Transgression of semantic boundaries by methodical terminology management

Application to the terminology and metrology of satellite based localisation systems

Lars Schnieder and **Marco Wegener**
Technische Universität Braunschweig
Institute for Traffic Safety and Automation Engineering
Braunschweig, Germany
{l.schnieder,Wegener}@iva.ing.tu-bs.de

Abstract— **In the future satellite based positioning systems will be applied to a broad application area ranging from civil to military use. This paper is focusing on the civil transportation sector and shall be a contribution to the consistent and domain-spanning use of satellite based localisatin systems with possible safety-related applications. Applying satellite based localization systems to surface transportation requires the bridging of the currently existing gap between the different terminological worlds. This includes a binding definition of the properties, characteristics and values further describing the meaning of the applied terms. On the one hand there is the terminology of aviation which formed the base for the specification of the satellite-based localization system, on the other hand there is the terminology of the application domain (here rail transportation respectively the concerned transportation domain). Especially for safety-related applications it is mandatory to ensure a clear understanding between the terminological worlds which shall also enable a certification process for receivers of satellite-based localization systems. This paper outlines an approach based on the linguistic concept of lexemes as well as a structured procedure in order to reach consensus on the terminology being applied in an increasingly interdisciplinary context.**

Keywords: specification, terminology, terminology management

I. INTRODUCTION

Interdisciplinary projects call for a cooperation of experts from different scientific disciplines and are in consequence often characterized by a Babylonian confusion. Communication is often hampered by the following aspects:

Ambiguity: often standards introduce one term which refers to a variety of different concepts as mental representations (so called homonymy or polysemy).

Inconsistency: often different standards (sometimes even within one and the same standard) contain definitions which state the opposite. Those opposites can not coexist at the same time.

Semantic vagueness: A concept's extension is not clear. This means that it can not be certainly stated if an object clearly belongs to a group of objects denoted with a specific term.

Context dependence: often terms are coined for one specific domain. Problems arise as soon as communication across the borders of one specific technical terminology is required.

A project that deals with interdisciplinary communication requires an interdisciplinary approach. In order to meet the currently existing problems stated above a new approach is required. For this reason engineers, computer scientists and linguists have designed an approach using an epistemological base, which is expected to accomplish more than any other terminology management system so far. The currently existing terminological confusion can be overcome if a methodical terminology management is introduced which comprises the following three elements:

Holistic Semiotic foundation: The methodical framework to overcome the existing problems is founded on a solid semiotic basis. Based on the novel trilateral model of a linguistic sign both the relationships between terms within terminologies and the clear separation of domain-specific language use can be achieved.

Terminology Management System (iglos): Currently existing terminology management systems are not suitable to master the challenges in conjunction with an interdisciplinary terminology. For this reason the semiotic model of linguistic sign is incorporated into the novel terminology management system iglos (acronym for intelligent glossary).

Terminology Engineering Process (tep): This process is a novel approach towards terminology management. It is the idea of this web-based approach to allow for a distributed and collaborative elaboration and approval of a commonly agreed-on terminology. This approach is based on the principles of consensus, openness and transparency known from international standardization procedures. The terminology management system iglos allows all users to enter new terms, change existing terms or delete deprecated terms during the Terminology Engineering Process.

II. METHODICAL TERMINOLOGY MANAGEMENT

The idea for the igloos developed from a close cooperation of the Institute for Traffic Safety and Automation Engineering and the German Linguistics Department of Technische Universität Braunschweig, Germany. The aim of this cooperation lies not only in developing an ordinary glossary but creating the possibility of modeling a highly interconnected terminological structure that helps to

overcome the currently existing linguistic barriers between separate disciplines and enables communication where before only incomprehension, misunderstanding and rejection existed. It is common linguistic knowledge that language cannot be perceived of as a set of words and rules. Accordingly, iglos is not only a dictionary: iglos models the language structure instead. Thus, scientific and technical terms are captured in their systematic contexts – with all their connections and dependencies on other language units. The key element of iglos are introduced in this section.

A. Holistic Semiotic Foundation

The basis of the iglos terminology management system is a new, trilateral sign model which for the first time understands the linguistic sign to be constituted by its special language context, and consistently models it according to this new understanding. Linguistic problem cases – like overlapping and imprecise terminologies, contradicting definitions and different application areas, synonymy as well as polysemy or homonymy etc. – are thus for the first time modeled precisely. Communication errors that previously showed only in the results can now be detected at an early stage. The trilateral sign model combines the following constituents:

The *designation* (signifier): A concept's (see figure 1) representation using linguistic (e.g. appellation) or other means (symbols, formulas). According to lexicography and lexicology the signifier (Saussure, 2006) or designation (ISO, 2000a) is also called lemma. (Lutzeier, 2003; Wolski, 1989).

