Some Controversial Issues of UNL: Linguistic Aspects

Igor Boguslavsky

Universidad Politécnica de Madrid (Spain)/IITP RAS (Russia)
igor@opera.dia.fi.upm.es

Abstract. We discuss several linguistic aspects of the Universal Networking Language (UNL); in particular, those connected with Universal Words (UWs), UNL relations, and hypernodes. On the one hand, the language should be rich enough and provide sufficient means to express the knowledge that might be required in the applications it is intended for. On the other hand, it should be simple enough to allow uniform and consistent use across languages and by all encoders. The major expressive device of UNL used for overcoming lexical divergence between languages is so-called restrictions. They have three functions, which are relatively independent of each other: the ontological function, the semantic function, and the argument frame function. We discuss various types of restrictions and propose new expressive means for describing UWs. Sample dictionary entries are given which incorporate our proposals. We propose several new UNL relations and discuss when and how hypernodes should be introduced.

1 Background

Among many problems that developers and users of a meaning representation language are facing, two somewhat conflicting requirements are standing out. On the one hand, the language should be rich enough and provide sufficient means to express the knowledge that might be required in the applications it is intended for. The more complex and knowledge-demanding the application, the more complex the design of the meaning representation language becomes. On the other hand, it should be simple enough to allow uniform and consistent use across languages and by all encoders. In the case of UNL, the latter problem is particularly serious, since the encoders work in different countries, belong to different linguistic schools, and have different linguistic traditions. Therefore, uniform understanding and use of UNL by all partners is difficult to achieve.

Since the start of the project in 1996, a large number of UNL-encoded documents have been accumulated that were produced by the project participants from 16 language groups each working on its native language. The analysis of these documents clearly shows two things: *UNL is still lacking means to express meaning adequately*, and *there is not enough uniformity in the UNL use among the partners*. To some extent, UNL has developed its own dialects. Despite the existence of the UNL Specifications, divergences between the dialects tend to grow. This tendency clearly manifests itself in the fact that all deconverters (=generators) are doing much better when dealing with the UNL documents produced by the authors of the deconverter than

Universal Network Language: Advances in Theory and Applications.
Research on Computing Science 12, 2005, pp. 77–100.

with those provided by other teams. If it goes on this way, the dialects will soon become hardly understandable by the deconverters and we will need special modules to translate from one UNL dialect to another.

These problems were raised at several discussions at the UNL workshops and working sessions. Of particular importance was the "Forum Barcelona 2004" project carried out in 2001 by the UNL groups from France, India, Italy, Russia and Spain. During this work a number of texts were encoded to UNL by project participants and each text was extensively debated. Participants of the discussion have been: Ramon Armada, Pushpak Bhattacharyya, Etienne Blanc, Igor Boguslavsky, Carolina Gallardo, Luis Iraola, and Irina Prodanoff. The results of this debate were presented in [1] and at the UNL conference in Suzhou, [2]. In this paper, I will summarize the understanding of UNL that took shape in the course of discussions and put forward some proposals on the linguistic aspects of UNL.

The paper is organized as follows. In section 2, some general remarks will be made concerning the requirements imposed on UNL representations. Section 3 will be devoted to Universal Words. In section 4, I will give some comments on the semantic categories of UWs which constitute upper levels of the UNL Knowledge Base. Aside from that, there will be no special discussion of the problems connected with the UNL Knowledge Base and Master Dictionary. Issues of UNL relations will be discussed in section 5. Finally, in section 6 I will speak about hypernodes (scopes).

I will not give any introduction to UNL. It can be found, for example, in [3], [4], [5]. It is expected that the reader have some preliminary knowledge of UNL, at least as far as the UNL Specifications are concerned [6].

2 General remarks

UNL representations (UNLR) can be evaluated from the points of view: correctness and adequacy. A UNLR is correct if it conforms to UNL specifications. To be adequate, the UNLR should contain enough information and be convenient for the applications it is intended to serve. UNL is conceived as a meaning representation language applicable in a wide range of applications – multilingual generation, machine translation, information retrieval, text summarization, question answering. I will discuss it mainly from the perspective of one of them – multilingual generation of UNL documents for the dissemination of information in the Internet. This is the application that received most attention in the UNL development so far and, at the same time, it is one of the most demanding.

To be adequate for multilingual generation, a UNLR should meet at least two requirements:

- it should preserve the meaning of the source text to a reasonable extent (i.e. without a significant loss);
- it should permit generation of the text bearing this meaning in all working languages.

Since the enconversion, i.e. transformation of the source text into UNLR, is not supposed to be fully automatic, we can address our encoding recommendations to a

human who will produce UNLRs with the help of special tools. These tools may range from more or less sophisticated editors (cf. for example the EditorUNL developed by the Spanish group, [7]) to semi-automatic enconverters (cf. for example the UNL module of the ETAP-3 system developed by the Russian group [8]).

The UNLRs need not be literal. They should not necessarily preserve the structure of the original sentence, nor its lexical composition. The only thing required of them is to represent the original meaning in a satisfactory way. To do it, the UNL writer may paraphrase the text in any way he/she finds convenient, provided the meaning of the original and its communicative intention remain intact. In particular, long sentences may be divided into several shorter ones. Language-specific syntactic constructions and idioms may be replaced with simpler constructions and non-idiomatic synonymous expressions, or an equivalent English idiom, should it exist.

To give a simple example, consider Spanish sentence (1):

(1) *Los estudiantes tenemos que trabajar mucho.*

Literally, the sentence reads: 'the students have to work much'. But this is not the whole meaning of the sentence. An idiosyncratic feature of this construction is that the predicate (*tenemos que* 'have to') has the form of the first person plural (= 'we have to') and therefore does not agree in the grammatical category of person with the subject (*estudiantes* 'students'). Due to this grammatical peculiarity, the meaning of (1) is 'we, students, have to work much'. What should be the adequate UNLR for (1)? A straightforward solution would be (1a) that directly reflects the structure of the source sentence:

(1a) aoj(must.@entry.@1-person, student.@pl)
obj(must.@entry.@1-person, work)
man(must.@entry.@1-person, much)

However, this UNLR should be discarded as too specific. It is the idiosyncratic property of Spanish to encode the information on the subject ('we') in the verb form. UNL should express this information in a less language-specific way:

(1b) aoj(must.@entry, we)
cnt(student, we)
obj(must.@entry, work(icl>do))
man(work(icl>do), much)
(1b = 'we being students must work much').

However, the freedom of replacing phrases with their paraphrases should be used with great caution. For example, special terms cannot be paraphrased and must be represented in the form in which they exist in English. For instance, *sustainable development* should not be represented as obj(sustain(icl>maintain).@ability, development.@entry). This UNLR, though it conveys a meaning close to the original phrase – "development that can be sustained", – is unacceptable as a representation for a term.

3 Universal Words (UW)

As an element of the dictionary, a UW consists of two major parts: the headword and the restrictions[1].

