Ontology-based understanding of scientific natural language texts

Prof. Dr. Winfried Schmitz-Esser
University of Applied Sciences, Hamburg

Abstract:
Foundational aspects of an onomasiological approach in constructing and maintaining terminology-based ontologies are discussed, as well as a transfer oriented mode in establishing cross-language implementation. Starting out from the concepts rather than from lexical primitives enables best choices as to the representational quality of the ontology while remaining open to any accepted linguistic way of expression encountered in the different languages. Means of how to ease usual restrictions as to expressiveness, consensuality, and conceptual orientation are focussed on, as well as questions of reliability, diachronic stability, and editorial engagement.

Introduction

Few linguistic ontologies have been built so far on the basis of grammatical units that are bigger than words, and where pluri-lingual ontologies exist, they usually focus on a one-to-one mapping of small lexical units [1]. In knowledge modelling, this state of the art sets narrow boundaries as to expressiveness, consensuality, reuse, and knowledge sharing and considerably reduces an ontology's potential of grasping the desired meaning in real world natural language texts.

An onomasiological approach in ontology construction, update and maintenance might be the answer to such problems, combined with a transfer approach on cross-language implementations. The layout and logical structure of the paper presented is shown in Fig.1.

1. What a software tool (or human interpreter) needs in order to understand natural language text
2. How the ontology must look like to allow statements (knowledge) to be modelled and formalized
    - 2.1.Arguments
        - 2.1.1.Concepts
        - 2.1.2.Instances
        - 2.1.3.Foundational aspects
    - 2.2 Semantic relations
        - 2.2.1. Foundational aspects
3. Modelling, and the functions of models
4. Foundations and principles of stipulating semantic statements

Fig. 1. Layout and logical structure of paper presented

1. What a software tool (or human interpreter) needs in order to understand natural language text

Assuming that some knowledge machine is capable of understanding scientific natural language text by means of a knowledge base constructed as an ontology, then what should we expect to be the foundations of such an ontology, and what are its principles of construction, update, and maintenance?

Spontaneously, one might prompt: Well, in the first instance, the ontology must know the specific language that is used in common communication between the members of a community on subjects of interest relating to a specific domain.

This seems simple, but it is not simple at all. Every single element in the equation is soft. What is a <specific domain>? Evidently, domains exist. But what are the frontiers? What is a <domain's specific language> as opposed to the specific languages of other domains, and how does this relate to those parts of some general language, authors would use anyway in their natural language texts? What are <subjects of interest> in a given domain, to-day, yesterday? And what may they come to be tomorrow? What is <common communication>? And what does it mean that somebody, or, in this case, a machine, < must know the language>?

Obviously, this is not a matter concerned solely with dictionaries. And it is not a matter of just a few search words on the Internet.

We have to plunge into the deeper waters of terminological communication [2] [3].

To understand the meaning of a message expressed in terms of a natural language (NL), a thorough knowledge of the respective World segment, what it is about, and how it works, is required on the part of the receiver, as well as an equally thorough knowledge of how all this usually is, or may be, addressed in terms of the respective NL. Much of this knowledge will then consist of tacit knowledge that comes from experience, observation, reasoning, inference, etc.

If we want a machine to reason and to understand the semantics of NL text, then we have to provide it with the knowledge required for that purpose. We have to do it by means of our natural language, and in an utterly explicit way [4]. This is what occurs in an ontology of our times. Such an ontology explicitly describes, or models, the knowledge with respect to a specific segment of the world. A suitable, machine-readable format, mechanism and repository are required for that purpose.

Since the condition is natural language, such modelling will find its limits within the possibilitiy of being expressed by means of natural language. We are on the well articulated, sometimes ambiguous, but practical grounds of first order logic, proven in centuries. In models of this kind, we are not talking about variables with their much higher degree of exactness.

The main condition is conceptualization. It is not possible to describe, or model some piece of knowledge by means of NL without having the possibility to express it by means of concepts that we expect to be known by other members of the language community.

This touches the focal question of the representational capacity of the spoken language. In an ontology conceived as a model of some section of the world, no authentic knowledge modelling is possible unless full, unequivocal representation of the meaning in terms of NL can be assured [5].

I do not want to enter the discussion about terminological representation here. My view in this dispute is a constructivist [6] view. Yes, such representation is possible, within, however, the limits of some serious constraints and conditions.

