The UNL Distinctive Features: Inferences from a NL-UNL Enconverting Task

Ronaldo Teixeira Martins[*], Lúcia Helena Machado Rino[**], Maria das Graças Volpe Nunes[***], Osvaldo Novais Oliveira Jr.[****]

[*]Núcleo Interinstitucional de Lingüística Computacional - NILC
Av. do Trabalhador São-Carlense, 400 - 13560-970 - São Carlos, SP, Brazil
ronaldo@nilc.icmc.sc.usp.br
[**]Departamento de Computação - Centro de Ciências Exatas e de Tecnologia - UFSCar
Rod. Washington Luiz, km 235 - Monjolinho - 13565-905 - São Carlos, SP, Brazil
lucia@dc.ufscar.br
[***]Instituto de Ciências Matemáticas e da Computação (ICMC) - Universidade de São Paulo
Av. do Trabalhador São-Carlense, 400 - 13560-970 - São Carlos, SP, Brazil
mdgvnune@icmc.sc.usp.br
[****]Instituto de Física de São Carlos (IFSC) - Universidade de São Paulo
Av. do Trabalhador São-Carlense, 400 - 13560-970 - São Carlos, SP, Brazil
chu@ifsc.sc.usp.br

Abstract

This paper reports on the distinctive features of the Universal Networking Language (UNL). We claim that although UNL expressions are supposed to be unambiguous, UNL itself is able to convey vagueness and indeterminacy, as it allows for flexibility in enconverting. The use of UNL as a pivot language in interlingua-based MT systems is also addressed.

1. Introduction

Machine Translation (MT) is one of the most controversial subjects in the field of natural language processing. Researchers and developers are often at odds on issues concerning MT systems approaches, methods, strategies, scope, and their potentialities. Dissent has not hindered, however, the establishment of tacit protocols and core beliefs in the area. It has often been claimed that[1]: 1) fully automatic high-quality translation of arbitrary texts is not a realistic goal for the near future; 2) the need of some human intervention in pre-edition of the input text or in post-edition of the output text is mandatory; 3) source language should be rather a sublanguage, and the input text should be domain- and genre-bounded, so that the MT system could cope with natural language ambiguity; 4) the transfer approach is more feasible than the interlingual one, since the latter, albeit more robust and economic, is committed to the somewhat insurmountable task of designing a perfect (universal) language, comprising any other one; 5) common sense and general knowledge on both the source and the target cultures are as important as linguistic information, like in Knowledge-Based Machine Translation Systems (Nirenburg et al., 1992); 6) existing human translations can be used as a prime source of information for the production of new ones, similarly to the Example-Based Machine Translation Systems (Furuse and Iida, 1992); 7) existing MT systems are not appropriate to monolingual users, although they can be used to facilitate, speed up or reduce the costs of human translation, or to produce quick and cheap rough translations that may help the users to get a very broad idea of the general subject of the text.

[1] Most of these assumptions can be extracted from the Survey on the State of the Art in Human Language Technology (Cole et al., 1995). Of special interest are the articles concerning multilinguality by Martin Kay (8.1, 8.2) and Christian Boitet (8.3, 8.4).

Many authors obviously do not endorse all the listed statements, specially the fourth one. Hozumi Tanaka (1993), for example, argues in favor of the interlingua-based approach, and so do the research and development groups involved in interlingua-based systems, such as ULTRA (Farwell and Wilks, 1993), KANT (Mitamura et al., 1993), or PIVOT (Okumura et al., 1993). These works, however, rather confirm the very general observation that commercially available MT systems (e.g., SYSTRAN, VERBMOBIL, DUET (Sharp), ATLAS I (Fujitsu), LMT (IBM), METAL (Siemens)) are primarily transfer-based.

The most serious arguments against the interlingua approach concerns its alleged universality and excessive abstractness (Hutchins and Somers 1992). In order to cope with multilinguality, the interlingua should put aside language-dependent structures (such as the phonological, morphological, syntactical and lexical ones) and work at the logical level, which is supposed to be shared by human beings. Even at such uppermost level, however, there seems to be cultural differences. Eco (1994) reports, for instance, the case for Aymara, a South-American Indian language which would have three truth values, instead of the two "normal" ones. Furthermore, it has been said that, even if one comes to find this kind of perfect language, it would be so abstract that it would not be cost-effective, since the tools for departing from natural language and arriving at the logical representation would be excessively complex.

