

Manual for Dictionary System DWS

Aleš Chejn

Translated into English by Ján Zaťko

December 2011

www.hvalur.org

Table of contents

1. Introductory information.....	4
1.1. Extension for Icelandic.....	4
2. System requirements.....	5
2.1. Web browser	5
2.2. MySQL.....	5
2.3. PHP.....	5
2.4. Free space on the server.....	5
2.5. FTP	5
2.6. Adobe Flash Player in the web browser.....	6
2.7. LaTeX.....	6
2.8. E-mail account.....	6
2.9. Optional / Recommended.....	6
3. Installation.....	7
4. Creating the new project	8
4.1. Import.....	9
4.1.1 Uploading the file to the server.....	9
4.1.2 Assigning the values to the fields.....	9
4.2. Adapting the application.....	9
4.2.1 Editing the language files.....	9
4.2.2 Editing the menu.....	9
4.3. Back-up.....	10
5. Introduction to the application Dictionary System.....	10
5.1. Exploring the working environment.....	10
5.2. List of functions in the Menu for the public.....	11
5.2.1 List.....	11
5.2.2 Information.....	11
5.2.3 Pronunciation.....	11
5.2.4 Additional notes.....	12
5.2.5 Download.....	12
5.2.6 About Dictionary.....	12
5.2.7 Login.....	12
5.3. List of functions in the Menu for the registered users.....	13
5.3.1 List	13
5.3.2 Word.....	13
5.3.3 Sorting.....	13
5.3.4 Declension.....	14
5.3.5 Multimedia.....	14
5.3.6 History.....	15
5.3.7 Links.....	15
5.3.8 Administration.....	15
5.3.9 Profile.....	16
5.3.10 Information.....	16
5.3.11 Help	16
6. Structure of the Dictionary	17
6.1. Structure of the printed Dictionary.....	17
6.1.1 Structure of the Dictionary	17
6.1.2 Structure of the entry	18
6.1.3 Structure of the meaning part.....	18

6.2. Structure of the online Dictionary.....	19
6.2.1 Structure of the Dictionary.....	19
6.2.2 Structure of the entry	19
6.2.3 Structure of the meaning part	20
6.2.4 Tables of declension and conjugation forms.....	20
6.3. Structure of the offline Dictionary	21
6.3.1 Structure of the Dictionary.....	21
6.3.2 Structure of the entry.....	21
6.3.3 Structure of the meaning part.....	21
7. Creation of the Dictionary.....	22
7.1. Adding a new headword.....	22
7.2. Headword editing.....	22
7.3. Deleting the headword.....	23
7.4. Manipulation with headwords.....	23
7.5. Manipulation with meanings.....	23
7.6. Declension / Conjugation.....	23
7.7. Entry history	24
8. Publishing.....	24
8.1. Publishing the online version.....	24
8.2. Publishing the printed version - LaTeX → PDF.....	24
8.2.1 Preparing the environment.....	24
8.2.2 Generating the printed version	25
8.2.3 Copying the photographs from the server to the local computer	25
8.2.4 Conversion from .tex to .pdf.....	25
8.2.5 Book format.....	26
8.3. Publishing the offline version - DSL format (for GoldenDict).....	26
8.3.1 Test of the export.....	27
8.3.2 Export.....	27
9. Structure of the application	27
9.1. Structure of the MySQL database.....	27
9.2. Structure of the directories.....	29
9.3. Structure of the web page.....	30

1. Introductory information

The application Dictionary System (DS) is a web application designed for creation of one-way bilingual dictionaries or encyclopaedias offering a working environment for creation of a dictionary and a web page which enables the general public to search in the dictionary. It is so-called DWS application (Dictionary Writing System¹) or DPS (Dictionary Production / Publishing System).

It is a web application and, therefore, it runs in any web browser (Firefox, Chrome, Opera, Safari, Explorer etc.), under any operating system (Linux, Windows, MacOs etc.).

The application DS supports creation of a multimedia dictionary. In the application, it is possible to add, edit and delete headwords and their meanings, to add photographs and sound files, to add, edit and delete declensions, to control the process of creation of a dictionary and to publish a dictionary in several versions. The final result of the application is a multimedia online version of a dictionary, a printed version of a dictionary typeset in a quality form for a book format in LaTeX (and then exported to PDF) and an offline version of a dictionary in DSL format (Abbyy Lingvo²) for offline programs like GoldenDict³ or StarDict⁴.

The application supports team work of lexicographers. The application monitors activities, helps to coordinate aims, records headword editing etc. It provides an easy-to-use user environment for lexicographers, WYSIWYG headword viewing and the access to the application through any browser. The full support of Unicode is guaranteed.

The application data is saved in the MySQL database which is automatically backed-up. The application DS is published under the General Public License GNU⁵, it is completely for free and it is possible to freely adapt the application to your own requirements.

1.1. Extension for Icelandic

The application contains the Extension for Icelandic. It is designed as a base for creation of an Icelandic(?) dictionary. The Extension includes a list of the most frequent Icelandic headwords (more than 22 000). Each entry contains the IPA transcription, pronunciation recorded by a native speaker (Jón Gíslason), the part of speech with word endings, word division, syntactic information on verbs. Noun, adjective and pronoun entries include headword declensions. Verb entries include headword conjugations. Adjective and adverb entries include degrees of comparison. Besides, the

1 http://en.wikipedia.org/wiki/Dictionary_writing_system

2 <http://www.lingvo.ru/>

3 <http://goldendict.org/>

4 <http://code.google.com/p/stardict-3/>

5 <http://www.gnugpl.cz/v3/>

Extension contains scripts for generating declensions and conjugations of new headwords and pronunciation of headwords. The Extension also contains approximately 10 000 synonyms and antonyms, about 8 300 example sentences with headwords, Latin names and photographs for botanical and zoological entries.

2. System requirements

2.1. Web browser

Supported web browsers: Mozilla Firefox >3, Opera, Safari, Internet Explorer. It is recommended to use Mozilla Firefox >4. The usage of Internet Explorer is troublesome. Internet Explorer does not support all possibilities of the application DS. More concretely, Internet Explorer cannot view the IPA characters in the editing form (on the other hand, it can view them on the search page). Therefore, we recommend you to use other browsers.