2. How the ontology must look like to allow statements (knowledge) to be modelled and formalized

Let us assume that the ontologic model can be based on a binary semantic relation of the type

Argument 1 - Relational Operator - Argument 2

The relational operator then is normally presented in two forms, as RelOp(a) and RelOp(b), expressing which of the two arguments stands on the active side and which one is on the passive side. In NL texts of the different single languages, however, the semantics of such a statement may be encountered as expressed in various other ways, e.g. as nouns or integrated as parts of nouns, or as modifiers to nouns or verbs. This is a major part of the problem (Fig. 2).

short-circuit - *Cause(s)* - black-out
black-out - *Is/AreCausedBy* - short-circuit

In English, equivalent expressions of the argument may range from “short circuit” to totally different constructs like: “induction”, or “lower resistance electric connection”.
Likewise, a relation saying that something is, or may be *“CausedBy”* another something, is equivalent to a number of other expressions in English NL texts, like “brings about”, “effectuates”, “leads to”, etc. This, still, is not exhaustive. Expressions may appear in other syntactic forms like: “a short-cut *induced* black-out”, or “thanks to a short-cut in our old transformer station we had a black-out yesterday”.

Fig. 2. A statement (enouncement formalized)

Having said this, and searching for the foundations of such an ontologic model, let us take a closer look on this kind of enouncements, and of its constituent elements, (1) the arguments and (2) the relations.

2.1.1 Arguments - Concepts

Maybe it is true “that languages affect the way we conceptualize the world”, as Hjörland. B. et Nicolaisen J. (2004) state [7]. Differences in word meanings among the different languages, (like wood [English] - Holz [German] - trae [Danish] -bois [French] etc.) are quoted here as obstacles to safe inter-language communication. This however holds true only on the assumption that conceptualized communication functions on the basis of what in early information retrieval was called a uniterm basis. But this is far from what we see in reality. A string of terms is much more precise than a uniterm. Besides, a string of words usually makes a much better discriminator in searches.

To model our conception of the world, we have to enlarge our conception of a concept. To take the above example: The single word Holz is considered a concept, it is a uniterm with its particular, valid meaning only in German. We cannot expect a uniterm with this same meaning to exist in English as well, or in Danish. When talking of concepts, we have to separate the idea of what we mean - the essence of the concept -, from the word or words that express it, in a particular language. Such expressions may be manifold for a given object.

What holds us back of treating any universal idea, or topic, simple or complex as it may be, as a concept in our ontology, irrespective of how it may be expressed in a particular language,

and irrespective of the number of single words of this language that may be in use to express its meaning? The only limits I see is practical feasibility in the face of an exploding complexity, and limited resources needed for an immense task. But that may change as the usefulness of ontologic modelling becomes more obvious, and the need for better ontologies arises.

The key to a valid ontology lies in how we construct it. We have to do it the onomasiologic way, as linguists are used to say. We do not start out from a given term asking for what it stands (the semasiologic way), but rather from the idea of what we want to express, of what we find expressed in a text. Then we have to ask how we find this particular idea/topic/universal expressed in other heterogeneous, valid, textual sources. This is the onomasiologic way.

In a recent critical analysis of the 17,000 concepts strong Thesaurus of the National Cancer Institute (NCI), which is the achievement of a spectacular attempt to homogenize different terminological resources, Ceusters, W. et al. (2005) [8] found many mistakes and inconsistencies, denouncing weaknesses and deficiencies that are seemingly unavoidable (and difficult to correct) in this and other semasiologic approaches.

The onomasiologic approach, in contrast, forces us to find the many variants of expressions used in communication for a given subject, including the most used paraphrases. As we find them in authoritative source texts, we see them as Authoritative Expressions (AE) that convey the sense. So, we are well advised to include them in the ontology.

In an Equivalence Chain of Expressions (ECE) yielded from such a procedure we can separate those AEs which are (A) clear, univoque expressions for the subject from those (B) with other meanings (polysems in a wider sense). Among the univoque expressions of the A-type, the curator/peer should then be able to stipulate the most fitting one to function as a proxy for the other univoke A-types which in my papers [18] are named "Additional Access Expressions - (AAE)". The proxy here is made up as a descriptor (DESC) very much along the traditional lines. In the ontologic model, the DESC represents the meaning. It is singular as ID, univoque and context-independent. The polysemic B-types in the ECE are the MULTIs (Fig. 3). They can be handled and used in other, appropriate ways.

