Towards a Systematic Process in the use of UNL to Support Multilingual Services

Jesús Cardeñosa, Carolina Gallardo, Edmundo Tovar

Departamento de Inteligencia Artificial- Facultad de Informática
Universidad Politécnica de Madrid
Campo de Montegancedo, s/n
28660 Madrid- Spain
{carde, carolina, etovar}@opera.dia.fi.upm.es

Abstract. The UNL Programme of the United Nations University (UNU) was launched in 1996 aiming at the elimination of linguistic barriers in Internet. Now, eight years later, UNL is not ready to support real applications due to several circumstances. This eight-year period can be divided in two: a first four-year period devoted to the formal definition of UNL as a formal language (under the sponsorship of the Institute of Advanced Studies (IAS) of the UNU) and the remaining four years devoted to the technical experimentation of UNL. A new period is starting right now, which could be the period of maturity at all levels, especially at technical and business levels. In this paper, the authors summarize the more significant experiences until now, their conclusions and the set of procedures to produce marketable multilingual services. This kind of work will be the work of the UNL consortium during the next two years before launching UNL to the market.

1 Introduction

The natural evolution of UNL as a project and as a Programme is the support of useful applications for a multilingual society. Apart from other uses of UNL, like cross-lingual information retrieval or support for ontologies, the more understable use and possibly the easiest application, is the support of multilingual services, that is, to represent contents written in any language and to generate any other language [1].

UNL is not conceived to become a (fully automatic) machine translation system (MT hereafter). Up to date, MT systems based on the transfer architecture have achieved reasonable results, always involving pairs of languages. These systems are somehow handicapped by their *language coverage*. In other words, a transfer based system involving N languages requires the development of $N \times (N–1)$ systems, which ends up with the consequent combinatorial explosion of the number of systems to be developed as the number of languages grows.

On the other hand, interlingua-based MT systems show, in principle, a highly attractive advantage over transfer systems: interlingua-based systems do not grow exponentially as the number of language increases since for a system to support N languages, only $2 \times N$ systems have to be developed. The ATLAS system [2] and the PIVOT system [3] in open domains, and Mikrokosmos [4] and Kant [5] in restricted

Universal Network Language: Advances in Theory and Applications.
Research on Computing Science 12, 2005, pp. 413–429.

domains are the most representative systems within the interlingua-based MT paradigm.

However, not everything is so easy and straight ahead in interlingua-based systems. In fact, currently there are not interlingua-based systems in open domains, nor Interlingua systems that have been able to penetrate in the market. One of the possible reasons to explain this fact is the practical design of the Interlingua itself and the pivotal role it plays in a MT system. The reason for this minor development of interlingua-based MT systems (specially in open domains) could be the difficulty in designing a formal language that simultaneously is far enough from the surface forms of natural languages (so that almost all languages can fit in the interlingua representation) and that is expressive and rich enough to convey the subtleties in meaning expressed in natural languages [6]. Thus, the proper design of the Interlingua will affect the overall behavior of the system in the analysis and generation processes.

UNL, in terms of Interlingua design, had to find the balance between a representation where linguistic meaning could be naturally expressed and a representation not devoted or inspired by a given natural language and, of course, not restricted to particular domains. After years of debates and discussions, it seems that this difficult balance was found. However, massive encoding experiences in the UNL context have given away a worrying aspect of UNL: the lack of common understanding of the specifications in almost all the components of the language (universal words, attributes and relations), possibly due to the incomplete definition of the language and codification procedures in the current version of the UNL specifications [7].

This incompleteness and imprecision in the definition of the specifications of the language provokes a wide variety of UNL code according to the encoder's understanding of the UNL language and even according to the source language of the contents to be encoded. Such a variety negatively affects the results of language generators (independently of the target languages and used systems). Not only should be pursued the interdependence among participants in the process of defining a uniform way to encode contents into UNL but also uniformity in the processes and methodology when working with UNL. That is, independently from low-level linguistic and codification considerations, the clear definition of both the working processes and the complete definition of the UNL as a language is indispensable if the development of services based on UNL is targeted at.

Some members of the UNL consortium have thoroughly considered these two aspects since time ago. The first experience with this purpose was in 2001 when, as a result of some conversations with the organizing committee of the international event Forum Barcelona 2004, it could be seen that the UNL lacked the necessary infrastructure to be able to provide multilingual services. During the encoding tasks of the Barcelona experience, it could be proved how the UNL specifications did not provide a clear answer of how to codify real texts (not just toy examples). The same applied to the definition of the Universal Words. Besides, there was neither a formal definition of the Knowledge Base nor how it has to be used, with the final result that even having the capacity to build a knowledge base for UNL, there was no way to do it. There were also no tools for UNL massive codification (the manual process is tedious, and with high risk of error), and moreover there was not a definition of the processes to be carried out in the production of UNL code and multilingual generation. From the point of view of the standards of technological development, in particular Software

Engineering, it could be confirmed that UNL was far from being considered mature when facing massive production [8].

