STANDARDIZATION OF THE GENERATION PROCESS IN A MULTILINGUAL ENVIRONMENT

Jesús Cardeñosa[1],Carolina Gallardo[2] and Edmundo Tovar[3]

Abstract *— Natural language generation has received less attention within the field of Natural language processing than natural language understanding. One possible reason for this could be the lack of standardization of the inputs to generation systems. This fact makes the systematic planning of the process of developing generation systems to become difficult. The authors propose the use of the UNL (Universal Networking Language) as a possible standard for the normalization of inputs to generation processes.*

Index Terms *— Natural language generation, standardization, multilingualism.*

1 INTRODUCTION

In natural language processing (from now on NLP) two areas can be differentiated: analysis and generation. However, one has not received the same attention as the other from the scientific community, that is why generation can be considered as the "poor brother" of the NLP. The reason for this minor development is the different nature of the input to the analysis and generation systems. The input to the analysis systems is always natural language, whose casuistic and phenomenology are known; while in a generation system, the output is always known, but not what it is going to generate from [1].

The input to a generation system varies depending on whether it is monolingual generation (dialogue systems) or a multilingual system (mainly machine translation systems). In dialogue systems it is difficult to establish appropriate characteristics common to all inputs, because "the problem" of generation is usually solved with solutions *ad hoc*, depending on the application and the system language. In machine translation systems, there are also many differences in the inputs to the generation subcomponents, conditioned by the nature of system architecture (transfer, interlingua, etc.), the kind of grammars being used (declaratives vs. procedural) [2], or the number of languages in the system.

This difference in the input to the generators makes a systematic planning of their development process impossible (main cause of the minor development of generation compared to analysis). It is necessary then, that the input to the "generator" can be supported with an appropriate model of contents representation, separated from the format or language that ensures a standard process for the development of generation systems.

In this article we propose the UNL as a possible standard for the generation inputs. To achieve this, in section 2 we will introduce the main generation architectures. Section 3 will describe in detail the UNL system, its qualities and basic architecture for generation. Section 4 will establish the conditions required by any technology in order to be considered a standard and which ones are fulfilled by the UNL. The article will end with the description of a real massively multilingual system (HEREIN) where UNL has been formally studied and proposed as a de-facto standard for generation of contents in natural languages.

2 GENERATION ARCHITECTURES

2.1 Dialogue Systems

Dialogue systems represent one of the main applications of natural language generation. This kind of systems have as their most important target "to present information to the users in an easy to understand format" [3] in very specific fields where the user generally interacts with the system in the same language. The user asks the system specific information; once obtained, the system can show it through an answer in natural language. This answer is very frequently obtained (with certain success) through the generation of a "built" language from a series of templates that keep a predefined relationship with the templates that support the questions [4]; this means the generation process takes as input a representation that depends on the way the user makes the question. It could be said that there is not a thorough analysis of the text, nor an abstract representation of the information that should be given to the user. The great dependency of the source language and the domain restrain the construction of multilingual dialogue systems and the reuse of these systems in other domains.

2.2 Machine translation systems

Machine translation systems (from now on MT) are essentially multilingual because their target is the "transformation" of a text written in language A into an equivalent text in language B. In this section main architectures of MT systems will be described, because each architecture sets a series of conditions over the appropriate characteristics of the inputs to the generation process.

[1] Artificial Intelligence Department, Universidad Politécnica de Madrid, 28660 Madrid. carde@opera.dia.fi.upm.es
[2] Artificial Intelligence Department, Universidad Politécnica de Madrid, 28660 Madrid. carolina@opera.dia.fi.upm.es
[3] Languages and Systems Department, Universidad Politécnica de Madrid, 28660 Madrid. etovar@fi.upm.es

2.2.1 Transfer systems

The basic tasks in a transfer system are analysis, transfer and generation. The analysis component produces a syntactic representation (sort of thorough) depending of the source language. This syntactic representation is the input to the transfer module whose task is to transform that representation into a closer structure to the target language. The output of the transfer module shapes the input to the generation system module which finally produces the phrase in the target language. In transfer systems, the components, inputs and outputs are strongly oriented to the source and target languages.