In source texts used for an ontology on public finances the following expressions for "public budget deficit" were encountered:

1) "public deficit"", (2) "scarce public means", (3)"hole in the federal budget", (4) "public deficit for a central government in a country", (5) "lack of public money"", (6) "federal budget deficit", (7) "gap in the federal budget", (8) "state budget deficit", (9) "budget gap" (10) "gap in the budget", (11) public budget deficit

The curator/peer chose/grouped/ stipulated:
(4) as DESC
(1), (3), (6), (7), (11) as AAEs
(2), (5), (9), (10) as MULTIs, if at all
(8) in a special way

Fig. 3. Three different types of elements in an Equivalence Chain of Expressions (ECE)

To the practitioner it is no surprise that perfect identity between any two words in a language is rare, and that on a level of phrases it is almost non-existent. Margotti, F. W. (2004) [9], in a

recent study, demonstrates the variety of algorithms modern languages offer to construct possible paraphrases, and that "the choice of one or the other paraphrase structure is never aleatoric". While identity of meaning is the exception, Margotti states, identity is possible in the sense of equivalence, "and equivalence is always gradual" (ibid).

If this is true, the occurrence of paraphrases for a given utterance in heterogeneous, valid source texts has something to say. We can use it in a statistical sense in ontology construction. We would have to accept the phrases, or chunks of phrases, as they come provided they occur in authoritative texts.

Equivalences within an argument encountered in English, French, and German authoritative texts, could then be formalized for use as arguments in the ontology the following way. (Fig. 4), presented in a redundant format, (MULTIs suppressed)

202, 0111, DESC, EN, breast cancer cell proliferation
202, 0111, AAE, EN, proliferating breast cancer cells
202, 0111, AAE, EN, motility of breast cancer cells
202, 0111, AAE, EN, proliferation of breast cancer
202, 0111, AAE, EN, breast cancer spreading
202, 0111, AAE, EN, breast cancer cells migrating
202, 0111, DESC, FR, prolifération de cellules cancereuses mammaires
202, 0111, AAE, FR, prolifération de la tumeur mammaire
202, 0111, AAE, FR, prolifération de cellules de la tumeur mammaire
202, 0111, AAE, FR, prolifération du cancer mammaire
202, 0111, AAE, FR, motilité du cancer mammaire
202, 0111, AAE, FR, prolifération de cellules de cancer du sein
202, 0111, AAE, FR, prolifération de cellules des tumeurs du sein
202. 0111, AAE, FR, motilité des cellules des tumeurs du sein
202, 0111, AAE, FR, prolifération des tumeurs du sein
202, 0111, DESC, GE, Proliferation von Brustkrebszellen
202, 0111, AAE, GE, Brustkrebsmotilität
202, 0111, AAE, GE, Motilität von Brustkrebszellen
202, 0111, AAE, GE, Proliferation der Brustkrebszellen
202, 0111, AAE, GE, Ausbreitung der Brustkrebszellen
202, 0111, AAE, GE, Brustkrebszellen wandern
202, 0111, AAE, GE, Entstehung neuer Brustkrebsherde
202, 0111, AAE, GE, Zellen des Mammakarzinoms verbreitet

Fig. 4. An example: Argument No. 202, an English, French, German set

---

Legend

202 = individual ID number of what is meant in the given ontologic model, whereby the first digit indicates the quality of the meaning as a (universally understandable) concept (as opposed, e.g. to an instantiation)
0111 = in humans (01), in vivo (11)
DESC = Descriptor, preferential expression in an Equivalence Chain of Expressions in a specific language
AAE = Additional Access Expression, equivalent in meaning to the DESC in a specific language
EN = Expression used in English NL texts
FR = Expression used in French NL texts
GE = Expression used in German NL texts
Last element in line = NL expression as found to actually occur in a valid, scientific NL text, unabridged, full forms

Set in a Boolean OR position, the NL fragments of each of the three language sets can be used for effective browsing in the respective NL text collections.