3.1 Headwords

As defined in the UNL Specifications, any English word, phrase or sentence can be a headword for a UW. UNL corpora abound in headwords consisting of more than one word, such as *Ministry of Foreign Affaires*, *Telecommunication Development Bureau, sustainable development, week-long feast,* etc. In our opinion, multi-word headwords should be introduced with much care. When a multi-word expression is compositional, i.e. when its meaning is representable as a combination of meanings of words it is composed of, it is better to represent it as a combination of UWs linked with appropriate relations and not as one multi-word UW.

Examples[2]:

(2) *sustainable development*
(2a) mod(development, sustainable)
(3) *week-long feast*
(3a) dur(feast, week)
qua(week, 1)

An example of a non-compositional phrase that could with good reason generate a multi-word headword is (4):

(4) *look for*
(4a) look for(icl>do,agt>thing,obj>thing).

However, even in this case a multi-word UW is not the only alternative. One can consider *look for* as a realization of a special lexical meaning of *look,* but in this case the meaning should be accordingly restricted:

(4b) look(icl>search>do,agt>thing,obj>thing).

The reason for avoiding multi-word headwords is obvious: if any free word combination can be made into a UW, one can hardly hope that other partners will have matching UWs in their dictionaries.

On the other hand, the idea behind the multi-word UWs is to express the fact that they denote a single concept. It might be useful to keep this information. Then, a convenient compromise might be to enclose the UNLR in a scope:

(5a) mod:01(ministry.@entry, affair.@pl)
mod:01(affair.@pl,foreign)

[1] As an element of UNLR, a UW can be supplied with additional pieces of information such as ID number and attributes.

[2] For simplicity's sake, here and in some other examples I will omit restrictions that are not directly relevant for the discussion.

Another solution (proposed by Ch. Boitet) is to allow UWs to have internal structure:

(5b) mod(ministry,affair.@pl)&mod(affair.@pl, foreign)

It should be noted, however, that the second solution requires a considerable modification of the specifications and of the EnCo/DeCo software.

3.2 Restrictions

UW restrictions have three functions:

- Ontological function: locate the UW in the Knowledge Base. This is needed, in particular, to ensure understanding of the UW in the case that it is absent in the dictionaries of some working languages and to help semantic inference.
- Semantic function: restrict the meaning of the headword. This is needed, in particular, to ensure disambiguation of the headword and selection of the translation equivalent.
- Argument frame function: provide the argument frame for the UW.

It is important to emphasize, that the requirements imposed by these functions do not always coincide. A restriction that is good for one purpose is not necessarily adequate for another. For example, restrictions of the type (icl>thing) or (icl>how) or (icl>do) are often very efficient for disambiguation, since they differentiate nominal, adverbial and verbal meanings from each other. At the same time, they will not help us much to translate a UW, if we don't have this UW in the dictionary.

On the other hand, the word *pern* is monosemic and does not need disambiguation. But if we don't have an exhaustive list of different varieties of birds in the dictionary, the restriction (icl>bird) will be very helpful to provide an understandable translation for this word. It can also be of help in other situations in which it is useful to know that the word denotes a bird.

The third function of restrictions – specification of the argument frame of the word – should also be clearly separated from other functions. One may wish to restrict the meaning by specifying some semantic relation (first function), but it does not necessarily imply that this relation makes part of the argument frame of the word. The English verb *to land* denotes reaching the land both from the sky (*The airplane landed on time*) and from water (*We landed on a lonely island in the middle of the ocean*). In these situations, Russian uses different verbs – *prizemljat'sja* and *vysazhivat'sja*, respectively. To construct UWs for these verbs, we need to restrict the meaning of *to land*. An obvious way to do so would be to indicate the initial point of the movement (src relation): *prizemljat'sja* = land(src>sky); *vysazhivat'sja* = land(src>water). However, these verbs do not have argument slots for the initial point of movement.

Restrictions on the basis of which the UWs are arranged in the KB will be designated **KB restrictions**. Restrictions oriented primarily towards the second goal will be called **semantic restrictions.** Restrictions which specify the argument frame will be referred to as **argument frame restrictions**. A restriction may serve more than one goal. For example, restrictions in the UWs orange(icl>fruit), orange(icl>tree), orange(icl>colour) can equally well differentiate three different meanings of the noun

orange and specify the KB position of each of them. However, we should keep in mind that in the general case semantic, argument frame and KB restrictions do not coincide.

The UNL dictionaries must have means by which we could distinguish between these three types of restrictions. KB restrictions are clearly separated from other types of restrictions, since they are only represented in the Master Dictionary and not in the UW dictionary. As a matter of fact, the difference between the Master Dictionary and the UW dictionary boils down to the presence/absence of the KB restrictions. As for the argument frame restrictions, in the present version of the UW dictionary they are represented very poorly and are not separated from semantic restrictions.

We will discuss semantic and argument frame restrictions in sections 3.2.1 and 3.2.2 respectively.

3.2.1 Semantic restrictions

As mentioned above, the function of semantic restrictions is to effectively separate the meaning of the UW from all other meanings which the headword may have. The major requirement imposed on semantic restrictions is as follows. Restrictions ascribed to a UW should not be equally applicable to other meanings of the same headword. For example, the UW people(icl>human) does not meet this condition, since the headword *people* has two different meanings, and both of them are covered by restriction (icl>human): 'persons' (as in *many people*) and 'nation' (as in *peoples of Africa*). Similarly, all meanings of the noun *operator* can be characterized as belonging to the "thing" category. Therefore, restriction (icl>thing) is too broad and should be narrowed down. *Operator* in the context (6a) corresponds to UW (6b), and in the context (7a) – to UW (7b).

(6a) *a long distance operator*
(6b) operator(icl>human)
(7a) *addition operator*
(7b) operator(icl>abstract thing).

In order to conform to this requirement, to be consistent and to ensure similar decisions as to what meanings an English word has, it is expedient that all the partners use the same one or two good English dictionaries, preferably available on-line.

In inventing semantic restrictions for UWs, we should adopt a certain procedure which would make it possible for different UNL writers to produce the same or very similar UWs for the same meanings. As a first step towards elaborating such a procedure, it is proposed to proceed along the following lines:

- If a headword is unambiguous in English, and the meaning of this English word expresses the required meaning with sufficient precision, no semantic restrictions are needed. Example: *September*. (NB: the absence of semantic restrictions does not mean that we should not supply KB restrictions in the master dictionary – September{icl>month}).
- If a headword has several meanings in English, and one of them corresponds to the required meaning with sufficient precision, we have to compose a restriction in such a way as to distinguish this meaning from other meanings of the headword.

For example: answer(icl>do) (for cases like *answer questions*) – answer(icl>be) (for cases like *answer expectations*) – answer(icl>thing) (for cases like *know the answer*)[3].

- If no English word exactly corresponds to the meaning of the headword we need, we have to find the closest more general English word available and restrict it accordingly. Example: Russian *zhenit'sja* – marry(agt>male), *vyxodit' zamuzh* – marry(agt>female).

