PLEDIT - A New Efficient Tool for Management of Multilingual Pronunciation Lexica and Batchlists

Damjan Vlaj*, Janez Kaiser, Ralph Wilhelm***, Ute Ziegenhain*****

* Research and Studies Center, University of Maribor
Razlagova 22, 2000 Maribor, Slovenia
damjan.vlaj@uni-mb.si
** Faculty of Electrical Engineering and Computer Science, University of Maribor
Smetanova 17, 2000 Maribor, Slovenia
janez.kaiser@uni-mb.si
*** Siemens AG, ZT IK 5
Otto Hahn Ring 6, 81739 Munich, Germany
ralph.wilhelm@mch.siemens.de, ute.ziegenhain@mchp.siemens.de

Abstract

The program tool PLEDIT - Pronunciation Lexica Editor - has been created for efficient handling with pronunciation lexica and batchlists. PLEDIT is designed as a GUI, which incorporates tools for fast and efficient management of pronunciation lexica and batchlists. The tool is written in Tcl/Tk/Tix and can thus be easily ported to different platforms. PLEDIT supports three lexicon format types, which are Siemens, SpeechDat and CMU lexicon formats. PLEDIT enables full editing capability for lexica and batchlists and supports work with multilingual resources. Some functions have been built in as external programs written in the C program language. With these external programs higher speed and efficiency of the PLEDIT have been achieved.

1. Introduction

Fast development in the area of speech technology requires the development of tools, which are able to satisfy the demands for fast and effective management of more and more extensive language resources. Various options of the speech technology systems use dictate the need for pronunciation lexica management tools, which will cover the already set demands to the highest extent and will enable efficient defining of new lexica with the use of already existing sources. The possibility of working with multilingual language resources is one of the important requirements. Recently, various tools for work with speech resources have been developed, however, they are mainly intended for work with speech databases (Ziegenhain et al., 1998; Draxler 1997; Grocholewski, 1997).

The tool PLEDIT, which we have developed, meets most of the demands mentioned above and also enables working with multilingual language resources.

The program tool PLEDIT is intended for work with three different types of data:

- Word lists, which are simple text files and are mainly used for creating new lexica.
- Pronunciation lexica, which can be edited and created from word lists or other lexica. The program tool supports Siemens, SpeechDat and CMU lexicon format types.
- Batchlists, which can be edited and transformed to pronunciation lexica. Batchlist file begins with a header followed by an arbitrary number of three-line packets. The packet consists of a relative name of the file, which contains the record of the sentence, and their orthographically and phonetically transcriptions.

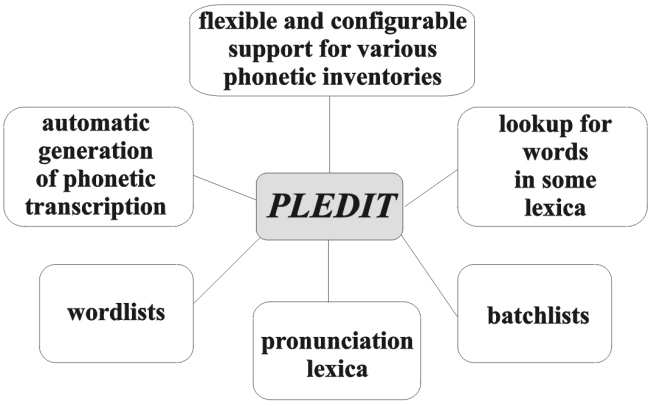

Figure 1: The structure of the program tool PLEDIT.