The main problem of the transfer systems is the almost impossibility to reuse the existing resources (transfer modules) and components in order to include new language pairs in the system. In fact, if it were necessary to increase the number of language pairs, a new system would have to be built. Generally, the great orientation of the "transfer" systems towards the target language involves great accuracy in the output, and a considerable difficulty in reusing components to include new languages in the system.

2.2.2 Interlingua based Systems

Interlingua based systems form the second great systems' paradigm of machine translation. Included into the systems based in interlingua are the "traditional" ATLAS-II [5], PIVOT [6] as much as the *knowledge-based* ones such as KANT [7] or Mikrokosmos [8]. Their defining characteristics are:

- **Unique intermediate representation**. The abstract representation, result from the analysis, "feeds" directly the generation module. This intermediate representation is the component named "interlingua".
- **Elimination of the transfer process**. The system carries out two basic tasks: analysis and generation.

The systems based on interlingua are oriented to cover the largest possible number of languages, given that the number of components that requires a system based in interlingua for n languages is 2*n, it is remarkably inferior to n*(n-1) that transfer systems require for the same number of languages.

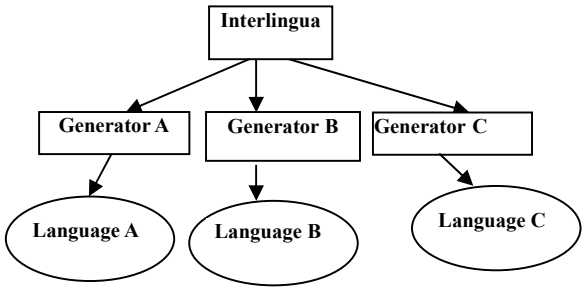

FIGURE 1. GENERATION IN INTERLINGUA SYSTEMS

The basic architecture of the generation in a interlingua system is shown in figure 1.

Interlingua based systems offer an important advantage over the transfer ones; the architecture facilitates the inclusion of new languages and are reusable. However, during the conversion process to the interlingua, it is possible that some significant grammar information for the generation may be lost, that is, the interlingua may have less information (grammatical, not conceptual) than a syntactic representation. To sum up, the systems based on interlingua offer a larger number of languages at the expense of lesser precision in the generated texts.

2.2.3 Fusion

Without any doubt, multilingualism is an added value for any generation system. The transfer-interlingua dichotomy seems to imply an opposition between precision vs. number of languages. To take advantage from every one, some transfer systems have "interlingued" their architectures to support a larger number of languages [9] [10]. The common characteristic in these systems is the existence of a deep syntactic representation that has some amount of independence from the source language. The process to combine the interlingua architecture in a transfer system requires the construction of a transfer module between the deep syntactic structure and an interlingua representation [11].

3 THE UNL APPROACH

3.1 The UNL system

UNL [12] is an artificial language designed to reproduce the content of texts written in any natural language. The UNL is provided with specifications that formally define the language. A UNL expression is an hyper graph consisting of:

- **Universal words**. They define the vocabulary of the language, i.e., they can be considered the lexical items of UNL. To be able to express any concept occurring in a natural language, the UNL proposes the use of English words modified by a series of semantic restrictions that eliminate the innate ambiguity of the vocabulary in natural languages. In this way, the language gets an expressive richness from the natural languages but without their ambiguity. Take, for example, the English word "construction" meaning "the action of constructing" and the "final product". Thus, the word "construction" will be paired with two different universal words:

  construction$_1$ → construction(icl>action)
  construction$_2$ →construction(icl>concrete thing)

where “icl” is the abbreviation for “included”.

