

Electronic Dictionary Notebook - Enhancing the Language Learning Process

Zhelyazko Zhelev and Vera Stoyanova

Universify of Sofia "St. Kliment Ohridski"
zzhelev@gmail.com, vera.stoyanova@yahoo.com

Abstract. Upon its entry to the European Union, Bulgaria was presented with the challenge of intensive interaction with foreigners, so, speaking fluently several languages became a must. This paper presents a software application aiming to improve the process of language learning. The proposed solution, specializing in vocabulary augmentation, supports entry of word meanings, image suggestions, conjugation, examinations and statistics, searching and filtering. Each feature is carefully selected to target a particular need while studying within a language course or on one's own. This paper also demonstrates how freely available resources on the Internet and a little programming effort can be combined to produce a useful tool for daily use.

1 Introduction

Globalization nowadays presents a challenge to the individual that needs to interact with an international environment. Whether a participant in a social network or a businessman that wants to trade on the international market, or just a student, one needs to read, write, and communicate in English, French, German, Spanish, Italian, etc. There are many resources, articles and books in different languages available on the global network, but in order to take any advantage of that, one should be able to know the particular language. That is why studying foreign languages has become a key factor for improving one's qualifications.

Information technology and the Internet have entirely changed the way one studies foreign languages. Paper books have been replaced by desktop software, online courses, international internet communities, teachers working from home with audio-video connection, etc. Current systems emphasize on studying in the convenience of one's own home and during their free time.

In spite of this global trend however, the language education in Bulgaria is still done mostly with paper textbooks. Commonly used language related software is available in some online and desktop dictionaries such as SADictionary and AFrancophone. Few software solutions are internationalized in Bulgarian and have freeware or shareware versions. The proposed in this paper, Electronic Dictionary Notebook (EDN) is a software alternative to the ordinary dictionary notebook used during a foreign language education course. Its purpose is not only to match the functions of an ordinary dictionary notebook, but also to enhance the whole process of vocabulary expansion. It is distributed as freeware and is available for download at <http://edn.leosoftbg.com/>. The target audience

is Bulgarians who are taking language courses or studying at home. The Cyrillic user interface helps, especially the ones that still speak only Bulgarian.

In this paper the need of free software for studying foreign languages and its features are described. In section two, a few similar existing products on the market are presented. Section three summarizes the main application functionality and some core features. Section four presents the EDN software architecture and the used technologies. Finally, as a conclusion, all benefits are summarized, and some directions for future development are outlined.

2 Existing software solutions

Before the design and implementation of EDN, several existing software solutions were reviewed and evaluated.

1. Byki (<http://www.byki.com/>)
2. MLN – My Language Notebook (<http://www.mylanguagenotebook.com/>)
3. jVLT - Vocabulary Learning Tool (<http://jvlt.sourceforge.net/doc/default/doc.html>)

2.1 Byki

Byki is a product created with the idea that to master a foreign language, one should learn as many words from it as possible. The authors of this application claim that the learning process with *Byki* is much more effective compared to the one with other products of similar functionality. *Byki* has both a free limited version and a commercial version with more functionality packed in it. Its main features can be summarized in the following way:

1. Learning foreign words with the help of flash cards
2. Separation of vocabulary in different categories (e.g. colours, plants, etc.)
3. Downloading and sharing of flash card files via the author's website (in the commercial version)
4. Ability to enter one's own vocabulary and building flash cards with it (in the commercial version)
5. Graphical representation of various statistics collected in the process of learning with the flash cards
6. Audio pronunciation of vocabulary
7. Multi-user support with a separate application context for each user
8. Support for 64 different languages

Some of the weaknesses of *Byki* from a functional standpoint are the following:

1. There is a separate version for every supported foreign language
2. Data can be imported from and exported to a proprietary file format
3. When entering new vocabulary, it can only be organized into a linear list used for the generation of flash cards

EDN's functionality is identical in some ways to the one of *Byki*, but with less limitations. *Byki*'s free version does not allow the entering of new vocabulary;

such can be downloaded. However, shared vocabulary is limited to some number of commercial users willing to do this. Compared to this, EDN supports the organization of words into sessions, which are easily shareable via a variety of file formats. EDN is also packed with a great deal of other features like a word game, machine translation of text, a dictation player and recorder, various filters, etc.