2.1.2 Arguments - Instances

With some minor modifications, the underlying construction can be used to build instantiations, i. e. descriptions of instances (individuals) which are phenomena singular in space and time: bodies by law or agreement, trade marks, events, the meteorologic highs and lows, typhoons, in short: unique cases of a class/category of universals, which in normal life all are given an individual name. Like concepts, the names of instances obey the linguistic laws of their respective languages.

2.1.3 Arguments - Foundational Aspects

To find out a multiple language set of the type shown above is more than a good translator's job. Both first class domain knowledge and an intimate command of the languages implemented on the ontology are required. The rules are those of a researcher and an interpreter combined, and the method is the one known among translators as "Transfer", or "transposition". This fits nicely into the onomasiological conception of the approach.

The choice of scientific texts that contain the knowledge for the knowledge base is subject to dispute in no other way than this is usual in traditional scientific reading and processing. The same applies to the choice of the single arguments (universals/concepts/categories and instantiations/individuals).

Enforcement by standard procedure should make sure that the extraction and transfer process is fully documented on all stages.

So, what are the foundations of such identification and transfer?

- ***Pragmatism, skill, and editor's responsibility.*** The editor of the ontology is responsible for the work of his team of curators/peers. It is their task to keep up the ontology, to control the choice of concepts and instances stipulated, to keep the ontology clear of idiosyncrasies, to warrant that the relevant part of the knowledge is modelled and organized in the ontology, and to execute this in conformity with the principles of good practice in the handling of terminology [10]. Any shades of meaning encountered as an equivalent or near to equivalent of the DESC should be considered a valid paraphrase as long as the meaning remains true (ILARI, R. e GERALDI, W. 1985) [11]. Pragmatism prevails in the handling of words borrowed from other languages: So, EN expressions found in German source texts will be labelled as GE expressions.

- ***Linguistics and Representation.*** Any full, unabridged, unaltered term string used by an author of a relevant source text to describe a relevant subject in a particular NL is considered an Authoritative Expression (AE). All AEs encountered are to be transferred 1:1 and included in their respective argument sets. Each act of transfer is to be controlled and fully documented in a scientific way. Within the Ontology, no bias or limitation is admitted from any lingware or other linguistic inference.

- ***Probability.*** Probabilistic approaches are admitted insofar as (1) the representational aspect of the sample source text collection needed in the beginning is concerned, and/or (2) the validity of the choice of current, new source texts needed for update, enhancement and extension of the ontology. Also (3) occurrence of structural forms

encountered in the source texts. No probabilistic approaches admitted on matters of meaning.

2.2 Semantic Relations

The semantic relations in focus are "Essential Relations" i. e. relations that exist on an inter-language, universal level [12]. They can be expressed in any language. As to their meaning, they are independent of a particular language, like *"Cause/Effect"* in the sample above.

Such relations may be of a general type like the 13 relations proposed for general use in the nineties by the German Committee on Classification and Thesaurus Research - KTF [13]. Others must be more specific to meet a corresponding need, such as in "Insulin-like growth factors IBFs *Regulate(s)* breast cancer cell proliferation". As an example, the definition of such a specific, universal relation is given in Fig. 5, plus an insight on the respective semantic/linguistic toolbox:

Fig. 5.1. - What the Semantic Relation "Enhancement" means

RelClass2-Type 33: Enhancement

In OntoStatements, the enhancing relationship expresses that Argument A which is the enhancer, enhances, or may enhance, Argument B which is the enhanced.

Roles of Arguments: *A is the Enhancer*
B is the Enhanced

Meaning of *enhance* as modelled in this Ontology: The enhancer (A) enhances, or may enhance, the enhanced (B), makes B greater as in value, activity, performance, desirability, but also as in negative values like fear, budget deficit, morbidity.