In what follows, we present some extra evidence towards the feasibility of interlingua-based MT. The Universal Networking Language (hereafter, UNL), developed by Uchida et al. (1999), brings some distinctive features that may lead to overcome some of the bottlenecks frequently associated to the interlingua approach. Although UNL was not designed as an interlingua, and MT is only one of the possible uses for UNL, it has been claimed that multilingual MT systems can use UNL as a pivot language. In this paper, some of the distinctive features of UNL are analyzed. We build

upon the experience in developing the Brazilian Portuguese (hereafter, BP) UNL Server, a bilingual MT system for translating Portuguese into UNL and vice-versa.

This paper is organized as follows. Section 2 provides a brief introduction to the UNL approach and some of its premises. In Section 3 we describe an experiment in which human subjects were asked to enconvert sentences from Portuguese into UNL. Section 4 brings the general results of the experiment. One of them is specially addressed in Section 5. Some issues arising from the results are presented in Section 6. Conclusions are stated in Section 7. The reader is supposed to have previous information on the UNL Project and knowledge on UNL Specification (at http://wwww.unl.ias.unu.edu) is considered mandatory.

2. The Universal Networking Language

The Universal Networking Language (UNL) is "an electronic language for computers to express and exchange every kind of information" (Uchida et. al., 1999, p. 13). According to the UNL authors, information conveyed by each natural language (NL) sentence can be represented as a hyper-graph whose nodes represent concepts and whose arcs represent relations between concepts. These concepts (called Universal Words or simply UWs) can also be annotated by attributes to provide further information on the circumstances under which they are used.

In this context, UNL is not different from the other formal languages devised to represent NL sentence meaning. Its structure is said to suffice to express any of the many possible meanings conveyed by any sentence written in any NL. This does not mean, however, that it is able to represent, at the same time, all the possible meanings conveyed by the very same NL sentence. Instead, UNL is able to represent each of them independently, and it is by no means able to provide a single structure coping with all of them. In this sense, there will never be a single UNL expression that completely suffices the meaning correspondence to a NL sentence. Or else: no UNL expression will be ever completely equivalent to a NL sentence, since the latter, but not the former, will allow for ambiguity.

In the following section, we report on results of a BP-UNL enconverting task that has been carried out by BP native speakers. In this experiment, we observe evidences that BP sentences must be disambiguated in order to be represented as UNL expressions.

3. The Experiment

In August 2001, we carried out an experiment on BP-UNL enconverting that involved 31 BP native speakers, all of them graduate and postgraduate students. Most of them (over 95%) were Computer Sciences students, aging 21 to 42 years old (90% of them were under 30 years old).

The experiment was split into training (steps 1-4) and test sessions (step 5), as follows: 1) a very general description of the UNL structure; 2) a general presentation of the definitions provided for five relation labels by the UNL Specification (1999), namely, 'agt' (agent), 'cag' (co-agent), 'obj' (affected thing), 'cob' (affected co-thing), and 'ptn' (partner); 3) an individual exercise on the use of the presented relation labels, in which subjects were asked to identify 50 different relations appearing in different BP sentences, indicating the corresponding UNL relation labels; 4) a public discussion on the exercise results; and 5) a final individual test in which subjects were asked again to identify 30 different relations appearing in different BP sentences, through their correspondence with the very same set of UNL relation labels. In Step 3 and 5, the subjects had also the option of pinpointing the impossibility of identifying either a relationship or its corresponding relation label, by choosing a "catch all" alternative (see option (a) in Figure 1). This exercise aimed at providing the means for the subjects to understand and explore BP-UNL enconverting, concerning the relation labels identification. This was then reinforced in Step 4, which was supervised by a UNL specialist. As it can be observed, these steps aimed at Step 5, the actual BP-UNL assignment, focusing on specific relation labels. In this step, some of the BP sentences presented to the subjects in Step 3 have been replicated.