2.2 My Language Notebook

My Language Notebook is a Java-based application for vocabulary expansion. It supports the organization of words into different categories. The application is free for personal use, however, a license is required for commercial users. Its core features are outlined below:

1. Ability to enter own vocabulary and organize it in user-defined categories
2. Sharing of application files from the application's website
3. Audio recording and playing of vocabulary and text paragraphs
4. Support for six languages

There are also several drawbacks from a functional point of view:

1. The application's sole purpose seems to be to supply a storage for the user's vocabulary, i.e. there are no word games, flash cards, or other features to aid the learning process
2. The application uses its own proprietary format for the importation and exportation of vocabulary
3. The application lacks any alternative views of the data that resides in its files, e.g. a tree structure of categories, flash cards, or something else

Basic functionality that exists in EDN is also missing in *My Language Notebook*, just like it was with *Byki*: exporting to and importing from several file formats, alternative views of user-entered vocabulary, recording and playing dictations, and word games that help the user in the learning process.

2.3 jVLT - Vocabulary Learning Tool

jVLT is another Java-based application, which puts the emphasis in the learning process on the vocabulary and not the grammar. It is distributed as freeware with the following functional characteristics:

1. Entering of new vocabulary and organizing it in user-predefined projects
2. Exporting of vocabulary to the CSV and HTML file formats
3. Linking of already entered vocabulary with external audio files that supposedly contain the pronunciation of the selected word or phrase
4. Advanced searching of vocabulary with the help of some user-selected filters
5. A word game from a pre-selected set of words

Just like with the other applications described above, there are a number of shortcomings with *jVLT*:

1. There is not an alternative view of vocabulary such as a tree view or flash cards
2. There is not a way to record the pronunciation of a foreign word
3. Using the application requires a heavy use of the computer mouse while going through a series of displays
4. There is not a way to modify the projects once they have been created.
5. The application interface is not consistent in the different screens it uses

Once again, like with the other applications described above, EDN's feature pack surpasses the one of *jVLT*. It has a consistent and intuitive user interface throughout the various screens. The vocabulary can be visualized in a treelike fashion or as a list. It can record and play WAV files and perform machine translation of text thanks to its integration to one of Google's public web services.

Table 1 below summarizes the features of fore-mentioned applications.

Table 1

Feature	Byki	MLN	jVLN	EDN
Organizing words in logical groups	yes	yes	yes	yes
Support of different views	no	no	no	yes
Learning sessions (flash cards, etc.)	yes	no	no	yes
Sharing word-lists through a common website	yes	yes	yes	no
Sharing word-lists with export/import of files (without common website)	yes	yes	yes	yes
Insertion of words (uses Google Dictionary to aid user)	yes	yes	yes	yes
Self-testing	yes	no	yes	yes
Statistics	yes	no	yes	yes
Audio recording of a word	yes	yes	no	no
Audio play of a word	yes	yes	yes	no
Multi-language support	yes	yes	yes	yes
Multi-user interface	yes	no	no	no
Filtering	no	no	yes	yes
Machine Translation (via Google Translate)	no	no	no	yes
Suggestions for meanings during word insert	no	no	no	yes
Association of a word with images (via Google Image Search)	no	no	no	yes
Recording dictations	no	no	no	yes
Replaying dictations	no	no	no	yes
Conjugation	no	no	no	yes
Free for use	Express edition only	For home users	no	yes

As it can be seen from the *Table 1*, the available, on the market software solutions have many useful features, but they also lack some important ones like dictations, conjugation, and word suggestions. Also, the free versions usually come with limited functionality. On the contrary, EDN is open source and free-ware and is still under active development. The features that it currently does not have are still under construction and will be soon available.

3 EDN Core Features

The presence of the current features of EDN was influenced by some careful studying of the foreign language learning process. In addition to this, similar to EDN products were reviewed and compared to it. The recommendations of several foreign language teachers about the set of features to be included in the application were also taken into consideration. A prototype of EDN was tested in a real language course, which provided useful information about its functionality. Some of the innovative features, as it becomes obvious from the previous section, are machine translation of text, association of vocabulary with images, dictations playing, dual view of the vocabulary in the application, and a few others. Most of the key features are common for all target foreign languages – insertion of a word, doing an exam, etc. However some of them cannot be granted by the application, as the conjugation for example. It relies on an external conjugator which may not be always available. As the software is still in development, the presented list is not complete and new features are already under evaluation and about to be included into the future versions.

Here follow the key features of EDN.

3.1 Inserting a Word

Studying a language has two main essential parts: vocabulary and grammar. Words are studied sequentially and are normally related to a paragraph of text or a lesson. One starts with simple words like "table", "pen", "wall", "door", and then gradually goes to abstractions like "freedom", "equality" and "brotherhood". Recording a word is generally a feature for an actual paper dictionary notebook. However, this feature is enhanced by EDN by utilizing an existing Internet connection to the public web service *Google Dictionary* and making suggestions for the correct meanings.

