

# Babel: A collaborative language learning system

Javier Albusac, Carlos Gonzalez-Morcillo Luis Jimenez-Linares

Oreto Research Group, University of Castilla-La Mancha (Spain)

{JavierAlonso.Albusac, Carlos.Gonzalez,  
Luis.Jimenez}@uclm.es

<http://oreto.inf-cr.uclm.es/>

**Resumen** Nowadays, plenty of people from very different countries are involved in the Free Software community. Therefore, communication among the individuals involves a big deal that we should take care of. Currently, Most of the existing language learning applications on the Internet are e-Learning tools which do not allow users to help and to share their knowledge with each other. These are a sort of applications where the teacher is the only one who produces information and users just can consume it. Babel is a free web-based content management system which provides users with synchronous and asynchronous communication tools and with the benefits of learning in decentralized communities of users. Our proposal is a common framework to learn and to practice different languages which may be used and adapted in any educational institution as the system is distributed under the GNU/GPL license.

**Palabras clave:** Free software; Language learning; Content management; Decentralized communities

## 1. Introduction

Over the last ten years, the use of new technologies and Internet have broken the communication barriers among people from different countries. This fact has facilitated the appearance of virtual communities with common objectives (e.g Free Software Community). If the goal is to learn, we talk about learning communities.

The fast-growing online learning communities is owed to the increasing use of cooperative distributed systems. In the case of learning communities, users collaborate in the knowledge construction process together with the adaptations and the corrections made by other members of the community. In this approach [4], the responsibility over the contents and the learning of the community is distributed among its members, so the distribution of knowledge and the growth of the database are reinforced. On the other hand, the interaction among the community members is carried out by means of synchronous mechanisms (videoconferences, voice IP, chats ) and asynchronous mechanisms (forums, mail, news).

Babel has been developed to offer a common framework to learn and practice different languages [5], which may be used and adapted in any educational institution as the system is distributed under the GNU/GPL license. All the related software is available at <http://raro.inf-cr.uclm.es/apps/babel/>

This paper proceeds as follows. Section 1 presents an introduction and some traditional teaching languages. Section 3 presents the objectives in the Babel design. In section 4, we present the architecture of Babel, where we emphasize the importance related to the behaviour of the State Machine, the communication tools, how the system searches the information, how the system use the dictionaries and a market research which justifies the chosen technologies. Section 5 presents how traditional teaching methods can be applied into the system. Finally, the paper concludes with the conclusions and directions for future research in this area.

### 1.1. Teaching languages

Flexibility on web systems(based on hypertext), in their non-sequential organization, allows us to improve the traditional methods of teaching, by adapting them in a more comfortable way to the natural mechanism of relationship among concepts. According to Sanchez Perez [2] and Richards [1], there are different models whose election will depend on the interest, the objectives or the learning environment of each person. Examples of these models are commonly used: *Grammar-Translation Method*, in which the aim is to reach the student ends up writing the language which he/she wants to learn. For this reason, the student should study the grammar and the vocabulary in detail. These activities are completed with many practical exercises. On the *Direct Method*, students learn a language through the oral-practical and reading exercises, reaching a great fluent linguistic. Unlike the method mentioned before, grammar is not studied at first place. The direct method is used by people who learn the native language at an early age. The *Audio-Oral and Audio-Visual Method* are based on the use of multimedia resources to ensure the knowledge by means of the continuing repetition of sounds and the association of images. Finally, the *Notional-Functional Method* structures the support material for the learning of a language according to the necessities of each people. For example, a computer specialist who wants to learn English to chat with other computer specialists from other countries, will study the vocabulary, technical terms, and expressions made in that language and related to the computer field.

### 1.2. Related works

To design Babel we analyzed the services offered by several web systems for the learning of languages. The following table resumes this analysis.

	MI	AF	H	AD	EAI	WS	LET	SS	LS	UE	ET	CA	Babel
Several Languages		✓	✓			✓							✓
Practical exercises	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Theoretical documents	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Automatic correction			✓	✓				✓	✓	✓	✓	✓	✓
User collaboration													✓
Internal tools			✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Games							✓						✓
Pay per use			✓	✓	✓			✓	✓	✓	✓	✓	

[MI] mansioningles.com

[AF] aulafacil.com

[H] home.es

[AD] auladiez.com

[EAI] e-aprender-ingles.com

[WP] worldspeaking.com

[LET] learn-english-today.com

[SS] speakspeak.com

[LS] learnspanish.com

[UE] usingenglish.com

[ET] englishtown.com

[CA] curso-de-aleman.de

After this study, some conclusions are observed. First, most of these systems are focused in the learning of only one language. Another important fact is all systems do not allow the users to contribute with their knowledge (there is no collaboration possibility). This is the case of the paradigm of the class imparted by the teacher in which the students are only passive receivers of information [3]).