- **Relations**. These are a group of 41 relations that define the semantic relations among concepts. They include argumentative (agent, object, goal), circumstantial (purpose, time, place), logic (conjunction, and disjunction) relations, etc. For example, in a sentence like “The boy eats potatoes in the kitchen”, there is a main predicate (“eats”) and three arguments, two of them are instances of argumentative relations (“boy” is the *agent* of the predicate “*eats”,* whereas “potatoes” is the *object*) and one circumstantial relation (“kitchen” is the *place* where the action described in the sentence takes place).
- **Attributes.** They express the semantic information resulting from the morphologic flexion and the functional elements of the phrase (auxiliary verbs, articles, etc.). They are put together with the universal words to complete their meaning when they appear in a specific context. The attributes include information about time or aspect of the event, number, polarity, modality, etc**.** In the previous sentence, attributes are needed to express plurality in the object (“potatoes”), definite reference in both the agent (“boy”) and the place (“kitchen”) and finally and special attribute denoting which UW is the head of the whole expression. (the *entry* node).

Formally, a UNL expression has the form of a semantic net, where the nodes (universal words) are linked by labeled arcs with the UNL concept relations. The graphical representation of the sentence “the boy eats potatoes in the kitchen” in UNL is shown in figure 2.

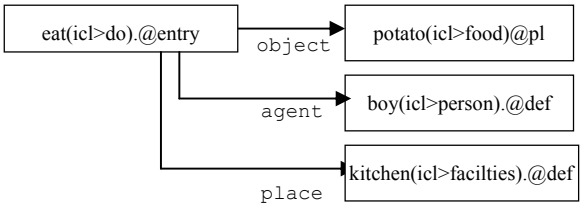

FIGURE 2: REPRESENTATION OF A UNL EXPRESSION.

This sentence is written in UNL in the following manner:

***agt*(eat(icl>do).*@entry*, boy(icl>person).*@def*)**
***obj*(eat(icl>do).*@entry*, potato(icl>food).*@pl*)**
***plc*(eat(icl>do).*@entry*, kitchen(icl>facilities).*@def*)**

3.2 Basic characteristics of UNL

The UNL system represents a generic framework for the massive generation of multilingual contents. Its main goal is the contents’ representation of a document, web page, data base, etc., in a *consensual and normalized structure* that may be transformed into a text in a natural language. The defining characteristics of the UNL system are:

a) It is a system oriented to the generation of multilingual contents. A document written in the UNL has its “own identity” and can be stored in a document data base, etc.

b) The UNL does not involve the use of specific components or tools. The tools and components, as well as the processes that may be defined to accomplish the edition and the generation in the UNL vary from one language to another. The use of the UNL only involves the standardization of the input into a generation system [13].

In spite of the emphasis given to the language generation in the system, the UNL framework includes the editing process of natural language into the UNL, named “enconversion” as well as the generation into natural languages or “deconversion” (see figure 3).

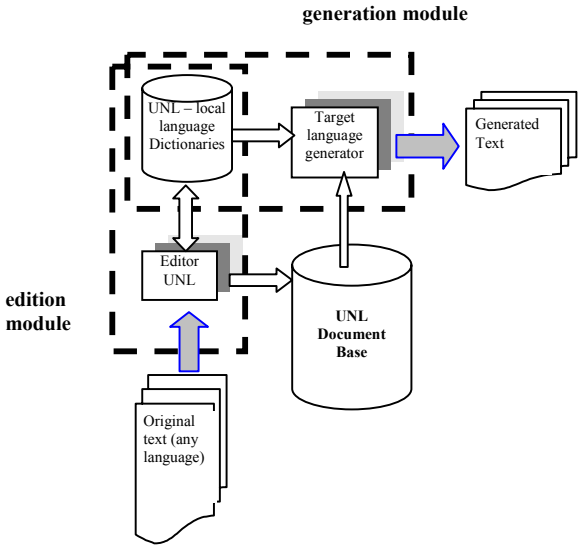

FIGURE 3: ARCHITECTURE OF THE UNL SYSTEM.