The actual *'concept'* (signified): According to (ISO, 2000a), terms are defined as "unit of knowledge formed by an amount of objects and determining their shared characteristics using abstraction". Terms are used for identifying objects, gaining a common understanding about objects, as well as organizing objects mentally (Arntz et al., 2004). In the trilateral model of a linguistic sign, the meaning of a term is represented by its definition (Bessé, 1997).

A *variety* is used for classifying concept and designation into a context of use (Dubuc et al., 1997), the respective language for special purposes. The variety is considered an essential constituent of the metalinguistic model of a linguistic sign.

Additionally, the actual concept (signified) can be described using its extension, intension and relations to other terms according to (ISO, 2000a). Those constituents of a term are usually put into words by definitions. Definitions require a proper knowledge of a concept's extension as well as its intension and relations.

Often definitions do not sufficiently explain the meaning of a terminological entry. Therefore the intension needs to be made more precise. In order to do this a concept's intension can be broken down into a hierarchical structure of attributes, the so called attribute hierarchy. This is done in accordance with Carnap's steps of concept formation (Carnap, 1966):

Properties relate to the perceivable qualities of an object. Following Carnap's logical empiricism everything knowable relates to sensory experience.

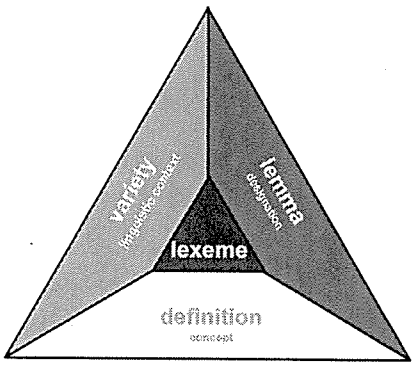

Figure 1. Trilateral model of a linguistic sign

A *characteristic* is an abstraction of a property of an object. A characteristic is seen as quantifiable which means it can be counted (in case of a topological measurement scale) or measured (in case of a cardinal measurement scale). Measurement is the assignment of numbers to objects or events in a systematic fashion. There is a relationship between the level of measurement (nominal, ordinal, interval, ratio or absolute) and the appropriateness of various statistical procedures.

A *physical quantity* is a characteristic, which can be quantified. This means it can be measured, calculated and expressed in numbers. According to Carnap this quantitative method of description has essential advantages over non-quantitative or purely qualitative methods. Firstly, it permits a more exact description of the separate facts. And secondly, it makes the elaboration of decidedly more effective general laws expressing connections between the values of various quantitative concepts with the help of mathematical functions possible.

A *value* is the output of a measurement. This quantity that is being determined by a measurement is also referred to as measurand.

B. Terminology Management System (iglos)

The aim is to develop a software application that can serve as a base for the shaping of different terminologies in various scientific disciplines and enterprises. Special consideration has been given to the relations between linguistic signs – e.g. synonymy, homonymy, overlap between different special languages, similarity of meaning, several possible translations etc. By intensive examination of different semiotic theories it was possible to develop a new sign model as well as a model of sign relations. This sign model has continuously been enhanced by taking practical problems as well as those of existing systems into account and is now the basis for the ongoing software implementation. Since the project consists of an interdisciplinary dialogue of linguists, terminologists, computer scientists, engineers, translators and users, it continuously gains complexity and new perspectives. The key functions of the iglos are as follows:

Relations: In contrast to currently existing terminology management systems the user can flexibly introduce the types

of relations required to adequately model the terminology of his domain. By means of those generic relation types specific relations can be introduced which means that the meaning of a terminological entry in the database goes far beyond the information given in its definition. Furthermore the meaning of a terminological entry can be derived from its position in the systemic context of the domain's terminology.

Versioning: All elements of the terminology – both the terminological entries with all their constituents (lemma, definition, variety) and the relations existing between them – will be versioned in the iglos terminology management system. Doing this the validity of a terminological entry for a given text can be reconstructed. This will become of utmost importance when a terminology is referenced in a contract or another normative document.

User glossaries: In projects the agreement on a common and shared terminology is done at the start of interdisciplinary collaboration. iglos supports this by allowing the user to define user specific or project specific terminologies. Doing this specific perspectives of the overall terminological base available in the iglos terminology management system are generated.

Web-Interface: By means of a web-interface the collaborative and distributed work on a common and shared terminological basis becomes possible. This ensures, that the user always can access the latest version of the previously agreed-on terminology. Based on the role of the user in the terminology management process (see next section) a role-based access to the iglos terminology management system is granted.