3.2 Organizing Words into Lessons/Sessions

Professionals agree that a word should not be studied in isolation from its context. Some words and phrases do not even have direct translation, and they have to be comprehended within its native environment. So words in EDN are organized in groups named "sessions". In an environment of an ongoing language course, they normally represent the lessons.

The display of the words within EDN is not alphabetical, but follows the order of the learning process. There is also a tree-view that shows sessions as nodes and words as leaves.

3.3 Searching and Filtering

Normally a person needs about 2000-3000 words in order to start speaking freely. When the vocabulary grows, some words are forgotten. One of the most frustrating things about the paper dictionary notebook is that manual, sequential searching is really slow and ineffective. With EDN users can easily find any forgotten word and also refer to the rest of the words from the same lesson.

The filtering feature in EDN is a way to show only a certain subset of words in the application at a given time. Thus, users can perform other actions from the application only to the selected subset of words, i.e. start a test session, export, etc. Currently this functionality extends to filtering by entry date, filtering by sessions, filtering by least known words (based on statistics from the exams).

3.4 Export to / Import from the XML Format

This feature is specially developed for people that take language courses together. It always happens to miss a lesson once or twice. Then the solution is to ask a fellow student for help and copy the words. Sharing words is made really easy with the export/import feature. Once the words are imported into the notebook, they can be distributed freely among the users of EDN. Another good use of this feature is when the teacher already has the words exported, and offers them to the students, who in this way can concentrate on grammar exercises instead of writing down every word. This would also avoid misspelling the word and give the study groups extra time for other language activities.

3.5 Export to PDF and Microsoft Excel

PDF export comes in two flavours here:

1. Printing out the words in a genuine dictionary notebook paper format. Genuine paper notebook format is not provided to replace the actual paper notebook, but to aid the student even more in learning the vocabulary. After taking a course, one can print out all the words and keep the printout in a folder as an archive. Parts of the notebook can be printed out at any time.
2. Easy preparation of flip cards. Flip cards are an old and proven method of learning new vocabulary. Manual preparation of flip cards, however, is a time consuming job. EDN handles this task easily and prints out a PDF, in which flip cards are generated.

4 Words Exam

A words exam is one of the main features of EDN. Trying to guess the meaning of the word by viewing it is the main idea of the flip cards. Flip cards, however, do not keep the count of the failed and the successful attempts like EDN does. The process of remembering a word takes several failed test attempts, and then, the user starts guessing the word each time it appears within the test. Testing produces statistics information explained in the next section.

4.1 Statistics

Statistics is also a key feature. Based on the results of the word exam, each word is ranked into GREEN, YELLOW, RED, WHITE groups (green – well known, yellow - not very well known, red – not known or hardly known, white – no statistics yet). Colored icons are displayed along with the words in the list view and tree view, so that the user can have a general idea about his actual vocabulary.

4.2 Conjugation

Most of the Latin languages have three groups of verbs with different conjugation, and often, the most used verbs are irregular. In order to speak fluently, one needs to study the conjugation and the exceptions. There are free resources for this on the Internet, and EDN refers to them when a user wants to conjugate a verb. The following table summarizes the most common languages studied in country-regionplaceBulgaria, and which URLs are being used in the conjugation process.

Table 2

Language	URL
English	http://conjugator.reverso.net/conjugation-english-verb-{} .html
German	http://conjugator.reverso.net/conjugation-german-verb-{} .html
French	http://www.wordreference.com/conj/FRverbs.asp?v={0}
Spanish	http://www.wordreference.com/conj/ESverbs.asp?v={0}
Italian	http://www.wordreference.com/conj/ITverbs.asp?v={0}

EDN can conjugate a verb on a right click using configurable URL like the listed above. The ”{0}” parameter is replaced with the verb to conjugate.

4.3 Word-to-Image Associations

Many people have a really good visual memory. They can easily remember what a word’s meaning is when an object is visualized, rather than simply translated. Google offers a convenient web service for searching for images by a keyword. Most common words, like ”*desk*”, ”*chair*”, or ”*tomato*”, have very accurate images in Google searches. EDN takes advantage of this functionality, and lets the users browse a list of images retrieved from the Internet from within the application.