## 2. Objectives

All analyzed systems in section 1.2 do not allow the users to collaborate in the growth of the base of contents. The users can only work with the contents exposed by the web site administrators. On the other hand, all systems always behave in the same way. The adaptation of these systems on an educational center can be complicated due to the difference of objectives and norms among different centers. Bearing in mind the results of section 1.2, we outline a series of initial objectives for the design and development of Babel.

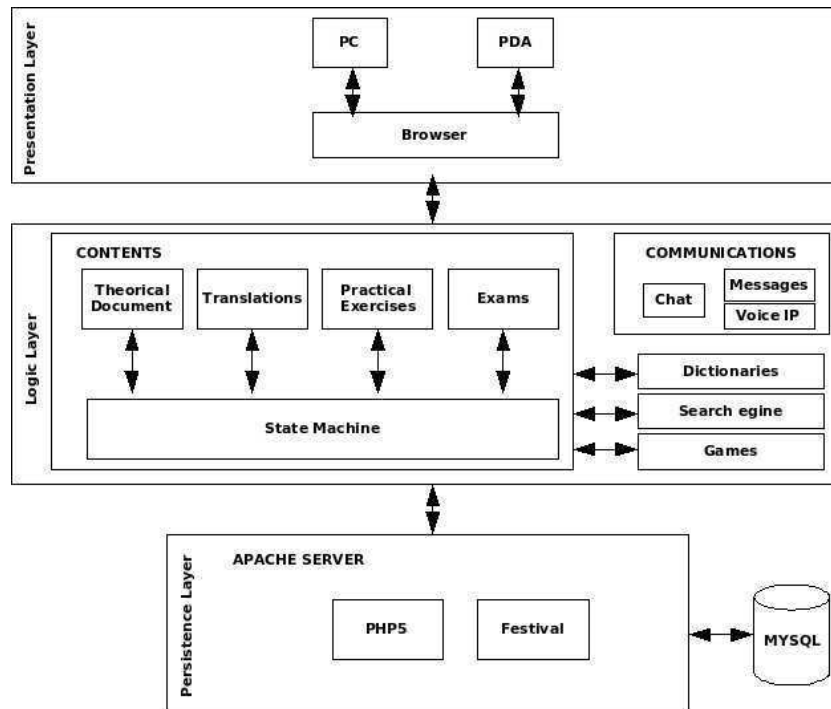
1. *Distributed content management.* The system has a common content-base in which the users may create, update, and reuse several kinds of contents. The information stored in the system will flow through different stages like acquisition, supervision, and finally, publication. The workflows allow more control over the information of the system.
2. *Dynamic behaviour of the system.* The system behaves according to the configuration of a state machine. This behaviour allows the system to be installed in several scenarios with different objectives and necessities.
3. *Concurrency control.* The concurrency control is necessary in systems in which a large community of users update the contents at the same time. In Babel, several access control mechanisms are used to avoid inconsistencies on the stored information.

4. *Search contents.* On systems in which the base of contents reaches a great size, achieving a fast access to the data is complicated, although the data are properly classified according to some criteria. The use of a search engine facilitates the information access to users by employing a set of keywords.
5. *Dictionaries and voice synthesis.* The system works with a dictionary for each language registered in the system. To facilitate the work (creation and update of the dictionary contents), standard formats are used.
6. *Student motivation.* The system contents not only are ordered according to the language below, but also it is ordered according difficulty level. On the other hand, the user has several points associated which determined the user's level on the system. The student can promote to superior levels if he/she completes practical exercises or exams proposed in the system. Playing the games or helping to other users is other way.
7. *Multilanguage Interface.* Because the system will be used by people from different countries to learn languages, the interface of the application is logical to be made in different languages. The system is able to add new languages in a simple way.
8. *Communication among users.* The System has different chat rooms with suitable communication tools for many users from different countries. The user may practice talking and improving his/her handwriting. All chat rooms are achieved according to native language and target language.
9. *Usability.* The application of different usability techniques [7] so that the system is the most intuitive possible and simple of using. The user should forget the tool and should center his/her efforts in learning. Due to the use of templates and global sheets of style, the interface consistency is also guaranteed; the tools are accessible from any place of the web and always in the same way.
10. *Multiplatform.* The device (PC,PDA ...) or Operating System (GNU/linux, Microsof Windows, Solaris ...) used is not a problem to access to the system because only a browser is necessary. The resolution of screen is not a problem either, given that CSS style is used to determine the width and height proportional to the page.
11. *Free Software* Babel is a system in wich all registered users can create new contents or to modify the existing contents. The created contents can be used freely by any user of the system. The system may also be used and adapted in any educational institution as the system is distributed under the GNU/GPL license.

