

Harmonisation and Formalisation of Nursing Terminology: a three-dimensional approach.

Dr. Werner CEUSTERS

Abstract: The use of computers and electronic documentation systems in the nursing environment has put forward new requirements on terminologies. In this paper, we present an introduction to the development of formal terminological systems along a cognitive, linguistic and communicative dimension. The goal is to arrive at formal terminologies that assist nurses in their documentation activities while minimising the burdens related to the use of traditional terminologies.

1. Introduction

Communicating information is an essential activity in nursing. Patients need to be informed to what diagnostic or therapeutic procedures they will be submitted in order to make them feel more comfortable in the often threatening environment of the hospital. Also nursing colleagues need to be informed on what happened to the patients there are responsible for before taking up their shift. It is mandatory that this exchange of information is done in an unambiguous, accurate and reproducible way. This is not always so easy because language itself - the prime vehiculum in information interchange - is difficult to use unambiguously. In addition, Europe is moving towards a global multilingual community in which from a functional perspective, national borders tend to fade. More often communication is required with colleagues speaking different languages, or having another cultural and educational background.

Given the rather limited capacities of the human brain in storing and retrieving large quantities of factual data, the same information must also be registered in nursing records for subsequent consultation. By using electronic nursing records, some additional functional requirements for this kind of "external memories" became apparent: in one way or another, the information has to be understandable by machines, such that linking to other applications or information sources can be achieved nearly automatically. Unfortunately, computers don't speak natural language (yet), and they also have little knowledge of medicine and nursing.

To overcome the problems related to the use of natural language in communication and clinical registration, coding and classification systems have been introduced as interlingua. Systems such as ICD, Snomed International, ICPC, CPT and many others are now widely used to register medical findings, diagnoses or procedures. Similarly, terminological systems such as NIC, NANDA, ICNP and others are proposed to be used as interlingua in a nursing environment.

The question of course is whether or not such systems are the right solutions to overcome the problems stated previously. After all, each of these systems is designed with a specific purpose in mind such as mortality and morbidity statistics, reimbursement, information retrieval to mention only three. They very seldom are detailed enough for a

faithful registration of all relevant clinical data. And at least in their current (paper) format, they are a burden to use.

In this paper, we present a non-technical introduction to formal terminological systems and how they can be designed. The purpose is to get the reader acquainted with the main ideas and principles, while more detailed information is to be found in the cited literature.

2. Terminology

In [1], terminology is defined as *the study and the field of activity concerned with the collection, description, processing and presentation of terms belonging to specialised areas of usage of one or more languages*. Central in this definition is the notion of *terms*, i.e. verbal representations of the things we speak or write about. Terminology differs from lexicology in the sense that only the terms pertaining to a specific domain are considered.

Three dimensions need to be considered when developing terminologies: the cognitive dimension, the linguistic dimension, and the communicative dimension. Also when existing terminologies are to be compared to be used in a specific environment, it is mandatory to keep these dimensions in mind.

In the *cognitive dimension*, the terms are related to their conceptual contents, i.e. the referents in the real world independent of their material or abstract nature. In this dimension, terms get their meaning fixed. In the *linguistic dimension*, the existing and potential forms of the terms are examined. Here term formation principles are studied. The *communicative dimension* finally looks at the use of terminologies. This dimension has to justify terminology work as such.

A rigorous method must be adopted when designing terminologies. Also, it is mandatory that the work is undertaken by a multidisciplinary team composed of skilled terminologists, linguists, and domain specialists. Usually, one starts by defining the area of usage, the application domain and the intended purpose. If a multilingual terminology is aimed for, also the source- and target languages need to be identified. As a first step, large corpora of documents need to be collated. These documents might be other terminologies developed within the domain under scrutiny - perhaps for a different purpose - or texts in which a high number of candidate terms can be found. This approach is justified when it is assumed that if a term is found in a document (the linguistic dimension), there must be a concept that it denotes (the cognitive dimension). On the basis of this material, a taxonomy of the terms can be set up, i.e. identifying generic relationships between them. Studying the taxonomy might in itself give clues for the existence of concepts for which no terms have been found in the initial corpus.

Special care needs to be taken when doing the work in a multilingual environment. It is always dangerous to translate terms directly from one language into another without giving careful thoughts at the concepts they denote. When the meaning of a term in language A is not exactly equal to the meaning of another term in language B, both terms should not be considered to be each other's translation.