The UNL is an interlingua in essence, that is, an appropriate language for the representation of the meaning in an independent way from the natural languages. The UNL is not restricted to a specific domain (as can be the KANT or Mikrokosmos interlinguas); the fact of not restricting the input in the vocabulary collection of the interlingua guarantees the UNL adaptation for the representation of contents in any language or domain.

3.3 Generation in the UNL framework.

There are several architectures for the generation of natural language from the UNL. Next, the two generation

architectures within the UNL framework will be described in detail.

3.3.1 Direct Generation

The UNDL Foundation (http://www.undl.org) supplies a module that carries out the generation process through a unique process. This module is known as DeCo (standing for DeConverter). This module is completely language independent, since all the necessary grammar knowledge for the generation of the target language is included in the dictionary and the rules' set proper of the language.

Given that this module directly transforms the semantic UNL representation into the morphological realization (that is, a sentence in natural language), the dictionary must contain the best detailed information in the following aspects:

- **Grammar category and subcategories**: the more organized by hierarchies the lexical level, the better quality will be expected from the generation.
- **Argument structure** and prepositions required by verbs, nouns, and adjectives.
- **Semantic information** that may be relevant for the syntactic configuration in the target language.

With the help of the information included in the dictionary, the generation rules have, as their main task, to transform the UNL expression into a phrase in the natural language. Basically the following tasks are being carried out:

- **Matching of the UNL relations with the grammar relations of the language**. In the previous sentence, the agent of the predicate in UNL corresponds to the grammatical subject in English or Spanish.
- **"Translation" of the UNL attributes into their appropriate morphologic or syntactic realization**. For example, the attribute "plural" has to be morphologically realized as a plural noun in Spanish. The attribute "definite reference" is translated into Spanish through the insertion of a definite article. Not always there is a direct translation between UNL attributes and morphological/pragmatic information in natural languages. For instance, when dealing with time, UNL only offers three possibilities (past, present and future). It would be "competence" of the generation rules of each natural language to correctly select the tense and verbal moods applicable to the languages that do not have this kind of time system (for instance, Spanish).
- **Generation of pronouns and anaphoric expressions**. The UNL expression is devoid of anaphoric elements, all concepts in UNL should be stated explicitly. It is the task of the generation rules to insert pronouns and other anaphoric elements in the generated texts.
- **Morphologic synthesis**. Finally, generation rules should tackle aspects such as agreement between verb and subject, or between adjectives and nouns, word order or the expression of the correct verb tense.

Figure 4 shows the architecture for direct generation, there it can be seen how the "bilingual" dictionary Natural Language-UNL and the generation rules feeds the DeCo module in order to carry out the generation of UNL text into a natural language text.

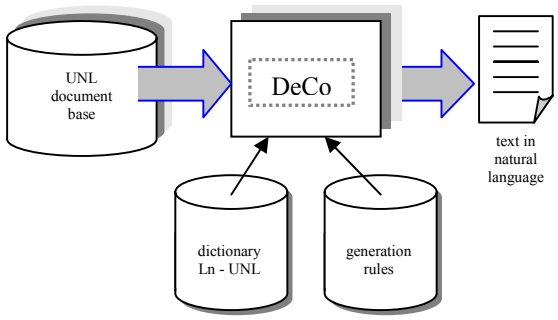

FIGURE 4: ARCHITECTURE OF THE DIRECT GENERATION IN THE UNL FRAMEWORK

3.3.2 Combined Generation (reuse of transfer components)

The treatment for Russian and French languages inside the UNL system is the perfect example of the *combined generation* within the UNL framework. Both teams have integrated the UNL system into their transfer systems, ETAP in Russian case [11], and Ariane for the French one [14].

These systems have chosen to reuse the available generators of the target languages and to develop an additional module that allows the conversion of the UNL representation into a friendly format through the generators of their "transfer" systems. An example of combined architecture would be exemplified in figure 5.