The PLEDIT allows efficient handling of all three types of data. Handling includes browsing, selecting, adding and modifying the data. It provides the user with the following capabilities (Horvat et al., 1999; Ziegenhain et al., 1998):

- automatic conversion of orthographically written words to their phonetic transcription,
- comparison of lexica,
- checking for the existence of words in a given lexicon,
- conversion between different formats of lexica,
- conversion of lexica to the word model description, which is used in HMM training,
- generating lexica from the batchlists,
- extraction of the batchlist entries that contain only words from the word list,
- flexible and configurable support for various phonetic inventories,
- checking of the lexicon for consistency with a given phonetic inventory,

- switching between different languages in a multilingual environment,
- automatic editing and refreshing of lexicon header,
- setting of the bookmarks, which enables the user to remember the entry, at which he stopped editing lexicon,
- support for headerless lexica and preselection of the phonetic inventory type when the headerless lexicon is read in,
- adding or removing the prefix in front of words or phonemes in *lexicon* window.

Entire functionality mentioned above can be used by means of buttons, menus and drag-and-drop actions between windows. Some of this capabilities will be presented in this article.

2. Working with the PLEDIT

After starting PLEDIT, the *main* window (Figure 2) is opened. All windows in the program tool consist of the following main parts:

- window bar with the name of the loaded data (word list, lexicon or batchlist),
- menu bar with menus containing commands for file handling, setting different options and performing actions such as checking for the existence of words in some lexicon and conversion of orthographically written words to their phonetic transcription,
- status bar at the bottom of the window with the number of items and the number of selected items in the currently opened data.

The *main* window is basically intended for work with word lists. There is always one *main* window opened. The default language and phonetic inventory can be set in this window.

2.1 Using the lexicon window

The program tool PLEDIT can work with pronunciation lexica which are written in the Siemens, SpeechDat and CMU lexicon format types. The program tool enables loading and saving of lexica in all lexicon format types, full editing capability and corrections of lexicon entries.

The outlook of the expanded *lexicon* window can be seen in Figure 3. The list of lexicon entries is divided into two fields. There are words in the left field and their phonetic transcriptions in the right field. Each entry is put in a separate line.

The Siemens lexicon format type begins with a header followed by words and their phonetic transcriptions. When the headerless lexicon is read in or when the head of the lexicon is read in incorrectly, the dialog box *Lexicon Header* will be opened. The program tool makes a guess about the phonetic inventory and preselects the inventory type field according to its guess. After choosing the right options from the given lists and clicking the OK button, the wanted lexicon is read in and can then be stored with the correct head.

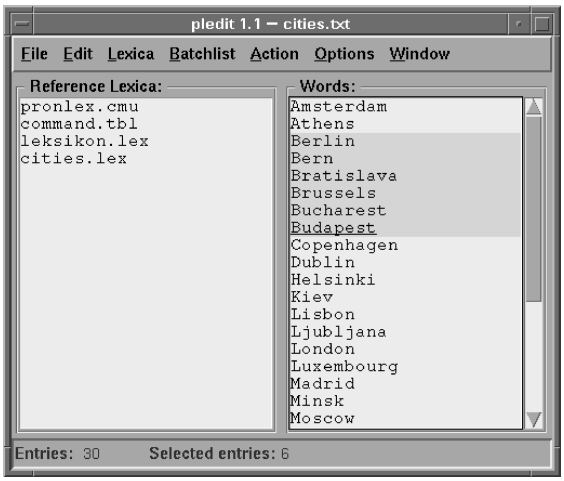

Figure 2: The *main* window of the PLEDIT.

Each entry in the lexicon consists of the word and its phonetic transcription, both of which can be manually corrected. The procedure is as follows. By double-clicking on the entry, which is to be corrected, the *lexicon* window will be expanded (if it is not already expanded) and its word and phonetic transcription will be loaded in the corresponding text fields. Both entries can then be corrected. Phonetic keyboard can be used in the *Phonetic transcription:* text field. Corrected word and phonetic transcription can either replace the old entry or can be added as a new one.

There are several actions which can be performed on words that the lexicon consists of. One of them is adding and removing prefixes from all the words in the *lexicon* window. The same can also be done with the phonemes in phonetic transcription.