As the last example shows, a restriction can be formulated in terms of any relation which can connect UWs in a UNLR (agt, obj, gol, etc.). Besides them, there are several other relations which can only be used to restrict meanings. These are: icl, pof, equ, ant, com. Relations icl and pof have been envisaged by the Specifications from the very beginning. Relations equ[4], ant and com are proposed for inclusion now. Some comments on these relations are appropriate.

UNL makes extensive use of two traditional types of paradigmatic relations: hyperonymy (class/subclass relation, icl) and meronymy (part/whole relation, pof). Examples: September(icl>month), month(pof>year). However, it is often difficult to find a more general term (hyperonym) that, on the one hand, could distinguish different meanings of the word and, on the other hand, is easy to understand. In this case, it is convenient to recur to a synonym. I think it is worth introducing to UNL the traditional distinction between synonymy and hyperonymy, which is obviously extremely useful for inference, for example.

As in the case of more general terms, restrictions based on synonyms should not be equally applicable to various meanings of the headword. For example, UW wealth(equ>richness) does not meet the above requirement. The words *wealth* and *richness* both have two meanings – 'having many valuable things at one's possession' (*wealth/richness of the nation*) and 'abundance of something' (*butterfly species richness - the wealth of rainforest resources*) – and this restriction alone does not differentiate them. Therefore, some other restrictions should be used, e.g. wealth(icl>well-to-do-ness) – wealth(equ>abundance, obj>thing).

Besides icl, pof and equ relations, we propose to use two more relations. One of them is the traditional antonymy relation, which in some cases may conveniently supplement synonymy. Example:

(8a) *poor quality* ⇒ poor(equ>bad),
(8b) *poor people* ⇒ poor(ant>rich).

Nevertheless, if one takes the task of distinguishing between close lexical meanings of the same word seriously, one will find that the available relations are not sufficient. In many cases, distinctions between the meanings cannot be naturally reduced to rigid categories of hyperonymy, meronymy, synonymy or antonymy. For these cases, we propose to introduce a new relation – com, standing for 'component'. We

[3] This example shows that it is often useful to give examples and/or comments, to make UWs more easily understandable. We will come back to this in 3.3.

[4] The equ relation, originally included in the list of relations, is absent in the latest version of the UNL Specifications (v. 3.2). However, even when it existed, it had a different meaning from what we propose now. It was only used to introduce a definition of an abbreviation: UNL(equ>Universal Networking Language).

will write A(com>B) if B is an (unspecified) important component of the meaning of A. Examples:

(9a) *seniority* ('being older', as in *He is chairman by seniority*) ⇒ seniority(icl>property, com>age);

(9b) *seniority* ('having higher rank by reason of longer service', as in *workers with less than 5 years' seniority*) ⇒ seniority(icl>property, icl>rank);

(10a) *sensational* ('causing intense interest', as in *The effect of the discovery was sensational*) ⇒ sensational(mod<thing, com>interest);

(10b) *sensational* ('very good or impressive', as in *You look sensational in this dress*) ⇒ sensational(mod<thing, icl>good)

(11a) *series* ('several events or actions happening one after another', as in *a series of years*) ⇒ series(icl>set>abstract thing);

(11b) *series* ('a number of connected social events (tournaments, lectures, TV-programmes)', as in *League Championship Series*) ⇒ series(icl>set>abstract thing, com>social).

3.2.2 Argument frame restrictions[5].

UNL as a meaning representation language should have an ability to draw a distinction between the argument and non-argument links of predicates. It is well known that for correct generation, as well as for a wide range of other NLP purposes it is essential to know the argument structure of the predicates and the way each argument is expressed in the sentence. This idea does not seem to require justification, yet it has not been implemented in UNL so far. Since there is no consensus in the UNL community as to what an argument of the predicate is, I will briefly present the problem as I see it.

A is an argument of predicate *L* if *A* is integral to the meaning of *L*. *A* is **semantically obligatory**. This means that *L* cannot be semantically defined, or explained, without *A* being mentioned. *A* is **not always syntactically obligatory**. This means that some arguments can remain unmentioned in a sentence. As an example, let us consider the verb *to borrow.* To define the situation of borrowing, four arguments are necessary.

X borrows Y from Z for W (e.g. *He borrowed a bicycle from his friend for a couple of days*) =

- 'Z owns Y'
- 'X makes Z to give him Y'
- 'X promises Z to give Y back after period W expires'.

All four arguments are semantically obligatory, since borrowing cannot take place without any one of them. None of them is syntactically obligatory. In (12a) *W* is not mentioned. In (12b) no arguments at all are represented.

(12a) *He never borrows money from his friends.*

[5] The problem of arguments in UNL has been raised on several occasions. Our presentation here is a further elaboration of the proposal outlined in [9].

(12b) *Borrowing is tempting but dangerous.*

Still, in both (12a) and (12b), a situation of borrowing is referred to which presupposes the existence of all the four arguments. To feel the difference between arguments and non-arguments better, note that any action has a certain duration, e.g.

(13) *He has been sleeping for three hours.*

Therefore, the duration role is assigned in the Knowledge Base to the topmost UW denoting an action (do{dur>time}) and is inherited by all UWs lying below, including *borrow*. On the other hand, as definition (11) shows, *borrow* has a semantic argument *W* with the role 'duration'. Two functions of the duration with respect to *borrow* (argument and non-argument functions) can be exemplified with sentences (14a) and (14b):

(14a) *John borrowed $10,000 for three years.*
(14b) *John has been borrowing money for three years.*

In (14a) it is a semantic argument and characterizes the terms of the loan. In (14b) it is a free adverbial and characterizes the period of time in which borrowings took place; the terms of each loan are not specified. It is obvious that the difference between arguments and non-arguments is important for semantic processing: (14a) can answer the question on the terms of the loan, while (14b) cannot do so. As a matter of fact, the semantic argument of duration and the adverbial modifier can very well coexist in a sentence: *He has been borrowing money until payday all his life*.

Another example: any object can be used for some purpose. For example, we can use a stone to drive a nail, if no hammer is available. Does it mean that *stone* has a purpose argument? No. A stone has no obligatory conceptual link with the purpose. On the other hand, *a method* has. A method cannot exist without a purpose. Therefore, seemingly similar phrases like (15a) and (15b)

(15a) *a stone for driving nails*
(15b) *a method for calculating taxes*

differ with respect to arguments.

The UNL dictionary does not contain explicit information on the argument structure. Neither semantic nor ontological restrictions are meant for this purpose. To come back to the example above, each object can be used for some purpose, and therefore the purpose relation (pur) is assigned in the KB to UW thing, and is inherited by all UWs lying below. Nevertheless, as we showed above, some of the things do have a purpose argument, while some others do not.