### 3. Architecture

In large information web systems, the architecture is usual to be divided into several layers. This multi-layer programming guarantees maximum clarity, easy debugging of errors, and code reuse. Specifically, the design of Babel has been divided in three layers: Presentation, Logic and Persistence. *The Presentation layer* involves how the data are presented. A great advantage is that the same

data can be shown in different modes (e.g many devices with different size of the screen), without modifying the behaviour of the system. *Logic layer* involves the behaviour of the system, in other words, the functionality of the system is implemented here. This layer processes and generates data and itself communicates with the Presentation layer to show the data. Finally, the *Persistence layer* is responsible for interacting with the database manager and storing the data on files. This layer responses to the logic layer requests and stores the data produced by the latter. The three layer are independent; a modification in a layer does not imply a modification in another layer. For example, if we decide to change the database manager, then the *Persistent layer* has to be modified, but the rest of layers remain intact. The architecture of Babel is illustrated in figure 1:



**Figure 1.** Architecture of Babel

Although the general architecture of Babel is generic and independent of the content types, the current installation (as platform for the learning languages) makes the system has the several types of contents: theoretical documents, exercises of translation, practical exercises, and exams. The system is designed to add new contents or new categories easily.

### 3.1. State Machine

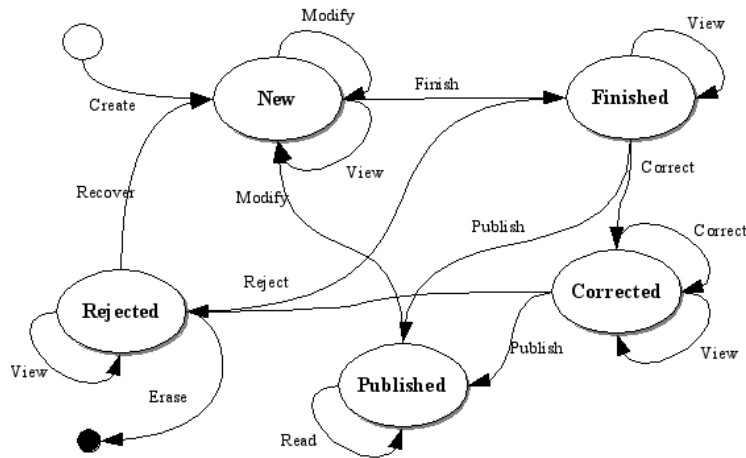
The contents of the system are managed according to the configuration of the state machine, obtaining a great configuration flexibility. The elements which form the state machine can be modified totally. Next, it is shown the elements belong to the state machine.

- *User Roles.* Systems in which to have the information under the biggest control possible, user's roles are necessary. The system of roles allows us to define different types of user groups. The actions which an user can carry out in the system are determined by the roles and the permissions.
- *States.* A state define the situation that a content is it. All the documents of the system are always in a certain state. When an user makes an action on the document, the latter can change state.
- *Actions.* The actions that an user can carry out on any content type are defined in the configuration of the state machine. It is possible to create new actions, modify them or delete those that already exist.
- *State Transitions (state-action-state).* A transition indicates a state change. In Babel, an user can undertake an action on a content (whenever the role system allows him to do it). This content is in a certain state, the action carried out on the document can provoke the content has a new state different to the previous one. When a content changes its state, the actions that can be carried out on it, can vary.
- *Permissions.* There are two types of permission in Babel: Role permissions and user permissions. Role permissions indicate the actions an user can carry out with his/her role, and user permissions indicate which actions an user can make. An user that has permission to make an action can make it although its role does not allow him/her to do it.