From that moment on, several partners of the UNL Consortium agreed on beginning to define such processes and at least some common guidelines for codification that will unify the procedures in order to assure a reliable production later on. The outcomes of such experience were encouraging. Initial guidelines for the codification were produced [9] and the first set of processes could be exposed [10].

Subsequently and up to now, there have been two more experiences trying to emulate the problems that may arise in massive codification scenarios. These are the so-called "HEREIN experience" [11] and UNESCO [12]. Both experiences proved that commercial production of UNL goes through the creation of huge amounts of contents in UNL and the concise definition of the involved processes, roles, techniques, tools and standards. Without all that, UNL would never surpass the theoretical limits of its possibilities.

This article presents a general methodology for multilingual generation in the UNL context. The article is organized as follows: section 2 summarizes a comparative analysis of the experiences carried out so far and their most representative drawn conclusions. In section 3, a working methodology will be presented. This methodology has been defined after the experiences of Barcelona, Herein and UNESCO and it is the first step in the staging of UNL as a support for multilingual services. Section 4 presents some advances in the definition of metrics, necessary to estimate costs and productivity. Without methodologies and processes it is impossible to evaluate costs in the development of applications based on UNL and, consequently, to evaluate possibilities of UNL in the market.

2 Experiences. A Comparative Analysis

The need to define and determine the involved procedures in the process of multilingual generation has lead to the UNL Consortium to undertake several experiences that will explore the processes of the complete cycle of production –that is, from contents written in a given language, to their enconversion and final deconversion into other languages. For the time being, the most general tasks in this process were:

- Lexicographic tasks: where UWs had to be defined and dictionaries updated with the new UWS.
- Codification task: once the UWs have been defined, the UNL code for the text is produced.
- Generation task: each source language must tune its generator to the new phenomena appearing in the text.
- Post-edition task: generated texts have to be revised by human post-editors, since no automatic translator or generator have (in this moment) enough quality to assure grammatical correctness and a natural and legible style.

These tasks were at the core of all the experiences so far. However, each of them has helped in one way or another to more concisely define the processes that are involved in multilingual generation and to bring into light some deficiencies of UNL.

2.1 Barcelona Experience (2001)

In the Barcelona experience, the original text was written in English and its approximate size was 3000 words. Lexicographic and codification tasks were shared among the four participant teams (Russia, France, Italy and Spain). There were continuous debates about definition of UW and codification issues among the teams. This process was fruitful for the most theoretical aspects of UNL (UWs and codification). The outcomes of such work were the definition of some common guidelines that will facilitate the unification of encoding styles.

However, the division and organization of work in this experience cannot be taken as a paradigm for competitive projects involving massive amount of contents, since the time and resources employed were out of any criterion of profitability. Certainly, experiences like Barcelona are extremely helpful to improve the bases for productivity and profit criteria. In the case of Barcelona, quality had priority over productivity.

2.2 HEREIN (2002)

Here the approach is different from Barcelona's. This experience tried to prove the UNL capacity for representing a big amount of contents coherently. The experiment was unilateral in the sense that the original text and the generated one involved the same language in order to update the rules of the language generator. The definition of UWs and UNL codification was undertook by one single team. In the codification work the guidelines produced during Barcelona experience was followed. The size of the text to be encoded was considerable around 12000 words dealing with many aspects of the cultural heritage of Spain.

An effective work requires a well trained team, and useful tools that could go from (semi-)automatic UNL editors to language generators. Work in Herein represents a borderline among what can be done and cannot with almost manual tools, dictionaries with reduced coverage and a generator with an acceptable quality, so that minor changes are required.

This time the novelty of the experiment lies in the fact that the contents were expressed in a complex type of language, resembling a legal style, which could occasionally yield more complex UNL representations that consequently would originate problems for deconversion. The produced UNL code in Herein, which was undertaken by just one team without intervention or consensus among other teams, could be posed difficulties to the generators of other languages, and even to any other expert codifiers. That is, the lack of uniformity in the process of codifying can yield UNL code not appropriate for real multilingual generation.

The main conclusion of this experiment is that the lack of agreement in the way to codify and the non existence of clear criteria for codification (like those following the spirit of the guidelines but more comprehensive) is the direct cause for an important loss of quality.

As a result of this experiment, it was established the need for the UNL teams to work together and cooperatively to define a definite Manual for Codification in UNL.

