For some fonts, one space is not enough, but two spaces is usually adequate.

The second problem can be solved by increasing the number of spaces used to separate the diacritic characters. To avoid clumping, enter at least two spaces before any overstriking diacritic when it is used out of a normal textual context (as in the Language Encoding Properties dialog box).

Q: How can I insert data that is in another font—for example, a Thai word—into an English field?

A: That is a problem for which there is no good solution. There are two possible workarounds.

Those working in non-Roman scripts may prefer a solution which will show the script. This involves interrupting the field in which you are working with another marker dedicated specifically to that language. For example, a comment on the etymology of a word would almost surely cite other forms. In that case, the **\ec** field (or whatever you use) might be interrupted by something like **\vw** for vernacular word and then continue with a continuation **\ecc** (e.g., for "etymology comment continued") immediately afterward:

```
\ec  xxxxxx xxx
\vw  yyy
\ecc xxxxxx xxxx …
```

The markers **\vw** and **\ecc** should both be marked as character styles in the Marker Properties dialog box. When they are exported as Rich Text Format (RTF) files, they will not start new paragraphs.

This is an awkward solution but does allow you to see the text using the appropriate script in Shoebox.

A better solution for exporting is to use subfield markers, which are converted to character styles by RTF export processes. The notation for a subfield is **|fv{***data in another language***}**, where **fv** is a marker and *data in another language* is the content of the subfield. The vertical bar and curly braces separate the subfield from the enclosing data field.

You add subfield markers in the Database Type Properties dialog box (as you do any other field markers). To indicate that a marker can be used in subfields and exported as a character style, choose Character as the Style to Export in the Marker Properties dialog box.

For example, in a Multi-Dictionary Formatter database there is a marker **fv** for vernacular language subfields. Its field name is f_vernacular, its language encoding is vernacular, and its style to export is character. If you were to embed the vernacular phrase *ana khwofw matamkwe* in a data field that contains English language data, it would look like this:

```
\ue  The phrase |fv{ana khwofw matamkwe} "I'm going to
     slap your face" is considered extremely harsh.
     Frequently this will result in bloodshed.
```

You can export the preceding data as an RTF file. When you open the exported document in Microsoft Word, the subfield text will have the f_vernacular character style. The name of the character style is determined by the field name of the corresponding marker. If you were to attach a document template defining f_vernacular to use bold font style, your data would appear something like the following:

> The phrase **ana khwofw matamkwe** "I'm going to slap your face" is considered extremely harsh. Frequently this will result in bloodshed.

If the various languages you must work with can be written using the same or similar fonts, there is no problem reading what you have enclosed in the brackets (as in the last example). But if the encompassing field uses a font significantly different from the font you need to embed, you may have to go to another field that uses this font, type the word there, cut the word, and paste it where you want it. The pasted words will be displayed using the font of the field where you pasted them and therefore not look right on the screen in Shoebox. However, the underlying character codes will be correct, and the text will be formatted properly with the RTF export option. In such a case, the solution proposed earlier might be preferable.

Database types

Q: How can I change the name of a database type?

A: You cannot rename a database type. You can copy a database type and give it a different name. If you *already* have databases assigned to the original database type, you must deal with them also. They will not be automatically changed to use the new database type.

1. Close any databases of that type that are open.
2. Choose Project, Database Types.
3. In the list, select the database type which you want to change.
4. Click Copy.
5. Specify the name you intended, and click OK.
6. Return to the list of database types by clicking OK.

You now have two database types which are identical except for the name.

7. Select the original database type and click Delete.
8. Click Close.
9. Reopen the databases (File, Open). For each file you open, Shoebox will display a message warning you that the database type file could not be found.
   a. Click Import as a Different Type.
   b. Click OK to close the warning regarding import.
   c. Select the new database type.
   d. Click OK.

Q: How can I change the database type of a database file?

A: There is no way to change the database type of a file while it is open in Shoebox. To change the database type, you must remove the information at the head of a database file that assigns it to a type. Then you can reopen it and select a new type.

It is strongly recommended that you back up your data before attempting this.

Here are two approaches to removing the part of the file that contains the reference to the database type:

- Export (File, Export) the database as an SF file (entire file, all fields). If you wish to keep the same filename, close the database in Shoebox and then rename the exported file.
- Close the database in Shoebox. In a text editor, open the database. Delete the first line of the database (the line marked **_sh**). Save the database as text.

Once you have removed the database file header, follow these steps to reopen the database and assign it to a new database type:

1. Open the modified database (File, Open).

2. On the Import dialog box, select the desired database type.
3. Click OK.

Miscellaneous questions

Q: Does Jump always have to open a new window?

A: No. For more information, search for Jump Target in Shoebox Help. Also, see the *Jumping* section, p. 82.

Q: Is there an automatic save option?

A: No. To prevent losing data, try to remember to save periodically, perhaps after every five new records. The File, Save All command saves all modified databases plus the settings files. The File, Save command saves the current database.

Q: How can I make a shortcut icon to access my Shoebox project?

A: The Shoebox installation program associates all project files with the Shoebox program. If Shoebox is not already running, double-clicking a project file runs Shoebox and opens that project file. If you wish to avoid having to navigate to project files, you can create icons which are shortcuts to your project files and put them in a convenient location. Double-clicking a shortcut icon will have the same effect as double-clicking the project file itself.

Windows 95

To create a shortcut on Windows 95 and later,

1. Using Windows Explorer, locate and select your project file (.prj).
2. Choose File, Create Shortcut.

A shortcut icon will appear in the folder where your project file is. It will be selected.

3. Choose File, Properties.
4. Click the Shortcut tab.
5. In the Start in box, type the path of the folder containing your data. (This is the folder you will see when you choose File, Open in Shoebox.)

6. Click OK.
7. Drag the shortcut icon to your desktop or to another convenient location.

If you want to revise the name of the shortcut you have just created,

8. If it is not already selected, click the shortcut icon to select it.
9. Choose File, Rename.
10. Modify the text appropriately for this project and press ENTER.

Windows 3.1 To create an icon on Windows 3.1,

1. In Program Manager, open the Shoebox program group (or whatever group you want to copy the project icon into).
2. Run File Manager and arrange it so you can see both it and the program group you want to put the project icon into.
3. In File Manager, find your project file (.prj) and drag it into your program group in Program Manager.
4. Click the project icon and choose File, Properties.

Shoebox will display the Program Item Properties dialog box.

5. Enter a description appropriate for the project.
6. Make sure the Command Line includes the full path and name of your project file (e.g., C:\Myshset\Myproj.prj).
7. In the Working Directory box, enter the directory containing your data. (This is the directory you will see when you choose File, Open in Shoebox.)
8. Click OK.

Macintosh To create an alias on the Macintosh,

1. In the Finder, locate and select your project file.
2. Choose File, Make Alias.

An alias icon will appear in the folder where your project file is.

3. Drag the alias icon to your desktop or to another convenient location.

If you want to revise the name of the alias you have just created,

4. If it is not already selected, click the alias icon to select it.
5. Click the text label of the icon. (The entire text within the label box will be highlighted.)
6. Modify the text appropriately for this project and press RETURN.

Notes on interlinearization

Whole-word parses

Underlying forms can show how whole words are to be parsed. For example:

```
\lx   reruns
\u    re- run -s
```

For a morpheme to be recognized as a prefix, it must have a space *after* the hyphen. Similarly, the suffix must have a space *before* the hyphen.

Longest-match principle

In Shoebox, the parser tries to make longer pieces take precedence over shorter ones. This means that if you import a set of conjoined affixes from Shoebox for MS-DOS, the affix combinations in your list should take precedence over the smaller pieces. For example, on the word *linguistically,* if *-ically* is in the lexicon, it will take precedence over the separate pieces *-ic, -al,* and *-ly* in the lexicon. In this way *-ically* could have different morpheme breaks or different forced glosses than the smaller pieces. Similarly, if you have *linguistic* in the lexicon, it will take precedence over the pieces *linguist* and *-ic* in the lexicon.

However, parsing is done from the outside of the word in, so if there are multiple possible ways a word can be parsed, the longest affix can sometimes win over the longest root. For example, if the lexicon contains all the morphemes just mentioned, the longer affix *-ically* will combine with the shorter root *linguist* instead of the shorter affixes *-al* and *-ly* combining with the longer root *linguistic*. The result might be the same, since the morpheme breaks from *-ically* may be the same as those produced by the shorter pieces. But if there is a distinction in cuts or in forced glosses, you might not get the result you expect. You can work around this by adding whole words or longer combinations of root and affixes to the lexicon as necessary.

The longest-match principle can incorrectly limit valid ambiguities. An example of this is the word *does*, which can be parsed as either *doe -s* or *do -s*. Shoebox will always select the latter parse because *-es* (an alternate form of *-s*) is the longer affix. This means Shoebox will never let you select *doe -s* as the parse, even when it is correct.

There are several ways to solve this problem. Choose the one which best suits you. The first way is to express the ambiguity explicitly in one of your lexical entries. For example, for the lexeme *do* you could add these lines:

```
\lx  do
\a   does
\u   do -s
\u   doe -s
```

This forces Shoebox to give two possible parsing choices for *does*, and it also clearly shows you the ambiguity (one alternate form comes from two underlying forms). But notice that this method makes reference to *doe*, which has no relation to the lexeme *do*. If this seems odd to you, you could split this up and put one explicit parse in each lexeme, *do* and *doe*:

```
\lx  do
\a   does
\u   do -s

\lx  doe
\a   does
\u   doe -s
```

This has the advantage that each relevant parsing is put with its lexeme, but has the disadvantage that you no longer can see the ambiguity (i.e., it is not clear when looking at the *doe* entry why *doe -s* needs to be specified). The danger is that at some later time, you might see this, think you do not need it, and delete it.