The figure 2 samples the default configuration when the system is introduced for the first time. The control on the contents is critical in distributed and collaborative environments. The system administrator can configure the state machine, determining the workflow among the users of the system according to theirs roles. The number of states determines in great measure the control on the information and the number of workflows. A high number of states implies several things: a bigger control on the information, a slow process until the publication and a bigger number of workflows among users. On the contrary, a low number of states implies another questions: a quick publication, few workflows among users, and scarce control on the information.

Note that the adopted solution configures an intermediate option. This solution allows us to have control on the information and it avoids the time of publication of the contents to be too high. Next, we show the values that the elements of the state machine take initially. As we have mentioned previously, these values can be modified by the system administrator according to the necessities of the place where it is introduced.

- *Roles:* anonymous, editor, reviewer, administrator, manager.



**Figura 2.** Initial state machine

- *Actions:* create, read, modify, view, finish, correct, publish, recover, reject, erase.
- *states:* new, finished, corrected, published, rejected.
- *state transitions:* the state transitions are exposed in the figure 2

When we create new actions in the system, it is necessary that we develop new source code, only in this case.

### 3.2. Communication tools

The asynchronous private messaging is one of the communication mechanisms to be employed by the user. Each user registered in the system has his own mailbox, which works in a similar way to the electronic mail. The users can also communicate among themselves in a synchronous way or real time by making use of the chat. There is a chat room for each pair of languages; the room “Spanish-English” will be formed by people who speak native Spanish or native English and who want to learn the other language.

The chat implemented in Babel has a series of features that make it specially interesting. It is possible to use dictionaries and to carry out corrections on the text written by another users. This way, the discussion in real time and the practice of the language with native users are facilitated.

Each user registered can specify its Skype account (any voice ip application can be used as long as this application has a web protocol with which an user can be called by other user from a web site). When an user enters in a chat room, an icon appears and any user can call it. Thus, the users can practice oral conversation.



**Figura 3.** Conversation written with corrections

### 3.3. Search of information

When coexisting different languages in a system, the management of a high number of dictionaries is necessary. Another added complexity is the combination of these languages (Spanish-English , English-German, English French ...) because each person needs the appropriate dictionary.

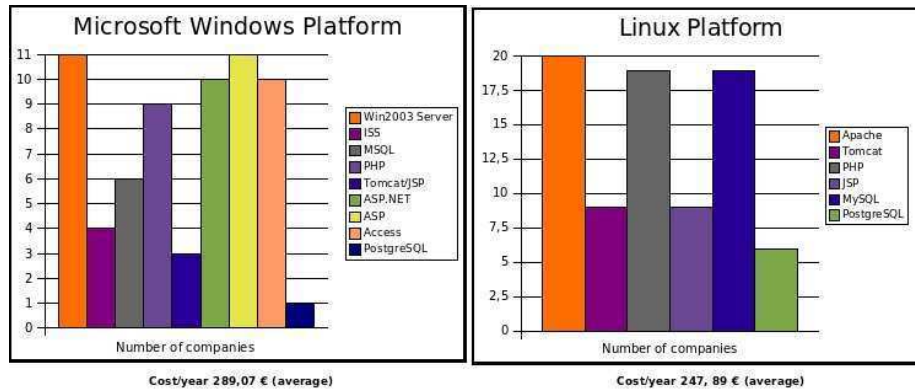
Babel is prepared for the direct importing of dictionaries in flat text format. These dictionaries are used in many tools for the search of terms. From the administration interface, new dictionaries can be added. Then, the base of terms is expanded. A more complete approach, based on model entity-relationship for the vocabulary learning can see in [6].

The search of contents allows the user to access to the information stored in the system starting from a group of keywords. The babel search engine allows the user to search on any content type. A previous processing of the stored contents is made with the purpose of extracting indexes and storing them in a data structure. The words which appear in a document are stored in the data structure together with the codes of the documents in which this word appears (each content in the system has a code that identifies). Therefore, the search is very fast because when we search a word, immediately the system recovers all the codes of the documents in which the word appears. The indexing of the contents, the same as in multitude of search engines, is made in schedules of low system activity.

