

## How to participate

Written by Ronaldo Martins

Saturday, 08 January 2011 20:25 - Last Updated Monday, 10 February 2014 13:43

---



The UNL<sup>web</sup> is a community of people and organizations who share a common interest or passion for the Universal Networking Language (UNL). It is open and free. See here how to participate.

## How to participate in the UNL<sup>web</sup>?

There are two main ways of participating in the UNL<sup>web</sup>: as an individual, or as an organization.

- **Individuals** work as language specialists or computer specialists. They may work either as volunteers (in non-funded projects or languages) or as freelancers (in funded projects and languages).

- **Computer specialists** create extensions, add-ons, plug-ins, engines, tools, applications and documentation for UNL-based systems. These resources are available in the **UNL<sup>for</sup>ge**

, a repository for downloading and developing UNL-driven technology.

- **Language specialists** insert entries in the dictionaries, create rules in the grammars, participate in the UNLization of documents, or provide documentation and training for any of these tasks, depending on their interests and accreditation. These resources are available in the **UNL<sup>arium</sup>**, a repository for

downloading and creating UNL-driven resources, and in the

[UNL](#)  
[dev](#)

, an environment for creating and testing grammars. Language specialists are classified in five different categories, depending on their expertise and academic background:

- Observers are allowed to browse and export dictionaries, grammars and corpora, but cannot add entries or grammar rules;
- Trainees are allowed to add entries, but only under supervision;
- Authors are allowed to add entries, but may edit only their own data;
- Editors are allowed to edit authors' data, but cannot edit other editors' data; and
- Revisers may edit editors' data, but cannot edit other revisers' data.

- **Organizations** work as partners, and may play linguistic, computational or social roles, depending on their vocation.

- **Language-oriented organizations** (linguistic departments, research institutes or associations) are normally eligible to become

### **Language Centres**

(LC's), which are responsible for managing the linguistic resources (dictionaries, grammars, corpora) created for a given language. LC's are headed by a Language Manager, who is responsible for defining the structure and the guidelines for the language, and for managing the work of individuals contributing to that language. Language managers may promote or demote individual collaborators, may propose new projects, and participate in the Linguistic Board, an assembly of specialists called together for consultation, deliberation or discussion of

issues related to the linguistic aspects of UNL. A single language may have several different LC's: in this case, decisions concerning the language are taken by the Language Council, which comprises all language managers associated to that language.

- **Computer-oriented organizations** (computer departments, research institutes or associations, software companies and labs) are normally entitled to become a

### **UNL Lab**

, which is responsible for research and development of UNL-based technologies. As in the case for computer specialists, products of UNL labs are normally hosted in the UNL

dev

, along with the corresponding documentation, licenses and source codes, in case of open-source software. The UNL

dev

, as the SourceForge, the Android Market or the Apple Store, is the place for downloading and developing back-end and front-end UNL-based software.

- **Other organizations** (corporations, governments, NGO's, international organizations, charities, cooperatives, etc.) normally support our work either financially (as donors or investors) or by any other means (infrastructure, training and education, communication, etc.). Financial support includes four main types of sponsorships:

- Languages, when the organization provides the funds for creating linguistic resources (dictionaries and grammars) for a given language or group of languages.

- Projects, when the organization provides the funds for a linguistic project (e.g., the UNLization of a given book) or a computational project (the development of a specific tool based in UNL).

- Research, when the organization provides a fellowship for a researcher to investigate a UNL-related topic.

- Nationals, when the organization provides the funds for paying citizens of a given country for their linguistic work in the UNLarium or their computational work in the UNLdev.

**What are the advantages of making part of the UNL<sup>w</sup>  
eb  
?**

- **Social responsibility.** The UNL is a property of the United Nations and, therefore, a patrimony of all of human kind. As our main goal is to enable all peoples to have access to information and to share their ideas in their own native language, contributing to the UNL is contributing to eliminate language barriers and to preserve language and cultural diversity.

- **Openness.** The resources produced in the UNL<sup>arium</sup> (dictionaries, grammars and corpora) are released under an attribute-share-alike license, which means that anyone is able to use them as they want, provided that any derivative work is released under the same license. Strongly formalized and heavily standardized, they constitute a valuable resource not only for UNL-related technologies, but for natural language processing in

general, and may be used in several different types of projects, domains and environments.

- **Education.** The UNL<sup>web</sup> contributes to the dissemination of the knowledge on languages and language structures. It comprises an e-learning facility (VALERIE) that simplifies the contact with linguistics and linguistic categories, and the daily work in dictionaries, corpora and grammars provides a concrete and tangible experience of the complexities involved in languages and in natural language processing.