If it is not possible to install other browser, we recommend you to read [how to properly view the IPA characters in Internet Explorer](#). Attention: the page of the application will not display optimally, but it will be possible to edit the IPA characters. Note that if the lexicographer does not edit the IPA characters, he can work in Internet Explorer and edit headwords with almost no problems.

2.2. MySQL

The application requires MySQL 5.0 and higher.

2.3. PHP

The application requires PHP 4.4.9 and higher.

2.4. Free space on the server

The application requires 30 MB of free space on the server. The Extension for Icelandic requires another 2 GB on the server.

2.5. FTP

FTP connection to the server for uploading the application files on the server and for downloading published files.

2.6. Adobe Flash Player in the web browser

Adobe Flash Player⁶ is freely available for all browsers.

2.7. LaTeX

The LaTeX environment serves to publish the printed version. LaTeX⁷ is freely available for various operating systems. The list of libraries required for publishing is indicated in detail in the section Publishing.

2.8. E-mail account

It is necessary to have an e-mail account on gmail⁸ (it is for free) or to have access to your own e-mail server. The application uses e-mails for sending confirmation on registration and the forgotten password as well as for sending out messages to all registered users.

2.9. Optional / Recommended

If you have access to the service Cron⁹, you can use automatic database back-up. The database can be backed-up also manually, through PhpMyAdmin¹⁰, for example.

Recommended applications for the administrator to adapt DS to his own requirements (national language requirements, entry editing requirements etc.):

1. text editor – Jedit¹¹, Gedit, Notepad, Writer, Word etc.
2. PhpMyAdmin – access to the MySQL database
3. PDF file browser – controlling the published final PDF file
4. offline dictionary program (GoldenDict, StarDict) – controlling the final offline version in DSL

6 <http://www.adobe.com/products/flashplayer.html>

7 <http://en.wikipedia.org/wiki/LaTeX>

8 <http://mail.google.com/>

9 <http://en.wikipedia.org/wiki/Cron>

10 http://www.phpmyadmin.net/home_page/index.php

11 <http://www.jedit.org/>

3. Installation

After downloading the ZIP file with the application, unzip all to the local computer and copy it to the web server (e.g. www.your-dictionary.com). In the browser, enter the address of the server to which you copied the application and add **/install/install.php**

(e.g. www.your-dictionary.com/install/install.php)

The web installation guide opens. In the following steps, follow the instructions of the guide.

Step 1 – select the installation language (Czech or English).

Step 2 – read the license conditions and check that you agree with the conditions of the General Public License GNU.

Step 3 – the installation guide controls the rights to the directories and to the file `connection.php`. If the rights are not correct, set them up with FTP Client to `chmod 666`. As soon as all the rights are correct, continue the installation process.

Step 4 – enter the name of the MySQL server, the username and the password to access the MySQL server. Enter the name of the database as well. The database has to exist. If you do not have the database, create it and then, continue the installation process. At the end of the installation, the values are saved to the file `connection.php` in the home directory.

Step 5 – the installation guide tests the connection. If it is correct, continue to Step 6. If it is not correct, go back to Step 4 and correct values for the MySQL connection.

Step 6 – (optional) enter the address of the e-mail server, the username and the password along with the e-mail address from which messages will be sent out to registered users; enter also your real name which will be displayed in the e-mail. The values are saved again to the file `connection.php` in the home directory.

Step 7 – (optional) enter the address of the FTP server, the username and the password to access FTP as well as the real absolute path to the backed-up MySQL file on your server and the name of the MySQL file (e.g. `backup-mysql.sql`). If you do not know these values, you can enter them later by editing the file `connection.php` in the home directory.

Step 8 – values for MySQL, Mail, FTP are saved to the file `connection.php` in the home directory on the server. The rights of the directories `/audio/uploaded_files/`, `/images/uploaded_files/`, `/tmp/`, `/import/uploads/`, `/import/tmp/`, `/backup/` and of the file `connection.php` are controlled.

Step 9 – by clicking on the hyperlink, you install MySQL tables to your MySQL database. When a pop-up window appears, confirm the installation in it and control the process. After finishing the installation process, close the window. The installation of MySQL tables may take more time – please, be patient. If the installation of MySQL tables fails for some reason and you want to try it again, at first, you have to clear your database.

Step 10 – choose your administrator username and create a new password (min 6 characters without diacritics). For verification, re-enter the password and enter your e-mail address through which users will be able to contact you. The e-mail address will be available only to registered users after login. Now, these values are saved to the MySQL database and the passwords are encrypted.

Step 11 – the installation guide controls the password length, whether values are identical and whether they are saved to the MySQL database. If the password is not correct, go back to Step 10.

Step 12 – choose the project password. The project password is a password which has to be entered in order to enable the user to register to the project (the application DS). Only the project administrator should know the project password. The project password can be later changed directly in the application. The project password can be changed only by the administrator. The project password is encrypted.

Step 13 – the installation guide controls the project password length and whether values are identical and saves them to the MySQL database. If the password is not correct, go back to Step 12.

Step 14 – the installation is successfully finished. Now, it is recommended to delete the entire directory **/install/** from the server. Continue by entering the address of your dictionary (e.g. www.your-dictionary.com) and by registering to the application.

4. Creating the new project

You can either import the new project – your dictionary – from the CSV file¹² or add it word-by-word. Before you start to work on the dictionary, it is necessary to adapt web pages to your requirements.

¹² http://en.wikipedia.org/wiki/Comma-separated_values

4.1. Import

4.1.1 Uploading the file to the server

The import function helps you to import your dictionary to the application DS. At first, export your dictionary to CSV format in Unicode. This format is supported by the majority of the database programs (Excel, MySQL etc.). CSV format should have the following form - "fallegur","adj","hezký","fagur" (double quotation marks around the text and columns separated by comma). Each line in the file represents one line in the table. The file is saved to the directory /DISK2/WWW/hvalur.org/hvalur/import/uploads on the server. The temporary file is saved to /DISK2/WWW/hvalur.org/hvalur/import/tmp/. It is necessary to set up the rights to both directories for writing. If the tables of the application already contain some values, the application asks whether to clear the tables or whether to import only non-duplicate values.