The third solution is to create another lexical database (of the same type as your main lexicon), and use this new lexical database as a storage place for all of these ambiguous "whole word" parses. In this new database, you would add an entry like this:

```
\lx  does
\u   do -s
\u   doe -s
```

You would then add this database to the search path for the Parsing process. (To do this, go to Database, Properties, Interlinear tab. Select the parsing process and click Modify. Then click the Lexicons button. Select the database file you just created and click Insert. Click OK repeatedly to close all dialog boxes.) Now if Shoebox were to interlinearize the word *does*, it would find this lexeme in your new (ambiguities) database and give you the choice between the two possibilities.

The advantage of this method is that it keeps your lexicon free of the bookkeeping details needed only for getting Shoebox to parse properly. The disadvantage is that it requires a bit more work to set up. If you use this method, remember to include this ambiguities database when you back up your data.

Another problem with the longest-match principle is that sometimes it can cause Shoebox to prefer an invalid parse over the valid parse. The primary way to solve this is to explicitly provide combinations of morphemes that are longer than the invalid morphemes. This can be done for parts of words (affix combinations) as well as for whole words.

For example, if a language has affixes like: *a-, ab-, bc-,* and *c-*, words that look like *abcroot* could be parsed in the following two ways:

ab- c- root
a- bc- root

Because of the longest-match principle, Shoebox will always choose the first one. But if in this language the combination *ab- c-*

is meaningless, you would need to create the following lexical entry to force Shoebox to select the correct parse:

```
\lx   abc-
\u    a- bc-
```

In this way, when Shoebox sees an affix combination like *abc*-, it will know that the only valid parse is *a- bc-*.

Like the *does* example given earlier, you might want to keep entries such as these in a separate database (possibly with the whole-word combinations).

Word formulas

When Shoebox produces both valid and invalid parses for a word because of limitations in its basic morphological parsing features, it can use word formulas to select the valid parses automatically. Using word formulas can reduce the number of ambiguities in parsing.

A common problem in Shoebox is when a word is parsed into an invalid sequence of morphemes (especially involving wrong homographs). You can solve it by writing formulas that define valid sequences of morphemes (usually in terms of part of speech categories or glosses in lexical databases).

Word formulas can eliminate invalid parses, but they do not enable any additional valid parses. Note that Shoebox does not have sentence formulas that would use a word's surrounding context to resolve ambiguities in its meaning when there are multiple valid parses.

The formulas are kept with the rest of the Parse properties in the database type for interlinear texts. In the Parse dialog box, select Enable word formulas. Click Formulas to access the Word Formulas dialog box. From there, click Help to learn more. There is a sample project in the WordForm folder. (In the folder in which Shoebox was installed you will find the folder called Samples. WordForm is in the Samples folder.)

Forced values

A forced value is when the content or value of the next line (e.g., the gloss line) is overtly specified earlier in the interlinear process (rather than found by a typical lookup process). When a forced value is given, the interlinearizer does not check for other possibilities but is forced to use the value specified. But a forced value is checked by the next process to be sure that it is valid. If it is not found in the lexicon, a message to that effect is given, and a failure mark is shown instead of the forced value.

Combinations of morphemes can use forced values to further reduce ambiguity. If the underlying form of a morpheme is ambiguous, then a forced value can be used to select the correct value for that morpheme in the current word.

For example, the suffix *-s* in English is ambiguous since it marks both *PL* (plural) and *3s* (third-person singular subject verb suffix). But in the combination *-tion -s*, it is always plural.[56] To reduce the number of times you have to disambiguate the *-s,* you could create the following entry:

```
\lx   -tions
\u    -tion –s{PL}
```

This tells Shoebox to do two things:

- Parse *-tions* as the suffix *-tion* followed by the suffix *-s*
- Look up the entry for the suffix *-s* that has the gloss *PL*.

The *{PL}* code is the forced value. The default characters for marking forced values are curly braces { }, but these can be changed in the Forced gloss start and end text boxes in the Parse dialog box.

Multiple forced values can be given. For example, a forced part of speech could be added after a forced gloss if desired:

```
\lx   -tions
\u    -tion –s{PL}{Nsuf}
```

[56] In verbs like *rations,* the *-s* marks *3s,* but since the root is *ration*, not *ra,* Shoebox will not parse this with the suffix *-tion*. Rather it will parse it *ration -s*, and the *-s* will need to be disambiguated. The *-tions* forced-gloss solution also does not help disambiguate words like *nations*.

In this example, *PL* is used on the line immediately under the parsed line, and *Nsuf* (noun suffix) is used on the line following that. If the gloss line were below the part of speech line in your interlinear setup, the two forced values would have to be reversed.

Marker hierarchy can help eliminate the need for most of your multiple forced values. This is because once the first value has been specified, it restricts the possible choices for the second value to only those choices that are in the same hierarchical section of the entry. Do not bother with multiple forced values unless the first one still leaves the second one ambiguous.

The marker hierarchy helps most if you set up the interlinear process with the gloss line *above* the part of speech line, as Quick Setup does. (This may be the opposite of what you have seen in some publications.) The reason this works is that a gloss is much more restrictive than a part of speech. For example, it is likely there is only one *PL* in your lexicon, but several *Nsuf* forms. Once you have picked *PL*, Shoebox knows it is an *Nsuf*; that is not necessarily so the other way around.

Reduplication

Reduplication is the repetition of a word, morpheme, or syllable to produce an inflectional or derivational form. For example, many languages use reduplication to form noun plurals (e.g., in Indonesian *buku* means *book* and *buku-buku* means *books*).

In Shoebox, reduplication is represented by a special entry in the lexicon.[57]

- The field used for the underlying form (in the following examples , this is **\lx**) must contain a label which includes *dup*, *Dup*, or *DUP*. When Shoebox interlinearizes a word with reduplication, it places this label on the morpheme line under the reduplicated part.
- The field normally used for the surface form must contain one or more variables (see *Variables*, p. 216).

[57] You might not want reduplication to appear as an entry in a printed dictionary. If you prefer, you can put these reduplication entries in a separate lexical database and include it as a database to be searched in the interlinear setup.

Shoebox disregards hyphens (and all other morpheme break characters) when it matches reduplication. Therefore, it can match either *bukubuku* or *buku-buku*.

Shoebox checks for reduplication before starting to parse each word and at every parsing step. Therefore, Shoebox can match reduplication even in words that contain morphophonemic changes, as long as you have accounted for the changes.

General reduplication

In Shoebox, general reduplication refers to exact repetition of one or more whole morphemes. It is represented by the […] variable. There are three kinds of general reduplication in Shoebox:

- General prefix reduplication
- General suffix reduplication
- Full reduplication

Prefix and suffix reduplication produce a single word, with a repeated part that is a prefix or a suffix. For example, in Selaru, *wer* means *water* and *werwer* means *wet*.[58]

General prefix reduplication is expressed in Shoebox as:

```
\lx   Dup-
\a    […]-
```

With this entry, Shoebox would parse *werwer* as:

Dup- wer

General suffix reduplication is expressed in Shoebox as:

```
\lx   -Dup
\a    -[…]
```

With this entry, Shoebox would parse *werwer* as:

wer -Dup

[58] As in this example, if you cannot tell whether the repeated part is a prefix or a suffix, choose one analysis until additional data makes it clear.

Full reduplication produces two words. It is expressed in Shoebox as:

```
\lx  Dup
\a   […]
```

Template reduplication

In Shoebox, template reduplication refers to reduplication that matches a particular sound pattern. It is represented by user-defined variables. The repeated part does not have to be a whole morpheme. Template reduplication produces a single word, with a repeated part that is a prefix or a suffix. For example, in Selaru, the verb root *oban* means *hit* and the noun *oboban* means *hitting stick*.

Template prefix reduplication of a vowel-consonant pattern is expressed in Shoebox as:

```
\lx  VCdup-
\a   [vowel][cons]-
```

With this entry, Shoebox would parse *oboban* as:

Dup- oban

Template reduplication patterns can contain fixed letters:

```
\lx  aCdup-
\a   a[cons]-
```

With this entry, Shoebox would parse _*aboban_ as:

Dup- oban

Here are the rules for matching template reduplication:

- Each variable must match a letter in the repeated part and the same letter in the main part of the word. For example, in *oboban*, [vowel] matches *o* and [cons] matches *b*, both in the repeated part *ob*- and in the verb root *oban*.
- Each fixed letter must occur in the repeated part but matches any letter in the corresponding position in the main part of the word. For example, in _*aboban_, *a* occurs in the repeated part *ab*-, but *o* occurs in the verb root *oban*.

Infixes

An infix is an affix that occurs inside a stem, not at its boundaries. Infixes are marked in Shoebox with a hyphen on both sides of the form. An example from Tagalog is:

```
\lx  -um-
```

Infixes can be found anywhere in the word. They can be found before or after other affixes are removed, but the program tries to remove other affixes before it searches for infixes. For example, if the root is *bilih* ‘buy’:

```
\t   bumilih
\m   -um- bilih

\t   bumibilih
\m   -um- Dup- bilih
```

Notice in the second example the *-um-* infix was inside the reduplicated *bi-* portion of the stem *bibilih*.