Visualization: The terminological entries modeled in iglos with their semantic relations can be visualized as a semantic net. This visualization of this network of semantic relations allows an intuitive navigation within the terminology base incorporated into the iglos terminology management system (see figure 2).

C. *Terminology Engineering Process (tep)*

Consistently maintaining, sharing and applying a terminology within the organization (and beyond) calls for a coherent methodical framework (Hurst, 2009). Because of the omnipresent language change this is always an iterative process. The elaboration of a shared terminology is performed in the following steps:

Term excerption: Existing terms and definitions are collected from different sources. On the one hand side this can be new term proposals which have been submitted by users through a web-client. On the other hand side existing texts can be analyzed using term extraction tools. The terms that are used most frequently will be a first starting point for the subsequent steps of the terminology management process (Wright, 1997; Zerfaß, 2008a).

Term validation: The critical examination and editing of the extracted terms will be done by terminologists (Haberland, 2008). The terminologists need to carefully evaluate the results of the term extraction tools – the terms used most frequently may not be the most important ones (Zerfaß, 2008b).

Assignment of a variety to a lexeme: The list of terms which is now available is not yet a complete lexeme with respect to the trilateral model of a linguistic sign introduced in the previous section. The list of lemmas available after the first steps needs to be amended with the varieties for which they are valid. Varieties can be further differentiated into sub-domains until a sufficient level of detail is reached. After a first critical review of the list of lemmas a first list of relevant varieties for the project can be compiled. Often a 1:1 correspondence of a lemma to a variety can be achieved as is the case with the lemma "control element" which is valid in the variety "control technology". In other cases one lemma may exist in different varieties as ist he case with the lemma "virus" which is both valid in the variety "microbiology" and "computer science".

Assignment of a definition to a lexeme: After the lemma has been assigned to a variety the next step can be taken. At this stage definition can be assigned to lexemes. Definitions can be derived from relevant glossaries and domain-specific standards. The definitions will be assigned to the already existing pairs of lemmas and varieties. It can happen, that several potentially identical definitions can be found for one pair. In this case different distinct lexemes are modeled..

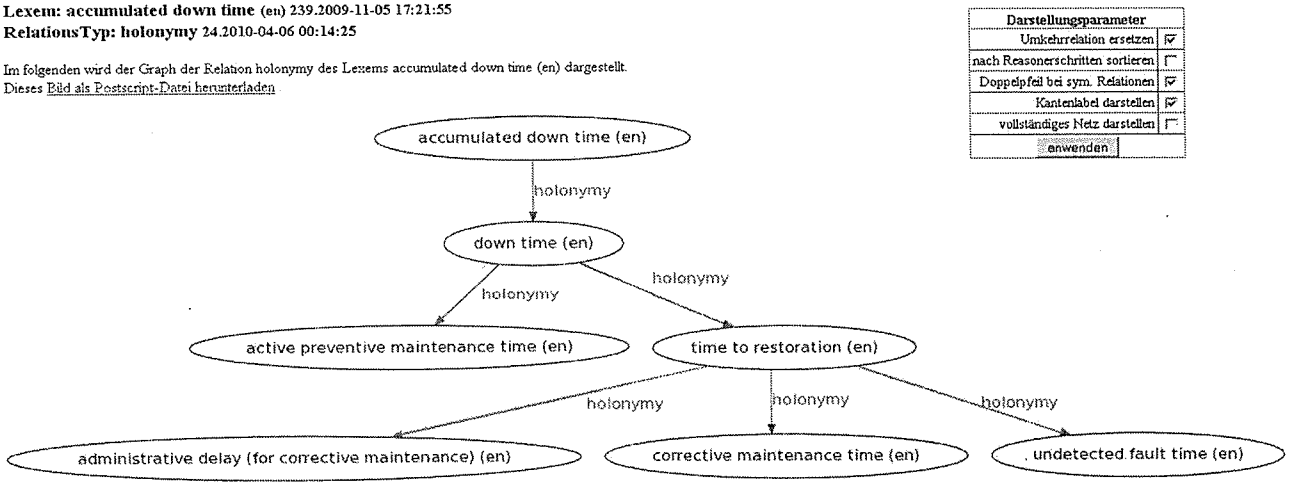

Figure 2. Visualization of semantic relations in a network structure

Project-specific monosemation: The lexemes available in the iglos terminology management system are only descriptive and depict the diversity of language use within one domain. Up to now no lexemes have been chosen to represent the preferred language use within the project. The descriptive view on language may still be contradictory and ambiguous (because of hononyms or polysemes). For this reason fort he project the available terminological base shall be transferred into a prescriptive mode. This process is referred to as monosemation. Lexemes with identical lemmas are referred to as "lexeme groups". The goal of monosemation is to chose one (!) valid lexeme out of each lexeme group. The result of this process step is a monosemated project-specific glossary.