How can arguments be introduced into UNL? First of all, argument structures should be assigned to all those UWs that have arguments. It can be done by means of restrictions, but argument frame restrictions should be clearly differentiated from semantic and ontological ones. One possible way to achieve this is to supply argument frame restrictions with a special symbol (@A, @B, @C). Then, the UW for *borrow* will look as follows:

(16) borrow(icl>do,agt.@A>volitional thing,obj.@B>thing,src.@C>volitional thing,dur>@D>time)

However, in the general case, the marking of the argument frame in the UW is not sufficient. In some cases, the same relation can attach to a UW both an argument, and a free adjunct – cf. (14a)-(14b) above. I will give another example to show that this situation is not unique. Emotional states (*be angry, be afraid, be surprised,* etc.) have an argument denoting the cause of the state. In sentence (17)

(17) *She is afraid to go out alone at night*

going out alone at night is the cause of her being in the state of fear. Therefore, relation rsn (='cause, reason') between *afraid* and *go out alone at night* is appropriate. On the other hand, *afraid* can have a non-argument cause, as in (18):

(18) *She is afraid (to go out alone at night), because this area is not very safe.*

Even if *afraid* is assigned a cause as one of the arguments, we should know whether or not a rsn-link in the UNLR denotes this argument. This means that in order to generate correct text, it is not sufficient to know the semantic role of word A with respect to B. One also needs to know whether or not A is an argument of B.

A possible solution would be to mark the argument relation in the UNLR with a special label. Then, a relevant fragment of sentence (18) will be represented as (19)

(19) rsn.@A(afraid(rsn.@A>uw), go_out)
rsn(afraid(rsn.@A>uw), safe)

Obviously, it only makes sense, when the relation in question can in principle fulfill both functions. If a relation is unambiguously argumentative (as agt or obj), this label is superfluous.

This example shows also that the difference between arguments and non-arguments is essential for correct deconversion, since they can be expressed differently. In English, the rsn-argument of *afraid* cannot be expressed by preposition *because of*, which is typical for this role:

(20a) *She is afraid of darkness*.
(20b) **She is afraid because of darkness.*

3.3 Samples of UW dictionary entries.

As of now (end of 2004), UNL partners have collected large UNL dictionaries, that is sets of UWs linked with words of their languages. The value of these resources is impaired by several facts:

1. UWs do not sufficiently differentiate between different meanings of the headword.
2. There is no systematic information on the arguments.
3. Some restrictions are difficult to understand.
4. Dictionaries of different groups are not harmonized.

Ways to solve the first and the second problems have already been discussed above. The third shortcoming can be overcome if the dictionary entry is supplied with examples and/or comments that illustrate and clarify UWs in non-obvious cases. The fourth problem requires that all the UW dictionaries be put together and made a uni-

fied UNL lexical resource[6]. The table below shows what this resource could look like. It presents a group of words beginning with the letter *L*. For the reader's convenience, examples and comments to UWs are given in a separate column. Translation equivalents are only given for Russian and Spanish, but obviously other working languages should also be added.

Symbols outside the Specifications:
@ex – example
@com – comment
uw – any UW
* – a string of characters
icl>adj – restriction for all types of adjectival UWs (see 4.2 below)
asp – aspect relation (see 5.3 below)

Table 1. Samples of multilingual UNL dictionary entries.

UW	Examples&Comments	Russian	Spanish
label(icl>conctrete thing)	@ex: *a luggage label* @ex: *Label the diagram* (obj) *as shown. The file* (obj) *was labelled "Top secret"* (cob) @com: cob>*: * is used because not only UWs are possible here but any string of symbols	ярлык	etiqueta
label(icl>write>do, agt.@A>volitional thing, obj.@B>thing, cob.@C>*)		поме-чать	etiquetar
label(icl>name>do, agt.@A>volitional thing, obj.@B>thing, cob.@C>*)	@ex: *the newspapers* (agt) *labelled him* (obj) *a troublemaker* (cob)	назы-вать	nombrar
labora-tory(icl>institution>orga nization, pur.@A>uw)	@ex: *The National* (mod) *Renewable Energy* (pur) *Laboratory; laboratory for renewable energy research and development* (pur)	лабора-тория	laborato-rio
labor_day(icl>holiday> date)		День труда	Día del trabajo
labour_intensive(icl>adj)	@ex: *labour intensive*	трудо-емкий	que nece-sita mu-cho traba-jo
laborious(icl>difficult> adj)	@ex: *laborious task*	трудный	laborioso

[6] The idea to construct a multilingual dictionary with UWs serving as interlingual index has been put forward within the PAPILLON project in [10].

laborious(icl>slow>adj)	@ex: *laborious progress* @com: done slowly and with difficulty	медленный	penoso
labour_union(icl> institution>organization)		профсоюз	sindicato
labour(icl>work>action, agt.@A>person)	@ex: *building involves a lot of manual labour, his* (UW=he) (agt) *labour*	Труд	trabajo
labour(icl>person>living thing)	@ex: *skilled labour, labour shortage*	рабочая сила	mano de obra
labour(icl>event> abstract thing, agt>living thing)	@com: a process in which a baby is born	роды	parto
labour(icl>do, agt.@A>person>living thing, obj>thing)	@ex: *to labour at a task* (obj), *over the report* (obj)	трудиться	afanarse
labour(icl>party> organization)	@com: the British Labour party	лейбористская партия	Partido Laborista
labour(icl>adj)	@com: connected with the British Labour party	лейбористский	laborista
lack(icl>abstract thing, aoj.@A>thing, obj.@B>thing)	@ex: *lack of food* (obj); *their* (aoj) *lack of patience* (obj)	нехватка	falta
lack(icl>be, aoj.@A>thing, obj.@B>thing)	@ex: *he* (aoj) *lacks courage* (obj)	недостать	faltar, carecer
lacking(icl>adj, obj.@A>thing)	@ex: *the crew is lacking in beef* (obj)	лишенный	carente (de)
lag(icl>period>time, obj.@A>time)	@ex: *a time* (mod) *lag of one month* (obj)	отставание	retraso
lag(icl>occur, equ>lag behind, obj.@A>thing, asp.@B>thing)	@ex: *Britain* (obj) *was lagging in the space race* (asp)	отставать	quedarse atras
lag behind(icl>occur, equ>lag, obj.@A>thing, asp.@B>thing)	@ex: *they* (obj) *worked badly and lagged behind; lag behind in development* (asp)	отставать	quedarse atras, retrasar-se
land(icl>area,ant>sea, ant>air)	@ex: *to travel by land* (via)	суша	tierra
land(icl>ground>thing)	@ex: *fertile land* @com: mostly when used for farming or building on	земля	tierra
land(icl>country>region)	@ex: *native land, visit distant lands*	страна	país