In the parsed field, infixes appear at either the beginning or end of the root or stem in which they are found, depending on which option you choose.

Note: Often people call a suffix or a prefix an “infix” if it comes closer to the root than subsequent suffixes or prefixes, but to Shoebox these are not infixes. For example, in the Indonesian word *mempersilakan* ‘invite’ is parsed as *mem- per- sila -kan.* Notice that the *mem-* and *per-* prefixes peel off like layers of an onion. Because the *per-* is not inside the root *sila*, it is not considered an infix.

Compound stems

Compound stems are stems containing more than one root. By default, Shoebox can parse compound stems. To turn this option off, clear the Allow multiple roots check box in the Parse dialog box.

In English, the compound stem *blackbird* can be parsed as *black* - *bird*. On the morpheme line, Shoebox separates the roots with a hyphen. For example:

```
\t   blackbird
\m   black - bird
```

This capability will sometimes cause words that are not in the lexicon to be parsed incorrectly. For example, *afar* would parse as *a - far* if *afar* is not in the lexicon. If *afar* is in the lexicon, it will be parsed correctly because of the longest-match principle.

Hyphens in the lexicon

Some people put hyphens on roots in their lexicon. They use the hyphens to show that the roots are bound forms that must always have an affix added. For example, in the sample Selaru lexical database in the Shlex project, all verbs have hyphens in front of them. An example is *-abuk,* which means 'call.' But the hyphen makes Shoebox see the form as a suffix. It will not parse a word that appears to be made of all affixes.

To deal with this, add an **\a** field to each root with a hyphen showing the root without the hyphen. You can sort by part of speech to group the ones that need the extra field; then make a copy of the field and remove the hyphen from the copy in the **\a** field.

The parser recognizes prefixes, suffixes, and infixes by the presence of the appropriate hyphens or other morpheme break characters. You must put the appropriate hyphens on all affixes for parsing to work correctly.

If you have left hyphens off some of your affixes, you can sort your lexicon by part of speech to group the affixes together. Then add a hyphen to each affix. If you have many affixes to correct and have a macro recorder available, you can use the macro recorder to add the hyphen and move to the next record.

Morpheme break characters

Different morpheme break characters can be used on different types of affixes. For example, you can mark derivational affixes with an equal sign and inflectional affixes with a hyphen. Because morpheme break characters are carried down unchanged from the parse line into the lower lines, this can make it easier to distinguish between the different types of affixes:[59]

```
\t   relations
\m   relate  =tion  -s
\p   V       =nom -PL
```

However, the distinction may be lost in morphophonemics. If the morpheme break character is not included in the output of a morphophonemic process, the first morpheme break character in the list is used.

For example:

```
\a   =tion
\u   e+tion
```

Assume hyphen is before the equal sign in the list of morpheme break characters. After this entry adds an *e* onto the end of the root, it outputs *-tion* instead of *=tion.* To remedy that, add the equal sign to the morphophonemic output:

```
\u   e+=tion
```

An explicit morpheme break character is always allowed in the underlying form. If present, it must come after the + morpheme break for suffixes or before the + for prefixes. Be sure you do not enter it like this: *=e+ish.* This would cause the morphophonemic process to put an equal sign into the root, causing the root to fail as a valid lexeme. For a review of the use of the + morpheme boundary character, see p. 188.

[59] These morpheme break characters must be defined in the Parse dialog box (choose Database, Properties, Interlinear tab; select a Parse process line; and click Modify).

Multiple parse processes

It is possible to parse a second time (e.g., to break stems apart):

```
\t   employees
\m   employee     -s
\p   n            PL
\m2  employ -ee   -s
\p2  n      -suf  PL
```

To do this, it is best to make a separate stem lexical database, complete with glosses and parts of speech as needed. This stems database should be included in the first parsing process but not the second. (Or you could put the stem information in a stem field in your main lexical entries and include the stem marker in the first parse but not the second.) The first parse process will see the stems and keep them together, and the second will not see them and so will break them apart.

Multiple words

Glosses and other fields inserted during interlinearization can contain multiple words. For example, the Selaru root *betik* means ‘go outside.’ This could be entered into the lexicon as:

```
\lx  betik
\ps  v
\ge  go outside
```

Shoebox will read the entire gloss phrase and use this as the gloss for *betik*. This multiple-word output capability should not be abused as it can make the interlinear text harder to read. Some glossing conventions require a period or underscore character here (e.g., *go.outside* or *go_outside*). These conventions work in Shoebox as well.

Shoebox can match multiple words. This can be helpful in glossing phrases that have non-literal meanings. For example:

```
\lx  right away
\ge  now
```

Shoebox does not automatically parse affixes on multiple-word stems. You can manually enter the correct parse in the lexicon. Enter the entire phrase with all affixes as a lexical entry or alternate form. Enter the parse as the underlying form.

Multiple glosses

A character such as a semicolon can be used as a gloss separator. The gloss separator character is defined in the Lookup dialog box.[60]

```
\lx  sey
\ps  n
\ge  house; hut
```

Occurrences of the gloss separator divide a gloss field into multiple possibilities, which will be treated as ambiguities by the Lookup process.

The Lookup dialog box also has a Stop at separator check box. If this option is checked, only the part of the gloss field up to the first gloss separator is used. Anything after the gloss separator is available for reversal or as comments. If the gloss separator is at the front of the field, then nothing will be used from the field.

If an interlinear Lookup process finds more than one output field in a record, it treats them as ambiguities. But it uses the marker hierarchy to limit how far it looks. For example:

```
\lx  ser
\ps  v
\ge  be
\ge  exist
\ps  n
\ge  being
```

If the part of speech is higher in the hierarchy than the gloss, once you select *exist* as the gloss, *v* will be chosen automatically as the appropriate part of speech.

Spell check

Spell Check runs the first interlinear process (usually the parse process) on fields in the current record which have the correct language for that process. This verifies that every word can be formed out of morphemes found in the lexicon.

[60] To access the Lookup dialog box, choose Database, Properties, Interlinear tab; select a Lookup process; and click Modify.

For Spell Check to work, there must be an interlinear setup in the database type of the file to be checked. If there is not, you can use Quick Setup to make one (described in the *Interlinear setup* section, p. 159). You can test it by interlinearizing sample words.

To run Spell Check, put the insertion point where you want to begin checking and choose Tools, Spell Check. If a word is found that cannot be parsed, it is selected and Spell Check stops. At that point, it can be ignored, changed, or entered into the lexicon:

- To ignore the word, press a right arrow to move the insertion point off the word (at least one space away) and start Spell Check again.
- To change the word, type the correct word, or move the insertion point into the word and change it.
- To enter the word into the lexicon, use Jump Insert (CTRL+right-click, Macintosh: CONTROL+OPTION+click). To enter the root of the word in the lexicon, select the root and use Jump Insert.

If no spelling problems are found, Shoebox displays a message saying that Spell Check is complete.

In interlinear text, the top line is checked for spelling, but none of the other lines are verified.

To check multiple records in a database quickly, use shortcut keys to move to the next record and to check the record.

Related language adaptation

An advanced feature of Shoebox is its ability to process and convert text from one dialect or related language to another. This capability is called related language adaptation. It is similar to the DOS CARLA (or CADA) software.

A series of sample projects demonstrating the adaptation capability of Shoebox can be found in the Adapt folder. (In the folder in which Shoebox was installed you will find the folder called Samples. Adapt is in the Samples folder.) Open the Adapt.prj project in that folder for more information on how to use these samples.

Editing interlinear text

You can edit words in interlinear text. Alignment will be maintained. This allows you to adjust parses and glosses by hand.

You can manually enter an aligning field with annotations that do not come from the lexicon. Use a Given interlinear process to specify the field marker. As you enter the annotations in the field, type a space to move to the next aligning position.

Shoebox performs interlinear alignment of new annotations based on the immediately preceding line. Therefore, word-level annotation fields should be inserted above the morpheme line. Morpheme-level annotation fields should be inserted below the morpheme line.

Copying interlinear text

You can copy interlinear text from Shoebox and paste it into other programs (e.g., Microsoft Word). In the Shoebox screen display, interlinear text is aligned even when the font has proportional spacing. However, the interlinear text is really stored in a fixed space format. Therefore, in other programs, correct alignment of the pasted interlinear text requires a font that has fixed spacing.

PART 3

APPENDIXES

Appendix A: Importing from MS-DOS

This appendix explains how to import data and settings files from Shoebox for MS-DOS. If you have data to import that is not from Shoebox, or you prefer not to convert your settings files, follow the (rather terse) instructions in the *What if I do not need the information in the SHA.SET file?* section of this appendix.

Please review the *Shoebox settings files* chapter of the *Tutorial*, p. 113. It describes projects, language encodings, and database types. You will need to understand these concepts.

This appendix will show you how to run the conversion program that will extract the database type and language encoding information from your SHA.SET file.

What does this have to do with me?

In Shoebox for MS-DOS, most of the settings information was kept in the SHA.SET file. (The SHB.SET contained some project information and display information such as the color layout.) By taking this information and separating it into different files, Shoebox allows for generalization of the information. Language encoding information can be shared among files of different database types in different projects. More than one project may open the same database and thus make use of the same database type information (and the same language encoding information).

Only you can determine which parts of the SHA.SET file belong with which language encoding or database type. However, we provide a tool to make it easier.