3. Towards formal terminologies

Traditional terminologies (nomenclatures, thesauri, classifications, etc.) are designed to be used by humans. Even electronic versions of these systems, in which it is possible to browse through the hierarchies of the terminology, are still intended to be used by humans, the computer just being there as a replacement for the book. A major problem for such naïve electronic versions is that they cannot take advantage of the knowledge implicitly available in the terms (or the rubrics in classification systems), but that they must rely on the limited knowledge available in the generic links between terms. Finding specific terms

requires a priori knowledge by the user on how the system is structured. With flat terminologies, in which large quantities of narrower-terms depend from one broader-term, the computer is even seen as a burden, because only a limited number of terms can be seen at the same time on the screen. A second disadvantage is that the terminologies only can be viewed in their original structure, and that reclassification of the terms, following different criteria, cannot be realised.

In order to overcome these problems, terminologies must be expressed in a formal way. When doing so, the three dimensions of terminology should not be forgotten.

4. Formalisation along the cognitive dimension: the Galen approach

The purpose of the GALEN project is to develop language independent concept representation systems as the foundations for the next generation of multilingual coding systems [2]. At the heart of the project is the development of a reference model for medical concepts (CORE) supported by a formal language for medical concept representation (GRAIL) [3]. A particular characteristic of the approach is the clear separation of the pure conceptual knowledge from other types of knowledge, including linguistic knowledge [4], in order to arrive in the future to application-independent medical terminologies [5]. Hierarchies in Galen-models are strictly formed around formal subsumption: what is true for a father, is true for a child. Galen-models apply to the closed-world assumption: what is not in the model, is not valid.

The hierarchy of the alpha-version of the ICNP is close to a subsumption hierarchy (Table 1). Each concept at a lower level of the hierarchy is explicitly defined as “a type of” of the father-concept. The differentiating criteria are clearly indicated in the definition. As such, the structure of ICNP is very clear, though one might not agree with the exact way in which it is populated. Also some operational guidelines on the use are not provided. Taking the definitions of “positioning” and “manipulating” into consideration, it is not obvious whether “manually putting a part of the body in a certain position” is to be considered positioning, manipulating or both.

Although its rather unambiguous structure, ICNP is hardly to be called a formal system. The generic hierarchy is represented in the codes of the terminology while the differentiating criteria are not. One of the requirements of a formal terminology is indeed that all information is in the structure.

Table 1: Part of the hierarchy of the International Classification of Nursing Procedures

<i>ICNP-Concept</i>	<i>Subsumption</i>	<i>Differentiating criteria</i>
Nursing intervention	a type of <u>intervention</u> with the following specific characteristics:	actions taken by nurses in response to nursing phenomena
Performing	a type of <u>nursing intervention</u> with the following specific characteristics:	doing a technical task
positioning	a type of <u>performing</u> with the following specific characteristics:	putting something in a certain position
manipulating	a type of <u>performing</u> with the following specific characteristics	manually moving a part of the body
mobilizing	a type of <u>performing</u> a task with the following specific characteristics:	rendering something movable

5. Formalisation along the linguistic dimension

5.1 *There is language and language ...*

Formalising terminologies along the conceptual dimension is “all” that is needed to allow computers to make properly use of them. It is however not sufficient if communication is required between computers and humans, and certainly not for interpersonal communication. The former requires a mapping from the formal language to a language understandable by humans and vice-versa, while the latter requires the unambiguous use of natural language amongst humans. Here we are in the domain of natural language processing and understanding.

While formal language and natural language are at the two most extremes of an axis representing the understandability of a language for a machine or a human respectively, there are two kinds of languages that more or less can bridge the gap. The first kind encompass “sublanguages”, i.e. natural languages used in a particular domain, f.i. nursing, and for a particular task, f.i. communicating or documenting nursing interventions. The second one are known as “controlled languages”. A controlled language is a precisely defined subset of a natural language, on the one hand constrained in its lexicon, grammar and style, and on the other hand possibly extended by domain-specific terminology and grammatical constructions. Both controlled languages and sublanguages have in common that they differ from “general” natural languages by being restrictive, deviant and preferential with respect to vocabulary, syntax, semantics and pragmatics [6, 7, 8, 9]. The main difference is however that sublanguages evolve naturally within a community while controlled languages are artificial adaptations of a language that are tried to be kept as natural as possible. Controlled languages are not to be mixed up with “controlled vocabularies” that are (possibly hierarchically) structured sets of certified terms that are verbal canonical representations of concepts. The aspect of control in a controlled vocabulary is related to the position of a specific term in the vocabulary as a whole, the choice of a particular term as canonical form, and the requirement that only terms from within the vocabulary are to be used in an application. The terms themselves are however not written in a controlled language.