- **Science.** The UNL<sup>web</sup>, as a distributed and integrated development environment for creating linguistic and computational resources, is a place for improving, sharing and disseminating the technology on languages, in a way that it will be accessible to all, from any country and any language.

**I'm a language specialist. How much money may I make working in the UNL<sup>web</sup>?**

This depends largely on your work and

expertise, and on the availability of funds for your language and project. Beginners (A0) receive only USD0.25 for each new entry inserted in the dictionary or for each new paradigm created in the grammar, but experienced users (C2) receive four times more (USD1.00) for the same work. Our average user has been inserting 30 new entries per hour (which makes USD7.50 for a beginner, or a USD30.00 for an experienced user), and we have been paying the average of USD300.00 a month per user (but some make more than USD2,000.00). It is important to observe, however, that many projects and languages have not been funded yet.

**Which are the projects available for freelancers and how can I start working within the UNL<sup>web</sup>?**

The list of open funded projects is

available at [www.unlweb.net/wiki/projects](http://www.unlweb.net/wiki/projects). In order to start, see [Getting Started](#)

.

**I represent an organization.  
How does it cost to support the  
UNL initiative?**

This depends mostly on the type of sponsorship. In general, the equation is  $(W * L * k)$ , where  $W$  is the number of the words of the source document;  $L$  is the number of funded languages; and  $k$  is the coefficient of repetition, which may range from 0.5, in case of specific texts, to 3.0, in case of no text at all (i.e., in case of general dictionaries). In that

sense, to have a book in UNL normally costs (  $W * L * 0.5$  ), whereas to have a language costs (  $W * L * 3.0$  ). Our main case study, *Le Petit Prince*, the famous novel by Antoine de Saint-Exupéry, contains 14,304 words, and has been costing USD7,152.00 per language, which is approximately the same that would be paid for an ordinary translation, with the difference that the resources will be used not only for translation, but also for indexation, summarization, simplification, etc., and will be available to all as the basic input for other systems. As for languages, to have a 30,000-word dictionary in French has been costing USD90.000,00, because every entry is double-checked and must include the corresponding grammatical



rules (such as inflectional paradigms and subcategorization frames).

## **How long does it take to finish a project?**

This depends on the funded languages. There are languages for which we have several users; for others, however, we have only a couple of users. In general, at least for the 20 most spoken languages, it normally takes  $W/100$  days, where  $W$  is the number of words of the original. To cite again our main case study, the whole UNLization of Le Petit Prince took around five months ( $14,304/100 = 143$  days).

# **What is the expected result of a project?**

This depends on the project. UNLization projects (i.e., to have a given document in UNL) provides the UNL version of the source document and the corresponding language resources (dictionaries, grammars, knowledge bases, example bases, memory bases). Dictionary projects provide dictionary entries and the corresponding grammatical features (inflectional paradigms and subcategorization frames, for instance).

## **What is the use of the UNLization of a book?**

The UNL version of a text is a knowledge-based graph representation of it. This version, because semantically-oriented, may be used for several different purposes: to generate a new version of the text, to generate an abstract of the text, to generate a summarized version of the text, to generate a simplified version of the text, etc., and all these to any of the languages for which the corresponding resources are available. Consider, for instance, the case for Le Petit Prince. After having the whole text UNLized, we are able not only to retell the story in English, Portuguese, Russian, Armenian, and even in French (although this was the source language), but also to generate several different versions of

the text, in different sizes and different levels of complexity.

## **Why not simply using other existing translation tools, such as Google Translate?**

In first place, because UNL is not only about translation. The results of the UNLization of a text may be used for translation, of course; but they may also be used for indexation, navigation, summarization, simplification, information extraction, semantic reasoning, etc. In second place, because UNL is committed not only to the most spoken languages, but also with less-resourced and endangered

languages, which are not subject to the statistical treatment normally adopted by the current technology in machine translation. At last, because UNL intends to be diversity-preservative, which means that we intend to go from Swahili to Tamil without using English or any other language (other than UNL itself) as a pivot language.

## **Who are you?**

The UNDL Foundation is a non-profit organization based in Geneva, Switzerland, which has received, from the United Nations, the mandate for implementing the Universal Networking Language (UNL). The UNL is an artificial

language that has been used for several different tasks in natural language engineering, such as machine translation, multilingual document generation, summarization, information retrieval and semantic reasoning. It has been, since 1996, a unique initiative to reduce language barriers and strengthen cross-cultural communication in the framework of the UN.