What if I do not need the information in the SHA.SET file?

If you conclude that you do not need the information from your old settings file, treat your data like non-Shoebox data for importing.

Open your own data files using File, Open. If you do not already have an appropriate database type, create one as described in the section *Creating a new database type*, p. 131. Provide a name, record marker, and description that is appropriate to your data.

Also, you may set up some language encoding files, described in the *Creating a new language encoding* section, p. 134.

Importing a sample project

This section is a guided tour of converting and importing files from MS-DOS. This process is demonstrated using a sample project. Working through the step-by-step procedures on a computer is much more effective than merely reading the information.

The sample database and settings files are from *The Text-Oriented Approach* in Chapter 5 and *Setting Up the Conjoined Affix Parser* in Chapter 13 of the *Users's Manual* for Shoebox for MS-DOS. The sample project includes a lexicon, a parse database, and an interlinear text file. The data is in French and English.

The first steps of the conversion process are done by a separate conversion program outside of Shoebox. If you are running Shoebox, please exit. You will be creating new language encoding and database type files, which Shoebox will need to find when it starts.

Macintosh file conversion

Macintosh users should convert their SHA.SET and data files into Macintosh files before running the conversion process. All the CC tables and other processes (such as handling of file names) assume that the conversion has been done.

Step 1: The Shoebox 2.0 conversion program

The conversion program is quite simple to run.

1. Run the Sh2 Converter program (this is found in the Shoebox\Sh2Conv folder).

You will see a welcome message explaining the purpose of the program. (Don’t worry that it says it will convert your settings to

Shoebox 3.0 format. The current version of Shoebox can read version 3.0 settings files.)

2. Click Help.

You will see the step-by-step guide. If you like to read an overview of a program before you run it, click the first step and read through all of the steps. If you prefer to get help only when you need it, go to the explanation of each step by choosing Help on the dialog box of that step.

Also, if at any point you feel you have made a serious error and wish to start over, click Cancel and it will allow you to exit with nothing performed.

3. If you would like the step-by-step guide to stay in view, click back in the welcome box. Otherwise close the help box.
4. Click OK in the welcome box.

Project You will see the Choose directories and project dialog box. It asks for three paths (the third is optional).

The first path requested is the location of the SHA.SET file.

1. Click Browse beside the Path of sha.set file box.
2. Navigate into the folder Sh2_text. (In the folder in which Shoebox was installed you will find the folder called Samples. Sh2_Text is in the Samples folder.)

You should see SHA.SET and SHB.SET in the list of files.

3. Select SHA.SET and click OK to return to the Choose directories and project dialog box.

The second path requested is the location of the project file. The rest of the settings files which you are generating from the SHA.SET file will also be placed in the same folder.

4. Click Browse beside the box labeled Path and file name.

5. Navigate into the Samp_set folder. (In the folder in which Shoebox was installed you will find the folder called Samples. Samp_set is in the Samples folder.)
6. Type `Frinttxt.prj` (for French Interlinear Text) as the file name.
7. Click OK to return to the Choose directories and project dialog box.

Next, Windows users need to specify a Consistent Changes conversion table that will convert the DOS upper ASCII French characters to ANSI for Windows. This is needed to convert characters in places like the sort sequence and case associations.

8. Click Browse beside the box labeled Path for character conversion table.
9. Navigate to the file IBM_ANSI.cct in the user settings folder (by default, My Shoebox Settings on Windows 95 and later and MyShSet on Windows 3.1).

If you do not have special characters to be converted, leave this box blank when you import your own data. If you do not have a conversion table, contact your local computer center.

10. Click OK to return to the Choose directories and project dialog box.
11. Click OK.

Language encodings

A new dialog box will appear asking you to enter your language encodings. This project has French and English. There is already a language encoding file for English, but in order to illustrate using multiple language encodings, you will specify both.

1. Type `French` and press ENTER (Macintosh: press RETURN or click Add).
2. Type `Eng` and press ENTER (Macintosh: press RETURN or click Add).
3. Click OK.

The next dialog box is titled Associate punctuation. It is displaying a list of punctuation from the SHA.SET file. This punctuation is not specifically French but is general for both languages.

4. Click French.

5. Click Eng so that both French and English are highlighted.

6. Click OK.

Next is the first of the sort orders, the Default sort order. Ignoring the **\srt** markers which the SHA.SET file puts along the side, you can see that this is the same form of the sort order as you would see under Globals, Sort Orders in Shoebox for MS-DOS.

7. Click Eng to highlight it.

8. Click OK.

Next it will display a sort order named French.[61]

9. Click French to highlight it.

10. Click OK.

Next, it will display a case association.[62]

11. Click French to highlight it.

12. Click OK.

After this will come several variables, displayed one at a time. Variables are associations of characters. The first variable includes the French vowels. Since this variable also includes the English vowels, include it for English as well. (However, later you will edit it to remove the extra vowels.)

13. Click French to highlight it.

[61] On a Windows system you can see some odd characters. They do not look right because the conversion needed from ASCII characters has not yet been done.

[62] Again, on a Windows system, the Çç and Éé will not look right.

14. Click Eng to highlight both English and French.

15. Click OK.

The variable named cons contains the consonants, including the C-cedillas.

16. Choose both French and Eng for the cons variable.

17. Click OK.

The variable named nasal contains only characters shared by both languages.

18. Choose both French and Eng for the nasal variable.

19. Click OK.

This completes all the information that will go into the language encoding files.

Database types

The next dialog box will ask for the database types. There are three database types in this project: a lexicon whose record marker is **\le**, a parse database whose record marker is **\wd,** and an interlinear text file whose record marker is **\ref**. Database type names do not have to be a single word (nor do language encoding names).

1. Add `French Lexicon, Parse DB,` and `Interlinear Text` as three database types.
2. Click OK.

The next dialog box will show the first of several filters.[63] Looking carefully, you can decipher the first filter to be an N as a “word” by itself in the **ps** field.

The filter for nouns and the next one for adjectives clearly belong to the lexicon, since they refer to the part of speech.

[63] In Shoebox for MS-DOS, variables in filters are enclosed in chevrons. On Windows, the converter program displays the filters using an ANSI font. Therefore, the chevrons are displayed as ® and ¯.

3. Select French Lexicon and click OK for each of these two filters.

The next filter is on the key field. This is less obvious since it could be used for either the lexicon or the parse database.

4. Select both French Lexicon and Parse DB for the *ab* filter.

Next is a search filter.

5. Do not put this in any database. It has been replaced by the ability of the Find command to search the entire database.
6. Click OK without selecting a database type.

Next is a field filter.

7. Select French Lexicon and click OK.
8. The date filter and the following non-unique filter can be applied to both the French Lexicon and the Parse DB.

Next is a column filter (as you can determine by the With command). This goes with interlinear text.

9. Select Interlinear Text and click OK.

Next is a filter which refers to the **mb** marker and the beginning of some word. This was a one-time filter which is no longer relevant.

10. Skip it, click OK without specifying a database.

Next you will see a dialog box labeled Associate Range Sets. Ignoring the **\rng** markers on the left, you can see that this is the form which you used when you specified a range set in Shoebox for MS-DOS. Following the first **\rng** is the **\ps** marker. This is the field to which the range set applies. This belongs to the French Lexicon.

11. Select French Lexicon and click OK.

Next, you will see something like the following on the screen:

```
\bx1 Lexicon
\au1 Y
\tp1  \gl
\tp1  \ps
```

You will see from the title bar that it is named Associate Template. This is the template for Box 1 (**\bx1**) which was named Lexicon. Ignore the **\au1** line (autoload information). The actual template is on the lines which begin with **\tp1**. So the template itself is simply:

```
\gl
\ps
```

Since it has the box label, there is no doubt which database this is for.

12. Select French Lexicon.

13. Click OK.

The next dialog box contains a display of a similar type:

```
\bx2 Parse
\au2 Y
\tp2  \m
```

Hopefully, you can see that this is the template for Box 2, named Parse. The template itself consists of only the **\m** marker.

You should notice something else about this: the French Lexicon is not in the database list. There is only one template allowed per database. Since one has already been chosen for the French Lexicon, it is now omitted from the list to avoid confusion.[64]

14. Select Parse DB.

15. Click OK.

[64] The same thing is done with case associations—only one per language encoding—but since there was only one in this example, we did not comment on it earlier.

There will be a pause. Then you should see on your screen a message telling you that you have successfully upgraded your SHA.SET file.

16. Click OK to end the program.

Step 2: Importing the data files

This step will open all the data files by opening the project. The files which are specified in the Shoebox project file will open automatically (whether or not they were to be autoloaded in Shoebox for MS-DOS).[65] If you have moved them, Shoebox will ask for the path. You will have to select the appropriate database type for each of the files.

1. Run Shoebox.
2. Choose Project, Open.
3. Navigate into the Samp_set folder. (In the folder in which Shoebox was installed you will find the folder called Samples. Samp_set is in the Samples folder.)
4. Select FrIntTxt.prj.
5. Click Open.

Shoebox will display a message indicating that some files could not be found.

6. Click OK.

Shoebox will display the Open a File which has been moved dialog box. Notice that the name of the file it is looking for is visible in the file box, but there is no path. It is looking for the first file, called Sh2Lex.db.

7. Navigate into the Sh2_text folder. (In the folder in which Shoebox was installed you will find the folder called Samples. Sh2_text is in the Samples folder.)
8. Click OK.