4.4 Dictations

One of the difficulties that exist when studying a foreign language is writing correctly. This ability is trained with dictations. However, without a teacher, it is difficult to make a dictation because this requires the help of someone, who speaks correctly, slowly, and has a clear diction. Having the ability to conduct or create dictations in EDN presents an alternative to the regular teacher-led

dictations. The users can record their own voice and replay it several times. A feature planned for the future improvement of EDN is the import/export of dictations (text and sound), so that a teacher or a native speaker can record a dictation and spread it around. This would overcome the pronunciation problems one may have.

5 Architecture and Used Technologies

EDN uses well-known technologies and publicly accessible resources. The application architecture follows the MVC model, and most of the features are implemented with the help of some Apache libraries, public RESTful web services, and a simple file-based database. The application is written in Java 1.6 and uses the following libraries and technologies (*Figure 1*): Java Swing, MVC architecture built upon the Spring framework, XML, XSLT, XSL-FO, JAXB and Apache FOP Engine, POI, along with Hibernate, HSQL, RESTful Web Services with JSON as the data transport format, and the Java Sound API.

The application is designed such that it can easily be extended. It consists of a core part that wraps the basic functionality, and modules that do not affect its runtime. *Figure 1* illustrates EDN’s internal structure. The core part consists of the classes that generate the visual layout (based on Java Swing), the data model, the persistence layer (implemented with XML files and HSQL), and a few controller classes that handle the communication between the GUI and the persistence layer.

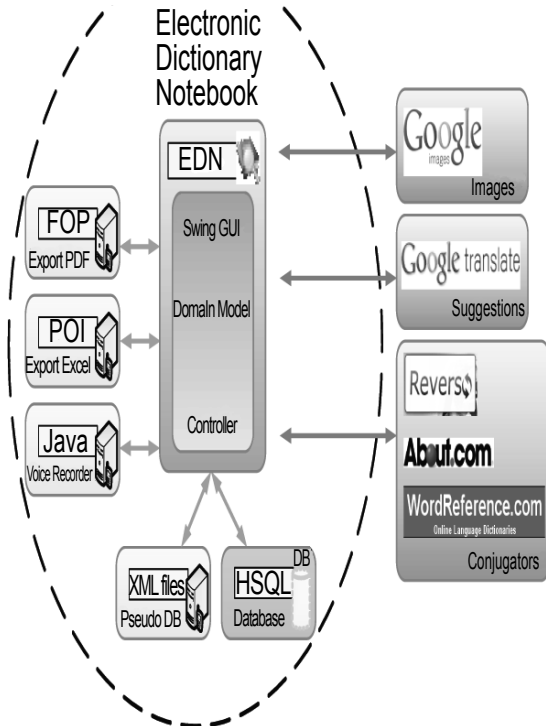


Fig. 1. EDN.

5.1 Modules

Module for Conducting and Recording Dictations

The audio processing library consists of classes, whose functionality is based on the Java Sound API, which is part of Java 1.6. The module is designed in such a manner that users can organize their dictations all by themselves. Currently, the only supported audio format is WAV, which was selected for its platform independence. Users can record their dictations and manage them from within the application.

Module for Machine Translation of Text

This module wraps the functionality of the public RESTful web service *Google Translate*. Users can translate any text from one language to another. Currently, the supported number of languages is 53. Of course, one should keep in mind that this is a machine translation of random text, and no arguments should be presented about the quality. The functionality of the module, however, is powerful enough to provide an understandable translation of any random text.

Module for Discovering Word-to-Image Associations

This part of the application relies on the RESTful web service *Google Image Search*. What it does is pretty straight forward. It uses the API provided by the web service to search the Google database for images that are related to the word or phrase the user has selected to see the word-to-image association for.

Module for Adding Vocabulary

This is probably by far the most important module in the application. If this module was not created easy and enjoyable to use, then people would probably give up on using the EDN application. Since a lot of the EDN functionality depends on the vocabulary entered by the user, it was decided that this should happen with a minimum amount of steps. Therefore a lot of effort was put in the visual layout and keyboard navigation in it. Moving around the visual components follows the conventions employed by the "Tab" key. In addition to this, *Google Dictionary's* RESTful web service was integrated into the module. When new words are entered, their meanings are automatically discovered and offered to the end user. Thus, people save a lot of time from typing several meanings to a single foreign word or phrase.

Export / Import Module

The export / import module works with the help of XML and XSLT transformations. Anything to be exported is converted into or comes from an XML file with the help of JAXB (an XML-to-object mapping library), which generates Java classes based on XML schemas and then serializes to or deserializes from Java objects. This allows for the export/import of the application state or its entire data set.

The export-to-PDF feature is implemented with the Apache FOP (Formatting

Objects Processor). Basically, it relies on XSL transformations to convert an XML file to a PDF file.