The so called "UNL transfer module" is with no doubt a

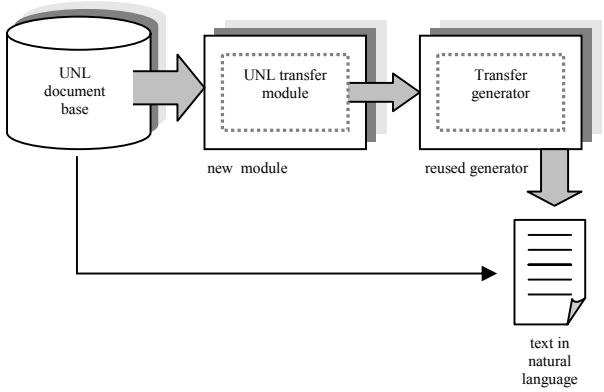

FIGURE 5: COMBINED GENERATION

new component to develop. However, the experience in the already mentioned systems has shown that the development costs of this module are cheaper than the costs for

developing a new generator that could have the UNL code as its direct input.

4 CAN THE UNL BE A GENERATION STANDARD?

4.1 What is a standard?

If we try to avoid the formal definitions for "standard", it could be said that a standard is a set of rules, criteria and recommendations that allow to build a product or to design and offer a service in a proper way that assures:

- The universalization of the work, that is, a unique way of doing something that at the same time can be independently evaluated, no matter who does it or when.
- The quality. When products or services have been carried out following a standard, there is a certainty that the processes are well implemented and the product quality is not at risk.
- The assessment of the product or service provisions, meaning that it could be determined through a unique way, when a product or a service fulfills the specifications it has been designed and built for.

Many more could be enumerated, but we are focusing in these three that may be the most intuitive. As it has already been mentioned, the lesser development of some products (in this case, language generators) is due to the lack of standards that could assure these characteristics. The diversification and extremely disperse casuistic of the inputs to a generator cause that the output become the only way to assess it to establish a subjective evaluation.

Although there are some researchers that have not neglected this side of the generation [15], this standard has not been yet established, neither formally nor de facto.

4.2 The UNL as a standard

Technically speaking, the lack of uniformity in the inputs to language generators is almost the only reason that restrains a bigger development. Therefore, the support systems to multilingual services see their action limited only to specific languages where translation services may be offered, either automatic or not. However, the language expansion is an unapproachable road with these methods. If the input to language generators is not standardized, this problem will not be solved in a global way. The only standardization would then be the choice of a content support that could express itself in a unique way, with a specific language. Actually, this concept has existed for many years, and it is the Interlingua concept. It is within this context where the UNL can play a role. The UNL has not been conceived as an interlingua, but it can be used as one. The interlinguas had their historic moment when they faced the same problems as the other systems created for machine translation during the 80's. At the beginning of the 90's it was clear that the subject of the languages was much more complex than it seemed during the technological development of the 80's and the exaggerated optimism of the time.

It is not the purpose of this paper to describe the economic advantages of an interlingua over the traditional systems of machine translation regarding many languages (a traditional machine system requires 90 systems to support 10 languages, while one based in interlingua requires only 20). In fact, the crossing point between systems takes place at three languages. For more than three languages interlingua is cheaper.

However, historic matters at the beginning of the 90's buried the interlinguas (mainly those developed in Japan and the USA) because while the interlingua based systems were not well defined, the "transfer" machine translation systems began to offer more positive results. Even so, within the group of language technologies, machine translation became kind of discredited. At the end of the 90's, the United Nations opted for models based on interlingua approximations to define the multilingual support systems for the Internet. The result is the today's named UNL, already described in this chapter. Apparently, it would be the ideal system to solve the problem of the absence of a standard input to language generators. Nevertheless, a standard is something else than a technological solution. It could be summarized like this: a standard is evaluated through the maturity concept that to sum up means that it would be associated to the organized and organizational maturity, that is, there has to be an organization behind the standard that may be able to maintain, modify, allow the study of its acceptance and real use for it, and other factors. Currently, it could be said that the UNL has weak and strong points to formally become a standard [16]:

Weak points:

- Relatively recent technology
- Not too much implemented
- Quality system not implemented

Strong points:

- Worldwide organization behind (dissemination assured)
- Business expectations increased by the incorporation of minority languages
- Quality system defined

However, independently of the global factors, the technological approach is nowadays the only one able to solve the problem of automatic multilingual generation systems. Regarding the business approach, the expansion of multilingual systems in the Internet requires much more than traditional systems of machine translation. This is why the UNL is not just an interlingua, but a language to support knowledge repositories, different ontological approaches, and other matters. Summarizing, the UNL (or something similar to it) is necessary and needed by others.