The other function is the bookmark function which enables the user to remember the entry, at which he stopped editing the lexicon. By using this option the user marks the place where he ceases his work and can thus start working at exactly the same entry next time. The bookmarks of the all opened lexica can be seen in the *Bookmarks* dialog box.

Another function, which is supported by the program tool, is comparison between two lexica. This function results in the creation of a new *lexicon* window, which contains all the items that which are not the same in both lexica.

The last, but certainly not least important actions that are supported by the program tool are automatic conversion from SAMPA to Spicos phonetic inventory and conversion from lexicon to a word model description. In both cases a new *lexicon* window is created.

2.2 Using the batchlist window

The next window type used in a program tool is a *batchlist* window (Figure 3). The program tool PLEDIT enables loading and saving of batchlists, moreover, it is also possible to correct the batchlist entries.

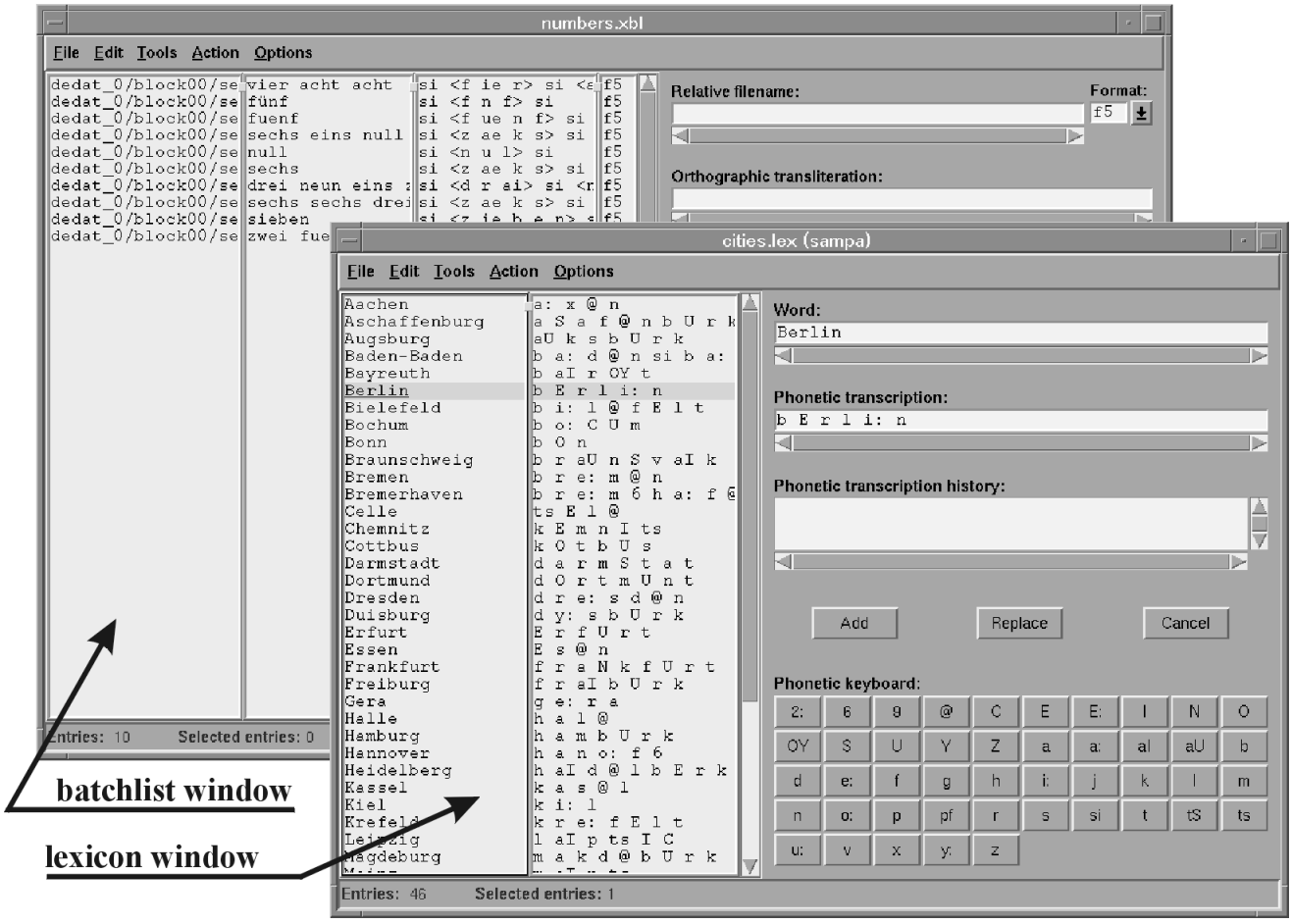

Figure 3: The expanded *batchlist* and *lexicon* window.

The outlook of the expanded *batchlist* window can be seen in Figure 3. The list of batchlist entries is divided into four fields. In the first field, there is the list with the relative names of the files, which contain the records of the sentences, that are written in the second field. The third field contains the phonetic transcription of these sentences and the fourth one provides the information about the format of each line in the batchlist.

The program tool enables correction of the batchlist head in the *Batchlist Head* dialog box.

The phonetic transcriptions can be generated from the orthographic transliteration in the *batchlist* window. With this action, all sentences are transformed into phonetic transcription. Generation of the phonetic transcription is a two step procedure. The first step is to look up for words in the reference lexica and the second one is to produce phonetic transcriptions for words, which have not been found.