land(icl>property> abstract thing)	@ex: *his lands extend for several miles*	земля	terreno
land(icl>do, agt.@A>thing, plc.@B>thing,src>sky)	@ex: *the plane <we>* (agt) *landed at the Geneva airport* (plc)	приземляться	aterrizar
land(icl>do, agt.@A>volitional thing, obj.@B>functional thing, plc.@C>thing,src>sky)	@ex: *the crew* (agt) *finally landed the plane* (obj) *on the soft part of the runway* (plc)	сажать	aterrizar
land(icl>do, agt.@A>volitional thing, plc.@B>thing,src>water)	@ex: *land on a lonely island* (plc)	высаживаться	llegar a tierra
land(icl>do, agt.@A>volitional thing, obj.@B>living thing, plc.@B>thing,src>water)	@ex: *land somebody* (obj) *on a lonely island* (plc)	высаживать	poner en tierra
land(icl>do, agt.@A>volitional thing, obj.@B>concrete thing, plc.@C>thing)	@ex: *to land containers* (obj) *on the shore* (plc)	выгружать	poner en tierra
last(icl>recent>adj)	@ex: *last night, last edition, last harvest*	последний	pasado, ultimo
last(icl>adj,ant>first)	@ex: *last page, last bus*	последний	ultimo
last(icl>occur, obj.@A>abstract thing, dur.@B>period>time)	@ex: *the hot weather* (obj) *lasted for the whole month* (dur)	длиться	durar
lay(icl>put>do, agt.@A>living thing, obj.@>concrete thing, plt.@C>thing)	@ex: *lay the dress* (obj) *on the bed* (plt)	класть	poner
lay(icl>set>do, agt.@A>person, obj.@B>table)	@ex: *lay the table*	накрывать	poner
lay(icl>fix>do, agt.@A>person, obj.@B>thing, plt.@C>thing)	@ex: *lay the carpet* (obj) *on the floor* (plt), *lay bricks, pipelines* (obj)	укладывать	poner
lay(icl>produce>do, agt.@A>bird, obj.@B>egg)	@ex: *lay eggs*	откладывать	poner

Revising the UNL lexical resources along the lines suggested above is, in my opinion, the most important task facing the UNL community at the moment. It can only be solved if all the partners join their efforts.

4 Semantic categories of UW

Semantic classification of UWs is embodied in the Knowledge Base. This is a very large topic, which I cannot discuss at full scale. Here I will only touch upon the upper levels of this classification. All UWs are divided into four major classes: verbal, nominal, adjectival and adverbial concepts. Of these classes, I will only deal with two – the verbal and the adjectival concepts.

4.1 Verbal concepts

In linguistics, there are various classifications of predicates based on their fundamental semantic properties. The most important classes of predicates are:

(a) **actions**: they have an active initiator – an agent (normally, a human) that performs the action as a step to achieving some goal. Most of the actions have a natural limit - a point in its development at which the goal has been achieved and after which the action cannot continue. Examples: *kill, write, eat, solve.*

(b) **activities**: they denote a set of actions, often heterogeneous, that have a common goal. Examples: *work, trade, cooperate*.

(c) **events**: they have no agent and denote a situation in which something happens to an object. Examples: *the bridge broke, an accident happened, the stone fell.*

(d) **processes**: they have no agent and denote a situation that occupies a certain time span in which an object undergoes a change. Examples: *the tree grows, the temperature rises.*

(e) **states**: they differ from the processes in that they are homogeneous (do not denote a change). They characterize a thing during a certain period of its existence. Examples: *see, hear, ache, know, want, wait, hope, proud.*

(f) **properties**: they differ from the states in that they are atemporal, i.e. they normally characterize things during the whole period of their existence. They are often expressed by adjectives. Examples: *blind, red, clever.*

(g) **relations**: they differ from the properties in that they do not characterize a thing but a relation between two or more things. They are often expressed by nouns. Examples: *love, hate, equal, friend, father.*

In UNL, not all of these semantic types are distinguished – only three. All verbal concepts group into three classes designated by restrictions (icl>do), (icl>occur) and (icl>be).

Class (icl>do) contains actions and activities. They are initiated by some active force which can be either a voluntary human (or autonomous mechanism, as e.g. computer) or some inanimate factor: *He solved the problem. The storm broke the tree. The silence frightened the child.*

Class (icl>occur) consists of events and processes, which are not regarded as initiated by an active force.

Class (icl>be) is composed of states, properties and relations.

Some examples. *Include* is an action in (21a) but not in (21b):

(21a) *I included* (icl>do) *his name in the list.*

(21b) *The list includes* (icl>be) *his name.*

Open is an action in (22a), but not in (22b)

(22a) *I opened* (icl>do) *the door.*

(22b) *The window opened* (icl>occur),

because in the latter case no initiator is necessarily implied. Even in the sentence

(23a) *The forum opened*

we are dealing with an (icl>occur) verb, because it does not mean exactly the same as

(23b) *The forum was opened* (icl>do).

(23b) definitely says that somebody opened the forum, while (23a) doesn't say anything to this effect and in this sense is similar to (22b). If a UNL writer wishes to ignore this difference, he may choose any option.

It is natural that the semantic type of the predicates should agree with semantic relations that link them to their main argument[7]. Obviously enough, the main argument of actions is an agent (agt), events and processes require obj-relation, while states, properties and relations attach their main argument by the aoj-relation. For this reason, predicates like *know* and *regret* which denote a state and not an action cannot be heads of the agt-relation.

4.2 Adjectival concepts.

According to the UW Specification, all adjectival concepts are divided into two classes. The first class is characterized by restriction (aoj>thing) and the second by restriction (mod<thing). The difference between these classes is explained in the following way: "(aoj>thing) is for expressing a predicative concept, whereas (mod>thing) is for expressing a restrictive concept" [11]. This formulation introduces an opposition "predicative" vs. "restrictive" which is based on heterogeneous criteria. This is logically unacceptable. Let us consider the facts with some detail.

We are dealing here with two different properties of adjectives:

(a) a syntactic property: it is the question of whether the adjective is used predicatively (*Greeks are wise)* or attributively (*the wise Greeks*);

(b) a semantic property: which shows what the adjective means when used attributively: restriction or qualification.

We should clearly distinguish between the syntactic construction in which a modifier is preferably used (attributive vs. predicative) and the meaning it conveys (restrictive vs. qualificative). I will begin with the meaning.

[7] By way of simplification, one can say that the main argument is the one that normally corresponds to the syntactic subject.

This semantic difference was discussed at least as far back as 1933 by O. Jespersen [12]. This is what we find in a modern English grammar: from the semantic point of view, «the modification can be restrictive or non-restrictive [= qualificative – IB]. That is, the head [the modified noun - IB] can be viewed as a member of a class, which can be identified only through the modification that has been supplied (restrictive). Or, the head can be viewed as unique or as a member of a class that has been independently identified (for example, in a preceding sentence); any modification given to such a head is additional information which is not essential for identifying the head, and we call it non-restrictive». [13, 13.3]. For example, the adjective *wise* in the sentence (24) can be understood both restrictively and non-restrictively.

(24) *Wise Greeks diluted wine with water*

(24a) restrictive interpretation: 'Those Greeks who were wise diluted wine with water. Silly ones didn't'.

(24b) non-restrictive (qualificative) interpretation: 'Greeks were wise. They diluted wine with water'.