[65] If you do not want a file to autoload in the new Shoebox, close it. Only those files which are open when the program is exited are automatically loaded the next time.

Shoebox will display the Import dialog box.

9. In the Database Type list, select French Lexicon.

Notice the Back up original as *.ORI check box. This option preserves your original file by changing its file name extension to .ori. This gives protection in case something goes wrong and your file is corrupted. The default is to make the backup. This would also be a good time to archive your original data files on a removable medium (e.g., CD-ROM, floppy disk, tape, etc.) before you import them. Making archives regularly and at times of transition can prevent loss of valuable data.

Convert the special characters in the lexicon file with the same Consistent Changes table you used in the conversion program to convert the settings.

10. Click Browse.
11. Navigate to the My Shoebox Settings folder (MyShSet for Windows 3.1) and select the file IBM_ANSI.cct.
12. Click OK to return to the Import dialog box.

The Preview box displays the first several lines of the file. Note that the record marker is **le**. In a moment, you will need to tell Shoebox to use this as the record marker for the French Lexicon database type.

13. Click OK to import Sh2Lex.db.

You will be informed that there is no record marker specified for this database type.

14. Type `le` as the record marker.
15. Click OK.

Next, Shoebox will try to open the parsing database. Since you have already told Shoebox where the Sh2Lex.db file is, it will find the Sh2Pars.db file there too. You will have to specify the database type to which it should be assigned. Note that the record marker is **wd**.

16. In the Database Type list, select Parse DB.

Notice that the Consistent Changes table (CCT) you selected last time is still activated. Shoebox will continue to remember a specified CCT for subsequent file imports. If you do not need a CCT for a subsequent import, be sure to clear the Use Change Table check box.

17. Click OK to import Sh2Pars.db.

You will be informed that there is no record marker specified for this database type.

18. Type `wd` as the record marker.

19. Click OK.

Next Shoebox will try to open the interlinear text file.

20. In the Database Type list, select Interlinear Text.

21. Click OK to import Sacrific.it.

22. Type `ref` as the record marker.

23. Click OK.

You will see the three databases that were open in Shoebox for MS-DOS. They are cascaded in the order of the boxes in which they were opened.

Setting the languages of markers

It is important to assign the correct language to each marker, especially those that will be used in sorting. As an example, set the language of the record marker of the lexicon.

1. Go to the Sh2Lex.db window.
2. Choose View, Browse (or click the Browse View button).
3. Scroll down and observe that *à* is the last record in the database.
4. Choose View, Browse to go back to record view.
5. Right-click (Macintosh: OPTION+click) the **\le** marker to display the Marker Properties dialog box.

6. Change the Language Encoding of the marker to French.
7. Click OK.
8. Choose View, Browse again and observe that *à* is now the first record.

In your own databases, the same process should be used to set the language of any field that is not the default language. Besides making the sort sequence correct, this also assigns the correct font to the field, allows languages to be distinguished by style and color, and allows the Find command to see all fields of the correct language.

Cleaning up

You may recall that when you were allocating the variables during the conversion process, you accepted variables for the English language encoding which contained French characters. Now is an appropriate time to deal with these.

1. Choose Project, Language Encodings.
2. Select Eng.
3. Click Modify.
4. Click the Variables tab.

This will display a list of the variables and part or all of their contents. By scanning the list you can see that the variable Cons contains the C-cedillas and that Vowel contains the French vowels.

5. Select Cons.
6. Click Modify or double-click Cons.
7. Edit the characters to remove the C-cedillas.
8. Click OK to return to the list of variables.

Similarly, remove the French vowels from Vowel.

9. Click OK to return to the list of language encodings.
10. Click Close to close the dialog box.

In your own data, you might need to clean up variables, a sort order and case association list, or some filters. (Filters belong to the database type.)

Step 3: Interlinearization

This step will set up your text databases for interlinearization. Notice that the words in the interlinear text, Sacrific.it, are not aligned. This is because the interlinear setup has not been done and so the markers are not yet being treated as interlinear.

1. Be sure the interlinear text window is the active window.
2. Choose Database, Properties.
3. Click the Interlinear tab.
4. Click Quick Setup.

You will see the Quick Setup Interlinear Text Markers dialog box. You can probably see the interlinear text markers still showing to the left of the dialog box.

5. Type `tx`, `mb`, `gl`, and `ps` as the markers.
6. Click OK.

You will see the Quick Setup Lexicon Markers dialog box.

7. Click Insert to put Sh2Lex.db in the Databases to Search list.
8. Type `gl` as the Gloss and `ps` as the Part of Speech. Leave the Alternate form and Underlying form as they are.
9. Click OK.

You will see the Interlinear tab now showing a Parse process and two Lookup processes.

10. Click OK.

You will see that the interlinear text is now aligned in most cases. The exceptions are places like the gloss *father-PL* where there is no space between the morphemes. Shoebox always puts a space between morphemes and uses these spaces as alignment points. (It also uses the space to distinguish between prefixes and suffixes in Lookup, which Shoebox for MS-DOS did not do.) The places that do not look aligned can be left alone in existing interlinear text, or they can be adjusted by re-interlinearizing the word. (Do not do it yet.)

Using a smaller point size on all-caps fields

You will also see that some words in the **ps** line are too long for the space allowed for them. For example, *PRONPOSS* is too long, and *PLSUF* is partially hidden by the following *N*. This happens because the letters in these all-uppercase words are wider than the average width of lower case letters. To remedy this you could enlarge the interlinear spacing setting, but the better way is to use a smaller point size for the part of speech.

1. Right-click (Macintosh: OPTION+click) the **\ps** marker to see its properties.
2. Click to clear the Use Language Font check box.
3. Click Choose Font.
4. In the Size box, select 10.
5. Click OK to return to the Marker Properties dialog box.
6. Click OK.

The part of speech is still very readable but does not take so much space.

Setting the language for interlinear markers

For interlinearization to work, the correct language must be set for each line. This is because the default sort sequence for the language is used to determine which characters are parts of words and which are punctuation.

1. Right-click (Macintosh: OPTION+click) the **\tx** field marker to see its properties.
2. Change its language to French. Click OK.
3. Change the language of the **\mb** field to French also.

The language of the other two fields can be left as default for this example.

Converting chevrons in the parse database

In your own data, if you use forced glosses in the parse database, you will need to convert them to something else before importing the file. Shoebox uses curly braces for marking forced glosses. The ParsPrep.cct Consistent Changes table changes chevrons to curly braces. If you are already using curly braces in your glosses, you will want to use something else for forced glosses. If you like the idea of curly braces for forced glosses, you can change the way you use curly braces. To use something other than curly braces for forced glosses, edit ParsPrep.cct in the Sh2Conv folder (in the Shoebox folder) and change the curly braces to something else.

Note: The ParsPrep.cct must be run *before* any other character conversion table, because it assumes the chevrons are still ASCII codes.

If you are using some other conversion table, you should either run the CC program with ParsPrep.cct to process the file first, or import the file using ParsPrep.cct and then do the other character conversion as follows: export the file as Standard Format text to the name Parsing.sfm, close the file Parsing.db, copy Parsing.sfm to Parsing.db and then open it, which brings up the import dialog box. You then enter your code conversion table (e.g., IBM_ANSI.cct) as the consistent change table.[66]

This step is not necessary for the example you are importing here, because it has no forced glosses in its parsing database.

Changing conjoined affixes to a database

To use the conjoined affixes from Shoebox for MS-DOS, you must convert them to a database. This is done during import with the Consistent Changes table ShAff.cct. This table adds markers and also converts chevrons to curly braces. If you plan to use something other than curly braces for forced glosses, edit ShAff.cct and change the curly braces to something else.

[66] If you are familiar with the CC program, you can do this by closing the Sh2Pars.db file, running CC on it to a different name, and copying the output back to the name Sh2Pars.db.

1. Open Sh.aff (in the Sh2_text folder, which is in the Samples folder). You will see the Import dialog box.
2. Select SHAff as the database type (it has been provided for you).
3. Click Browse.
4. Navigate to the Sh2Conv folder (which is in the Shoebox folder).
5. Double-click ShAff.cct.
6. Click OK to import the conjoined affixes.

This set of affixes does not contain any of the French special characters so does not need character conversion. If your own data needs a character conversion, you can export your data using a Standard Format (SF) export process and import the resultant file using the consistent change procedure described above.

SH2 style parse

For this sample conversion, you will set up parsing like the old SH2 style parse. Later there is a brief discussion of how to convert to the new style of parsing in Shoebox.

1. Change to the window showing Sacrific.it.
2. Choose Database, Properties.
3. Click the Interlinear tab.
4. Click Modify to edit the parse process.
5. Select the SH2 style parse check box.

Two new buttons will appear: Conjoined Affixes and Parse Database. First set up the conjoined affixes.

6. Click Conjoined Affixes.
7. Select Sh.aff as the database to search. (First remove any other databases.)
8. Select **aff** as the marker to find.
9. Select **u** as the marker to output.

10. Click OK to return to the parse box.

Next set up the parse database.

11. Click Parse Database.
12. Select Sh2Pars.db as the database to search. (First remove any other databases.)
13. Select **wd** as the marker to find.
14. Select **m** as the marker to output.
15. Click OK to return to the parse box.

Next, adjust the morphophonemic boundary to match Shoebox for MS-DOS.

16. Delete the + that is there and type # as the morphophonemic boundary.
17. Click OK to return to the Interlinear tab.
18. Click OK.