The program tool enables extraction of those batchlist entries, which contain only words from the word list. The extracted entries will be displayed in the new *batchlist* window.

If needed, the relative filenames as well as orthographic transliteration can be transformed into the lowercase or to the uppercase.

2.3 Producing new lexica

The program tool supports generation of new lexica, which can be done in two different ways. The usual way to produce a new pronunciation lexica is as follows: first, load a word list, second, search for words in the existing lexica, which are loaded as reference lexica, and last, automatically create phonetic transcriptions for the words which were not found in the reference lexica.

The same programs can be executed within the batchlist window, which means, that the window does not have to be changed, if someone wants to create new lexica, where he is working with batchlists.

The neural network which was taken for the basis of our text-to-phone converter is based on a method used and described in the *Stuttgarter Neuronaler Netz Simulator* SNNS (Zell, 1994) which provides different training methods for a variety of applications.

3. Configuration of phonetic inventories

The program tool PLEDIT enables flexible and configurable support for varios phonetic inventories.

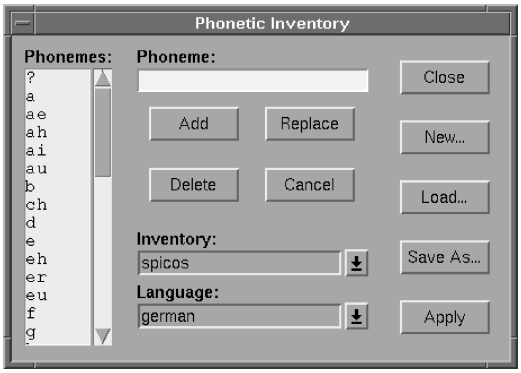

Figure 4: The *Phonetic Inventory* dialog box, which enables full control over phonetic inventories, which are present in the program tool.

The program tool supports management of all phonetic inventories, which are supported in the PLEDIT. The phonemes of the current phonetic inventory can be seen in the *Phonetic Inventory* dialog box (Figure 4). This dialog box enables adding, replacing and removing phonemes from the phoneme list of the current phonetic inventory. Another phonetic inventory can be seen by selecting a different inventory or language. The *Phonetic Inventory* dialog box also enables generation of the new phonetic inventory. The phonetic inventory can also be loaded from an external file. All changes of the phonetic inventory can be saved in an external file and will be seen in the program tool (phonetic keyboards and menus) after applying the changes in the program tool.