4.1.2 Assigning the values to the fields

After uploading the CSV file to the server, the application will select the first line from the imported file. The values are assigned to the fields of the dictionary. It is necessary to assign one value to the field Headword. Only one value can be assigned to each field. After clicking on the button, the import will be initialized – please, wait until the import is finished. After the import is successfully finished, it is possible to start editing.

4.2. Adapting the application

4.2.1 Editing the language files

Above all, it is necessary to edit the language files in the directory **/language/**. Each file has to be edited separately for each language version. Basic information on the dictionary (its name, authors, information) is in the file **language.php** (full path for the English version is **/language/en/language.php**). Information which has to be changed is in the upper part. Put the texts between double quotation marks (as it is in the original version). If you need to put double quotation marks into the text, you have to add the symbol `\` in front of each quotation mark (like this `\`). As soon as you edit the file, upload it to the server and test it in the browser.

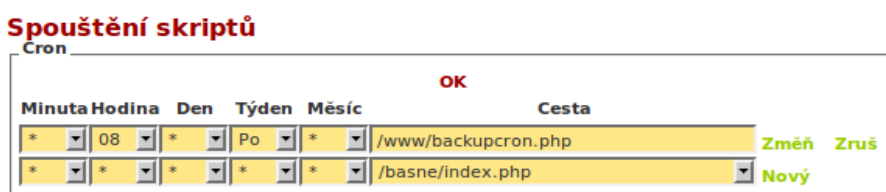
4.2.2 Editing the menu

Another possibility how to adapt the web pages to your requirements is to edit the file

menu.php in the root directory. For in the menu, there are all the hyperlinks to the pages in the application – that means, to all the functions the application provides. For example, if you do not want to use declensions, you will simply remove the entire menu item from the menu. You can add new pages and functions in the same way.

4.3. Back-up

It is recommended to back up the database regularly. It is possible to back it up manually through phpMyAdmin, for example, or automatically through Cron, if the host server enables to back up the database automatically. The path to the backup file and the name of the file are in the file connection.php. You can edit these values in the file manually. Then, set up the automatic back-up for once a week, for example. The file which initializes copying of the backup file to the /backup/ directory is called backupcron.php and it is located in the root directory. Example of the set-up:



5. Introduction to the application Dictionary System

5.1. Exploring the working environment

The application consists of two parts. The first part is for the public and it is freely available without registration. It contains the function enabling to search in your dictionary, information on your dictionary, pronunciation, resources that you used to create the dictionary, the authors, the license etc. The menu helps to navigate the pages of the dictionary. On the right side of the menu, there is also the function Login. After the registered users (lexicographers, creators of the dictionary) log in, they will get to the second part of the application. Unlike the first part, the second one enables to edit and publish your dictionary. This part is available only to the users who registered and logged in to the application. Again, the menu helps to navigate the pages of the application, nevertheless, it differs from the menu for the public, because it contains dozens of

functions enabling to edit the dictionary. Information for the public is on the right side of the menu, in the section Information.

Navigating the menu is easy, because it does not change its content. If the function is not available when needed (e.g. it is not possible to delete declension if no headword is chosen), it means that the menu item is not active. The menu which never changes its content, guarantees easy orientation in the application.

5.2. List of functions in the Menu for the public

The main menu of the application for the public contains the following items:

5.2.1 List

The item List contains two sub-items: Own list and Alphabetic list. The function Own list enables to go back to the headword list that the user created by searching headwords through the search field. The function Alphabetic list displays on the left side of the screen the alphabetic list of headwords. The headword that you choose, differs from other headwords in the list by its red colour.

The arrows help to navigate the list.

5.2.2 Information

The item Information contains three sub-items: Information on the Dictionary, Guide to the Dictionary and Tips for searching in the Dictionary. Information on the Dictionary include information on your dictionary and the list of versions of the dictionary. The Guide to the Dictionary indicates in detail the way the data is set in the dictionary. The item Tips for searching in the Dictionary presents various ways of searching.

5.2.3 Pronunciation

The item Pronunciation contains six sub-items: Information on pronunciation, Pronunciation of vowels, Pronunciation of consonants, List of phonemes, Generation of pronunciation and Sound game. Information on pronunciation includes the set of information on pronunciation in the dictionary. Pronunciation of vowels includes the list of vowels with the IPA transcription and example words. Pronunciation of consonants includes the list of consonants with the IPA transcription and example words. List of phonemes includes the list of phonemes with examples.

Through the item Generation of pronunciation, you can generate the IPA transcription of any word along with explanatory notes on the rules of pronunciation which were applied and the reasons why they were applied. Through Sound game, the user listens to the sound file (of one of headwords) and then, he tries to write the headword correctly. Both the correct and incorrect answers are counted. If the user's answer is incorrect, he can replay the sound file again. It is possible to reset the game and to start counting correct and incorrect answers once again.

5.2.4 Additional notes

The item Additional notes contains three sub-items: Abbreviations in alphabetical order, Abbreviations listed according to usage and Resources. The item Abbreviations in alphabetical order includes the alphabetic list of abbreviations used in the dictionary. The item Abbreviations listed according to usage includes abbreviations which are ordered in the lists according to usage (Primary grammatical information, Secondary grammatical information, Field categories, Language categories). The item Resources includes the list of all the materials and resources used by the authors of the dictionary.

5.2.5 Download

The item Download contains the list of files that the user can download. After clicking on the hyperlink Download the file, the user has to copy down the values from the picture. After he successfully fills in the control field, he downloads the file to his computer.

5.2.6 About Dictionary

The item About Dictionary contains two sub-items: Authors and Licence. The item Authors introduces the authors of the dictionary as well as the authors of web page translations. The item Licence includes information on the licence under which the dictionary is published.

5.2.7 Login

The item Login contains the form for registered users enabling them to log in to the application DS. There are also links to the pages Forgotten password and Registration. After entering the username and the e-mail address on the page Forgotten password, a new password is sent to the user's e-mail account. On the page Registration, the unregistered user can register by entering the username, the password, the e-mail address and the project password that he obtained as a part of the project administrator's invitation.