Assignment of terminological relations: In a next step the isolated (trilateral) lexemes are related to each other by means of terminological relations. This allows to model a domain's terminology as a language system. The meaning of a term becomes more clear when it is embedded into the overall structure of a domain's terminology.

Terminology agreement: The draft terminology (lexemes as well as relations existing between them) prepared by the terminologist will be approved by an interdisciplinary board which is composed of all stakeholders (e.g. product managers, technical writers, marketing, ...). The process of terminology agreement can be supported by a web-application which informs all stakeholders of changes in the terminology and requests them to cast their ballot for term acceptance or deprecation. Possible review comments are incorporated into the terminology until the terminological base is finally approved and released for further use.

Terminology dissemination: The terminology will be made available in all required languages to all users (possibly in different data formats). It is now available for quality assurance in the process of document preparation and translation (Haberland, 2008)..

III. Summary and Outlook

The methodical approach towards terminology management is currently being tested intensely within the scope of two research projects in the field of satellite navigation. In order to do this currently existing terminologies are integrated into the iglos, are cross-linked and validated by domain experts. The result is accessible to all persons involved in the project via a web interface and can be continuously commented, amended and improved. The feedback from industrial practice is thus used to improve both the iglos terminology management system in conjunction with the Terminology Engineering Process (tep) as well as the interdisciplinary terminology incorporated in it. Based on a commonly agreed-on terminological basis it will be possible to master the challenges of successful specification, implementation and certification of satellite-based navigation systems in an interdisciplinary context.

In addition to the current application of the iglos terminology management system and the Terminology Engineering Process in research projects for satellite-based localization systems the approach is currently discussed with the terminology office DIN-TERMKONZEPT of the German standardization body DIN (Deutsches Institut für Normung e.V.). Standardization products and services are to be flawless and of high quality. The processes and instruments used by standards bodies should be analysed and optimized and the organization should make quality assurance a key feature of their business policies and implement it consistently. These efforts towards improved quality of terminology standards go hand in hand with ensuring the participation of "everyone". It is the aim to simplify participation in terminology standardization through appropriate instruments to ensure continued access for all without discrimination. The Terminology Engineering Process might be such an open and transparent approach within standards committees dedicated to terminology standardization which allows for a fair and balanced participation of all stakeholders in standardization. Using a web-based platform, as stipulated with the terminology management system iglos access to the results of terminology standardization can be facilitated significantly.

References

Arntz, R.., Picht, H., and Mayer, F. (2004). *Einführung in die terminologiearbeit*. Hildesheim: Georg Olms Verlag.

Bessé, B.(1997). Terminological definitions. In: Wright, S.E.; Budin, G. (eds): *Handbook of terminology management – volume 1: basic aspects of terminology management*: pp. 63-74. Amsterdam: John Benjamins Publishing.

Carnap, R. (1966): *Physikalische begriffsbildung*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Dubuc R.; Lauriston, A. (1997). Terms and contexts. In: Budin, G.; Wright, S.E. (eds.): *Handbook of terminology management. Volume 1*. Amsterdam: John Benjamins Publishing Company.

Haberland, K. (2008). Unternehmensterminologie bei Siemens Energy Sector. *eDITion* 4(2), 8-9, 2008.

Hurst, S. (2009). Wake up to terminology management. In: Communicator – the quarterly journal of the ISTC; Spring 2009, pp. 14-16.

Lutzeier, P.R. (2007). *Lexikologie*. Tübingen: Stauffenberg.

Saussure, F. (2006). *Grundfragen der allgemeinen sprachwissenschaft*. Berlin: de Gruyter.

Wolski, W. (1989). Das Lemma und die verschiedenen Lemmatypen. In: Hausmann, F.J.; Reichmann, O.; Wiegand, H.E.; Zgusta (eds.): *Handbücher zur sprach- und kommunikationswissenschaft, band 5: ein internationales handbuch zur lexikographie, teilband 1*,.Berlin: de Gruyter.

Wright, S.E. (1997). Term selection. The initial phase of terminology management. In: Budin, G.; Wright, S.E. (eds.): *Handbook of terminology management. Volume 1*. Amsterdam: John Benjamins Publishing Company.

Zerfass, A. (2008a): Tools in der Terminologiearbeit. In: *Tagungsband tekom-Jahrestagung* 2008, pp. 287-289. Wiesbaden.

Zerfass, A. (2008b). Terminology Extraction. In: *Tagungsband tekom-Jahrestagung* 2008, pp. 303-304. Wiesbaden.