The phonetic inventory can be extracted from the lexicon or from the HMM Reference File. The result of the extraction is represented in the *Extraction of the Phonetic Inventory* dialog box. Extracted phonemes are displayed in the first column of this dialog box, whereas the calculation of the phonetic inventory histogram from a lexicon is displayed in the second column. Extracted phonemes can be saved in an external file and applied in the program tool.

The program tool PLEDIT supports management of the inventory mappings. The inventory mappings can be loaded and displayed in the *Inventory Mappings* dialog box (Figure 5), where they can be corrected as well. All changes can be saved as an instruction file for the convert-filter or as an instruction file for HModify.

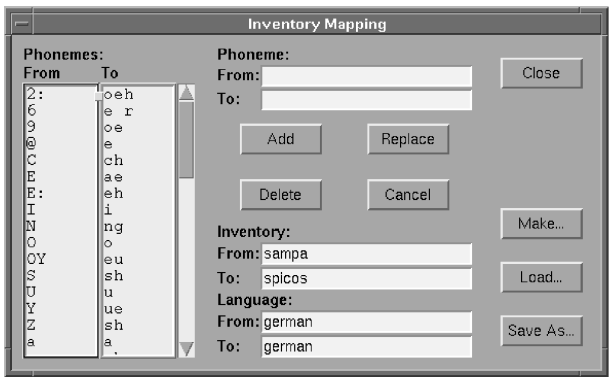

Figure 5: The *Inventory Mappings* dialog box.

The program tool PLEDIT also enables generation of the inventory mapping list by comparing two different lexica. The result is displayed in the *Inventory Mappings* dialog box.

4. Conclusion

The Pronunciation Lexica Editor PLEDIT is a program tool, which enables efficient work with pronunciation lexica and batchlists. PLEDIT enables flexible and configurable support for varios phonetic inventories.

The program tool is written in the Tcl/Tk/Tix language, which is suitable for visual programming. By including some programs written in the program language C, substantial reduction of the time, required for some operations, has been achieved.

The program tool PLEDIT supports working with pronunciation lexica, batchlists and multilingual resources. It also supports generation of phonetic inventory mappings, which are needed as input data for the convert-filters, the task of which is to convert one phonetic inventory to another.

At this moment the conversion of lexica from SAMPA to Spicos phonetic inventory is integrated. However, in the future we want to integrate a convert-filter, which will be able to use inventory mappings for all possible phonetic inventories.

5. References

Draxler C. (1997). WWWTranscribe - a modular transcription system based on the world wide web. *In Proceedings of the 5th European Conference on Speech Communication and Technology - EUROSPEECH'97*. (pp. 1691 - 1694). Rhodes. Greece.

Grocholewski S. (1997). CORPORA - Speech Databases for Polish Diphones. *In Proceedings of the 5th European Conference on Speech Communication and Technology - EUROSPEECH'97*. (pp. 1735 - 1738). Rhodes. Greece.

Horvat B., Kačič Z., Vlaj D. & Kaiser J. (1999). *The Pronunciation Lexica Tool Final Report, Report Version 3*. Faculty of Electrical Engineering and Computer Science. University of Maribor. Maribor. Slovenia.

Ousterhout J. K. (1993). Tcl and the Tk Toolkit. ISBN 0-201-63337-X. Addison-Wesley. University of California Berkeley. California. USA.

Welch B. B. (1997). Practical Programming in Tcl and Tk, Second edition. ISBN 0-13-616830-2. Prentice Hall. New Jersey. USA.

Zell, A. (1994). Simulation Neuronale Netze. Boon, Paris. Reading, Mass., Addison-Wesley.

Ziegenhain U., Harengel S., Kaiser J. & Wilhelm R. (1998). *Creating Large Pronunciation Lexica for Speech Applications*. In Proceedings of the First International Conference on Language Resources & Evaluation. (pp. 1039 - 1043). Granada. Spain.