Controlled languages are designed to make natural language processing more feasible both for humans and machines. Texts written in a controlled language contain no ambiguities and are easier to read. They can also be processed by a machine, such that translation or formal representation of meaning can be realised automatically. The drawback is that they are a little more difficult to write as some naturally occurring phenomena in language are not allowed.

In [10], we proposed the use of a controlled language to reduce ambiguity in the terms or rubrics of medical nomenclatures, vocabularies, and coding and classification systems (Table 2). This was based on the many inconsistencies and ambiguities that were found in Snomed International [11] (Table 3).

Table 2: Some basic recommendations for controlled language usage in term formation for clinical nomenclatures

-
1. Avoid using the same word in different meanings and with different parts of speech.
 2. Use prepositions in such a way that they (preferably uniquely) identify the thematic role or object-relation.
 3. Use double or triple prepositions for expressing meaning with greater precision.
 4. Maintain normal word order as indicated by the general grammar of the language in which the terms are expressed.
 5. Limit term length to what (at least) a skilled human reader can easily understand.
 6. Use co-ordination with extreme care.
-

Table 3: Phenomena reducing the understandability of terms in Snomed International

-
1. Inappropriate use of synonymy
 2. Misleading use of homonyms
 3. Complexity of noun groups or noun clusters
 4. Long-distance dependency and cross-modification of term constituents
 5. Ambiguous use of co-ordinated constructions
 6. Different (and unpredictable) semantics of the word “and”.
-

5.2 Cognitive versus linguistic modelling

When formalising terminologies along the cognitive dimension, an *ontology* has to be defined, i.e. a representation - to be used in computer systems - of what concepts exist in the world, and how they relate to one another. Ontologies are often viewed as strictly language independent models of the world, especially in the medical informatics community [12], though the need for an ontology in natural language processing applications is generally well accepted [13]. This is not to say that knowledge structuring based on a linguistic approach leads to the same result as when opting for a conceptual approach. A typical example is the ontological distinction between *nominal* and *natural kinds* [14], that in no language is grammaticalised just because the difference is pure definitional [15]. This again does not mean that such distinctions are not useful in a natural language processing applications.

Situated ontologies - i.e. ontologies that are developed for solving particular problems in knowledge based applications [16] - that have to operate in natural language processing applications, are better suited to assist language understanding when the concepts and relationships they are built upon, are linguistically motivated [17]. In the perspective of re-usability, two dimensions have however to be explored: (relative) independence from particular languages and (relative) independence from particular domains. Linguistic semantics based analyses allow us to separate f.i. entities from events and property concepts, a rather crude distinction being the fact that in most languages these concepts are respectively grammaticalised by means of nouns, verbs and adjectives [18]. Linguists are concerned on how these concepts give overt form to language, while from a computational point of view, these concepts also have to be “anchored” in a *linguistic ontology*.

While working on the language engineering aspects of Galen-In-Use, numerous examples were found where linguistic principles were in conflict with conceptual principles [19]. Physicians want to see medical concepts organised in a framework that reflects their clinical way of thinking. As an example, the Galen model categorises the concepts of “filling” and “injecting” as specialisations of a “LiquidInstallingProcess” that itself is a child of “InstallingProcess”. This categorisation is useful from a clinical perspective where from the place in the hierarchy it can be derived that the concepts of injecting and filling have to do with the installation of liquid (though not necessarily exclusively as the Galen model supports multiple parents). This categorisation does however not line up with the linguistic structures that (at least in European languages) are used to express installing, filling and injecting events. From a language understanding perspective, it would be better to categorise these motion events according to the way the thematic roles of *goal* and *theme* may surface in sentences expressing these events.

As can be seen from Figure 1, there can no straightforward relationship be drawn between the two categorisations.