This opposition is only relevant for the attributive position (*the wise Greeks*). The predicative one (*The Greeks are wise*) only adds a characteristic without restricting the extension of the noun.

Which of these two properties is captured by means of restrictions (aoj>thing) and (mod<thing)? Preferential ability to be used in the attributive vs. predicative construction or preferential type of interpretation in the attributive construction? Even though these properties are correlated, they are quite different.

If UWs are to reflect the first opposition, it is not clear why we should wish to incorporate into UWs a syntactic difference between English words. Why should we treat this difference at the same level as the fundamental semantic difference between actions and states? This position is evidently untenable.

If we wish to capture the second opposition (which is much more reasonable), we should first of all take into account the distribution of adjectives between these classes. Some adjectives (such as *many*) can only be restrictive or are restrictive in the majority of cases:

(25) *Many dogs have curly hair.*

Some other adjectives (such as *damned* or *dear* – in the sense presented in (26b)) can only be non-restrictive:

(26a) *Get those damned dogs out of the room!*

(26b) *Dear colleagues!*

However, the overwhelming MAJORITY of adjectives can easily have BOTH interpretations. If we choose to convey this opposition by means of restrictions, we will have to split all these adjectives in two concepts, which is obviously rather strange. But this is not the most important shortcoming of this description. After all, it is technically possible to postulate two concepts for every adjective. The crucial fact is that the opposition restrictive/non-restrictive is not only relevant for adjectives, but also for other types of modifiers, such as relative clauses or prepositional phrases:

(27a) *The students(,) who are sitting in the corner(,) are waiting for the professor.*

(27b) *The students in the corner are waiting for the professor*.

The phrase *(who are sitting) in the corner* can be either restrictive (= 'those of the students who are sitting in the corner are waiting for the professor; others are not') or non-restrictive ('the students are waiting for the professor; they are sitting in the corner'). If we wish to mark this opposition for the adjectives, there is no reason not to do so for other types of modifiers.

Moreover, for such phrases it is even more important than for the adjectives, because in some languages restrictive and non-restrictive relative clauses have different punctuation and therefore should be differently treated by the deconverters. For example, in English and in Spanish restrictive relative clauses are not marked with commas, while non-restrictive necessarily are. Cf. synonymous English and Spanish sentences (28a) - (28b) and (29a) – (29b).

Restrictive:

(28a) *The old people who came a long way were tired.*

(28b) *Los viejos que habían venido de muy lejos estaban cansados.*

Non-restrictive:

(29a) *The old people, who came a long way, were tired.*

(29b) *Los viejos, que habían venido de muy lejos, estaban cansados.*

Thus relative clauses and other types of modifiers share with the adjectives the capacity to have restrictive and non-restrictive interpretations and should be treated in the same way. However in relative clauses and prepositional modifiers there is no UW to which a restriction can be assigned.

Therefore, I propose to renounce from the division of adjectives into (aoj>thing) and (mod<thing). In order to account for the opposition between restrictive and non-restrictive modifiers, two attributes are introduced (@restr, @non-restr) which can optionally be added to any modifier (an adjective, a prepositional phrase, a relative clause), if the UNL writer wishes to mark the restrictive or non-restrictive interpretation. As a general adjectival restriction, I would propose to introduce the one that is neutral to the restrictive/non-restrictive distinction, e.g. (icl>adj).

There are some more arguments to support the attribute solution:

(a) Attributes reflect the point of view of the speaker in the current situation and not the permanent property of the word. It is just the case with restrictive vs. non-restrictive interpretation of modifiers. It is the property of the given sentence and not the inherent property of the modifier. True, some of the adjectives cannot be used in one of these interpretations and for them this is a permanent property (see (25) – (26a,b)). But this does not in the least undermine the statement made above. Simply, these adjectives cannot be assigned one of the attributes @restr or @non-restr. It is the same with the nouns that have no plural form: they simply do not accept attribute @pl.

(b) The attribute is optional and need not be assigned if the UNL writer does not wish to specify his point of view. It is in fact not always easy to decide, whether or not a modifier is used restrictively. If we have two differently restricted UWs for an adjective, the UNL writer will always have to make a choice, very often irrelevant for the meaning he wishes to convey.

5 Relations

Currently, UNL disposes of 41 relations listed in the UNL Specifications. This set of relations has been tested in various encoding experiments and showed relative stability. However, the analysis of texts reveals three kinds of problems connected to UNL relations. First, some relations seem to be weakly differentiated and therefore difficult to use consistently. Second, the opposition between some relations seems to be based more on the semantic class of UWs than on the semantic relation that holds between them. Such distinctions should be avoided in a language designed for meaning representation. Third, some relations seem to be lacking. This topic deserves a special investigation that will be carried out later. Here I will only give several examples and, on a pilot basis, formulate some proposals.

5.1 Weakly differentiated relations

Example: gol (final state) – plt (final place); src (initial state) – plf (initial place).

According to the UNL manual (sec. 4.10) , examples like

(30) *John went to Brussels*

can be described both with gol and plt. The difference between the two is that gol characterizes Brussels as the final state of John, while plt – as the final place of the whole event "John went to Brussels". To put it mildly, it is difficult to understand what could be the final place of a movement as opposed to the final place of the moving object. The same applies to relations src and plf.

5.2 Distinction determined by the class of UWs

Example 1: mod (modification) – man (manner).

Both relations are very general and cover a wide range of situations which are not described by any specific relation, such as tim (time), plc (place), ins (instrument), etc. In practice, the difference between them boils down to the semantic class of the starting point of the relation: mod applies to things while man applies to situations.

(31a) *answered politely* (man)
(31b) *a polite answer* (mod)
(32a) *meet often* (man)
(32b) *frequent meetings* (mod)
(33a) *wrote in Japanese* (man)
(33b) *a letter in Japanese* (mod)

In my opinion, the difference between *to answer* and *an answer*, *to meet* and *a meeting*, or between *to write* and *a letter* has no bearing on the semantic relation in pairs (31a-b) – (33a-b). Relations man and mod can be safely merged into one relation. Any semantic difference between them, if it existed, is derivable from the context.

Example 2: plt (final place) – to (destination); plf (initial place) – frm (origin).

It is difficult to find any singnificant difference between relations in these pairs. They seem to differ only as to the type of the starting point of the relation: in case of plt it should be an event (action, process or state) while in case of to it should be a thing:

(34a) *The train is bound for Edinburgh* (plt).
(34b) *the train for Edinburgh* (to).

This difference does not seem to be fundamental enough to constitute different relations. The same is true for relations plf and frm.

Example 3: mod (modification) – agt (agent) / obj (object) / gol (final state) / …

According to the UNL Specifications, nominal UWs cannot be starting points for many argument relations, such as agt, obj, gol and some others. All these arguments are connected to nominal UWs by means of the mod relation. This approach is motivated by syntactic factors more than by semantic considerations. Phrases like (35a)

(35a) *arrival of the minister*

are described by means of the mod relation, obviously under the influence of the surface *of*-construction (cf. *decision of great importance*). Due to this, UNL fails to express the indentity of semantic relations in phrases like (35a) and (35b):

(35b) *The minister arrived*[8].