Altogether, this experiment has taken 1 hour and 40 minutes, considering a 20-minute interval between the training and test sessions. Steps 1 and 2 have last 20 minutes, and so has Step 3 alone. Step 4, the longest one, has taken 40 minutes. Step 5, the actual test, has taken another 20 minutes. The interval between training and test aimed at allowing for the subjects settling on UNL specification, since test has been totally unsupervised. This also justifies our replication of some of the BP sentences used in training.

An English version of the task proposed in Step 3 is presented in Figure 1 below.

Considering the information presented in the first part of this experiment, identify the following:
1) If the relation depicted between the words signaled in each of the sentences below belongs to the five-relation set discussed previously; and
2) If so, which relation label would most suitably describe the involved relationship.

Use, for reference, the following code:
a) if NO label describes the relationship between the signaled words;
b) if the label AGT (agent) is the most suitable one;
c) if the label CAG (co-agent) is the most suitable one;
d) if the label COB (affected co-thing) is the most suitable one;
e) if the label OBJ (affected thing) is the most suitable one;
f) if the label PTN (partner) is the most suitable one.

Figure 1. Instructions for identifying and classifying relations.

The 30-sentence set used in the test session, along with its corresponding English translation, is shown in Figure 2.

	SENTENCES
1.	A crise quebrou o empresário >> ???(quebrou, crise) *The crisis broke the business man.* >> ???(broke, crisis)
2.	A crise quebrou o empresário >> ???(quebrou, empresário) *The crisis broke the business man.* >> ???(broke, business man)
3.	A farsa acabou. >> ???(acabou, farsa) *The farce is over.* >> ???(is over, farce)
4.	A neve caía lentamente. >> ???(caiu, neve) *Snow felt slowly.* >> ???(felt, snow)

5.	Alugam-se casas. >> ???(alugar, casa) *Houses are rented* (also: *Someone rents houses*) >> ???(are rented, houses)
6.	Choveu canivete ontem. >> ???(choveu, canivete) *It rained knives yesterday >> ???(rained, knives) (Brazilian Idiom)*
7.	João jogou o vaso com Maria contra Pedro. >> ???(jogou, Maria) *John threw the bowl with Mary against Peter.* >> ???(threw, Mary)
8.	João jogou o vaso com Maria contra Pedro. >> ???(jogou, Pedro) *John threw the bowl with Mary against Peter.* >> ???(threw, Peter)
9.	João lutou com Maria para vencer a doença. >> ???(lutou,Maria) *John fought with Mary to win the disease.* >> ???(fought, Mary)
10.	João não teve filhos com Maria. >> ???(ter, João) *John did not have children with Mary.* >> ???(have, John)
11.	Maria esqueceu o dia do aniversário da filha. >> ???(esquecer, dia) *Mary forgot her daughter's birthday.* >> ???(forgot, birthday)
12.	Maria foi despedida. >> ???(despedir, Maria) *Mary was fired.* >> ???(fire, Mary)
13.	Maria lembrou Pedro do horário. >> ???(lembrou, horário) *Mary remembered Peter about the schedule.* >> ???(remembered, schedule)
14.	Maria morreu com a falta de oxigênio.. >> ???(morreu, falta) *Mary died with the lack of oxygen.* >> ???(died, lack)
15.	Maria namorou Pedro. >> ???(namorou, Maria) *Mary flirted (with) Peter.* >> ???(flirted, Mary)
16.	Maria não foi ao cinema com a vizinha. >> ???(foi, vizinha) *Mary did not go to the cinema with her neighbor.* >> ???(go, neighbor)
17.	Maria não quis matar Pedro! >> ???(matar, Maria) *Mary did not intend to kill Peter.* >> ???(kill, Mary)
18.	Maria não se sentiu bem. >> ???(sentir, Maria)
	Mary did not feel well. >> ???(feel, Mary)
19.	Maria nunca conquistou Pedro. >> ???(conquistou, Pedro) *Mary never conquered Peter.* >> ???(conquered, Peter)
20.	Maria parece cansada. >> ???(parece, Maria) *Mary looks tired.* >> ???(looks, Mary)
21.	Maria se esqueceu de João. >> ???(esquecer, João) *Mary forgot John.* >> ??(forgot, John)
22.	Maria se matou. >> ???(matou, Maria) *Mary killed herself.* >> ???(kill, Mary)
23.	O filme deu origem a muitas controvérsias. >> ???(deu, filme) *The movie raised many controversies* >> ???(raised, movie)
24.	O frio congelou o pássaro. >> ???(congelar, frio) *The cold froze the bird.* >> ???(froze, cold)
25.	O medo da morte provoca insônia. >> ???(provoca, medo) *Fear of death causes insomnia.* >> ???(causes, fear)
26.	O pai com os filhos matou a mãe. >> ???(matou, filhos) *The father with the children killed the mother.* >> ???(killed, children)
27.	O pássaro congelou com o frio. >> ???(congelar, frio) *The bird froze (i.e., was frozen) with the cold.* >> ???(froze, cold)
28.	Os carros se chocaram na estrada. >> ???(chocaram, carros) *The cars crashed each other on the road.* >> ???(crashed, cars)
29.	Pedro se parece com a mãe. >> ???(parece, mãe) *Peter looks like his mother.* >> ???(looks, mother)
30.	Precisa-se de funcionários. >> ???(precisar, funcionários) *Employees are needed.* (also: *Someone needs employees*) >> ???(need, employees)