5 A REAL EXPERIENCE: HEREIN AND UNL

5.1 Herein and standardization of form and structure.

The Herein system (IST-2000-29355) [17] is a perfect example of a massively multilingual environment. It constitutes an Internet-based facility for improving cultural heritage management methods at the European level. Among the main tasks of the project, participant countries must compose a report providing detailed information about all aspects regarding cultural heritage.

Due to the large number of countries participating in the project (almost 30) and the huge variety of topics that comprise cultural heritage (legislation, preservation, dissemination, etc.), there was an urgent need to standardize both the format and the structure of the contents that each country should provide. A definite structure was established and every country involved in Herein had to integrate its particular contents into such structure. Eventually, this structure turned out to be a de-facto standard for the description of the cultural heritage issues of a country since it met the two basic conditions for a de-facto standard, which are:

a) it has been actually used in a real and working environment
b) its application has been universalized: it proved valid for almost 30 countries.

Furthermore, a de-facto standard requires a support for its physical representation. In Herein's case, XML was the chosen support. The standardization of format/structure is twofold in Herein: structure has been normalized as a de-facto standard, whereas format has been normalized with a canonical standard (XML). Figure 6 shows the appearance of a typical report in the Herein project.

the verbalization of such contents: that is, the linguistic aspect in Herein.

If the contents' side has been solved, the linguistic aspect has not. Although there are almost 30 countries involved in the project, the Herein web site and produced resources are far form being truly multilingual: only three languages are official (English, French, and Spanish), therefore all documents and resources created in Herein can only be accessed in these languages.

The reasons for such a dramatic reduction of languages are simple and straightforward:

a) Translation costs: If HEREIN were a really multilingual environment, one document of a given country will require around 24 translation into the other involved languages. For all documents of all languages, the number of required translations will be 25*24, that is, 600 translation works. Providing that only the Spanish national report counts with 10.000 words, costs for translating the Spanish report into the other languages ascends to the translation costs of 240.000 words.
b) The availability of translators in all pairs is not the same. Obviously, availability of translators for the pair English – French is higher than availability of translators for Dutch – Croatian, which can be really difficult to find.

Both reasons are enough for desisting from human translation in massively multilingual environments.

There is only one alternative to this approach, and it is the use of an interlingua. Previously we have briefly described an interlingua as "a common intermediate representation" between languages, and have postulated two conditions that interlinguas should meet, namely:

```
<?xml version="1.0" encoding="UTF-8" ?>
<!DOCTYPE rapport (View Source for full doctype...)>
<rapport id="1.3" pays="ES" langue="es">
        <theme id="1">
            <titre>PERSPECTIVAS DE CAMBIO EN EL PATRIMONIO</titre>
        </theme>
         <stheme id="1.3" contenu="COMPLET">
            <titre>Prioridades a corto y medio plazo</titre>
          <para> Con carácter general son 3 las prioridades básicas:
            <liste type="PUCE">
                <elem>  1. Documentación.
                    <para>
                        <liste>
                            <elem>
                                A) la llamada Iniciativa info XXI "Una sociedad de la Información para todos". Esta
                                   iniciativa en materia de patrimonio tiene como objetivos básicos:
                                <para>
                                    <liste>
                                        <elem>Obtener un catálogo colectivo de los bienes integrantes del patrimonio
                                              histórico español, que sirva por un lado como instrumento efectivo para
                                              su protección y por otra parte como base para su difusión a través de
                                              Internet.</elem>
                                    </liste>
                                .....
```

FIGURE 6: EXAMPLE OF SPANISH CONTENTS IN XML STRUCTURE.