Fig. 5.2. - How the Semantic Relation is expressed in the different languages

>English NL Relational Expressions
English NL expressions for RelClass2-Type 33: Enhancement to be activated for searches in English NL text collections:

Phrases****** *A enhance(s)* Phrases B
Phrases A *activate(s)* Phrases B
Phrases A *increase(s)* Phrases B
Phrases A *accelerate(s)* Phrases B
Phrases A *intensify/intensifies* Phrases B
Phrases A *promote(s)* Phrases B
(and all reverse)
Open for inclusion of more candidates when encountered in valid, real world texts
---
**)Phrases are either (1) given, defined, "known" NL denominations of the argument (universals or instances) inferred from the Ontology, or (2) searched, "unknown" NL character strings of variable length expected to occur in NL text corpora when searching.<

>French NL Relational Expressions
French NL expressions for RelClass2-Type 33: Enhancement to be activated for searches in French NL text collections:

Phrases A *augmente(nt)* Phrases B
Phrases A *agrandit/agrandissent* Phrases B
Phrases A *renforce(nt)* Phrases B
Phrases A *fait hausser/font hausser* Phrases B
Phrases A *stimule(nt)* Phrases B
Phrases A *intensifie(nt)* Phrases B
Phrases A *accellère(nt)* Phrases B
(and all reverse)
Open for inclusion of more candidates when encountered in valid, real world texts<

>German NL Relational Expressions
German NL expressions for RelClass2-Type 33: Enhancement to be activated in searches in German NL text collections:

Phrases A *erhöht/erhöhen* Phrases B
Phrases A *vergrößert/vergrößern* Phrases B
Phrases A *verstärkt/verstärken* Phrases B
Phrases A *intensiviert/intensivieren* Phrases B
Phrases A *beschleunigt/beschleunigen* Phrases B
Phrases A *stimuliert/stimulieren* Phrases B
(and all reverse)
Open for inclusion of more candidates when encountered in valid, real world texts<

Fig. 5.3. - How Semantic Relations inter-relate

As to: RelClass2-Type 8: Beneficial
RelClass2-Type 33: Enhancement *Is/AreNarrowerConceptOf* RelClass2-Type 8: Beneficial
RelClass2-Type 8: Beneficial *Is/AreBroaderConceptOf* RelClass2-Type 33; Enhancement

As to: RelClass2-Type 9: Detrimental
RelClass2-Type 33: Enhancement *Is/AreNarrowerConceptOf* RelClass2-Type 9: Detrimental
RelClass2-Type 9: Detrimental *Is/AreBroaderConceptOf* RelClass2-Type 33; Enhancement.<

Remarks
On a higher level of abstraction, ENHANCEMENT OF something may be judged to be BENEFICIAL or DETRIMENTAL, respectively, depending on the respective value attributed to it. Enhancement is value-free. This suggests that the same type of interrelations may happen to occur between both Beneficial and Detrimental.

Fig. 5. Striving to grasp as many NL expressions for the "Enhancing" relation as possible and its potentially related subjects or objects

2.2.1 Foundations and Principles of Relation Stipulation

So, what are the foundations, and principles, of stipulating semantic relations?

Pragmatism again. The set of semantic relations admitted to the ontology must fit the purpose. There are no established theories about how to find out, and define, task or domain-related inventories of semantic relations. Wherever workable sets of them have been tackled, we see them as a result of trial and error. Pragmatism prevails. In a way, the task resembles that of a translator, i. e. insofar as true, valid, understandable expression is the target of what a relation of relevance means.

But the meaning of a particular relation must be defined in a general way, and it must be semantically related in a widely intersection-free manner with respect to the meaning of each of the other relations of the set. This is a task of a linguist and information scientist.

It is obvious that ideally the inventory of semantic relations should be ready at the beginning of semantic modelling. Practice, however, shows that a given set must always be open for further adjustment. Some lack of exactness seems unavoidable. All in all, the choice of the relations admitted should guarantee fair coverage of the knowledge organized/to be organized.

A clear-cut, widely intersection-free definition of each relation is of prime importance, as well as a reasonable restriction in number, right from the beginning. This is important to ease handling and to maintain a sharp consensual profile. Insofar as specific, task-oriented relationships are admitted, it is wishful that they are put in a formal relationship to a restricted number of known, general relations. This, then, may help to obtain higher recall in searches if wanted in given cases.

Experience from two major projects (EXPO 2000 [14], SERUBA [15]) shows that with only half a dozen of well-chosen, general relationships *recall as well as precision* can be improved.

3. Modelling, and the Functions of Models

The instruments outlined allow the modelling, i. e. the construction of a Statement, saying that

Insulin-like growth factor IGF-I *enhance(s)* **breast cancer cell proliferation**
the same as (produced by algorithm)
breast cancer cell proliferation *Is/AreEnhancedBy* **Insulin-like growth factor IGF-I**

and to express this in all three languages, in all paraphrased shades detected as occurring in valid source documents, and stipulated in what can be assumed are its most used/accepted, relevant structural forms.