### 3.4. Technologies used

We have used several technologies to develop babel: PHP5 like programming language, Apache like server of applications web, MySQL like database manager system and Festival like the voice synthesizer. Previously to the development of the system, we made a market research which has influenced in a great measure in the decision of the technologies used for the development. For this reason, we analyzed the hosting services offered by 23 companies, with platforms GNU/linux and platforms Microsoft Windows. Each one of these companies offers different

proposals, in which the proposals of more cost offer better services. For this reason, we have only take into account the proposal with minimum services of each analyzed hosting servers. The figure 4 summarizes a part of the obtained results:



**Figure 4.** Part of the market research result

We have obtained important conclusions starting from our market research:

- Most of the servers provide support for PHP and MySQL, in GNU / Linux servers and Microsoft Windows servers.
- The number of servers which allow us to install applications implemented in Java and JSP is not very numerous.
- Of the 23 companies, only 11 of them work with GNU / Linux servers.
- Only 3 companies work exclusively with Windows servers.
- The GNU / linux servers suppose an annual smaller cost than the Microsoft Windows server.

#### 4. Application with the traditional teaching methods

To apply different traditional teaching methods was one of the objectives that we sought to get with the creation of the system. If we remind the methods of the section 1.1, the *Grammar-Translation method* can be applied with the usage of many theoretical and practical documents. The user could improve his ability to write the language which he/she wants to learn correctly. The *Direct method* can be applied through the communication in real time with the rest of users of the system. The communication can be established by means of the writing process (chat) or by means of the speech (voice IP). The objective of this method is to reach a great linguistic fluency. On the other hand, it is possible to associate files (images, sounds ...) to the different types of contents. Therefore,

the user will be able to practice the *Audio-Oral* and *Audio-Visual* methods. Finally, the *Notional-Functional method* can also be used because all contents belong to thematic categories.

## 5. Conclusions

Most of countries consider important to learn one or two languages besides the native language, because this learning has important advantages at various levels (personal level, professional level, ...). Until some years ago, it was only possible the learning of languages in educational specialized centers by means of the study of books or travelling to foreign countries. The appearance of new technologies and Internet like a communication mechanism enlarges the number of possibilities in the process of learning language.

In the environment of the learning of languages, most of existing systems on the network have a series of static contents which are upgraded in a way centralized by the administrators of the web site. These systems delete the wealth of the collaborative learning among users. With the creation of Babel, we propose a free tool for the support in the learning of languages which the users can use to contribute freely with their knowledge to the growth of the community base contents.

The state machine regulates the behaviour in the process of acquisition and data handling, allowing the system to adapt to the necessities and preferences of the environment in which is implanted. Also, the system is supported by communication tools, dictionaries, voice synthesis, search engine of contents and games that make it more complete. The modular design of the system will allow the user to add new types of contents and new modules in the future easily.

## 6. Future works

In our current work, we concentrate on three research lines. First, is the enhancement of the dictionaries, that is to say, dictionaries which allow the semantic search and the definition of ontologies. Another line of work is the incorporation of an access mechanism more versatile than the based on semaphores. The use of a low-level layer to control the versions by means of CVS or SVN would allow a better parallelism in the work with the system documents. Finally, we are creating an interface in which the system administrator could see the current configuration. This way the administrator could modify it in a visual way. Currently, the setup of the state machine is shown by means of tables. This fact can hinder the administrator to have a general perspective of the setup.

## 7. Acknowledgments

This work has been funded by the *Consejería de Ciencia y Tecnología* and the *Junta de Comunidades de Castilla-La Mancha* under Research Projects PAC-06-0141 and PBC06-0064.

## Referencias

1. Richards, C., Rodgers, T. S.: Enfoques y métodos en la enseñanza de idiomas. Cambridge Press, (1998).
2. Sanchez Perez, A.: La Enseñanza de Idiomas. Principios, problemas y metodos. HORA, S.A.(1982).
3. Parker A.: Interaction in Distance Education: The Critical Conversation. Educational Tecnology Review, 12. p. 13-17. (1999).
4. Rogers J.: Communities of Practice: A framework for fostering coherence in virtual learning communities. Educational Technology & Society. p. 384-392. (2000).
5. Romiszowski, A.J.: Web-Based Distance Learning and Teaching: Revolutionary Invention or Reaction to Necessity?. Web-Based Instruction – Englewood Cliffs. (1997).
6. Vaquero A., Saenz F., Barco A.: Improving the Language Mastery through Responsive Environments. Computers and Education – Towards an Interconnected Sopciety. Ed. Kluwer Academic Publishers. (2001).
7. Steve Krug.: Don not make me think! A common sense approach to web usability. Pearson Educacion. Madrid. 2001