5.3. List of functions in the Menu for the registered users

The main menu of the application for the registered users contains the following items:

5.3.1 List

The item List contains two sub-items: Own list and Alphabetic list. The function Own list enables to go back to the headword list that the user created by searching headwords through the search field. The function Alphabetic list displays on the left side of the screen the alphabetic list of headwords. The headword that you choose, differs from other headwords in the list by its red colour.

The arrows help to navigate the list.

5.3.2 Word

The item Word contains four sub-items: Add a new word, Add a new meaning, Edit the meaning and Delete the meaning. The item Add a new word enables to add a new headword to the dictionary. The item Add a new meaning enables to add a new meaning to the chosen headword. Through the items Edit the meaning and Delete the meaning, the user gets to the headword editing page where he can edit or delete meanings of the chosen headword.

5.3.3 Sorting

The item Sorting contains three sub-items: Order meanings, Change the order of meanings and Change the headword number. The item Order meanings enables to change numbers of the order of meanings and to order them. E.g. if the headword has four meanings, this function assigns to these meanings numbers of the order of meanings from 1 to 4. The item Change the order of meanings displays the arrows (up and down) in the headword preview and by clicking on the arrows, it is possible to change the order of meanings. After re-clicking on the item Change the order of meanings, the arrows disappear. Attention – if the headword has only one meaning, the arrows are not displayed. The item Change the headword number enables to change the number of the headword in the entire database. In order to maintain the consistency, the names of photographs and sound files related to the chosen headword will be changed as well. If the dictionary contains a homonym for the headword (the same word with a different headword number), the application controls it in order to avoid overwriting another headword and, if necessary, it increments the headword number by one. E.g.: if there were two headwords "hestur" with the headword numbers 1

and 2 and if the user wanted to change the headword number 1 to 2, the application would change that number to 3 – the application would prevent overwriting already existing headword number 2 by omitting the number 2.

5.3.4 Declension

The item Declension contains four sub-items: Add – generation, Add – empty, Edit and Delete. The item Add – generation generates headword declension or conjugation. Generation is based on information on the part of speech and on declension or conjugation endings which have to be in the correct form (see Manual for headword editing). At first, it is necessary to add this information on the page Headword editing and then, to generate declension tables. The item Add – empty enables to add empty declension tables. Before adding the empty declension tables, it is necessary to enter information on the part of speech on the page Headword editing. The item Edit enables to edit declension or conjugation forms of headwords. The process of declension or conjugation editing for different parts of speech is the same – save edits and then, save changes to the word form database. Moreover, there is a page concerning adjective declension and verb conjugation where the lexicographer can determine which declension tables exist and which do not exist. The existing tables can be edited and they are also displayed in the dictionary. The item Delete deletes declension or conjugation of the chosen headword after the user confirms it. Deleting declension and conjugation is irreversible.

5.3.5 Multimedia

The item Multimedia contains two sub-items: Add a picture and Add a sound. If the entry includes a sound or a picture, the items Delete a picture and Delete a sound are displayed and ordered according to number. The item Add a picture uploads a picture from the local computer to the server, creates its miniature form and names both the files according to the following example: `ds_images_heslové-slovo_číslo-heslového-slova_pořadové-číslo-obrázku.jpg` (name of the miniature picture is the same as that of the original one, but it begins with `th_`). Pictures are saved to the directory `/images/uploaded_files/`. While adding the picture, it is necessary to indicate the author's name and the license. The item Add a sound uploads a sound file from the local computer to the server and renames it according to the following example: `ds_heslové-slovo_číslo-heslového-slova_pořadové-číslo-zvuku.mp3`. Sound files are saved to the directory `/audio/uploaded_files/`. If you want to delete a picture or a sound, you must confirm it. The action is irreversible.

5.3.6 History

The item History contains two sub-items: Save to the history and Show the history. The item Save to the history saves the chosen headword to the history. A new record is saved to the headword history and the script for writing compounds is launched as well. The script goes through the entire database and finds compounds which use a word form of the headword to create a compound. Subsequently, it saves all found compounds to the item Compounds which is displayed while viewing the headword. The item Show the history shows the headword history.

5.3.7 Links

The item Links contains hyperlinks to resources which are useful for the creation of the dictionary. In the file menu.php in the home directory, the administrator edits manually the set of links contained in the item.

5.3.8 Administration

The item Administration contains nine sub-items: Import, Phonetics, Usage category, Resources, Messages for unregistered users, Messages for registered users, Scheduled tasks, Dictionary publication and Project administration. The item Import contains instructions on how to import the created dictionary to the application. The item Phonetics enables to edit the phonetic rules used while generating the IPA transcription. The item Usage category enables to create, edit and delete usage categories. The item Resources enables to create, edit and delete resources and materials used in the dictionary. Resources are divided into five groups (Dictionaries, Online dictionaries, Other related and used material, Other internet sources and Sources in programming). After pressing the button Edit resources, there will be displayed the green ring for editing and the red cross for deleting. The item Messages for unregistered users enables to create, edit and delete messages for unregistered users. The messages are displayed on the homepage of the dictionary and they contain important information on evolution of the dictionary. The item Messages for registered users enables to create, edit and delete messages for registered users. The messages are exclusively for fellow creators. They are displayed on the homepage after login. The item Scheduled tasks enables to create scheduled tasks. Registered users select the task and after accomplishing it, they state the action was successfully done. Through the item Dictionary publication, the project administrator publishes the dictionary for printed and offline version. Through the item Project administration, the administrator sends information e-mails to registered users, changes the project password and has direct access to adding new tasks and writing messages for both registered and

unregistered users.

5.3.9 Profile

The item Profile contains five sub-items: Own profile, Users, Statistics, Last queries and Log out. The item Own profile enables to edit user's personal data like, for example, the password, the e-mail address etc. The item Users displays the list of users where it is possible to view each user. The item Statistics displays statistical information on evolution of the dictionary along with the current number of headwords etc. The item displays also activities of the user / users enabling to monitor activities of individual users or these of the entire team. The user's activity graph helps to monitor editing activities of the user. It registers the number of records in the headword history – the user who edits headwords and saves them to the history, does not have any record in the graph. The item Last queries displays the list of last searched and viewed headwords ordered according to date of search; there is also the graph for the chosen month showing the number of searched headwords per day. Through the item Log out, the user logs out of the application. The last viewed headword is saved – after the user re-logs in to the application, the headword is displayed.