The Galen view	The linguistic semantic view
ResourceManagementProcess	
InstallingProcess	→ To install <theme> [in <goal>]
LiquidInstallingProcess	
Filling	→ To fill <goal> [with <theme>]
Injecting	→ To inject <theme> [in <goal>]
	→ To inject <goal>

Figure 1: Differences in linguistic and conceptual categorisation

Also concerning part-whole relationships, there are differences in categorisation and actual expressions. Clinicians want to have the fingernail classified as part of the upper extremity, following a long chain of transitivity over “distal phalanx”, “finger”, “hand”, “lower arm” and “arm”, while they would never actually say that “a fingernail is a part of the upper extremity”.

5.3 Unifying the cognitive and linguistic dimension: the Cassandra approach

The Cassandra approach is a tagging technique used to project a semantic representation of a phrase onto the phrase itself, without changing the original order of the words. The technique is successfully being used for the processing of clinical terminologies in the Galen- In-Use project [20].

The goals of the Cassandra tagging within the Galen-In-Use project are multiple. First, the tagging makes the relationships between the constituents of the phrases in clinical terminologies explicit. Second, it connects “linguistic” concepts and relationships to the “conceptual” representation of Galen. Third, it projects the conceptual representation on the surface structure of the expressions. And fourth, it allows afterwards to generate automatically lexicons, grammars and even a conceptual-linguistic cross-categorisation scheme on the basis of the tagged corpus. As such, it combines the advantages of the pure conceptual approach (clean categorisation of medical concepts) with more corpus-linguistic oriented approaches [21].

At the heart of the Cassandra tagging technique is a bracketing and encoding convention that relates the surface structure of a sentence to a linguistic representation and a conceptual representation. As an example, the sentences “*excision of cicatrix of skin*” and “*debridement of skin*” are respectively tagged as

- (1) (excision)35 {[of]111 ((cicatrix)2120 {[of]216 (skin)474}0)0}0
- (2) (debridement)82 {[of]142 ({palmar}1785 (skin)474)0}0

where the different types of brackets categorise a sentence constituent as referring to a concept, a link (i.e. conceptually, or a thematic role linguistically), or a criterion (i.e. a link applied to a concept). This notation provides a fairly adequate bridge between the “topic-attribute-value” paradigm adhered to in Galen, and the predicate paradigm on which our linguistic engineering work is based. The figures refer to a semantic lexicon that, restricted to the phrases presented above, can be represented as in Table 4.

Table 4: Semantic lexicon used in the Cassandra tagging technique

RefId	Prototype	Conceptual repr.	Linguistic repr.
35	excision	excising	excising
82	debridement	debriding	debriding
111	of	ACTS_ON	THEME
142	of	ACTS_ON	SOURCE
216	of	HAS_LOCATION	SOURCE
474	skin	skin	skin
1785	palmar	[IS_PART_OF](palm)	[LOCATIVE](palm)
2120	cicatrix	cicatrix	cicatrix

Some additional conversion rules are needed to generate the desired representations from the tagged sentences as not always (not to say seldom) a direct structural correspondence between the two representations is attainable. Clinicians for instance want to express that they “operate on pathologies that are located somewhere in the body”, while they don’t care about motion events and thematic roles at all even if they express it in that way (Figure 2).

Conceptual representation	Linguistic representation
excising	excising
ACTS_ON cicatrix	THEME cicatrix
HAS_LOCATION skin	SOURCE skin

Figure 2: Conceptual and linguistic representations of “excision of cicatrix of skin”.

The Cassandra technique has also some particular features to cope with special phenomena such as “semantic gapping” as occurs in noun concatenation, e.g. the use of the asterisk in:

(division)49 {[of]84 ((joint)129 []217}0 (cartilage)511{[of]217((foot)983 @and#622 (toe)984)0}0}0*

6. Formalising along the communicative dimension

It is often stated that concept systems in health care must be language- and purpose independent, and that they should be formally described in a powerful and expressive formalism on which computationally tractable algorithms can be applied. However, our analysis of the relevant literature in the domains of medical informatics, computational linguistics and philosophy has shown that these requirements cannot be fulfilled at the same time [22]. As explained in the previous section, language - independence cannot completely be achieved as structuring the knowledge domain and building the concept system is a matter of thematic sublanguage analysis and of subcategorisation which itself only can be performed by using the information provided in a given language. In different languages, the same concept may be subcategorised on different criteria or features.