* Students were presented only to the original Brazilian Portuguese sentence. In the translation from Portuguese into English we tried to preserve the Portuguese syntactic structure as often as possible, even when the resulting English sentence sounds agrammatical.

Figure 2. Test corpus.

4. Results

The results of the experiment were the following:

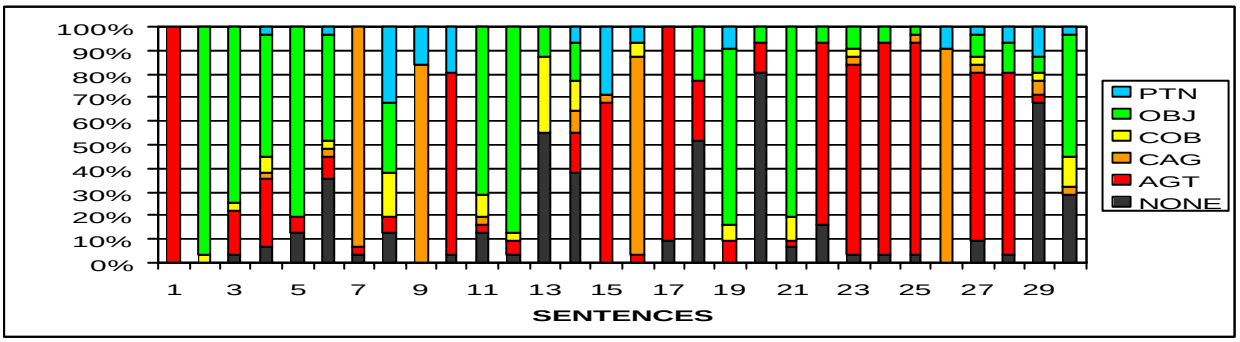

Figure 3. Distribution of BP-UNL enconvertings by subjects, with respect to the 5-relation labels set

Figure 4 below groups the results according to the agreement among enconverters.

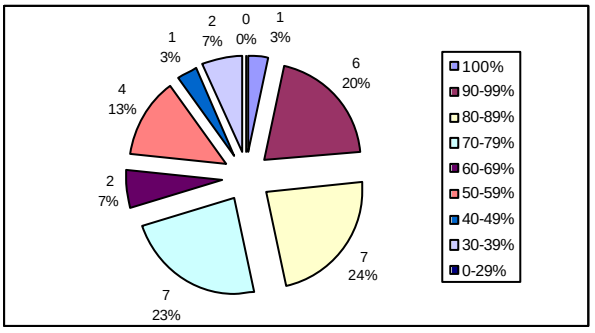

Figure 4. Agreement among enconverters.

A single relation (between "crise" (*crisis*) and "quebrou" (*to break*) in sentence 1: *"A crise quebrou o empresário" (= The crisis broke the business man*) led to an agreement of 100% among enconverters: they all used the 'agt' label in this case. There was an agreement between 90% to 99% on labeling relations in 6 sentences. Enconverters also agreed between 80% to 89% in assigning labels in 7 sentences. Other 7 sentences involved 70% to 79% agreement. In the remaining 9 sentences, agreement among enconverters was lower than 70%.