Adding hyphens to the affixes in the lexicon

In Shoebox, all affixes in the lexicon must have hyphens. If they do not, they are not recognized as affixes. The easiest way to identify the affixes in a lexicon is by part of speech. Sort the database by part of speech to group them together.

1. Go to the Sh2Lex.db window.
2. Choose Window, Duplicate.
3. Choose Database, Sorting.
4. Select **ps** as the first sort field. Leave **le** as the second.
5. Click Sort first field from end.
6. Click OK.
7. Choose View, Browse Fields.
8. Select fields of **ps, le,** and **gl** in that order.
9. Click OK.

This gives a view in which it is easy to see the parts of speech grouped, and you can see which ones look like affixes. The affixes in this sample are all suffixes and have *SUF* in their part of speech.

10. Highlight the first *SUF* record and change out of browse view.
11. Add a hyphen in front of the suffix in the **le** field.
12. Change back to browse view.
13. Do this for every entry that has *SUF* in the part of speech.
14. When finished, close the second lexicon window where you were working.

Setting the jump path

For best results, set the jump paths for the interlinear text to refer to fields in the lexical database(s). This allows you to jump quickly to the lexicon entries for words and to insert new words easily.

1. Change to the interlinear text window, Sacrific.it.
2. Choose Database, Properties.
3. Click the Jump Path tab.
4. Add a new jump path from the **m** and **tx** fields (these are the Source Fields) to the Sh2Lex.db **le** field, the Sh2Pars.db **wd** field, and the Sh.aff **aff** field (these are the Destination Fields).
5. Click OK.
6. For each of the databases on the jump path, go to its window and choose View, Jump Target.

To make jumping more effective, make it treat the hyphens the same way that interlinearization does.

7. Switch back to the interlinear text window and make sure the insertion point is at the top of the record.
8. Choose Edit, Jump To.

9. Make sure that the jump path you added in step 4 is chosen as the primary jump path.
10. Under match characters, choose Even those normally ignored.
11. In response to the message that appears, click Yes to save the matching criteria as the default for this jump path.
12. Click OK. If the insertion point was at the top of the record, Shoebox will jump to the word *sacrifice* in Sh2Lex.db. If it did not match something, click Cancel.

Verifying annotations

In your own text there will be no need to re-interlinearize words that are already correctly interlinearized. You can go on to new text or correct words that are not correct. Shoebox does not verify existing annotations. It replaces them, and if a word is ambiguous it asks again for disambiguation. You can do some verification of existing text with Spell Check. It checks to be sure that every word in the top line can be parsed, but does not verify the other interlinear lines.

To confirm that this text parses as it did in Shoebox for MS-DOS, re-interlinearize every word in this text.

1. Switch back to the interlinear text window, and make sure the insertion point is at the top of the record.
2. Interlinearize (click the button).

You will see that annotations under the first word do not change. The insertion point stops at the beginning of the second word *nos*.

3. Interlinearize again.

You will see a space appear after *PRONPOSS*.

4. Interlinearize again.

You will see a space added before *-PL*, and it will become correctly aligned.

5. Keep interlinearizing up to just before the word *aux* and observe that each word gets the same annotations as before.

6. Interlinearize the word *aux* and observe that the morpheme *-le* fails to gloss.

This shows an important difference between Shoebox for MS-DOS and Shoebox for Windows and Macintosh.

7. To see what causes this failure, jump to *le* (not *-le*).

Observe that it is a root in the lexicon and has no hyphen. Shoebox for MS-DOS ignored hyphens when looking up affixes in a lexicon, but now Shoebox pays attention to the hyphens. This causes Shoebox to show fewer ambiguities.

The word *aux* actually needs to be parsed into two roots. (Shoebox for MS-DOS did not allow this but Shoebox for Windows does.) To fix this, change the parse of *aux* into two roots instead of a root and suffix.

8. Jump to *aux*.
9. Change the hyphen in the **\m** line to a SPACE.
10. Go back to the interlinear window and interlinearize *aux* again.

Observe that it now parses correctly.

11. Interlinearize the rest of the words.

They should all give the same annotations as before.

The interlinear text is probably too wide for the window it is in, so you can reshape it.

12. Put the insertion point somewhere in the first **\tx** field at the top of the record.
13. Choose Tools, Reshape.

You will see the whole sentence reshaped to fit in the window.

Converting to the new style of parsing

You have now seen how to set up and use Shoebox to do parsing similar to that of Shoebox for MS-DOS. For your own work, you might want to convert your setup and data to use the new style of parsing that Shoebox offers. This will eliminate the need for separate conjoined affix and parsing databases.

You need not take time to do this with the Frintxt.prj sample, but with your own data you would perform the following steps:

1. Turn off the SH2 style parse setting to allow multiple affixes to be matched.
2. Delete all conjoined affixes that do not have morphophonemic changes.
3. Move each affix that does have morphophonemics to the lexical entry of its underlying form.
4. Delete from the parse database all entries that can be parsed from the lexicon.
5. Move each of the others into the lexicon under the appropriate root or as a separate word if appropriate.

Importing your data files

This section lists issues that you will need to consider before you begin converting and importing your data files.

How many different projects are you working with? A project corresponds to the SHA.SET file.

1. Make a list of them and where they are located.
2. For each SHA.SET file, create a new Shoebox project file.
3. Make a list of the languages that are involved in each SHA.SET file you are converting.

These will probably include the vernacular language in which you are working, possibly the national language, and probably another major language such as English, French, or Spanish.

Presumably, most projects share a common set of languages. More than one of them may use the same lexicon or other database.

4. Examine each of your Shoebox for MS-DOS setups.[67]

Do you treat a particular language basically the same in each setup? Do you have different sort orders or variables for special purposes?

5. Choose the best treatment of each language and note which project it is in.

You will want to deal only once with each language, if possible. However, if two different treatments have different strengths, make note of that. They can be combined after they are imported.

Which languages are covered by the default sort order? Do any other sort orders cover more than one language?

Shoebox does not require you to set up separate language encodings for languages which can use the same sort order. However, some of the functions in Shoebox make a distinction based on language. Will this distinction be useful for you, or will it be a hindrance? (For advantages of separate language encodings, see *Language encodings*, p. 113.)

You can always change the language of a marker, so your decision need not be final. However, a wise decision now will mean less work later.

6. Make a list of your User Variables.

Which ones correlate with which languages? Which ones are explicitly for one language? Which ones are shared? Which ones have the same names but different contents for different projects?

7. Make a list of the different databases in each project.

Are some of them shared? Which ones are the same type of data? For example, you may keep proper nouns in a separate file from

[67] For some, printing out the SHA.SET file will be a reasonable approach. The file can be viewed as a summary of the project. However, others will find this file unintelligible. They will prefer to bring up the project and examine the settings using the Shoebox for MS-DOS interface, i.e., by viewing the Globals, Sort Orders, etc.

your main lexicon, but it is a lexicon file nonetheless. Other separations may include affixes in a separate file, loan words, or other special vocabulary. Or perhaps you felt that as it grew, keeping it all in one file was unwieldy, and you broke it into alphabetic sections.

8. Make a list of your filters.

Some people have invested much time and effort making complex filters. Some of these will not be necessary in the new program, but many will still be useful. For more information, see *Filter changes,* p. 290. With which databases are these filters associated? (Many filters contain markers which connect them to a particular database. If you have filters in which you regularly changed the markers, you can note that they go with all your databases if that is the case.)

Interlinear setup is part of the database type in which the interlinearizing is done. If you interlinearize your illustrative sentences in your lexicon, you should include this information with the lexicon. Usually, however, interlinear information belongs to the database type of your texts.

Do you use any special characters (anything beyond plain A-Z) as part of your orthography?

9. Is there a character conversion table available?

If you have moved data from Microsoft Word, you probably have such a table. If not, contact your branch computer people to see if such a table is available. If not, the table IBM_ANSI.cct is available; however, it covers primarily the European characters in the upper ASCII set. Characters in the line-drawing section of upper ASCII have not been converted. IBM_ANSI.cct and its inverse ANSI_IBM.cct are available in the Shoebox\Std_Set folder.

10. Do you already have a font for your special characters?

11. Do you already have a Keyman keyboard?

After answering each of these questions, you should have a good idea of what is required to move your own data to Shoebox so that it works very much as it did in Shoebox for MS-DOS.

Appendix B: Comparison to MS-DOS

Improvements

Here are some of the major improvements since Shoebox for MS-DOS:

- You can have more than one window open on the same database (as you might have expected from a Windows program).
- Each window can be sorted and filtered differently and can show the browse view as needed.
- You can sort by *any field,* not just the record marker field, and you can sort by multiple fields.
- You can jump to *any field,* not just the record marker field.
- If you sort by a field which has multiple occurrences within a record, that record will appear multiple times as you step through the database.
- Different fields can display different fonts, and the Keyman program can work with Shoebox to switch keyboards when you put your insertion point in a different field.
- There is a list of markers which links them to language encoding information so the program will know which sort order to use when you sort by a particular field and which font and keyboard to use. This list is fully modifiable but can also be configured as a standard for others to use.
- You can use the Find command to search the entire database, not just the current record. Find can also be restricted to a particular field or a particular language.