5.3.10 Information

The item Information contains the following sub-items: Information on the Dictionary, Information on pronunciation, Pronunciation of vowels, Pronunciation of consonants, List of phonemes, Generation of pronunciation, Abbreviations in alphabetical order, Abbreviations listed according to usage, Download, Authors, Licence. These sub-items are for the public and they are in the online version for the public. Nevertheless, they are also in the part for registered users.

5.3.11 Help

The last item Help contains four sub-items: Guide to the Dictionary, Tips for searching in the Dictionary, [Manual for Dictionary System DWS](#) and Manual for headword editing. The first two items are for both the public and the registered users. The item Guide to the Dictionary indicates in detail the way the data is set in the dictionary. The item Tips for searching in the Dictionary presents various ways of searching. [Manual for Dictionary System DWS](#) is this text file and it explains how to use the application Dictionary System. Manual for headword editing indicates the character and function of information and the way it is organized within the entry. Individual parts of the Manual are on the pages Edit the headword and Add a new headword and they are realized by clicking on the hyperlink.

6. Structure of the Dictionary

The dictionary has three structures. The first one is the printed version of the dictionary. The second one is the online version of the dictionary and the third one is the offline version of the dictionary designed for offline browsing in dictionary programs like GoldenDict.

The printed version is designed for book printing or for normal printing. It is limited by place (e.g. it is not possible to include complete declension / conjugation tables) and by medium (it is not possible to use sound files).

Unlike the printed one, the online version enables to get complex information because it is not limited by place. Therefore, the online version contains more information (e.g. sound files with headword pronunciation, complete declension / conjugation tables etc.) and it is also possible to find the headword by entering any word form.

The offline version enables to browse the dictionary without Internet connection. Its advantage is high-speed headword searching (depending on the computer speed), but it has some limits in comparison to the online version. For example, it does not contain declension / conjugation tables and it is not possible to search headwords by entering any word forms. However, unlike the printed version, it contains sound files and, through hyperlinks, it enables to move quickly among headwords.

At first, we introduce you the structure of the printed version of the dictionary, then, the structure of the online version of the dictionary and at last, the structure of the offline version of the dictionary.

6.1. Structure of the printed Dictionary

6.1.1 Structure of the Dictionary

1. Name, authors
2. Introductory information
3. Guide to the Dictionary
4. Phonetic rules
5. List of abbreviations
6. Entry part
7. Resources

8. List of the authors of photographs
9. Licence

6.1.2 Structure of the entry

1. headword (with additional marks)
2. headword variant
3. IPA transcription
4. grammar – part of speech
5. grammar – endings
6. grammar – additional information
7. etymology
8. meaning no. 1, no. 2, no. 3 ...
9. photograph, author of the photograph, licence

6.1.3 Structure of the meaning part

1. phrase group
2. indicator
3. synonym
4. synonym – detail
5. word phrase
6. secondary indicator
7. word phrase grammar
8. field category
9. language category
10. meaning – translation
11. meaning – detail
12. link to another headword
13. Latin term
14. antonym
15. example
16. example – translation

6.2. Structure of the online Dictionary

6.2.1 Structure of the Dictionary

1. Introductory information
2. Guide to the Dictionary
3. Tips for searching in the Dictionary
4. Phonetic rules
5. Generation of the IPA transcription
6. List of abbreviations
7. Entry part
8. Download the printed version and the offline version
9. Resources
10. Authors
11. Licence

6.2.2 Structure of the entry

1. Left column
2. Right column

Left column

1. headword (with additional marks)
2. headword variant
3. grammar – part of speech
4. grammar – endings
5. grammar – additional information
6. *frequency*
7. etymology
8. meaning no. 1, no. 2, no. 3 ...
9. *headword synonyms (hypertext)*
10. *headword in another headword examples (hypertext)*
11. photograph, author of the photograph, licence
12. *headword in usage categories (hypertext)*
13. *thematically related headwords (hypertext)*
14. *compounds derived from the headword (hypertext)*

Right column

1. IPA transcription
2. *headword pronunciation – sound file*
3. *tables of declension / conjugation forms*

6.2.3 Structure of the meaning part

1. phrase group
2. indicator
3. synonym (hypertext)
4. synonym – detail (hypertext)
5. word phrase
6. secondary indicator
7. word phrase grammar
8. field category
9. language category
10. meaning – translation
11. meaning – detail
12. link to another headword (hypertext)
13. Latin term
14. antonym (hypertext)
15. example
16. example – translation

6.2.4 Tables of declension and conjugation forms

1. Noun declension (no article and definite article)
2. Adjective declension and degrees of comparison
3. Pronoun declension
4. Numeral declension
5. Verb conjugation – active voice, middle voice, imperative mood, past participle, present participle declension
6. Degrees of comparison of adverbs

6.3. Structure of the offline Dictionary

6.3.1 Structure of the Dictionary

1. Information on the Dictionary and the authors, licence
2. Entry part

6.3.2 Structure of the entry

1. headword (with additional marks)
2. headword variant
3. IPA transcription
4. grammar – part of speech
5. grammar – endings
6. grammar – additional information
7. etymology
8. headword pronunciation – sound file
9. meaning no. 1, no. 2, no. 3 ...
10. photograph, author of the photograph, licence

6.3.3 Structure of the meaning part

1. phrase group
2. indicator
3. synonym (hypertext)
4. synonym – detail (hypertext)
5. word phrase
6. secondary indicator
7. word phrase grammar
8. field category
9. language category
10. meaning – translation
11. meaning – detail
12. link to another headword (hypertext)
13. Latin term
14. antonym (hypertext)
15. example

7. Creation of the Dictionary

7.1. Adding a new headword

Headwords are added through menu/word/add a new word. If the headword has homonyms, choose the headword number higher than 0, otherwise, do not change it (0 is not displayed). At the minimal level, enter only the headword and the headword number and press the button Add. Afterwards, the new headword can be edited on the page Headword editing.