In the Herein experience, productivity increased but the overall quality decreased.

2.3 UNESCO (2003–2004)

This experiment was the first one that was developed in the laboratory context but under a contract that will demand results. It was the first contract for multilingual production using UNL. Apart from multilingual generation, the contract also included the measurement of productivity and associated costs. The objective was to establish a benchmarking that would allow for the establishment of some general definition of the processes of production and of the maximum costs associated to each process in any language. Taking into account the multilateral nature of Barcelona and the unilateral nature of Herein, this project was defined in between, as the closest model to achieve productivity in the medium term.

More concretely, the tasks for UWs production and UNL codification were assigned to a single team (with the associated risks of lack of consensus). The tasks for local dictionaries and generation along with post-edition were carried out by the other teams. The volume of contents was also considerable (15000 words) in the domain of World Heritage. For the first time, the codifying team used a UNL Editor that substantially accelerated this process and increased productivity up to the point of starting to define business models based on the use of UNL. In this case, there was neither debate nor consensus in principle but the produced UNL code could be improved with the feedback of other teams. The use of the tool for UNL edition was essential also for revision of errors (reaching 1 minute per sentence as average in the revision process, quite a distant measure from manual revision and codification).

The objective of UNESCO was the establishment of metrics for productivity in every process and task on the one hand; and on the other, specifying the processes that needed improvement and what sort of improvement. The results of this experience have been positive, although still they somewhat incomplete. The main issues that need to be improved in the nearby future are:

- A consensus should be reached when codifying into UNL as an essential condition for massive production.
- A higher degree of automatization in the lexicographic and codification process is indispensable. They require for clear standards in production that will help to alleviate the error rate in these two processes.
- A standardization of the processes that will allow for measuring costs and will make compatible the processes in different languages.

During both the Herein experience and the UNESCO experience the Spanish Language Centre attempted to measure the employed time in all the processes involved in multilingual generation. The processes are depicted in detail in the next section, whereas the obtained metrics and the results will be the topic of section 4.

3 Methodology

3.1 Overview: Context, Roles and Goals of the Methodology

This section contains a description of a general methodology for multilingual generation within the UNL system. This methodology is mainly derived from the multiple

experiences involving UNL codification and targeting at multilingual generation carried out by the UNL consortium.

The purpose of the methodology is to show the main processes involved under the broad concept of "multilingual generation". For the sake of generality, these processes have been described avoiding concrete procedures that depend on particular applications and technologies.

The common context where this methodology applies is that of a given customer (be it an institution or any particular customer) providing a document or set of documents in a specific natural language. For each document, it is required:

1. The UNL codification of the document (that is, a UNL document)
2. The generation of the UNL document into the number of natural languages that the customer establishes (multilingual generation *per se*).
3. The resulting bilingual Natural Language – UNL dictionaries of the involved languages (multilingual lexical resources).

In order to carry out these three main tasks, the methodology distinguishes two types of participants, according to the roles they play.

- **Coordinator**: The co-ordinator supports direct communication with the providers. The coordinating team will receive the original documents that will be codified into UNL and lately generated into several natural languages. Normally, the "working" language of the co-ordinator will coincide with the language of the provided documents. The reason for this equality in the language is simple: the co-ordinator is in charge of creating the relevant UWs and the UNL codification of the document.
  The general tasks that the co-ordinator carries out are:
  - **Vocabulary extraction** from the original documents (in the original language)
  - **Construction of the list of UWs** belonging to the complete vocabulary of the document (they are pairs of words)
  - **Codification** of the original document into UNL.
  - Distribution of aforementioned materials (UWs and UNL code) to the rest of participants.
  - Finally, elaboration of the project **documentation**, if needed.
- **Local Teams:** They communicate with the coordinator. Local teams are defined according to the language they work on. So, if generation assignments are required in three languages (say English, French and Spanish) there will be three local teams: English team, French team and Spanish team.
  The tasks of local teams are three-fold:
- **Creation of the pairs** (Headword-UW) according to the UWs provided by the co-ordinator. (local dictionaries).
- **Generation** of the provided UNL document into the local language.
- **Post-edition** of the generated language.

Please note that if one of the involved languages is the own language of the co-ordinator, these tasks also apply to the co-ordinator team. For example, if one of the involved languages is Spanish, being Spanish the "working" language of the co-ordinator team, the co-ordinator team will have to follow all the processes described for local teams.