Purpose - independence seems to be the most problematic goal to achieve as orientation towards a purpose is required for (1) identifying what concepts should be represented, (2) deciding on what should be introduced in the concept system as a concept or as a role, (3) eliminating unnecessary complexity of the concept system's structure by avoiding unneeded subcategorisations, and (4) limiting the depth of the terminology in order to avoid the problems associated with the computational intractable property of many formal terminological systems. The interest-relativity of conceptual systems is due to the fact that descriptions tend to have a particular explanatory role. When describing objects, answers to

particular questions are implicitly given. What is accepted as an interesting answer, is usually a context-sensitive matter [23].

The communicative dimension of terminologies is both related with the maintenance of terminologies, and the purpose(s) for which they are designed. As a consequence, problems such as how to guarantee that a (formal) terminology is properly used for what it is designed for, how can it be put in practice, how can it be maintained, and what is needed to allow co-existence with other systems, need to be accounted for. To all these questions, there is one common answer: there must be a general computational framework upon which various terminological tools and applications can be built. Such a framework has been specifically designed for graph- and network operations such that it can be considered to be a database manager for knowledge represented in the form of a semantic network. Within the environment, a network programming language has been developed, together with a compiler for extremely fast code execution. API's can be developed to integrate the system in front-end applications.

An important aspect of the communicative dimension of a specific terminology is its relationship with other terminologies in the same or a related domain, be it possibly developed for different purposes. ICNP being a terminology for nursing procedures, its relationship with systems such as NIC and NANDA has to be considered. Quite often, mapping tables are set up as a means to go from one terminology to another. Ideally however, all systems should be represented formally according to a common framework. This has the advantage that mapping tables are an automatic by-product of such an effort. Moreover, there is no reason why such activities should be restricted to the nursing domain. The final goal should be a formal terminology that can be used within any particular healthcare domain.

7. Towards the beta-version of ICNP

The next version of ICNP is going in the direction of a compositional system. To represent clinical information, elements may be picked from different axes and combined into a specific format. To represent an observation such as "risk for disturbed sleep", three distinct elements must be combined: "risk for", "disturbed", and "sleep". According to the provisions of ICNP-2, this is done as outlined in Table 5.

Table 5: ICNP-2 representation of "risk for disturbance of sleep"

focus :	sleep
judgement:	disturbed
likelihood:	risk for

Compositional systems have the advantage that out of a (limited) set of basic concepts, more complex concepts can be composed, hence increasing the expressive power of the terminology. Compositional systems are also better suited for a formal representation. When additional requirements are put forward on allowed combinations of concepts, they even can become generative in nature such that only valid and sensible concepts can be generated out of the basic building blocks.

Care must however be taken to get the design right from the very beginning. It would be wrong for instance to consider the "focus" as the semantic head of a nursing observation, and the other elements such as likelihood, judgement, severity, body site, topology and others as being merely attributes of the focus. In the example of Table 5, there is in fact a deeper structure. The likelihood in this case is not linked to the focus, but to the judgement. Only a deeper semantic representation can capture these details properly.

8. Conclusion

The development of a nursing terminology - as for any terminology - should follow a three-dimensional approach. In the cognitive dimension, the basic building blocks of the terminology are to be identified and put together in a general framework. In the linguistic dimension, terms are anchored to the concepts identified. These terms act as verbal representations of the concepts. Ideally, these verbal representations should be formed according to the provisions of a controlled language. In the communicative dimension, one has to guarantee that proper mechanisms are set up to use and maintain the terminology, and to link it to other terminologies in the same or related domains.

The three dimensions should not be looked at as being independent from each other. Identifying concepts (cognitive dimension) is best achieved by examining terms and phrases used in the nursing domain (linguistic dimension), in particular in the context for which the terminology is going to be designed for (communicative dimension).

It is highly recommended that formal methodologies and tools are used for each step in the development process. By doing so, it is possible to develop terminological systems that can be used by computers. It is indeed a misperception that terminologies a priori are to be used by humans. At a time when no computers were available, that was indeed the case, simply to guarantee that information was correctly understood. Formal terminological systems, fully specified along the three dimensions, and properly integrated in computerised documentation systems, can take the burden of using terminologies away. As such, nurses can concentrate on their most important task: giving care to patients.