In reuse, queries like e.g. "What *Is/AreEnhancedBy* IGF-I? can then be prompted right away from the internal Basic Semantic Reference Structure (BSRS) of the ontology. Easy- to-understand NL answers to the query can be generated in EN, FR and GE. They will be given on the basis of their respective DESCs, in their canonic forms (which are supposed to be applicable free of any other context, by definition), without their paraphrased AAEs, however, since some contextual pollution would be unavoidable as shown in Example 3 above.
.
The other part of the search result would go beyond that. It would consist of a number of uncontrolled text strings detected in NL texts in either of the three browsed text collections. These should be strings of variable length that either precede *"EnhancedBy",* or follow "enhance" or "enhances" (this covering all stipulated shades of meaning) in the three languages. That type of search one day may allow to detect tacit, henceforth unknown knowledge in NL texts.

On top of that, such a query could be extended by taking an upposted, broader view , like: "What *Is/AreFavouredBy* IGF-I"?, which is a specific shade of the "Beneficial" relation as recommended for use by the former KTF mentioned above.

All this is possible also on a cross-language basis without further formal restraints. An information seeker may put the query in German and receive texts from divergent sources written in English, French or any other language implemented in the ontology (Cross-Language Information Retrieval - CLIR). The same applies for reused knowledge streaming directly from the reasoning within the BSRS. This, then, would be something really new: Cross-Language Knowledge Reuse - CLKR.

The consensuality command poses restraints for sharing: Another pragmatic aspect is that of sharing. Since the most specific and most valuable part of knowledge organized in ontologies is drawn from the statements, and since the main element of the statements are semantic relations, the semantics of semantic statements are of crucial importance in sharing

knowledge among different ontologies. In any attempt to share knowledge, consensuality potentially is at stake.

To illustrate this: If in ontology A the enhancing relation is admitted and used as explained in Chart 5 above, and in ontology B nothing more specific exists than the beneficial relation, then a sharing of knowledge is possible between the two ontologies. This is due to the inter-relational facility foreseen in ontology A, but it is possible only at a lower level of expressiveness which is dictated by ontology B. If B would force A further down to a still broader level, say, a loose, undefined isness relation, then the usefulness of sharing statements would be drastically reduced.

This leads us to the following recommendation: Stipulate relations in accordance with the required exactitude of expressiveness or level of exploitation (which in the end is coming to the same) while purveying interrelated relations down to the more common, broad levels! Provide the possibility of segmentation according to the criteria of such relational hierarchy! In all future sharing operations, you will then find yourself on the strong side.

4. Foundations and principles of stipulating semantic statements

Pragmatism also, but… The stipulation of semantic statements will always be governed by a pragmatic, target-oriented approach. The act of transfer and transposition of enouncements from the source texts can be made fully transparent and widely controlled. Curators/peers are supposed to be bilingual (at least) and experts in their proper disciplines. Machine-aided support must be provided in many ways, and at different points of the stipulation process. The ontology is transparent, not open to uncontrolled content, the sources are well documented, the BSRS and the vocabulary is under full machine-aided, but in last instance, intellectual control, consensuality among curators/peers is high and can be enforced, the editor's responsibility for the message is established. So, what is the problem? Are there any?

There are some [16]. Just to mention a few:

Representation. No single ontology will ever represent a domain in full. This may be in contrast to what users expect from ontologies. Exact sampling of source texts will be crucial, especially so in the early stages of ontology development. A more process-oriented approach may take over as the ontology matures. The explosion of the demand for constant, daily update alone will quickly impose serious economic constraints with gaps and rebates on reliability as a possible consequence. Task ontologies are better off inasmuch as tasks can be exhaustively defined beforehand.

Epistemology. Transferring statements and terminology from heterogeneous textual sources on to an axiomatic model under strict, general, defined conditions, but of limited expressiveness, forces the ontology constructor to take an epistemologic stance that cannot always be expected to be up to what the author of the source wants to express.