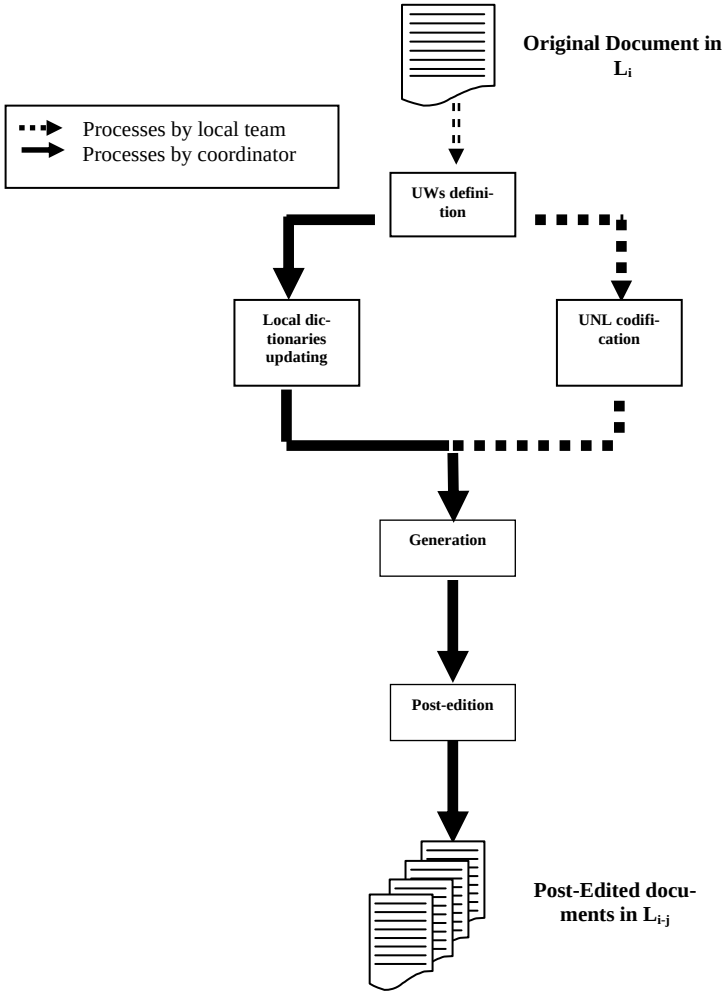

Fig. 1. Overview of the methodology.

The remaining subsections specify each process and subprocess that conform the methodology. For each process (or subprocess is the process is decomposable), the objectives, input and expected outputs are specified. As have been mentioned, no explanations or hints about how to perform these processes are included in the methodology, since such procedural information depends much on the state of the technology available for every language and for every local team.

The fact that this "know-how" information is not included does not mean, of course, that processes are to be performed without the help of specialized tools and software. In fact, some processes can be done automatically with the use of adequate tools. For example, some tools may be designed *ad hoc* to perform some processes like lexical extraction and lemmatisation (in Process 1) or instead the process can be

done manually. Other processes (especially Process 3, language analysis) tackle very well known problems in the area of Natural Language Understanding, and thus the availability of tools and specialised software may vary from language to language and from team to team. For this reason, the methodology is not defined according to a given language processor or analyser, to the extent that the process could be performed with no machine aid at all. The same applies for Process 2 (updating of dictionaries), that heavily depends on the specific dictionary physical support and design of each team.

However, two issues should be pointed out for processes 4 and 5. Process 4 is fully automatic (that is, generation should be fulfilled automatically and with the lest amount of human interaction). On the other hand, Process 5 (as it will be explained) is a complete manual activity.

Finally, for clarity reasons some conventions has been used when referring to documents and different languages. These are the following:

- The document (or set of documents) provided by the customer will be referred to as Original Document.
- Such document is written in a specific language, referred to as Language A, or L_A as an abbreviation.
- The different natural languages involved in the methodology (those of the local teams) will be referred to as Local Languages or L_N as an abbreviation.

A general overview of the first level processes of the methodology is shown in Figure 1**.** A concise description of each process will be included in the remaining of the section, from section 3.2 to section 3.6. The presentation of both the general methodology and specific processes will be done according to the following schema:

- A description of process or subprocess.
- A table detailing the input and output of process or subprocess.
- A graphical representation of the process, showing the workflow, input and output.

3.2 Process 1: Definition of Universal Words

This process is decomposed in the following 3 subprocesses.

