Upon Comparison Of Some Online UNL Modules

Yenok Grishkyan
Yerevan State University, Armenia
eng83@rambler.ru

ABSTRACT
The purpose of the present article is to try to present some comparative study of the UNL international modules, their development with in a two years of range from the point of view of the UNL-Armenian Module.

Keywords
Computational linguistics, applied linguistics, informatics, semantics, languages, translation.

The mechanical translation has always been an important human step towards better understanding between various cultures of different nations. Many attempts have been made to facilitate conversation and make the translation more rapid. One of this translating systems is the UNL (Universal Networking Language) system.

During the past two years we have worked to develop the system's Armenian Module, many a times comparing it with other languages online working Modules and have made observations, a survey, which we'd like to present. . But first of all, let's mention that many of the UNL Language Centres of different languages never functioned online within the past two years, but a couple of them.

The best working Module online is the UNL-Russian Module Server, with double translation possibility: The UNL-Russian Module has an easy interface and a big variety of options for translation. The server system translates from either a UNL or a natural language sentence. In case of UNL expression translation we can have an output text either in English or Russian. In case of direct translation from a natural language system (in our case Russian or English) we receive a clear UNL expression. But this sentence may be different form the original English text provided by the UNDL which has been converted into the UNL expression earlier. Let's observe the following changes.

	UNL original sentence	**UNL-Russian Module**
Original Text	Water mobility links upstream activities in a catchment with downstream problems and opportunities.	
UNL expression	{unl} agt(link(agt>thing,gol>thing,obj>thing):0H.@entry, mobility(icl>reference):08.@topic) gol(link(agt>thing,gol>thing,obj>thing):0H.@entry, with(obj>thing):1M) scn(link(agt>thing,gol>thing,obj>thing):0H.@entry, catchment(icl>action):1C.@indef) obj(link(agt>thing,gol>thing,obj>thing):0H.@entry, activity(icl>action):0W.@pl) aoj(upstream(aoj>thing):0N, activity(icl>action):0W.@pl) obj(with(obj>thing):1M, :01) and:01(opportunity(icl>state):2F.@entry.@pl, problem(icl>abstract thing):22.@pl) aoj:01(downstream(aoj>thing):1R, problem(icl>abstract thing):22.@pl) mod(mobility(icl>reference):08.@topic, water(icl>liquid):02) {/unl}	{unl} mod(mobility(icl>quality>thing,ant>immobility), water(icl>binary_compound>thing)) agt(link(icl>relate>do,cob>thing,agt>thing,obj>thing).@entry.@present,mobility(icl>quality>thing,ant>immobility)) mod(activity(icl>act>thing,ant>inactivity).@pl,upstream(icl>adj,ant>downstream)) obj(link(icl>relate>do,cob>thing,agt>thing,obj>thing).@entry.@present,activity(icl>act>thing,ant>inactivity).@pl) obj(activity(icl>act>thing,ant>inactivity).@pl,catchment(icl>structure>thing).@indef) mod(problem(icl>task>abstract_thing,obj>thing).@pl,downstream) and(opportunity(icl>possibility>thing).@pl,problem(icl>task>abstract_thing,obj>thing).@pl) met(catchment(icl>structure>thing).@indef,opportunity(icl>possibility>thing).@pl) {/unl}
Translation	իուր շարժողունակություն գործուներություններ հոսանքին հակառակ կապ)>gol դրենավորման մեջ:	No translation is generated possibly due to server problems

Let's have a closer look on the UNL expressions in the following table. Here we can observe a clear difference within them.

	UNL original sentence	**UNL-Russian Module**
Common	agt(link(agt>thing,gol>thing,obj>thing):0H.	agt(link(icl>relate>do,cob>thing,agt>thing,obj>thing).@entr

relations	@entry, mobility(icl>reference):08.@topic) obj(link(agt>thing,gol>thing,obj>thing):0H.@entry, activity(icl>action):0W.@pl) and:01(opportunity(icl>state):2F.@entry.@pl, problem(icl>abstract thing):22.@pl) mod(mobility(icl>reference):08.@topic, water(icl>liquid):02)	y.@present,mobility(icl>quality>thing,ant>immobility)) obj(link(icl>relate>do,cob>thing,agt>thing,obj>thing).@entry.@present,activity(icl>act>thing,ant>inactivity).@pl) and(opportunity(icl>possibility>thing).@pl,problem(icl>task>abstract_thing,obj>thing).@pl) mod(mobility(icl>quality>thing,ant>immobility),water(icl>binary_compound>thing))
Not coincided relations	gol(link(agt>thing,gol>thing,obj>thing):0H.@entry, with(obj>thing):1M) scn(link(agt>thing,gol>thing,obj>thing):0H.@entry, catchment(icl>action):1C.@indef) aoj(upstream(aoj>thing):0N, activity(icl>action):0W.@pl) aoj:01(downstream(aoj>thing):1R, problem(icl>abstract thing):22.@pl) obj(with(obj>thing):1M, :01)	obj(activity(icl>act>thing,ant>inactivity).@pl,catchment(icl>structure>thing).@indef) met(catchment(icl>structure>thing).@indef,opportunity(icl>possibility>thing).@pl) mod(activity(icl>act>thing,ant>inactivity).@pl,upstream(icl>adj,ant>downstream)) mod(problem(icl>task>abstract_thing,obj>thing).@pl,downstream)