It is also possible to export into a Microsoft Excel file. This implementation is developed with the help of the Apache POI Java API.

Word Exam Module

This is a no-frills module that tests the users about their knowledge of randomly-chosen words from the data set of the application considering any filters a user may have applied to the data set. This module is also the one that feed data into the statistics functionality.

5.2 Deployment

EDN is distributed in a Java Archive (JAR) file. Here is the layout of the JAR file:

```
edn-2.0.0-release.dir /root directory/
| -edn-2.0.0.jar /the main EDN jar/
| -settings /settings files/
| -readme /readme file/
| -lib /libraries used by EDN/
| -fonts /Cyrillic fonts/
```

The application is platform independent and is started from the *edn-
<<version>>.jar* file.

6 Conclusion

The purpose of EDN from the very beginning was to provide an easy, intuitive, and pleasant to use application for the process of studying foreign languages. It uses contemporary audio-visual and network technologies which put a great deal of features at the fingertips of the users, but at the same time, it also hides the complexity of the combination of such technologies. The advantages that come with the application are numerous: every user can decide on his/her own pace of learning, users do not become demotivated from being slow-learners, and so on. Most software applications related to the process of studying a foreign language tend to attempt to replace the actual language courses. EDN, however, is designed to support the language courses, and not to replace them. Following this line of thought, the main direction for future work is related to the improvement of the existing features and adding more user-friendly features that minimize the end-user's efforts in using the application, but at the same time maximize the end-user's satisfaction. Of course, there is a lot of room left for the implementation and integration of different study-games and grammatical exercises and tests. Covering both lexical and grammar area will allow EDN to handle an even greater part of the process of learning a foreign language.

As of this writing, the software is being tested by a small group of people, studying English (upper-intermediate level). There is a bug tracking system

open for contribution of ideas for new features, as well as for bug reports. It can be found at <http://mantis.leosoftbg.com>. Several cycles of testing will be done with different groups, different languages and levels in order to adjust the software to be user-friendly and bug-free as much as possible before showing it to the general public.

Acknowledgements. We would like to express our gratitude and thanks to our Associate Professor Sylvia Ilieva from the Universia of Sofia, who shared her ideas with us and gave us her time and comments. Also, EDN could not have been written the way it is without all the efforts invested in it by our colleagues Irina Marudina and Elitsa Yaneva, also from the University of Sofia.

References

1. Byki: What is it? Learn Spanish, French, German, Italian and over 70 languages with Byki software. Transparent Language. Web. 05 Dec. 2009. <http://www.byki.com/byki_descr.html>
2. My Language Notebook. My Language Notebook. JMGS Software. Web. 05 Dec. 2009. <<http://www.mylanguagenotebook.com/>>
3. jVLT - Documentation. jVLT - a vocabulary learning tool. Henning Burchardt. Web. 05 Dec. 2009. <<http://jvlt.sourceforge.net/doc/default/doc.html>>
4. Trail: Sound (The Java Tutorials). Developer Resources for Java Technology. Web. 12 Dec. 2009. <<http://java.sun.com/docs/books/tutorial/sound/>>
5. Kayne, R. What is a WAV File? wiseGEEK. Conjecture Corporation, 14 11 2009 r. Web. 5 Dec. 2009. <<http://www.wisegeek.com/what-is-a-wav-file.htm>>
6. Fielding, Roy Thomas. Architectural Styles and the Design of Network-based Software Architectures, Doctoral dissertation. Irvine : University of California, 2000
7. Richardson, Leonard Ruby, Sam. RESTful Web Services. Sebastopol, USA: O'Reilly, 2007. 0596529260
8. Costello, Roger L. Building Web Services the REST Way. xFront. Web. 18 Dec. 2009. <<http://www.xfront.com/REST-Web-Services.html>>
9. Pautasso, Cesare, Zimmermann, Olaf Leymann, Frank. RESTful Web Services vs. Big Web Services: Making the Right Architectural Decision. CityplaceBeijing, country-regionChina : 17th International Web Conference, 2008 r.
10. Aziz, Atif Mitchell, Scott. An Introduction to JavaScript Object Notation (JSON) in JavaScript and .NET. MSDN: Microsoft Development, MSDN Subscriptions, Resources, and More. 1 Feb 2007 r.
11. Spring. Spring source community. Web. 23 May 2010. <<http://www.springsource.org/about>>
12. Reference for XML related languages Zvon.org. Web. 23 May 2010. <http://zvon.org/index.php?nav_id=references&mime=html>