However, the contents and their structuring are just one side of the problem in the HEREIN system, the other side is

- Independence from any natural language.
- Same semantic expressiveness as a natural language.

The UNL takes these two conditions as its defining characteristics. The elements that compose the UNL are all based on semantic notions, detached from any residue of morpho syntactic categories found in natural languages. These elements and the way to compose them in order to form valid and meaningful UNL expressions are completely defined and formalized in the UNL specifications [12]. But the main potential of the UNL for achieving the same expressiveness of a natural language lies in its vocabulary (the universal words). The UNL profits from the richness of natural language vocabulary (universal words are based on English lexical items) while devising a system of semantic restrictions that eliminate the ambiguity and vagueness inherent to lexical units of natural languages. In this case, the UNL perfectly fits in the definition of an interlingua or a "pivotal language".

Alongside with its adequacy for being used as an interlingua, the UNL also satisfies the conditions for its qualification as a potential standard for generation. These two characteristics have been already exploited in the Herein project, as it will be shown in the next section.

5.2 The UNL approach in HEREIN

As an initiative of the Ministry of Education and Culture of the Spanish government, representative institution of the Herein contents in the Spanish language, and in collaboration with the Spanish Language Center (representative and responsible of the Spanish language in the UNL program), the complete report of the Spanish cultural contents was codified into the UNL.

This UNL code has been capable of being embedded into the XML structure common to all reports, as if the UNL were another "natural language"(see figure 7).

```
<?xml version="1.0" encoding="UTF-8" ?>
<!DOCTYPE rapport (View Source for full doctype...)>
<rapport id="1.3" pays="ES" langue="unl">
<?xml version="1.0" encoding="UTF-8" ?>
<!DOCTYPE rapport (View Source for full doctype...)>
        ....
<stheme id="1.3" contenu="COMPLET">
<titre>{unl}
            mod(priority@, term(icl>time))
            mod(term(icl>time), short(mod<thing))
            and(short(mod<thing), long(mod<thing))
            {/unl}
</titre>
<para>      {unl}
            obj(exist(icl>be).@entry,priority(icl>thing).@def.@pl)
            mod(priority(icl>thing).@def.@pl,basic(aoj>thing))
            qua(priority(icl>thing).@def.@pl,3)

            {unl}
<para>
```

FIGURE 7: UNL TEXT EMBEDDED IN A XML DOCUMENT

The difference lies in the fact that the aforementioned contents can be "captured" by the generators of any language. After generation of the UNL, the corresponding contents (now in the "form" of a natural language) will be reinserted in the XML structure of the document. The result, at the internal level, is visualized as shown in figure 8 and 9 for the English and Russian language.

```
<elem>
this initiative regarding heritage have the basic
following objectives.
  <liste>
    <elem>      a collective catalogue of the goods
                the Spanish historical heritage is
                integrated protection diffusion thro
                Internet is obtained.
    </elem>
    <elem>      the structure of the information and
                the manner identify, describe and
                to classify the goods of the catalogue
                is normalized.
    </elem>
</liste>
```

FIGURE 8: OUTPUT OF ENGLISH GENERATOR

```
<elem>
У этой инициативы относительно наследия есть
основные следующие цели
      <liste>
        <elem>  Получить коллективный каталог
                этого товара, который служит, как
                эффективный инструмент для
                защиты этого товара и основа для
                товара, который интегрирует
                испанское историческое наследие,
                распространения
                посредством Интернета..
        </elem>
        <elem>  Нормализовать структуру
                информации и способ
                идентифицировать, описывать и
                классифицировать товары каталога.
        </elem>
      </liste>
</elem>
```

FIGURE 9: OUTPUT OF RUSSIAN GENERATOR

6 CONCLUSIONS

It seems clear that the architecture of an "interlingua" system (based on a format of a unique input to all generators) supports the idea of formally defining the input to develop the "generator" component, according to some precise specifications (not existing until now). This precise specification would be the base of a standard generators development, creating an environment that may allow carrying out tests of reliability for this component, essential in the generation of multilingual contents.