Clearly we see that some relations are common one with the other, yet others are different and even the UNL-Russian Module has a shorter UNL expression and in the result of which less relations. So, what is a **aoj** /*aoj(upstream(aoj>thing):0N,activity(icl>action):0W.@pl)*/ relation in the original UNL sentence, it is a **mod** relation/*mod(activity(icl>act>thing,ant>inactivity)@pl,upstream(icl>adj,ant> downstream))*/ in the other with a somewhat shifted action. However, this is normal for the UNL system.

Some words representing a relation, have quite different references. Thus, in the original variant water is defined as a liquid /*water(icl>liquid)*/ meanwhile in the Russian Module it is defined as a binary compound>thing /*water(icl>binary_compound>thing)*/. In the same manner, mobility is an reference /mobility(icl>reference)/ in the original sentence, meanwhile in the Russian UNL expression it as a quality and the antonym of mobility /mobility(icl>quality>thing, ant>immobility)/.

Another UNL site online worth mentioning is the Spanish UNL Server of the Spanish Language Centre (Centro de la Lengua Española). Let's observe a UNL expression in Spanish System annotation and enconversion of the same sentence by the Russian UNL server.

	UNL original sentence	**UNL-Russian Module**
Original Text	This computer translates from English to Spanish.	
UNL expression	agt(translate(icl>do).@entry, computer(icl>machine)) mod(computer(icl>machine), this) src(translate(icl>do).@entry, english(icl>language)) gol(translate(icl>do).@entry, spanish(icl>language))	mod(computer(icl>machine>thing),this(icl>adj)) agt(translate(icl>do,com>language,src>thing,agt>thing,obj>thing, gol>thing).@entry.@present,computer(icl>machine>thing)) src(translate(icl>do,com>language,src>thing,agt>thing,obj>thing, gol>thing).@entry.@present,english(icl>adj)) gol(translate(icl>do,com>language,src>thing,agt>thing,obj>thing, gol>thing).@entry.@present,spanish(icl>adj))
Translation in system language	No translation is generated possibly due to server problems	Этот компьютер переводит с английского на испанский.
English translation	No translation is generated possibly due to server problems	This computer translates from the English into the Spanish.

In the second example, we bring an annotated Armenian text already translated from the same UNL expression into Spanish, Russian and Armenian. A more close translation to original Armenian text we observe in the Armenian deconverted text. Here we have some differences of word order, but a correct usage of declination system (correct usage of locative) and the number.

The Spanish Module has some errors in the word order, **las ciudades de centros** means cities of the centers and the system ought to have used the relative pronoun ***como*** (*like*) in place of the preposition ***de*** (*of*)to form the part – **las ciudades como centros**. Another relevant error is the usage of the past participle in a verb construction *estar + Participio* (in our case – **estar desarrollado** /*have developed*/). In place of estuvieron

desarrollado the system ought to have written **estuvieron desarrolladas**, because the Spanish has concordance by gender and number, as the verb denotes plurality (*estuvieron*/3ps,pl/) and the participle refers to the cities which in Spanish are of feminine gender (*ciudad*/f./).

The Russian Module has generated a better example referring to the morphological usage of the words, but no in case of word choosing from the dictionary. A simple example of it is the confusion of the word oriental (восточный), which the system has translated into ориен-тирует (orientated). Although the same thing as in the case of the cities of the centres has been conserved (города центров), it can be considered a better translation.

Armenian *(original text)*	**UNL Spanish Module** *(deconverted text)*	**UNL Russian Module** *(deconverted text)*	**UNL Armenian Module** *(deconverted text)*
Ð³ Ûª³ Ï ³ Ý µ³ ñÓñ³ í ³ Ý-¹ ³ Ï áðÙ »Õ³ Í ù³ Õ³ ù-Ý»ñÁ áñå »ë ³ é¨ ï ñÇ¨ ³ ñÑ»ëï Ý»ñÇÏ »Ýï ñáÝ-Ý»ñ ² ñ¨ Ùï Ûª Ý¨ ² ñ¨ »É Ûª Ý Ð³ Ûª ëï ³ ÝáðÙ ëÏ ëáðÙ »Ý ½³ ñ· ³ Ý³ É Ù.Ã.³ . III-I ¹ ³ ñ»ñÇ ÁÝÃ³ óùáðÙ£	**las ciudades de centros, en meseta, , estuvieron desarrollado en los siglos desde primer B.C. a tercer en Armenia occidental y Orient.**	Западный и Армения, в которой города центров как торговли и мастерства, которые существовали в армянском плато, развились в века в the процессе из – от 1 B.C. до 3 -, ориен-тирует.	³é»õïáõñ »õ ³ñÑ»ëï áñå»ë Ï»ÝïñáÝÝ»ñ ù³Õ³ùÝ»ñ »Õ³Í Ñ³Û µ³ñÓñ³í³Ý¹³ÏáõÙ 3Çó Ù.Ã.³. ÙÇÝã»õ 1 ¹³ñ»ñ ¿ ½³ñ·³Ý ³ñ»õÙïÛ³Ý »õ ³ñ»-õ»ÉÛ³Ý Ð³Û³ëï³ÝáõÙ