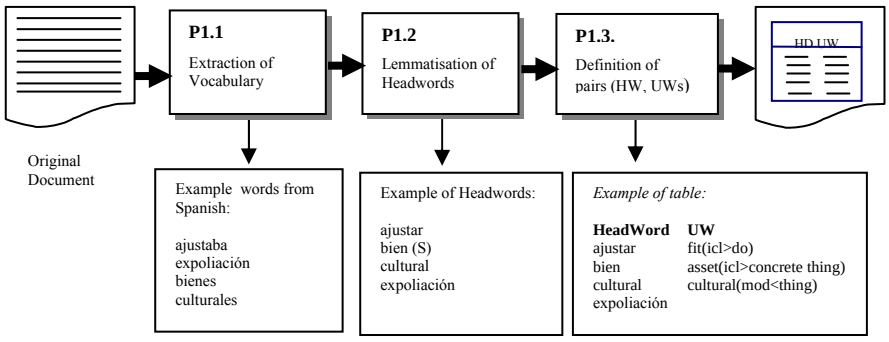

Fig. 2. Workflow for Process 1: Definition of Universal Words.

Sub-process P.1.1: Extraction of Vocabulary

Given a document, the relevant vocabulary (id est, lexical items or words) must be identified and extracted. For relevant vocabulary, it is understood lexical items that denotes concepts and thus have an equivalent Universal Words. Such lexical items are usually refers as "lexical categories" as opposed to closed-class categories (articles, auxiliary verbs, some prepositions, etc).

Input and expected output are detailed in Table 1.

Table 1. Input & Output of Subprocess P 1.1

INPUT	Original document in Language A.
OUTPUT	List of words belonging to the document that require a UW.

Sub-process P.1.2.: Lemmatisation

In the document, words appear inflected. That is, a verb may appear in the 3rd person singular of tense present in the subjunctive mood, or and adjective may appear in the feminine plural form. In this subtask, the inflected forms found in the document should be converted into headwords or lemmas. Lemmatisation is done in the following way:

1. For an inflected verb, convert it into the infinitive form.
2. For an inflected noun, convert it into the singular and nominative form (if case applies)
3. For an inflected adjective, convert it into the masculine singular noun.

Input and expected output are detailed in Table 2.

Table 2. Input & Output of Subprocess P 1.2

INPUT	List of words belonging to the document that require a UW.
OUTPUT	List of headwords that require a UW

Subprocess P.1.3: Definition of pairs

In this subtask, the pair (Headword L_A, UW) must be constructed. That is, for each headword of the list of headwords resulting from P1.2, the equivalent Universal Word must be identified.

Input and expected output are detailed in Table 3.

Table 3. Input & Output of Subprocess P 1.3

INPUT	List of headwords, output of P1.2
OUTPUT	Table with the pairs (Headword L_A, UW) for the whole list

3.3 Process 2: Updating or Building Local Dictionaries

This process is decomposed in 2 subprocesses.

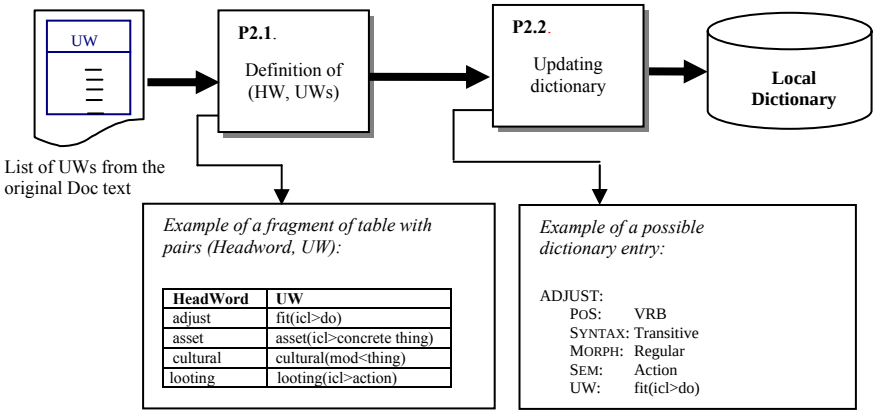

Fig. 3. Workflow for Process 2

Subprocess P.2.1: Definition of local pairs

In this subtask, each local team is provided just with the list of UW that has been resulted from the complete table, outputted in Process 1. The objective is to "find" the headwords belonging to the local team language that best fits into the UW. As a help, local teams can be also be provided with the original document and with the complete table with the pairs L_A – UNL. Note that this will be only helpful if the Language A is familiar to the local teams; otherwise, providing the original document and the complete table will have no apparent utility.

Input and expected output are detailed in Table 4.

Table 4. Input & Output of Subprocess P 2.1

INPUT	List of UWs belonging to the original document.
OUTPUT	Table with the pairs (Headword L_N, UNL)

Subprocess P.2.2: Updating or building the local dictionary

In this subtask, local teams must update their dictionaries and insert (or update) the adequate entries (the headwords identified in the previous table) together with the corresponding Universal Word.

Input and expected output are detailed in Table 5.