Given the characteristics of the UNL (independence from the natural languages and adaptation to express any content of a natural language) and the possible integration of the UNL system with any other existing generation system, it is appropriate to propose the UNL as a standard for the normalization of the inputs to natural language generation systems.

A standard must be supported by an organization that can assure its stability and maintenance. In this case, there is an organization that fulfills these requirements: the UNDL Foundation under the protection of the United Nations. Finally, it is important to mention that the UNL has been recently qualified as the first software patent of the United Nations.

7 REFERENCES

[1] Dale R., Di Eugenio B, and Scott D. (1998). "Introduction to the special issue on Natural Language Generation". *Computational Linguistics*, Volume 24, number 3.

[2] Whitelock, P. (1995): "Linguistics and computational techniques in machine translation system design". UCL Press Limited. London.

[3] Reiter E, y Dale R. (1995). "Building applied natural language systems". Cambridge University Press.

[4] Ballim, A and Payota V. (2001). "Weighted Semantic Parsing: A robust approach to interpretation of Natural Language Queries". *Flexible Query Answering Systems*. ed Larsen H., et al. Physica Verlag.

[5] Uchida, H. (1989). "ATLAS-II: A machine translation system using conceptual structure as an interlingua". *Proceedings of the Second Machine Translation Summit*. Tokyo, Japan.

[6] Muraki, K. (1989). "PIVOT: Two-phase machine translation system". *Proceedings of the Second Machine Translation Summit*. Tokyo, Japan.

[7] Nyberg E and Mitamura T. (1992). "The KANT System: Fast, Accurate, High-Quality Translation in Practical Domains". *Proceedings of COLING-92*: 15th International Conference on Computational Linguistics. Nantes, France.

[8] Beale, S., S. Nirenburg and K. Mahesh. (1995). "Semantic Analysis in the Mikrokosmos Machine Translation Project". *Proceedings of the 2nd Symposium on Natural Language Processing*. Bangkok, Thailand.

[9] Aikawa T, Melero M, Schwartz L and Wu A. (2001). "Multilingual Natural Language Generation". *Proceedings of MT Summit VIII*. Santiago de Compostela, Spain.

[10] Boguslavsky I, Frid N, Iomdin L, et al. (2000). "Creating a Universal Networking Language Module within an Advanced NLP System". *Proceedings of COLING 2000: 18th International Conference on Computational Linguistics*. Saarbrucken, Germany.

[11] Lavoie B, Kittredge R, Korelsky T; Rambow O. (2000). "A Framework for MT and Multilingual NLG Systems Based on Uniform Lexico-Structural Processing". *Proceedings of 6th Applied Natural Language Processing Conference*. ACL, 2000 Seattle, USA.

[12] Uchida, H. (2002). "The Universal Networking Language. Specifications". http://www.undl.org.

[13] Boitet C and Sérasset, G. (2000). "On UNL as the future "html of the linguistic content" & the reuse of existing NLP components in UNL-related applications with the example of a UNL-French deconverter". *Proceedings of COLING 2000: 18th International Conference on Computational Linguistics*. Saarbrucken, Germany.

[14] Boitet C and Sérasset G. (1999). "UNL-French deconversion as transfer & generation from an interlingua with possible quality enhancement through offline human interaction". *Machine Translation Summit 99*. Singapore.

[15] Hovy E. ; Reeder F. eds. (2000). *Proceedings of the International Workshop on Machine translation Evaluation*. AMTA 2000. Cuernavaca, Mexico.

[16] Cardeñosa J. and Tovar E.; "A descriptive structure to assess the maturity of a standard. Application to the UNL System". *2nd IEEE conference of standardization and innovation in information technology*. Boulder. Colorado. USA. International Centre for standard research. ISBN: 0-7803-9817-3

[17] HEREIN Project (IST-2000-29355). Final Report. European Commission.