5. Case Study: Sentence 14

Sentence 14 ("Maria morreu com a falta de oxigênio." (literally: "Mary died with the lack of oxygen.") can be taken as a typical example of those involving considerable disagreement among enconverters. The relation between the verb "morreu" (to die) and the noun "falta" (lack) was encoded in varied ways, as follows: a) as an agent one (16%); b) as an object one (16%); c) as a co-object one (13%); d) as a co-agent one (10%); e) as a partner one (6%); and f) as none of the previous five relations (39%).

The unavoidable issue that follows from the above is why UNL labels were used in such apparently fuzzy way. Several reasons could be pinpointed here: a) the lack of expertise (or even of attention) of human enconverters', for they could not have had enough knowledge of language, or motivation, to carry on the experiment (although they are BP native speakers and seemed to be willingly helpful and interested in participating); b) the lack of clarity of the UNL Specification itself, even though there had been considerable discussion in the training session, for the problems posed by the enconverters to be tackled; c) the structure of the experiment itself, which was indeed too brief and too shallow to properly evaluate the human enconverters' performance; and, finally, d) the ambiguity of test sentences.

The analysis of the enconverters' choices certifies that disagreements are due to the latter point. Although it is unlikely for a BP speaker to say that 14 above, out of context, could have many different colliding meanings, the experiment has proved that apparently unambiguous sentences are unambiguous only apparently. Although eventually invisible, NL vagueness and indeterminacy would be pervasive in ordinary language,

Actually, none of the labels assigned to the relation between "morreu" (to die) and "falta" (lack) in sentence 14 could be considered wrong. The lack of oxygen could be understood in many distinct ways, such as:

a) an agent ("agt"), or the "initiator of the action" of "Mary dying" (or "killing Mary");

b) a co-agent ("cag"), or a "non-focused initiator of an implicit event that is done in parallel", in the sense it was not the lack of oxygen that killed Mary but either b.1) the situation (or the person) that has provoked the suppression of Mary's air supply or, in a more precise way, b.2) the reaction provoked (mainly in the brain) by the lack of oxygen;

c) an object ("obj") for the event described by "dying", since it is somehow "directly affected" by it, as the conclusion that the oxygen was lacking might be said to come directly from the fact that Mary died, otherwise no one would perceive that oxygen was lacking;

d) an affect co-thing ("cob"), or as being "directly affected by an implicit event done in parallel", if the observation that the oxygen was lacking were said not to come directly from the fact that Mary died, but from the fact that her lungs stopped working, which caused her to die;

e) a partner ("ptn"), for it could be somewhat "an indispensable non-focused initiator" of the action of "Mary dying", as if the main responsible for Mary's death was Mary herself (or someone else) that turned the oxygen suply off.

Besides such illustrations, many other relations can be said to hold between 'lack of oxygen' and 'die', namely, "met" (method), "man" (manner), "ins" (instrument), and "rsn" (reason), all easily applicable to such a case.

Such a variety proves that sentence 14 was indeed vague. The syntactic relation between the BP verb and its adjunct can convey many different semantic cases. Nevertheless, the UNL expression – whatever it may be – will have, in turn, a single interpretation, because relation labels are not supposed to overlap. The relations agt(die,lack), cag(die,lack), cob(die,lack), obj(die,lack), ptn(die,lack), although applicable to that very same NL sentence, are expected to label different (albeit related) phenomena. Indeed, to say agt(die,lack) is not the same as to say cag(die,lack) or ptn(die,lack). No intersection between these relations is envisaged in the UNL Specification, since they are meant to be exclusive[2].

This makes clear that the UNL specification forces filtering possible interpretations for NL sentences, in the sense a UNL expression must provide a completely unambiguous representation for the source sentence. As a matter of fact, although UNL is intended to be as expressive as any NL, UNL expressions cannot convey, at least at the relation level, NL vagueness and indeterminacy. Like any other formal language, UNL is committed to disambiguate NL sentences and, hence, to impoverish their semantic power.