Table 5. Input & Output of Subprocess P 2.2

INPUT	Previous dictionary of the local Language - UNL UNL and table with the pairs (Headword L_N, UNL).
OUTPUT	Updated dictionary of the local Language - UNL

3.4 Process 3: Conversion into UNL

This process is decomposed in two subprocesses.

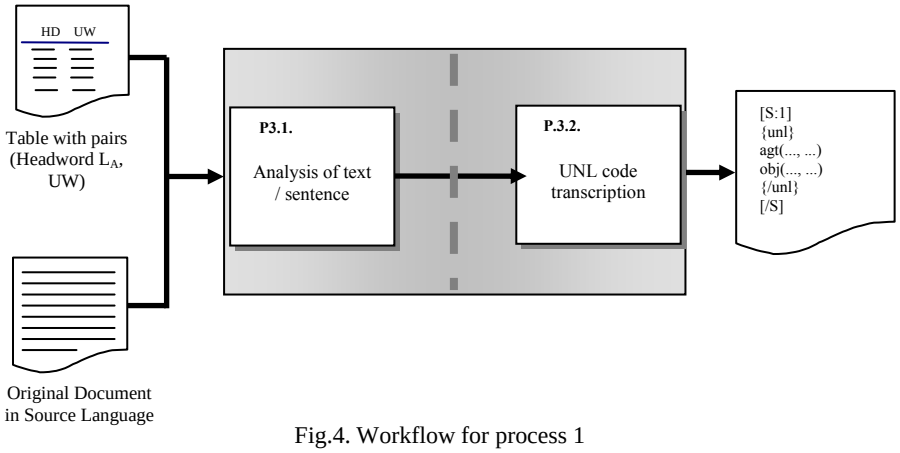

Fig.4. Workflow for process 1

Subprocess P.3.1: General Understanding of the text

This is quite an analytical task, the objective is to comprehend the meaning of the text and how this meaning is expressed in the sentence, that is, to "understand" the grammatical and semantic relations of the text. Since UNL expressions correspond to sentences, this subtask is performed iteratively sentence by sentence.

The borderline between subtask P3.1 and P3.2 is rather fuzzy. Analysis of the text may be guided by the UNL final representation or it can be done more independently from the final UNL representation, simulating NLP components that carry out the analysis tasks in the following traditional processes:

– Morphological and Lexical analysis
– Syntactic Analysis
– Semantic Analysis

Be it that as it may, there are two clear conceptual processes: and analytic one, and a "transforming" one: transform the meaning of the sentence into a UNL representation. Table 6 specifies input and output for this subprocess.

Table 6. Input & Output of Subprocess P 3.1

INPUT	Original document and list of pairs (Headword L_A, UNL)
OUTPUT	Abstract representation of the meaning of the sentence*

Please note that this subtask may not have a physical output, this "abstract representation" can be allocated in the head of the codifier.

Subprocess P.3.2: UNL ENCODING

This subprocess is the "transformation" of the abstract representation of the sentence obtained in P.3.1 into the UNL representation according to the UNL specifications and codification manuals if available. In this subtask, also document markers should be included in the final UNL document. Input and expected output are specified in table 7.

Table 7. Input & Output of Subprocess P 3.2

INPUT	– Abstract representation – UNL specifications
OUTPUT	UNL document (corresponding to the original document).

Figure 4 shows the workflow of process 3. The grey box in the graphic representation of the process simply gives account of such fuzziness in the separation of both processes.

3.5 Process 4: Generation into Local Languages

This process consists on the generation of the UNL document (output of P.3) into the local languages. This process is not decomposable, since generation is performed

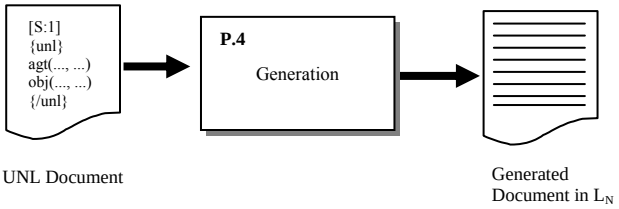

Fig. 5. Workflow for process 1

automatically. Each local team should be provided with language generators that will actually perform this task. Inputs and outputs to the process are presented in Table 8. The workflow of the process is illustrated in figure 5.