If you as the project administrator decide to import headwords, it is necessary to import the list of headwords to both table ds_1_headword and ds_2_senses. The items keyword and num_keyword identify that it is the same headword. These two items are used in other tables and they identify the headword (e.g. tables with declensions / conjugations, sounds, photographs, history etc.). For importing the existing dictionary, use Manual for importing in menu/administration/import.

7.2. Headword editing

For headword editing, at first, it is necessary to find and view the headword. Then, only click on the hyperlink for headword editing. If the headword has several meanings, click on the meaning that you want to edit. If you want to edit information related to the headword (not to the meaning), click on any meaning, because general information are displayed for all meanings in the upper part of the editing form.

The headword preview is displayed in the lower part of the page. It is the preview in the online version of the Dictionary. In the Manual for headword editing, you will find information on how to fill in the fields in the editing form. A pop-up window with help information is displayed after clicking on the field name. Close the pop-up window by clicking next to it. To navigate among meanings, use hyperlinks in the headword preview. In the headword preview, the red cross is displayed for individual meanings. By clicking on the red cross, you delete the chosen meaning; the action requires your confirmation and it is irreversible. The small number next to the red cross is only for information.

7.3. Deleting the headword

Deleting the headword is an irreversible action. The headword can be deleted only by deleting all its meanings. Deleting the last headword meaning deletes the headword from all tables in the database – from tables with headwords, meanings, declensions / conjugations, sounds, photographs, word forms, history, work. At the same time, you delete all sound files which were identified with the headword (from the directory /audio/uploaded_files/) as well as picture files and miniature pictures which were uploaded by the user (and which are in the directory /images/uploaded_files/).

7.4. Manipulation with headwords

You can change the headword number through menu/sorting/change the headword number. You choose always between two options – decrement by 1 and increment by 1. The script controls the action in order to not overwrite the headword if it already exists. As the headword number together with the headword is the identifier, it is overwritten in the entire database; sound and picture files on the server are renamed as well.

7.5. Manipulation with meanings

After the blue arrows Up and Down are displayed, you can change the order of meanings through menu/sorting/change the order of meanings. Place the meaning in the position you want. If the meaning cannot be relocated, at first, it is necessary to go to menu/sorting/order meanings.

7.6. Declension / Conjugation

The Extension for Icelandic contains tables with declension of Icelandic nouns, adjectives and pronouns, conjugation of verbs and degrees of comparison of adjectives and adverbs. Scripts enable to generate declension or conjugation based on information in Grammar – parts of speech and Grammar – endings. Each declension / conjugation editing requires to save declension / conjugation to the word form database.

If the new project focuses on another language, it is necessary to edit declension / conjugation tables as well as pages with declensions / conjugations. It is important to know that the application includes two types of pages. The first one functions only in one table (e.g. for nouns). The second one functions in several tables and, moreover, it contains the information table

indicating which tables exist (adjective and verb tables).

Declension / conjugation is the part of the dictionary which is included only in the online version. The project administrator decides whether he will include declension / conjugation tables along with the function of searching according to word forms in the new project.

7.7. Entry history

The entry history registers the current form of the entry together with the date of editing and the name of the user. The entry history enables to register the evolution of entry editing. Saving to the history also launches the script for updating compounds. Therefore, the rule is: after each entry editing, save all changes to the entry history.

8. Publishing

8.1. Publishing the online version

The online version is available for the public during all the dictionary creation process. Therefore, it is unnecessary to publish the online version.

8.2. Publishing the printed version - LaTeX → PDF

At first, generate the dictionary to the LaTeX file and then, convert it to PDF format from which you can print the dictionary. LaTeX is freely available and it works under the following systems: Windows (MikTeX¹³), Linux, MacOS (MacTeX¹⁴). In the environment of Ubuntu 11.04 the Natty Narwhal¹⁵, we will show you how to publish the printed version.

8.2.1 Preparing the environment

Preparing the LaTeX environment and the required libraries. Through the Synaptic Package Manager¹⁶ (or via command line), install the libraries and the related libraries.

texlive (texlive-base, texlive-latex-base, texlive-fonts-recommended, tipa, texlive-luatex) texlive-latex-extra

13 <http://miktex.org/>

14 <http://www.tug.org/mactex/2011/>

15 <http://releases.ubuntu.com/natty/>

16 <http://wiki.ubuntu.cz/Synaptic>

texlive-lang-czechslovak (for Czech)
texlive-lang-other (for Icelandic)
texlive-fonts-extra

Then, it is necessary to install these LaTeX libraries (if they were not installed during the installation of texlive-base and others):

fontenc (libl_{at}ex-encode-perl)
fix2col
l_{at}exsym

8.2.2 Generating the printed version

As soon as the LaTeX environment is prepared, the application controls whether the files **/tmp/dictionary.tex** and **/tmp/dictionary_test.tex** have rights set for writing (chmod 666).

Then, generate the test file **/tmp/dictionary_test.tex** in Administration/Dictionary publication and later, repeat the same process in complete dictionary generating to the file **/tmp/dictionary.tex**.

The file **dictionary_test.tex** / **dictionary.tex** will be sent to your e-mail account or you will find it on the server in the directory **/tmp/dictionary_test.tex** / **/tmp/dictionary.tex**. Save the file to your computer to the directory **/tmp/**.

8.2.3 Copying the photographs from the server to the local computer

For the correct conversion of the file **dictionary.tex** to the PDF file, it is necessary to copy from the server photographs in the directories **/images/biolib/** and **/images/uploaded_files/**. Keep the same path. In the generated file **/tmp/dictionary.tex**, edit real absolute path to photographs in your computer. E.g. `\graphicspath{{/home/autor/icelandic-czech_dictionary/images/biolib/}{/home/autor/icelandic-czech_dictionary/images/uploaded_files/}}`

LaTeX needs to know the path to photographs, if it does not find it, it reports error and does not display the photograph.

If you want to print without photographs, edit the file **fileprint.php** (part `action==latex`).

8.2.4 Conversion from .tex to .pdf

Open the terminal. Via command line, navigate to the directory **/tmp/**. By entering the command `pdflatex`, convert **dictionary_test.tex** / **dictionary.tex** to **dictionary_test.pdf** /

dictionary.pdf.