- The interlinear feature can refer to multiple fields within a record, not just the record marker field, so allomorphs can be part of their main entry and not scattered throughout the lexicon.
- Multiple affixes in a word can be parsed individually. You no longer need to list all possible combinations in a conjoined affix list. You may still list combinations.
- Prefixes can now be distinguished from roots, suffixes, and infixes, resulting in fewer ambiguities.
- Shoebox can now use all the computer's available memory.
- Shoebox can now save to disk partitions larger than 2.1 GB.
- Date stamping in Shoebox is now Year 2000 compatible.

Filter changes

Three types of filters which were in Shoebox for MS-DOS have been replaced:

- The Search filter is replaced by a Language Text filter using Any Lang. as the language.
- The Field filter is replaced by the Marker Text filter without any text specified.
- The Key filter is replaced by the Marker Text filter. No special distinction is made for the record marker.

The need for these and other filters is lessened because Shoebox for Windows and Macintosh can:

- Sort by fields other than the record marker.
- Use the Find command to search across the whole database.[68]
- Use the Jump To command to jump to fields other than the record marker field.

[68] Filtering causes an index to be built. If you will be moving back and forth in the set of records that match some criteria, you might want the speed of an index. If you want to look up a data item only once, Find is probably a better choice.

Shoebox no longer uses the square brackets and the vertical bar which were used by Shoebox for MS-DOS as part of the filter notation. The And, Or, Not, and parentheses are inserted using the visual user interface. Do not type them as part of the filter text.

Menu changes

This section lists the menu commands in Shoebox for MS-DOS and their equivalents in Shoebox for Windows and Macintosh.

File

MS-DOS	Windows and Macintosh
New	Same
Open	Same
Close	Same
Save	Same
Save As	Same
Options	
Box name	No equivalent (File name is displayed in the title bar of each database window.)
Date stamp marker	Database, Properties, Options tab
Indirect date stamp marker	No equivalent
Autoload	All files open when a project is closed will be reopened when the project is reopened.
Database Template	Database, Template

	Sort order name	Databases can be sorted by any field. Each field is associated with a language encoding. A language encoding can have multiple sort orders, one of which is the default. To sort the records in a window, choose Database, Sorting. To access a marker's language encoding, right-click the marker in the marker pane (Macintosh: OPTION+click).
	Key length	No equivalent (The "key" is the entire field up to 100 characters or the end of the first line.)
	Merge	Database, Merge Database
	DOS	Tools, Run Batch File (Windows only)
	Output	File, Export
	Exit	Same
Box	Box 1-7	Window, 1-9
	Textfile	No equivalent
Edit	Edit Key	The record marker can be edited directly like other fields.
	Margins	Tools, Set Wrap Margin
	Cut	Same
	Copy	Same
	Paste	Same
	Extended Copy and Cut	No equivalent
	Find	Same
	Repeat Find	Edit, Find Next
	Reshape	Tools, Reshape

	Undo All	Same (For incremental undo, choose Edit, Undo.)
	Jump To	Same (also Windows: right-click; Macintosh: OPTION+click)
	Jump Insert	Same (also Windows: CTRL+right-click; Macintosh: CONTROL+OPTION+click)
	Search Path	Edit, Jump To, Jump Path (Search and Insert Record are limited to the active window; only Jump To and Jump Insert use jump paths.)
Database	Next	Database, Next Record
	Previous	Database, Previous Record
	Top	Database, First Record
	Bottom	Database, Last Record
	Search	Same
	Browse	View, Browse (see also View, Browse Fields)
	Filters	Database, Filtering (see also Database, Properties, Filters tab)
	Tally	The status bar displays the number of records matching the active filter.
Record	Insert	Database, Insert Record
	Delete	Database, Delete Record
	Move	Database, Move Record
	Copy	Database, Copy Record
	Join	No equivalent (Use cut and paste)
Utility	Macros	No equivalent
	Record macro	No equivalent

Reversal	Sort window by the desired field; choose File, Export and select Current window's records.
Inventory	
Field markers	Database, Properties, Markers tab (List does not contain frequency count)
Words in field	Tools, Word List
Number Text	Tools, Break/Number text and Tools, Renumber text. To import and number plain text, use TextPrep.cct.
Extended Char	No equivalent (On Windows, you can use Character Map; On the Macintosh, you can use Key Caps.)
Spell Check	Tools, Spell Check

Globals

Interlinear Setup	Database, Properties, Interlinear tab
Conjoined Affixes	Affixes are stored in a database, usually the lexicon.
Begin Interlinear	Tools, Interlinearize
Display	No equivalent
Sort Orders and Variables	Each language encoding has its own sort orders and variables. To access a marker's language encoding, right-click the marker in the marker pane (Macintosh: OPTION+click); then click Language Properties.
Punctuation	Defined by the ! variable
Range Sets	To access a marker's range set, right-click the marker in the marker pane (Macintosh: OPTION+click); then click the Range Set tab.

	Insert Range Member	Edit, Insert from Range Set
Help	Help	Help, Contents
	Keys	No equivalent
	About	Help, About Shoebox

Appendix C: Uninstalling Shoebox

To uninstall Shoebox and its components,

1. If you have data that you intend to keep, this would be a good time to archive your data files on a removable medium (e.g., CD-ROM, floppy disk, tape, etc.).
2. Delete the Shoebox folder and all its contents. (You should not have data files in this folder or any of its subfolders. If you do and you wish to keep them, move them to an appropriate folder.)
3. Delete the user settings folder. (Usually this is Named My Shoebox Settings or MyShSet and can be found on the C: drive.)
4. On Windows, delete the Shoebox.ini file from the Windows folder. On the Macintosh, delete the Shoebox Preference file from the System Preferences folder.
5. Delete the shortcut icons from the desktop and start menu on Windows 95 and later, or from the Program Manager if using Windows 3.1. On the Macintosh, delete any aliases to Shoebox.
6. Remove the association between the .prj file name extension and the Shoebox program (Windows only).

When you upgrade to a new version of Shoebox, the old version is usually replaced. If you have an obsolete version of the Shoebox program in a separate folder, you can uninstall it by deleting that folder and all its contents. (You should not have data files in this folder or any of its subfolders. If you do and you wish to keep them, move them to an appropriate folder.)

Appendix D: Reinstalling the samples

If you have trouble with any section of the tutorial, you can start fresh at the next milestone. It is also possible to back up to the previous milestone, but to go back you must reinstall the sample files. (As you do the tutorial the sample files are changed. If you try to redo a section of the tutorial without renewing the sample files, steps you are told to perform will already have been completed.)

As long as your language data and settings are not located in the Shoebox program folder or any of its subfolders, it is safe to reinstall Shoebox in its entirety.

Reinstalling the samples for Macintosh

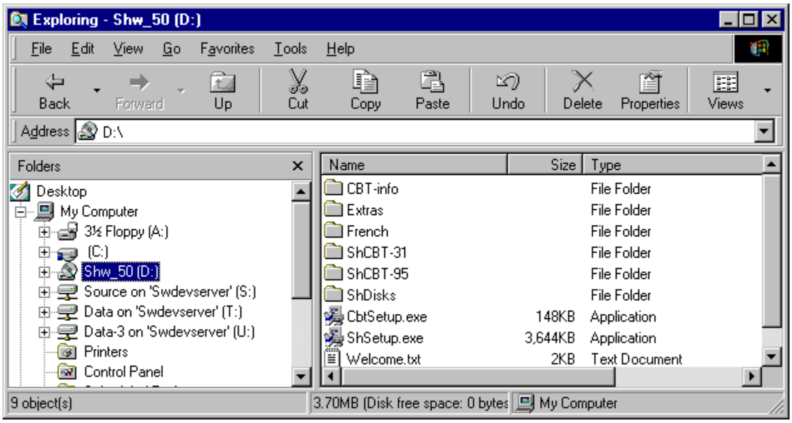
On the Macintosh, it is impossible to reinstall only the samples. You must run the Shoebox Installer to reinstall the program along with its samples.

Reinstalling the samples for Windows

On Windows, you can run Shoebox Setup and install only the sample files. The installation is similar to an initial installation except you must manually elect not to install some of the components of the default installation. You must also enter the folder information for the installation if you used something other than the default locations when you initially installed the program.

The following steps show the details of installing the CD-ROM edition of Shoebox on Windows 95 and later:

1. Insert the Shoebox CD-ROM into the CD-ROM drive.
2. In Windows Explorer, click the CD-ROM drive.



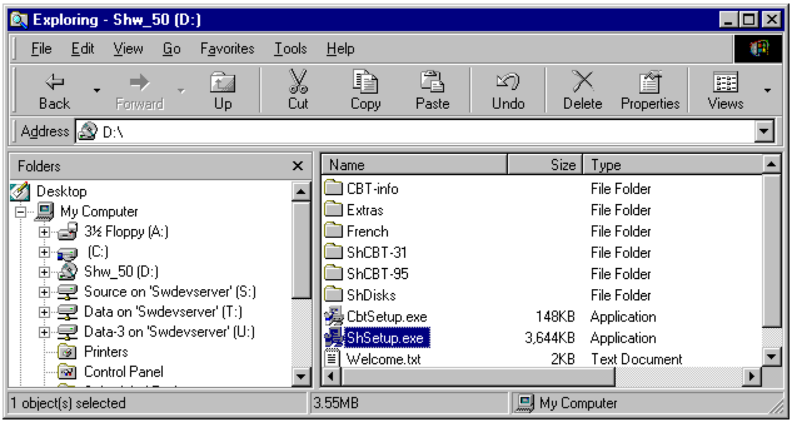
Name	Size	Type
CBT-info		File Folder
Extras		File Folder
French		File Folder
ShCBT-31		File Folder
ShCBT-95		File Folder
ShDisks		File Folder
CbtSetup.exe	148KB	Application
ShSetup.exe	3,644KB	Application
Welcome.txt	2KB	Text Document

9 object(s) | 3.70MB (Disk free space: 0 bytes | My Computer

Fig. 268. Shw_5.0, Shoebox for Windows 5.0, selected

Note: To see a similar view, select View, and then Details.