Table 8. Input & Output of Subprocess P 4

INPUT	– UNL document – Updated local dictionary
OUTPUT	Document with the raw generation of the original document in the local language

3.6 Process 5: Post-Edition

Since language generators may occasionally produce incorrect language, or at least, a

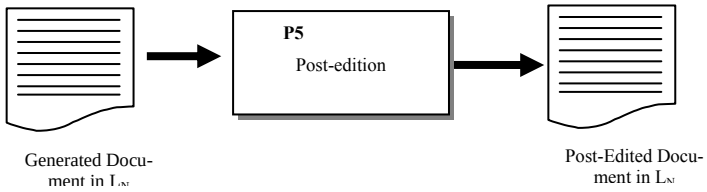

Fig. 6. Workflow for process 1

low quality language (incorrect style, non fluent language, etc.), texts are post-edited. Post-edition consists merely on giving "style" to texts, that is, making them natural. At this moment of the technology, this task is performed entirely manually. As usual, input and outputs of this process are gathered in table 9, whereas the graphical representation of the process is shown in figure 6.

Table 9. Input & Output of Subprocess P 5

INPUT	Generated Document in the local language.
OUTPUT	Post-edited document in the local language.

4 Results and Conclusions

These processes are necessary for establishing a benchmark in order to evaluate the productivity of global processes for UNL to become a firm candidate to support multilingual services in the market. However not only the definition and description of each involved process is required, productivity cannot be accounted for without defining explicitly its associated costs, measured (usually) in time. Thus the definition of metrics associated to each process in the global methodology is of paramount importance, so that there will be no business future in UNL without a way to evaluate costs, which inevitably involves measuring tasks.

Metrics, evaluation, validation, etc. are quite obscure fields in NLP; however there is not any engineering product or project that is thrown into the market that obviates metrics. UNL cannot be an exception.

There are several aspects that may hinder a straightforward establishment of metrics in the UNL contexts. These are:

– The non uniform nature of the UNL Consortium. We have different systems, different dictionaries, different generators, and different tools. At this point we could think that it is not comparable the time employed in creating a lexical entry in a x-uw.txt dictionary or in a dictionary in another system (say ETAP or Ariane). Likewise, analyzers and editors are different from team to team.

– The degree of expertise of the actors in charge of the processes. Obviously, a higher degree of expertise will reduce the extra load time for review in all processes.
– Until clear and definite instructions for building UWs and for codification into UNL is made, metrics for the overall UNL enconversion process will be flawed.

In spite of all this apparent drawbacks, the Spanish Language Centre noticed the urgency and need to begin establishing metrics for all the processes exposed in the methodology (section 3). Almost all processes were measured in time, especially in the following tasks:

– Construction of UWs
– Construction of dictionaries entries
– UNL codification
– UNL post-edition

Measures were taken in two different domains and experiences: Herein and UNESCO, with different actors showing different degree of expertise, and different available tools in the enconversion task. Let's have a look at the results.

4.1 Metrics in the Enconversion Process

The context of Herein is the following:

– No proper tools available for UNL enconversion, the available tools were either too rudimentary or not robust enough to undertake a massive codification task. Therefore the codification process was made mainly manually. This implies almost the same amount of time in reviewing the code (in particular reviewing syntactic aspects of UNL expressions).
– The degree of expertise in UNL encoding was acceptable (no need for prior training).

When measuring the employed time for codification, several decisions have to be taken: are we interested in measuring time to encode a text, a sentence, a paragraph or simply the number of words? Since these matters were not very clear, it was decided to take into account the time of enconversion per sentence, thus obtaining a correlation between sentence/time for codification.

The sentences extracted in Herein showed an average length of 20 words and an average time of 4'8 minutes per sentence. If counted on total values, the 16 sentences amounts to 322 words, and the total time to codify all the sentences was 77 minutes, which means 14'4 seconds per word.

At this point, it has to be remarked that the UNL code in Herein was produced manually, needing ulterior revision and requiring additional tools to catch up syntactic errors.

On the other hand, the UNESCO metrics differs in two main aspects: the degree of expertise and the available tools. In UNESCO, there exists data for a total of 116 sentences. The total amount of words in the 116 sentences is 3178. In this case, the average length of the sentences is superior to Herein, the sample of the sentences shows

27'4 words average length. The arithmetic average of time of codification per sentence is 9'95 minutes. When taking into account total facts (total of words and total of time), there results in 21 seconds per word. Data for UNESCO is summarized in table 10.

Table 10. Results of the metrics taken in the Unesco experience

Number of sentences	116 sentences
Total number of words	3178 words
Average length of sentences	27'4 words
Total time for enconversion	1155 minutes
Average time for enconversion of a sentence	9'95 minute /sentence
Time for codification of a word	21 seconds

As can be observed, there is a significant increase in time of codification per sentence. Common sense will make us predict that, due to the use of edition tools, there would a significant improvement in the time of codification; however, there is not. A possible reason for this is that the length of the sentence may interfere in the time of codification (being shorter sentences easier to codify than longer sentences) and the degree of expertise. That is, the difficulty in codifying may be related to the domain and type of language used in the domain.