Command:

```
pdflatex dictionary.tex / pdflatex dictionary_test.tex
```

Control the PDF file, whether its formatting is correct, whether there are all photographs etc.

Control the Guide to the Dictionary – if you want to edit the introductory part, edit the file **/language/cz/dictionary_instructions.tex** and generate the file again.

You have successfully prepared the PDF file suitable for normal printing. This PDF file is used in the following section concerning the PDF file for the book printing.

8.2.5 Book format

To prepare the PDF file for book printing (two pages on one sheet, 20 pages bound in the booklet form), use Java scripts [Multivalent](#). Copy the file /tmp/Multivalent.jar to the directory /tmp in the local computer or download the previous version (the latest one does not contain Impose tools) and save it to the directory /tmp

see <http://stackoverflow.com/questions/465271/gluing-imposition-pdf-documents>

Now, the directory /tmp should contain Multivalent.jar (or something similar). It is necessary to have the latest version of Java. If it does not work, visit forums - [e.g. here](#)

Via command line, navigate to the directory **/tmp/** in the local computer. Then, enter the following command to the command line, and, at the same time, set up the correct path to the directory **/tmp/** and, eventually, edit the name of the PDF file that you want to convert.

```
java -classpath /home/autor/your-dictionary/tmp/Multivalent.jar tool.pdf.Impose -dim 2x1 -verbose -paper-size "42.2x28cm" -layout "1,20,19,2,3,18,17,4,5,16,15,6,7,14,13,8,9,12,11,10" dictionary.pdf  
(Multivalent.jar and dictionary.pdf are in the same directory /tmp/)
```

(it prints two pages on one sheet, reverses it and typesets it according to the pattern for book format (20 original pages in each booklet) – in duplex printing mode, there are 5 pages)

8.3. Publishing the offline version - DSL format (for GoldenDict)

Generating the dictionary to DSL format (for offline browsing in the dictionary program GoldenDict, Stardict). In brief, it is necessary to:

- a) generate the dictionary to DSL format
- b) LINUX/Mac users – to convert the file from UTF-8 to UTF-16

You will find information on how to convert the file from UTF-8 to UTF-16 in the command line

on [this web page](#)

c) create the directory of the offline dictionary and to copy the converted file to it

d) copy all the files from **/audio/uploaded_files/** to the local computer and, subsequently, to convert them from .mp3 format to .wav format, for example, through the program [SoundConverter \(linux\)](#) and to copy them all to the directory of the offline dictionary

e) copy image views in **/images/biolib/** to the local computer and to copy these .jpg files to the directory of the offline dictionary

f) zip the entire directory (prepare it for download)

8.3.1 Test of the export

Control whether the file **/tmp/dictionary_test.dsl** has rights set for writing (chmod 666). By clicking on the link, you export a small part (1 000 headwords) to the file **/tmp/dictionary_test.dsl**. Install [GoldenDict](#) and control whether it works properly.

8.3.2 Export

Control whether the file **/tmp/dictionary.dsl** has rights set for writing (chmod 666). By clicking on the link, you export the entire dictionary to the file **/tmp/dictionary.dsl**. Attention – the process may take up to several tens of minutes – do not close the web browser.

All information on dictionary creation in DSL format are [here](#).

The result is the zipped file containing the directory with the generated .dsl file, picture views and sound files in .wav format.

9. Structure of the application

This section is for the project administrators who need to edit the application and to adapt it to their requirements; it also introduces the parts of the application Dictionary System.

9.1. Structure of the MySQL database

ban_ip - table contains a list of forbidden IP addresses

country_info – table contains a list of IP addresses with the names of the countries

ds_1_headword - table contains a list of headwords with information related to them

ds_2_senses - table contains a list of headword meanings with information related to them

ds_abb_grammar - table contains a list of grammatical abbreviations for parts of speech

ds_abb_grammar_additional - table contains a list of additional grammatical abbreviations

ds_bibliography - table contains a list of resources

ds_dec_adj_info - table contains information on degrees of comparison of adjectives – whether their forms exist or not

ds_dec_adj_3 - table with information on declension of adjectives - 3rd degree

ds_dec_adj_2 - table with information on declension adjectives - 2nd degree

ds_dec_adj_1 - table with information on declension adjectives - 1st degree

ds_biolib_full - table with complete information on photographs from Biolib.cz

ds_biolib - table with information on photographs from Biolib.cz

ds_dec_v_info - table with information on verb conjugation indicating whether there are verb forms in the active voice, the middle voice, the past participle; and past participle declension

ds_dev_info - table contains messages for unregistered users

ds_dec_v_4 - table with information on verb conjugation – past participle declension

ds_dec_v_3 - table with information on verb conjugation – the imperative mood and the past participle form

ds_dec_v_2 - table with information on verb conjugation – the middle voice

ds_dec_v_1 - table with information on verb conjugation – the active voice

ds_dec_pron_pers - table with information on personal pronoun declension

ds_dec_pron - table with information on pronoun declension

ds_dec_noun - table with information on noun declension

ds_dec_adv - table with information on degrees of comparison of adverbs

ds_images - tables with information on photographs which were uploaded together with the name of the author, the licence and their orientation (horizontal, vertical)

ds_history - table contains the headword history

ds_settings - table contains the project password

ds_sound - table contains information on sound file names

ds_spam_notice - table recording spam or other attacks on web applications (attempts to log in etc.)

ds_spec_field - table with abbreviations of field categories

ds_spec_usage - table with abbreviations of language categories

ds_queries - table recording searched word forms and IP addresses with date of search

ds_phonemes - table contains detailed rules of pronunciation of Icelandic phonemes

ds_message - table contains messages for registered users

ds_users - table contains a list of registered users with information on them (user passwords are encrypted)

ds_wordform - table contains word forms of headwords

ds_work - table recording user activities (e.g. login, editing and deleting headwords etc.)

ds_usage_category - table contains usage categories

ds_todo - table contains a list of scheduled tasks

pron_wow - table contains rules of vowel pronunciation

pron_cons - table contains rules of consonant pronunciation

phonems_list - table contains a list of phonemes with examples

9.2. Structure of the directories

The home directory contains all the files which are necessary for the correct functioning of the application.