3. Double-click ShSetup.exe, the Shoebox Setup program.

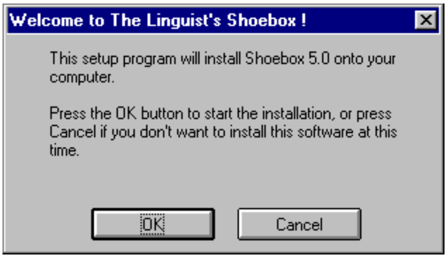


Name	Size	Type
CBT-info		File Folder
Extras		File Folder
French		File Folder
ShCBT-31		File Folder
ShCBT-95		File Folder
ShDisks		File Folder
CbtSetup.exe	148KB	Application
ShSetup.exe	3,644KB	Application
Welcome.txt	2KB	Text Document

1 object(s) selected | 3.55MB | My Computer

Fig. 269. Shoebox Setup program selected

Setup will display a welcome message.



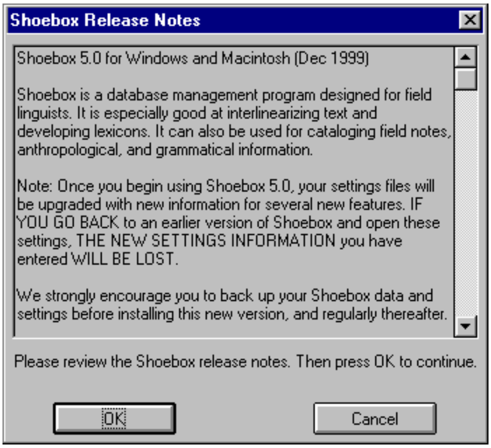
This setup program will install Shoebox 5.0 onto your computer.

Press the OK button to start the installation, or press Cancel if you don't want to install this software at this time.

Fig. 270. Welcome dialog box

4. Click OK.

Setup will display a window showing release information (the Read Me) for Shoebox.



Shoebox 5.0 for Windows and Macintosh (Dec 1999)

Shoebox is a database management program designed for field linguists. It is especially good at interlinearizing text and developing lexicons. It can also be used for cataloging field notes, anthropological, and grammatical information.

Note: Once you begin using Shoebox 5.0, your settings files will be upgraded with new information for several new features. IF YOU GO BACK to an earlier version of Shoebox and open these settings, THE NEW SETTINGS INFORMATION you have entered WILL BE LOST.

We strongly encourage you to back up your Shoebox data and settings before installing this new version, and regularly thereafter.

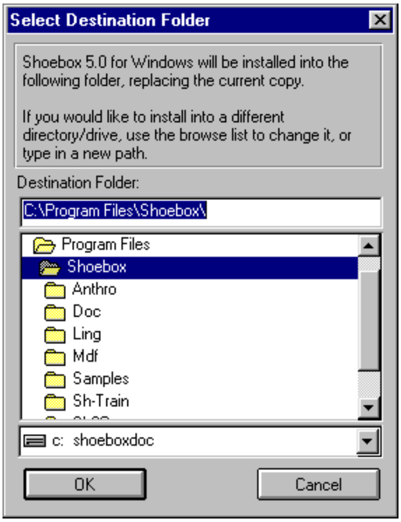
Please review the Shoebox release notes. Then press OK to continue.

Fig. 271. Shoebox Release Notes

The Release Notes give installation information and additional insights into the program that will help you get the most from Shoebox.

5. Review the release notes, then click OK.

Setup will then ask where you want to install Shoebox. Unless you wish otherwise, use the default destination folder.



Shoebox 5.0 for Windows will be installed into the following folder, replacing the current copy.

If you would like to install into a different directory/drive, use the browse list to change it, or type in a new path.

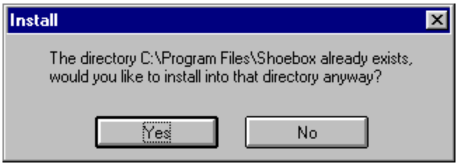
Destination Folder: C:\Program Files\Shoebox\

Fig. 272. Select Destination Folder dialog box

6. Click OK to continue.

If you used the suggested destination folder with your original installation, Setup will return there. If you originally installed Shoebox using a special folder, you must supply this information before clicking OK and moving on.

7. Click OK.



The directory C:\Program Files\Shoebox already exists, would you like to install into that directory anyway?

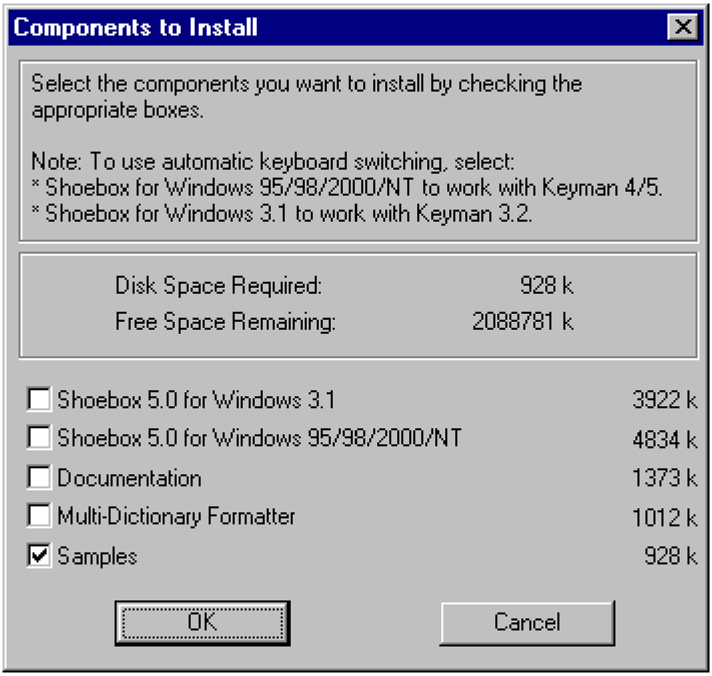
Fig. 273. Install dialog box

You will not actually replace the program, documentation, or Multi-Dictionary Formatter. In a subsequent step, you will select only the samples to be replaced.

8. Click Yes.

Setup will next ask you which parts of the Shoebox package you want to install. When the dialog opens, most of the components are selected.

9. Click to clear the components you do not want reinstalled. Leave only the Samples checked as shown.



Components to Install

Select the components you want to install by checking the appropriate boxes.

Note: To use automatic keyboard switching, select:
* Shoebox for Windows 95/98/2000/NT to work with Keyman 4/5.
* Shoebox for Windows 3.1 to work with Keyman 3.2.

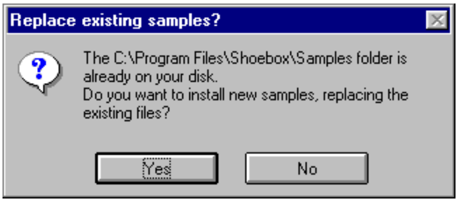
Disk Space Required:	928 k
Free Space Remaining:	2088781 k

□ Shoebox 5.0 for Windows 3.1	3922 k
□ Shoebox 5.0 for Windows 95/98/2000/NT	4834 k
□ Documentation	1373 k
□ Multi-Dictionary Formatter	1012 k
☑ Samples	928 k

Fig. 274. Components to Install dialog box

10. Click OK.

Setup will prompt you before it overwrites the Samples folder.

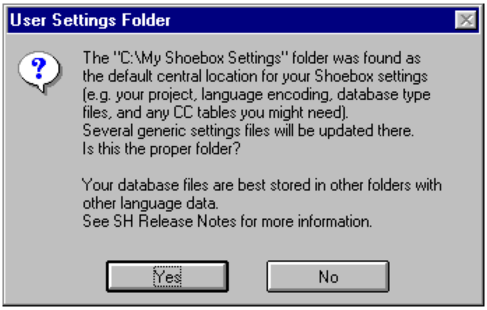


The C:\Program Files\Shoebox\Samples folder is already on your disk.
Do you want to install new samples, replacing the existing files?

Fig. 275. Replace existing samples dialog box

11. Click Yes.

The program installs the new samples and comes back to you with the following User Settings Folder dialog box.

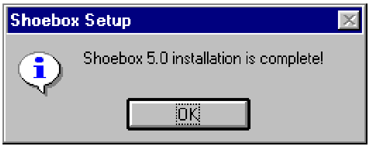


The "C:\My Shoebox Settings" folder was found as the default central location for your Shoebox settings (e.g. your project, language encoding, database type files, and any CC tables you might need).
Several generic settings files will be updated there.
Is this the proper folder?

Your database files are best stored in other folders with other language data.
See SH Release Notes for more information.

Fig. 276. User Settings Folder dialog box

12. Click Yes.



Shoebox 5.0 installation is complete!

Fig. 277. Shoebox Setup dialog box

13. Click OK to close the Setup program.

You have finished and should be able to restart the Shoebox program and open sample projects.

INDEX