Further, one does not have to forget that the UNL code obtained in UNESCO was syntactically, at least, correct. Whereas the UNL code obtained in Herein required subsequent syntactic revision.

4.2 Metrics in the Post-edition Process

Post-edition, as conceived in the UNL context, has to be carried out manually completely. In the metrics for the post-edition process there were involved two different actors and different types of domains as well. The actors varied in the degree of expertise, from a native speaker of a language to a professional translator.

Regarding the native speaker of the language to be post-edited, the average time to post-edit a sentence showed a striking uniformity: disregarding the domain, the average time for post-edition of a sentence is 1 minute.

The data collected by a professional translator is summarized in table 11, being the most significant conclusion a considerable descent in time.

Table 11. Specific data in the post-edition process by a professional translator

Number of sentences	164 sentences
Total number of words	4188 words
Average length of sentences	25'7 words
Total time for post-edition	120 minutes
Average time for post-edition of a sentence	45 seconds
Time for post-edition of a word	0'6 seconds

4.3 Metrics in Lexicographic Processes

For the construction of UWs, bilingual and monolingual dictionaries were used and the metrics obtained pertains to just one actor. The average time was 3 minutes for the construction of an UW and 1 minute for the construction of a lexical entry in a dictionary x-uw.txt type. This data applies both to Herein and UNESCO experiences.

5 Conclusions

For the time being we cannot say that we dispose of reliable, systematic and trustworthy metrics. As can be seen, there are a lot of parameters that influence in the final metrics. Some of them are expectable (like the degree of expertise or the use of tools) but other (like the linguistic particulars of a given domain) may be not so obvious, and even debatable. In such a heterogeneous context like the UNL consortium, all these hidden variables have to be made explicit and taken into account when establishing common metrics and common reference times for us all.

The metrics and times presented here are, of course, not definite. However, they hint at the possible maximal boundaries of the time to be employed in each process that should not be surpassed by any team in the UNL consortium in order to achieve a minimum degree of productivity. The objective of the metrics and of the definition of a common benchmarking is to determine the minimum time required for the several process so that a cost evaluation can be done. Such evaluation would be as a reference for the others. It is a very critic point for the exploitation of the UNL to acknowledge the most competitive costs we can have.

References

1. Jesús Cardeñosa; Luis Iraola; Edmundo Tovar; (2001) *"UNL: a Language to Support Multilinguality in Internet"* International Conference on Electronic Commerce. ICEC2001. Vienna. Austria, 2001, 9 pp.
2. Uchida, H. ATLAS-II: A Machine Translation System using Conceptual Structure as an Interlingua. Proceedings of the Second Machine Translation Summit, Tokyo, 1989.
3. Muraki, K. PIVOT: Two-phase machine translation system. Proceedings of the Second Machine Translation Summit, Tokyo, 1989.
4. Beale, S., S. Nirenburg and K. Mahesh. Semantic Analysis in the Mikrokosmos Machine Translation Project. Proceedings of the 2nd Symposium on Natural Language Processing. Bangkok, Thailand, 1995.
5. Nyberg E y Mitamura T. The KANT System: Fast, Accurate, High-Quality Translation in Practical Domains. En Proceedings de COLING-92: 15th International Conference on Computational Linguistics, 1992.
6. Arnold, D.J., Balkan, L., Meijer, S., Humphreys R.L., and Sadler, L. Machine Translation: an Introductory Guide, Blackwells-NCC, London, 1994, clwww.essex.ac.uk/MTbook/HTML/.
7. Uchida, H. The Universal Networking Language Specifications v3.2, 2003, www.undl.org.

8. Cardeñosa, J., Tovar, E. A Descriptive Structure to Assess the Maturity of a Standard: Application to the UNL System. Proceedings of the 2nd IEEE Conference on Standardization and Innovation in Information Technology. Boulder, Colorado USA, 2001.
9. Boguslasvsky, I. Some remarks on the UNL encoding conventions. Proceedings of the First International UNL Open Conference "Building global knowledge". SuZhou, China, 2001.
10. Cardeñosa, J., Iraola, L. and Tovar, E. Managing a real implantation of the UNL System. First International UNL Open Conference "Building global knowledge". SuZhou. China, 2001.
11. European Commission. HEREIN Project (IST-2000-29355). Final Report (2003)
12. Cardeñosa, J., Gallardo, C. UNESCO Project: General Methodology for UNL conversion and multilingual generation. Technical document". Spanish Language Centre. Facultad de Informática. UPM. Spain, 2003.