In another example, we can see word order differences in a sentence. The example of the English original text has the Armenian equivalent to it as the best result, although the modal verb can (**կարող են**) here is omitted. Then comes the Russian one, which has the modal verb (**могут**) but not in the proper place and misallocation of some words, too. The Spanish Module only gives some words, some again without concordance (*ciudades situada*), meanwhile there is a word that ought to have been two words (*vivdonde*). In all three target languages we can trace errors in comparison with the English original version. Spanish lacks of pronouns (*it* - **esto**), misallocates the word *where* (**donde**). Russian excesses the usage of the word phrase *where can* (**могут то, где**), Armenian uses infinitive forms of the verbs due to the system configuration and errors in reading rule priorities.

English *(original text)*	**UNL Spanish Module** *(deconverted text)*	**UNL Russian Module** *(deconverted text)*	**UNL Armenian Module** *(deconverted text)*
It determines where people can live and where cities can be located.	**define que gentes pueden ser vivdonde y ciudades situada donde.**	Это определяет, что люди могут то, где жить, и города могут то, где находиться.	³Û¹ áñáßáõÙ ¿ áñï»Õ ÅáÕáíáõñ¹Ý»ñ ³åñ»É »õ áñï»Õ ù³Õ³ùÝ»ñ ·ïÝí»É :

Syntactic rules are rather strong in Russian and Spanish modules[1], however, these rules sometimes make a bit of a mess. Let's compare the following two examples where the together to the UNL original passage are given the UNL Armenian and Russian Modules examples of enconverted and deconverted text:

UNL *(original text)*	**UNL Russian Module** *(enconverted text)*	**UNL Russian Module** *(deconverted text)*	**UNL Armenian Module** *(deconverted text)*
Since the mid-twentieth century, humans have faced an awesome realization.	The humans have faced an awesome realization since the mid century 20.	Люди встретят благоговейное осуществление с середины века 20.	երկյուղալի իրա-գործում մարդիկ ձեռնամուխ լինել կես-քսան դարից:

Here we can observe misallocation of words in the Armenian translation. This is due to non proper usage of the KB dictionary and lack of syntactic rules. We suggest creating syntactic rules based not on the basic computer syntax, but on language modeling. In the present version the UNL system in the UNL-Armenian Module has the following rules used in Armenian is the SVO model, as the Armenian corresponds to it, first of all writing down the subject, the verb and then the object of the sentence:

In this instance we come across some internal rules that are common to each individual language. Many synthetic languages use a GNP (General Noun Paradigm) system referring to the nouns, and a verb system common to nearly all Indo-European languages, where we have a different ending for each person in singular and plural and amply use the changing auxiliary verb, no matter whether this auxiliaries stand before or after the conjugated verb.

Besides comparisons, we have observed some errors within the system. Thus, one of the errors worth mentioning is the case of adverbs vs. verbs. Some examples brig forth the word ***mad*** as a verb, when really the verb is a phrasal one and is composed with the help of the auxiliary verb ***to be*** – **to be mad**.

[1] At the time of writing the article the Spanish version did not give any result at all of the upper mentioned example.

Another example is the word **freshwater** that is translated into Russian as an adjective (**пресноводный**) and not as a noun as it can be understood from the original word. Armenian translates it as a phrase consisting of an adjective and a noun (**å³ñ½ çáõñ**).

Sometimes word may be proper in English and simple in Armenian. A clear example of it is **Heaven**, a proper noun synonym to ***paradise***. But Russian and Armenian commonly translate this noun into a common noun – ***paradise*** (Russ - **рай**; Arm - **¹ñ³Ëï**).

Summing up the following articlewe have come to a conclusion, that many translations done for the UNL expression are generated from enconverters and deconverters not presented online, there is no possibility to compare their UNL expressions and output texts with ours and besides, there can be observed a number of errors between the original and the UNL expressions within the working ones online. It is highly recommended to revise carefully all translators of the UNL system adding some more rules and modeling the words in more details, which will help the system to proceed with better results in the future.

REFERENCES

1. Uchida H., Zhu M., "The Universal Networking Language (UNL) specifications" version 7, *UNDL Foundation*, June 2005."The Universal Networking Language (UNL) Specifications", v.3 edition 2, *UNDL Foundation*, 2003

2. UNL Armenian Module Server, www.unl.am, www.unl.am/webdeco/index.jsp

3. UNL Russian Module Server, www.unl.ru, www.proling.iitp.ru/deco

4. UNL Spanish language centre, www.unl.fi. upm.es/spanish/lg_test.htm