The model at best is an abbreviation, not the original. What finally is transferred may lag behind an author's message, it may not wholly meet his message, e. g. because it cannot be expressed by the instruments the ontology provides, or for lack of understanding on the side of the curator/peer, or, on the other hand, it may be the ontology that poses a demand for clear-cut expression which the source text cannot fulfil, because the author hesitates to be clear-cut in the matter, etc., etc.

Logic, and Conclusion. It is often postulated that knowledge representation systems must strictly follow the rules of logic, must be contradiction-free, and insofar must be consistent [17]. Such postulate cannot apply to that part of the ontology that deals with AAEs and MULTIs. It is easy to demonstrate that NL language often is all but logic. Isolating AAEs and MULTIs corresponds to that need. Principles of logic and truth, and the demand for consistency, however, do apply to what is represented by the descriptors in terms of words, and is stipulated as DESC. Whether we can go beyond postulating logic and truth to be the underlying principles when dealing with semantic relations, and thus with entire statements, I am not that sure. As Margotti [9] puts it: "Equivalence is always gradual" This seems to be especially true of the relations between relations.

However, can we expect a model to enable a better match than does its prototype?

References

1. GÓMEZ-PÉREZ, A. et al., 2004: Ontological Engineering. Springer, 2004

2. BENVENISTE, E., 1966: Problèmes de linguistique générale. Paris, Gallimard

3. MARTINET; A., 1985: Syntaxe générale. Armand Colin, Paris 1985

4. GIBBON, D., et al., 1997: Handbook of standards and resources for spoken language systems. Mouton de Gruyter, 1997

5. JACKENDOFF, R., 1996: How language helps us think. Pragmatics and cognition. 4/1. Amsterdam, Benjamins, 1996, 1-34

6. FUCHS, C., 1997: Diversité des représentations linguistiques : Quels enjeux pour la cognition ? In : Diversité des langues et représentations cognitives, Paris, Ophrys, 1997, 5-24

7. HJÖRLAND, B., NICOLAISEN, J., 2004: The epistomological lifeboat. http://www.db.dk/jni/lifeboat/

8. CEUSTERS, W. et al., 2005: A terminological and ontological analysis of the NCI Thesaurus. In: Methods of Information in Medicine, 4/2005. http://ontology.buffalo.edu/medo/NCIT.pdf

9. MARGOTTI, F. W., 2003: Sinonímia e paráfrase: Algumas consideracões a partir de dados do Atlas Lingüístico-etnográfico da Região Sul-ALERS. In: Linguagem em discurso, 3/2, Tubarão, Brasil

10. SCHMITZ-ESSER, W. 2000: How to cope with dynamism in ontologies. In: Dynamism and Stability in Knowledge Organization. Proc. 6th Int. ISKO Conf., Toronto, Canada. Ergon. 2000, 83-89

11. ILARI, R., GERALDI, W., 1985: Semântica. Sao Paulo, Atica 1985

12. SCHMITZ-ESSER, W., 2003: Meaning, understanding, and the organization of knowledge in a multilingual world - New tools for new tasks: Ontologies. In: Linguistic cultural identity and international communication. Vielberth, J., Drexel, G., Eds., AQ-Verlag, 2003, 149-171

13. SCHMITZ-ESSER, W., 1999: Thesaurus and Beyond: An advanced formula for linguistic engineering and information retrieval. Knowl. Org. 26 (1999) No. 1, 10-22

14. SCHMITZ-ESSER, W., 2000: EXPO 2000 - INFO 2000. Visuelles Besucherinformationssystem für Weltausstellungen. Springer 2000.

15. SCHMITZ-ESSER, W., 2000: SERUBA - A New Search and Learning Technology for the Internet and Intranets. Proc. 11th ASIS&T SIG/CR Classification Research Workshop. Chicago, IL., Nov. 12, 2000, 91-102

16. SCHMITZ-ESSER, W., 2005: Ontologien als Herausforderung an Publizisten und Wissensorganisatoren. Elbe E-Lectures, Broadcasting &Events, Hamburg, 2005 www.elbe-studios.de/list.php?pers=101

17. DE BRUIJN, J. 2003: Using ontologies. DERI Technical Report. DERI-2003-10-29. Innsbruck, Galway, 2003. http://www.deri.at/publications/teachpapers/documents/DERI-TR-2003-10-29.pdf

SCHMITZ-ESSER, W. 2002: www.schmitz-esser.de