Nevertheless, in no one of the above situations it is possible to say that a relation label is wrong, or that is completely inappropriate, although some of them may seem really unlikely to hold, depending on the context. The point is that the meaning of the sentence "Mary died with the lack of oxygen." is not encapsulated in the sentence itself but it is built out from the reading (and hence from the analysis) made by human enconverters. Since different enconverters have different underlying assumptions during their readings, the same BP phenomena can naturally imply different interpretations, which in turn lead to distinct UNL labeling. To conclude, it seems impossible to prevent subjectivity (or context-sensitiveness, or else, enconverter-sensitiveness) at that extent, no matter how univocal NL sentences seem to be.

6. Consequences

From the above it is possible to state that UNL should not seek for a straightforward correspondence between UNL expressions and NL sentences. It would be useless. As meaning is not encrypted in NL sentences but build through the analysis process, different enconverters will unavoidably propose different UNL expressions for the

[2] Accordingly, it is worthy to observe that the individuality of relations seems to be less strong when we consider other UNL relation labels set, e.g., that comprising "qua" (quantity), "nam" (name) and "pos" (possessor), which seems to be, to some extent and context, replaceable by "mod" (modification), implying that the latter can quite feasibly be at an uppermost level in a relation hierarchy. The same could be said of "met" (method) and "ins" (instrument), which seem to be under the scope of "man" (manner). Conversely, this does not mean that "mod" comprises any of "qua", "nam", or "pos", or that "man" embeds "met" and "ins". Instead, it does mean that both "mod" and "man" seem to share a comprehensive set of features with the relations that they replace. This is not the case of "agt", "cag", "cob", "obj", and "ptn", which seem to be in a more outstanding opposition.

very same NL sentence and many of these different expressions are legitimate.

Due to structure of UNL, UNL expressions cannot replicate NL sentence vagueness and indeterminacy. Enconverters are obliged therefore to choice a single interpretation among many different possible ones. This choice will be inevitably affected by the enconverters' context, which will be unreplicable itself by other enconverters. Once all these enconvertings will be valid, in the sense they are context-motivated, there will never be a one-to-one mapping between NL sentences and UNL expressions.

Accordingly, correctness, in UNL, instead of representing a (impossible) single possibility of enconverting, should rather be considered as fidelity to enconverters' intentions. UNL should clearly state that it would be up to the (human and machine) enconverter to decide what should the UNL representation be for a NL sentence. That is to say, the object of the UNL representation should be considered not exactly the meaning conveyed by the NL sentence but the *interpretation inferred by the enconverter from the use of that NL sentence in the enconverter's specific context*.

The fact that there could be more than a single (and adequate) UNL expression for the same NL sentence implies that UNL allows for flexibility in the enconverting process, although the UNL expression itself is not supposed to be flexible. It is up to the enconverter, and not the UNL specification itself, to decide which of the many possible interpretations is to be represented by a UNL expression. This is a significant UNL distinctive feature. Most formalisms do not allow for such variability and postulate that there should be a biunivocal relation between NL and its artificial representation. Otherwise, the formal representation would keep mirroring NL vagueness and indeterminacy, resulting useless.

The problem here is how to assure that enconverting flexibility will not prevent UNL from being a machine tractable language. As far as UNL expressions are dependent on the enconverter, there could be uncontrolled variations, which could blow out UNL into many different (and maybe mutually unintelligible) dialects.

This problem can be divided into two parts: 1) how to be sure that the UNL expression represents indeed what is intended by the enconverter; and 2) how to be able to generate, from such varied UNL expressions, NL grammatical sentences.

The first question is somewhat an educational problem. There are obviously misunderstandings and misuses of many relations. To say that it is up to the enconverter to decide which label should be used is not to say that the enconverter can do whatever he/she/it wants. The UNL Specification and other guidelines are to be followed. The relation "agt" must be applied to "a thing that initiates an action", and "ptn" should stand for "an indispensable non-focused initiator of an action". The relation "agt" cannot be used in a different sense: it would be wrong. Flexibility in encoding should not be mistaken for permissiveness. There are many correct UNL expressions for the same NL sentence, but there are also wrong UNL expressions.