Besides that, UNL is unable to disambiguate phrases like (36)

(36) *invitation of the minister*,

which has at least two interpretations:

(36a) the minister invited (somebody) (agt)
(36b) (somebody) invited the minister (obj)

Obviously, a meaning representation language should have the means to establish identity of relations in (35a) and (35b), as well as to detect ambiguity in (36). This will be ensured, if argument relations like agt, obj, gol, etc. are allowed to go out of nominal UWs.

5.3 Missing relations

As is known, there exist no well-established criteria for deciding how many relations it is appropriate to have and what their semantic load should be. There is often a liberty of choice between introducing a relation for some specific meaning and expressing this meaning by other means. For example, how could we represent the difference between *after* and *before* in sentences (37) and (38), given that UNL has a relation for time (tim)?

(37) *He left after dinner.*
(38) *He left before dinner.*

[8] Moreover, the UNL Knowledge Base does not establish any link between semantic derivatives of the type *to arrive – arrival*, but this problem is beyond the scope of the present paper.

We have at least two options:

- use the existing relation tim and convey the difference between (37) and (38) by means of UWs after(icl>time) and before(icl>time):
  (37a) tim(leave(icl>do), after(icl>time))
  obj(after(icl>time), dinner(icl>event))
  (38a) tim(leave(icl>do), before(icl>time))
  obj(after(icl>time), dinner(icl>event))
- introduce special relations tim-after and tim-before and do without UWs after(icl>time) and before(icl>time):
  (37b) tim-after(leave(icl>do), dinner(icl>event))
  (38b) tim-before(leave(icl>do), dinner(icl>event))

These options are equivalent, although the first one is obviously preferable. It is better to keep the number of relations at the reasonable minimum, while the number of lexical units may be unlimited.

Nevertheless, there is a class of situations in which it might be more adequate to somewhat increase the number of relations. This is the area of relations between predicates and their arguments (cf. 3.2.2 above). The number of roles adopted by different authors for representing argument relations ranges from 14 in [14] to 57 in [15]. The list of relations in [15] proposed by Jury Apresjan is oriented towards the needs of deep semantic annotation of texts. For the UNL purposes it seems to be too detailed. However, some of the relations from this list deserve to be adopted in UNL. For example:

- cont (content): *he ordered us to attack; he proposed that; I think that…*
- top (topic) : *He knows nothing* (cont) *about women* (top); *review of the paper; the paper on UNL*
- rec (recipient): *He sent Mary flowers; He told me* (rec) *to come* (cont). *He informed us* (rec) *of his arrival* (cont)
- mot (motivation): *punish for disobedience, praise for achievements*
- asp (aspect): *differ in quality, distinguished for strength.*
  This is a topic for further discussion.

6 Hypernodes: their links and attributes

A UNLR is a hypergraph, i.e. a graph whose nodes are either simple or compound UWs (hypernodes, scopes). A compound UW is a subgraph consisting of simple or compound UWs linked with UNL relations. The major contribution of hypernodes in UNLR is their ability to bear relations and attributes of their own.

Each graph and subgraph (compound UW) contains a special node called the entry of the graph. Informally speaking, it represents the “main” element of the graph which normally corresponds to a syntactic top node of the corresponding part of the sentence. For example, the entry node of the phrase *music in Polynesia* is *music*, because it is this word that links the whole phrase to other words of the sentence. A phrase

should be made a hypernode if its link to some element of the outer context is not semantically equivalent to the link of its entry node.

Situations where hypernodes are really necessary are rather rare. In the majority of sentences in which they are currently used, hypernodes are superfluous in the sense that their entry nodes effectively inherit their relevant properties. In other words, the replacement of a hypernode by a combination of simple nodes of which it consists does not result in any shift of meaning. Nevertheless, hypernodes are a useful and important formal device. In section 1 we saw one of the examples of its possible use for the representation of relatively fixed multi-word expressions. Below, I will show some situations when hypernodes are necessary as holders of relations and attributes.

6.1 Links of hypernodes

In the sentence

(39) *Music in Polynesia is extension of poetry*

there is no need to introduce a hypernode, because linking the phrase *music in Polynesia* to the verb is semantically equivalent to linking the noun *music* to this verb: 'music in Polynesia is extension of poetry' = 'music is extension of poetry; this music is in Polynesia'. The same is true for the sentence

(40) *Music and dance are extensions of poetry.*

It can only be interpreted in the sense that both music and dance are extensions of poetry. Therefore, there is no need to merge *music and dance* in a hypernode. The situation is different in the sentence

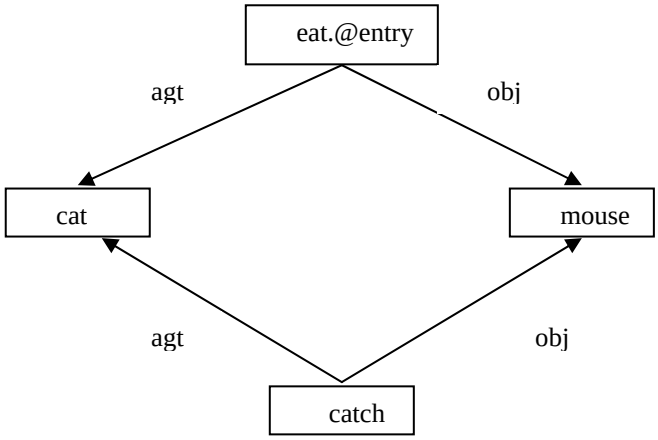
(41) *Music and dance in Polynesia are extensions of poetry.*

This sentence is ambiguous between two interpretations: 'music in Polynesia and dance in Polynesia' and 'music (in general) and dance in Polynesia'. Under the first of these interpretations, the phrase *in Polynesia* is linked to the whole phrase *music and dance*, while under the second one it is only linked to *dance.* Thus, to assure proper understanding of the sentence, one has to introduce a hypernode *music and dance* for the first interpretation.

(41a) aoj(extension(icl>abstract thing.@entry, :01)
obj(extension(icl>abstract thing.@entry, poetry)
and:01(dance(icl>activity).@entry, music(icl>abstract thing))
plc(:01, Polynesia(icl>place))

As seen in (41a), UWs which constitute a hypernode do not have any links with UWs outside this hypernode. All the links which external UWs might have with UWs inside the hypernode are established with the hypernode itself. In (41a) these are links aoj and plc (*music and dance are extensions* and *music and dance in Polynesia*). An important question is whether it is possible that UWs inside the hypernode have direct links with UWs outside the hypernode. This question is raised by E. Blanc in [16]. Naturally, it is preferable to exclude this possibility. However, E. Blanc demonstrates a case where it is desirable, if not inevitable, to allow such a link. His example:

(42)



This UNLR can be verbalized in two different ways:

(42a) *The cat eats the mouse it caught.*
(42b) *The cat which caught the mouse eats it.*

Although these sentences describe the same situation, in a certain sense they are not equivalent, and this difference should not be lost in UNLR. E. Blanc proposes to express this difference by means of hypernodes. In (42a), a hypernode will look like obj:01(catch, mouse.@entry), while in (42b) it will be agt:01(catch, cat.@entry). Note that the entry nodes of these hypernodes are nouns and not verbs. The price paid for distinguishing (42a) and (42b) is admitting links going from the inside of the hypernode to the outside: agt(catch, cat) in the first case, and obj(catch, mouse) in the second case.