The home directory also contains the directories **/audio/**, **/backup/**, **/download/**, **/images/**, **/import/**, **/install/**, **/language/**, **/scripts/**, **/tmp/**.

The directory **/audio/** contains the directories **/audio-player/** and **/uploaded_files/**. The directory **/audio/audio-player/** contains Flash audio-player for replaying sound files. The directory **/audio/uploaded_files/** contains uploaded sound files.

The directory **/backup/** contains the MySQL back-up file.

The directory **/download/** loads the files to download (e.g. the printed version of the Dictionary etc.).

The directory **/images/** contains the sub-directories **/autosuggest/** (contains pictures for viewing help while choosing the part of speech etc.), **/biolib/** (contains picture views with the sub-directory **/full/** containing original photographs from Biolib.cz), **/flags/** (contains the national flags of the states – they help the project administrator to add the new language of the application), **/help/** (contains pictures used in manuals and guides), **/securimage/** (contains additional pictures for the security picture on the page Download), **/uploaded_files/** (contains uploaded photographs – the miniature ones and the original ones).

The directory **/import/** contains the additional files for importing the created dictionary to the application.

The directory **/install/** contains the installation files. After finishing the installation process, it is recommended to delete the directory.

The directory **/language/** contains the language files in the following manner: the sub-

directories are always marked with the abbreviation of the language (e.g. /language/cz/ or /language/en/) and the files in these directories are named the same way (e.g. /language/cz/language.php and /language/en/language.php), they contain the same variables and they differ only by their translations. Therefore, a new language can be easily added by adding a new directory with the abbreviation of the national language, by copying the files from the directory for English, for example, and by translating language variables in the language files.

The directory **/scripts/** contains additional scripts which are important for headword formatting and other editing activities. E.g. the files beginning with view_ enable to view the headword.

After publishing the dictionary, in the directory **/tmp/**, there will be 4 files: the test and complete file for the printed version and the test and complete file for the offline version.

9.3. Structure of the web page

The structure of the web page used in the application is the same for all the files in the basic directory. After understanding the structure, the project administrator is able to write new pages and, thereby, to enrich his own project. All the files end in .php. We will show you the structure of the application file on the following example (the file licence.php containing the licence of the Dictionary).

```
<?php
//=====
// Dictionary system. Web-based application for development of bilingual dictionaries
// Version: 1.0
// Copyright (c) Ales Chejn, hvalur.org 2011
// All rights reserved
// This program is free software: you can redistribute it and/or modify
// it under the terms of the GNU General Public License as published by
// the Free Software Foundation, either version 3 of the License, or
// (at your option) any later version.
// This program is distributed in the hope that it will be useful,
// but WITHOUT ANY WARRANTY; without even the implied warranty of
// MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the
// GNU General Public License for more details.
// You should have received a copy of the GNU General Public License
```

```

// along with this program. If not, see <http://www.gnu.org/licenses/>.
// For support contact us at www.hvalur.org
//=====
ini_set('arg_separator.output','&');
include './start.php';
include './head.php';
?>
<body onload="setfocus ()">
<div id="wrapper">
<?php include 'header.php';
include 'menu.php';
echo $MAIN_MENU;
?>
<div id="content">
<div class="left_huge">
<h2><?=$lang_license_1;?></h2>

<br>
<?=$lang_license_2;?>
<br><br>
<?php if ($_SESSION["lang"]=='cz') { ?>
<a href="http://www.gnugpl.cz/v3/" target="_blank">
<?php } else { ?>
<a href="http://www.gnu.org/licenses/gpl.html" target="_blank">
<?php } ?>
<?=$lang_license_3?>
</a>
</div>
<div style="clear: both;">&nbsp;</div>
</div>
<div id="footer">
<?=$lang_footer;?>
</div></div><?php
include ('./html_end.php');
?>

```

As you can see, the file loads another files for displaying the content. These are: start.php, head.php, header.php, menu.php, html_end.php (in that order). In the file start.php, there are three basic files with which MySQL operates (connection.php, /scripts/mysqlclass.php, /scripts/query_function.php) and then, the IP address is controlled and the language files are loaded according to the chosen language. The file head.php contains the html head of the document. Sometimes, the file head_s.php is used instead of the file head.php. The difference between them is that the file head_s.php does not contain the closing tag of head tag (</head>) and, therefore, another javascripts can be added. The file header.php displays the name of the dictionary and the entire search field with all parts as well as the language menu. The file menu.php displays the navigation menu. The file html_end.php displays javascript for language selection and closing tags for body and html tags (</body>, </html>).

Here, you can see the sample of the final page in pseudo-source code:

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"
"http://www.w3.org/TR/html4/loose.dtd">
<html>
<head>
HTML HEAD
</head>
<body onload="setfocus ()">
<div id="wrapper">
<div id="header">
<div id="logo">
<div id="lang_selector">LANGUAGE SELECTION </div>
<h1>NAME OF THE DICTIONARY</h1>
SEARCH FIELD
</div>
</div>
<div class="menu_edit">MENU</div>
<div id="content">
<div class="left_huge">
<h2>HEADING</h2>
TEXT
</div>
<div style="clear: both;">&nbsp;</div>
</div>
<div id="footer">
FOOTER
</div>
</div>
</body>
</html>
```


The design of the page is realized by using Cascading Style Sheets (css). CSS styles are defined in the file style.css, located in the home directory. The structure of the design is the following:

```
<!DOCTYPE html PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN" "http://www.w3.org/TR/html4/loose.dtd">
<html>
  <head>...</head>
  <body onload="set focus ()">
    <div id="wrapper">
      <div id="header">...</div>
      <div class="menu_edit">...</div>
      <div id="content">...</div>
      <div id="footer">...</div>
    </div>
    <script type="text/javascript">...</script>
  </body>
</html>
```

As you can see, the marker wrapper wraps another css markers or classes, namely header (it displays the name of the dictionary and the search field), menu_edit (it displays the menu), content (it displays the content of the dictionary), footer (it displays the footer).