The solution to such a problem cannot be, however, to state a rigid (a culture-, language-, context- and even enconverter-independent) relationship between a NL and UNL, otherwise UNL will not suffice to cope with inevitable varying enconvertings. The fact that meaning is build through the enconverting process and its main consequence, the fact that different enconverters will propose different expressions for the same NL sentence, should be both considered starting points, instead of something that one can or should avoid.

The best solution is, thus, to trust the enconverter (and maybe to certify enconverters), and to be conscious that, as in any other translation activity, there are good and bad translations, and bad translations do not prove that translating is not possible or that it does not work. Only time and enconverters' expertise can make UNL expressions better.

Nevertheless, to trust enconverters may imply making deconverting extremely difficult and costly. The more UNL allows flexibility in enconverting, the more costly will be UNL-NL deconverting, since the UNL expression may contain unexpected relations.

This is, however, a false problem. Deconverters are not committed to generate back the source sentence enconverted into UNL. Instead, they should be supposed to generate a NL sentence corresponding to the UNL expression. The original source sentence is definitely lost as it has been enconverted into UNL; only one of its possible interpretations (the one carried out by the enconverter) is preserved. Deconverters should take then UNL expression as the new source sentence, instead of using it just as an intermediate expression.

Furthermore, deconverting seems to be easier than enconverting, since much of the eventual meaning gaps may be inferred from the context by a human being (which is supposed to be the final user), instead of a machine. There is a very fragile break-even-point, from which generation results become excessively degraded, but the extent to which this happens will depend on the architecture of the UNL System.

7. Conclusion

The main conclusion to be extracted from the previous section seems to be a paradox: in multilingual MT Systems, in order to be a pivot language, UNL should not be treated as an interlingua, but as a source and a target language, at the same level as any other NL. Flexibility in enconverting brings UNL to be just like any other NL, in the sense it would allow UNL for coping with NL vagueness and indeterminacy, without sacrificing, however, the explicitness and clarity of UNL expressions, which would continue to be univocal and machine-tractable.

Acknowledgments

The authors acknowledge Mr. Tadao Takahashi, for his management of the Brazilian branch of the UNL Project, and to CNPq (Brazil) for the partial financial support. This work has also been partially supported by the UNU/IAS and the UNDL Foundation. The opinions expressed in the paper, however, do not necessarily represent the view of the UNL mentors.

References

Cole, R.A.; Mariani, J.; Uszkoreit, H.; Zaenen, A.; Zue, V. (Eds.) (1995). *Survey of the State of the Art in Human Language Technology*. NSF/CEC/CSLU. Oregon Graduate Institute. November.

(http://cslu.cse.ogi.edu/HLTsurvey/HLTsurvey.html)
Eco, U. (1994). *La recherche de la langue parfaite dans la culture européene*. Paris, France: Editions du Seuil.
Farwell, D. and Wilks, Y. (1993). ULTRA: A Multilingual Machine Translator. In S. Nirenburg (Ed.), *Progress in Machine Translation*. Washington, DC: IOS Press.
Furuse, O. and Iida, H. (1992). Cooperation between transfer and analysis in example-based framework. In *Proceedings of the 14th International Conference on Computational Linguistics*. Nantes, France.
Hutchins, W. J. & Somers, H. L. (1992). *An Introduction to Machine Translation*. San Diego, CA: Academic Press.
Mitamura, T., Nyberg, E. and Carbonell, J. (1993). In S. Nirenburg (Ed.), *Progress in Machine Translation*. Washington, DC: IOS Press.
Nirenburg, S., Carbonell, J, Tomita, M. and Goodman, K. (1992). *Machine Translation: A Knowledge-Based Approach*. San Mateo, CA: Morgan Kaufman.
Okumura, A., Muraki, K and Akamine, S. (1993). Multi-lingual Sentence Generation from PIVOT Interlingua. In S. Nirenburg (ed.), *Progress in Machine Translation*. Washington, DC: IOS Press.
Tanaka, H. (1993). Multilingual Machine Translation Systems in the Future. In S. Nirenburg (ed.), *Progress in Machine Translation*. Washington, DC: IOS Press.
Uchida, H., Zhu, M. and Della Senta, T. (1999). *Universal Networking Language: A gift for a millennium.* Tokyo, Japan: The United Nations University.