References

- [1] Sager JC. A Practical Course in Terminology. John Benjamins Publishing Company, Amsterdam, 1990.
- [2] Rector AL, Nowlan WA, Glowinski A. Goals for Concept Representation in the GALEN project. In Safran C. (ed). SCAMC 93 Proceedings. New York: McGraw-Hill 1993, 414-418.
- [3] Rector AL, Glowinski A, Nowlan WA, Rossi-Mori A. Medical concept models and medical records: an approach based on GALEN and PEN&PAD. *Journal of the American Medical Informatics Association* 1995, 2: 19-35.
- [4] Rector AL, Nowlan WA, Kay S. Conceptual Knowledge: the core of medical information systems. In Lun KC, Degoulet P, Piemme TE, Rienhoff O (eds.). *MEDINFO 92 Proceedings*. Amsterdam: North - Holland 1992, 1420-1426.
- [5] Rector AL. Compositional models of medical concepts: towards re-usable application independent medical terminologies. In Barahona P & Christensen JP (eds.) *Knowledge and decisions in health telematics*. Amsterdam: IOS Press 1994, 133-142.
- [6] Kittredge R., Lehrberger J. (eds.) : *Sublanguage : studies of language in restricted domains*. de Gruyter, Berlin, 1982.
- [7] Harris Z. *Mathematical Structure of Language*. John Wiley & Sons. New-York, 1968.
- [8] Harris Z. *A theory of language and information*. Clarendon Press, Oxford, 1991.
- [9] Ceusters W, Spyns P, De Moor G, Martin W (eds.) : *Syntactic-semantic tagging of medical texts: the MultiTALE-project*. IOS Press, Amsterdam, 1997.
- [10] Ceusters W, Steurs F, Zanstra P, Van der Haring E, Rogers J. *From a time standard for medical informatics to a controlled language for health*. *International Journal of Medical Informatics* 1998, 48: 85-101.
- [11] Côté R.A., Rothwell D. (eds.), *Systematized Nomenclature of Medicine - SNOMED International*, College of American Pathologists, Chicago, 1993
- [12] Rector AL, Rogers JE, Pole P. The GALEN High Level Ontology. In Brender J, Christensen JP, Scherrer J-R, McNair P (eds.) *MIE 96 Proceedings*. Amsterdam: IOS Press 1996, 174-178.

Ceusters W. Harmonisation and formalisation of nursing terminology: a three-dimensional approach. In RA Mortensen (ed.) *ICNP and Telematic Applications for Nurses in Europe*, IOS PRESS Amsterdam, 1999;:164-173.

- [13] Bateman JA. Ontology construction and natural language. In *Proc. International Workshop on Formal Ontology*. Padua, Italy, 1993, 83-93.
- [14] Kripke S. Naming and Necessity. In Davidson D & Harman G (eds.) *Semantics of natural language*. Dordrecht: Reidel, 1972, 253-355.
- [15] Welsh C. *On the non-existence of natural kind terms as a linguistically relevant category*. Paper presented at the Linguistic Society of America, New Orleans, LA, 1988.
- [16] Mahesh K & Nirenburg S. A situated ontology for practical NLP. In *Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing, IJCAI-95*. Montreal, Canada, 1995.
- [17] Deville G, Ceusters W. A multi-dimensional view on natural language modelling in medicine: identifying key-features for successful applications. Supplementary paper in *Proceedings of the Third International Working Conference of IMIA WG6*, Geneva, 1994.
- [18] Frawley W. *Linguistic Semantics*. Hillsdale, Hove and London: Lawrence Erlbaum Associates, 1992.
- [19] Ceusters W. *Language Engineering as an Enabling Technology for Clinical Terminology Harmonisation*. In: CEC-DGXIII (ed.) *Important Issues in Today's Telematics Research, TAP'98 Conference Barcelona*, 1998, 168-173.
- [20] Ceusters W, Waagmeester A, De Moor G. *Syntactic-semantic tagging conventions for a medical treebank: the CASSANDRA approach*. In van der Lei J, Beckers WPA, Ceusters W, van Overbeeke JJ (eds.): *Proceedings MIC'97, Veldhoven, The Netherlands*, 1997, 183-193.
- [21] Marcus M, Santorini B, Marcinkiewicz MA. "Building a large annotated corpus of English: the Penn Treebank." *Computational Linguistics*, 1993, 19: 27-45.
- [22] Ceusters W, Deville G, Buekens Ph. *The Chimera of Purpose- and Language Independent Concept Systems in Health Care*. In Barahona P, Veloso M, Bryant J (eds.) *Proceedings of the XIIth International Congress of EFMI*, 1994, 208-212.
- [23] Buekens F, Ceusters W, De Moor G. The Explanatory Role of Events in Causal and Temporal Reasoning in Medicine, *Met Inform Med* 1993, 32: 274 - 278.