This proposal certainly solves the problem, but the price seems to be somewhat too high. The impermeability of hypernodes with respect to links from and to the outside is a property that is worth preserving as long as possible. To save this property, E. Blanc proposes to split one of the nodes in two identical and coreferential nodes. One of them would stay outside the hypernode, while the other would be included into it. This solution, however, implies a serious modification of the UNL specifications and the EnCo/DeCo software.

I would propose another solution which permits to distinguish interpretations (42a) and (42b) without violating the scope impermeability requirement and within the current specifications. What is the difference between (43) and (44)?

(43) *A girl is holding a peach*
(44) *a girl holding a peach*

These phrases describe the same situation but the meaning is organized in different ways. In (43) it is presented as an assertion, and in (44) as an object. How can we represent this difference in UNL? The first method is to declare *a girl* to be the entry node of (44). This inevitably leads to postulating a hypernode, since (44) may make part of a sentence where the predicate is also marked as entry (*I admired.@entry a*

girl holding a peach). This will be exactly the type of a hypernode we saw in (42a) - (42b).

The second method of representing the difference between (43) and (44) is based on the fact that participles in the attributive position have dual nature. Participle *holding* in (44) conveys two different messages. The first message is purely semantic: a girl is the agent and a peach is the object of holding. The second message concerns the communicative organization of the meaning: the fact that a girl is holding a peach is presented as something that characterizes the girl, something that serves as a qualifier of the girl. In other words, *a girl* is the agent of *hold*, and at the same time *hold* is a modifier of *girl*. **This fact can be directly represented in UNL without recurring to hypernodes:**

(44a) agt(hold, girl)
mod(girl, hold)

(44a) makes explicit the dual nature of the attributive participle and thus effectively distinguishes (43) and (44). This approach can as well be applied to sentences (42a,b). In (42a), the link mod(mouse, catch) will be added, and in (42b) – mod(cat, catch).

6.2 Attributes of hypernodes

Example (41) shows that hypernodes may have links of their own which are not reducible to the links of their inner nodes. Now I will illustrate the situation where a hypernode has an attribute that cannot be assigned to any of its inner nodes.

The meanings that express the speaker's attitude towards the situation, such as 'not', 'can', 'must', etc. are expressed in UNL by means of attributes ascribed to a UW. For example, the meaning 'they do not sleep' is represented in UNL as aoj(sleep.@not, they). Consider sentences (45) and (46) which look similar but are opposed semantically:

(45) *They (do not sleep) because of the noise.*
(46) *They do not (quarrel because of money).*

(45) means 'noise is the cause of their not-sleeping', while (46) means 'money does not make them quarrel'. These readings differ in the scope of the negation, which we show by means of brackets. To express this difference in UNL, it is necessary to be able to attach the negation attribute to a hypernode:

(45a) rsn(sleep.@not.@entry, noise)
aoj(sleep.@not.@entry, they)
(46a) rsn:01(quarrel.@entry, money)
agt:01(quarrel.@entry, they)
:01.@not.@entry

Conclusion

I hope that interpretations and proposals presented here will be discussed by the participants of the UNL project, both at the UNL Workshop (Mexico, 2005) and at the

forum for discussions. After that, two tasks seem to be the most important: revision of the UNL dictionaries according to the solutions taken during discussions and compilation of a corpus of UNL documents which incorporate all enconversion conventions which we will arrive at[9].

References

1. Boguslavsky, I. Guidelines for writing UNL expressions. FB2004: a showcase of UNL deployment. Technical document. Spanish Language Centre. Facultad de informática. UPM. Spain (2001)
2. Boguslavsky, I. Some remarks on the UNL encoding conventions. Proceedings of the First International UNL Open Conference "Building global knowledge". SuZhou, China (2001).
3. Uchida H., Zhu M., Della Senta T. A Gift for Millenium. (1999) http://www.undl.org
4. Boguslavsky, I., Frid, N., Iomdin, L., Kreidlin L., Sagalova I., Sizov, V. "Creating a Universal Networking Lnguage Module within an Advanced NLP System". Proceedings of the 18th International Conference on Computational Linguistics, 2000, p. 83-89.
5. Boguslavsky I., Cardeñosa J., Gallardo C., Iraola L. The UNL Initiative: an Overview. CICLING 2005 (in print).
6. Uchida, H. The UNL Specifications, v.3.2. (2003) http://www.undl.org
7. De la Calle Velasco G., Gallardo C., Cardeñosa J. Manual de instalación y uso del EditorUNL. Technical document. Spanish Language Centre. Facultad de informática. UPM. Spain (2003)
8. Boguslavsky, I., Iomdin, L, Sizov, V. Interactive enconversion by means of the ETAP-3 system. Proceedings of the International Conference on the Convergence of Knowledge, Culture, Language and Information Technologies, Convergences'03. Alexandria (2003)
9. Boguslavsky I. Some Lexical Issues of UNL. Proceedings of the First International Workshop on UNL, other interlinguas and their applications, Las Palmas, 2002, 19-22
10. Sérasset G., Mangeot-Lerebours M. Papillon Lexical Database Project: Monolingual Dictionaries & Interlingual Links. Proc. NLPRS'2001 The 6th Natural Language Processing Pacific Rim Symposium, Hitotsubashi Memorial Hall, National Center of Sciences, Tokyo, Japan, 27-30 november 2001, vol 1/1, (2001), pp. 119-125.
11. Uchida, H. The UW Specifications, v.2.0 (2002) http://www.undl.org
12. Jespersen O. Essentials of English Grammar. London: Allen & Unwin, 1933
13. Quirk R., Greenbaum S. A University Grammar of English. Longman, 1973, xi, 484 p.
14. Fry J., Bond F. Semantic annotation of a Japanese speech corpus, Proceedings of the 18th International Conference on Computational Linguistics (COLING 2000), 2000.
15. Apresjan Ju., Iomdin L., Sannikov A., Sizov V. Semantic Tagging in a Deeply Annotated Corpus of Russian. (in Russian, in print)
16. Blanc E., About and around the French enconverter and the French deconverter. In this volume.

[9] This paper is based on the research supported by the RFBR grant N 04-06-80148 and the Ramon y Cajal program of the Spanish Ministry of Science and Technology, whose assistance is gratefully acknowledged.