*Rainer Osswald & Hermann Helbig*

# Derivational Semantics in HaGenLex – An Interim Report

## 1 Introduction

Natural Language Processing (NLP) becomes more and more important for all sorts of intelligent processing of textual information, be it text retrieval, question answering, or natural language interfaces to databases. A central component of an NLP system coping with such tasks is a large lexical database. Moreover, any approach going beyond mere statistical association and pattern matching calls for a lexicon that provides not only the morphosyntactic characterization of words but also information about their meaning, i.e. their semantics.

The focus of this article is on the lexical semantics underlying the derivational relations between lexical entries and in particular on the semantic relation between verbs and deverbal nouns. In question answering tasks, for instance, derivational semantics is indispensable for drawing appropriate inferences. Suppose a document contains sentence (1a).

(1) a. The phonograph was invented by Thomas Alva Edison in 1877.

b. Who is the inventor of the phonograph?

In order to answer a question like (1b) the system needs to know that an inventor is somebody who invents things.

### 1.1 Morphosemantic Links in WordNet

The importance of derivational information in the lexicon has been recently addressed by Fellbaum and Miller (2003), the designers of the lexical-semantic database WordNet. Fellbaum and Miller point out that representing derivation as a purely morphological relation, i.e., as a relation between *words*, is not enough for NLP tasks in general. They illustrate this point by the verb *digest* and its derivative noun *digestion*. Although the core sense of *digest* is physiological, there is also a psychological sense (of absorbing or assimilating mentally), which is represented in WordNet by

a separate entry. The noun *digestion* has two corresponding senses (cf. *op cit*, p. 72):

(2) a. He saw a doctor about his digestion.

b. His appetite for facts was better than his digestion.

Fellbaum and Miller conclude that in order to be useful for NLP applications, WordNet must represent both senses of the derivative noun and, in addition, each of the two senses must be linked to the corresponding sense of the base verb.

WordNet (version 2.0) provides only one type of morphosemantic link.[1] This restriction gives rise to the following kind of problems: The only way to find out that *inventor* refers to the agent of *invention* is to use the fact that *inventor* is positioned below *person* in the hyperonym hierarchy of WordNet. But this only works for verbs whose arguments have different selectional restrictions. In symmetric cases like *examine* there is no semantic distinction between the agent and the undergoer – both participants are persons (or institutions). The hyperonym hierarchy for *examiner* and *examinee* in WordNet, which has the form sketched in (3), does not provide any formal clues as to which word refers to the agent and which to the undergoer of an examination.

(3) a. examiner, tester, …⇒ inquirer, questioner, …⇒ speaker, talker, … ⇒ articulator ⇒ communicator

b. examinee, testee ⇒ respondent, responder, answerer ⇒ communicator

Obviously there is no upper level in WordNet for differentiating between the agent role and the undergoer role. This is no surprise since a taxonomic hierarchy is not well suited for representing information of this type. Typically enough, the hyperonym synset of {*examinee*, *testee*} shown in (3b) consists solely of nouns referring to the agent of an action.

### 1.2 Desiderata for Language Understanding Systems

We can agree with Fellbaum and Miller that in order to be useful for an NLP system, derivational relations in the lexicon should be sensitive to sense distinctions. If language understanding is the issue, we regard it as

[1] According to Fellbaum and Miller (2003), WordNet 2.0 contains 21,000 morphosemantic relations. GermaNet (version 4.0), the German pendant of WordNet (cf. Kunze and Wagner 2001), contains about 1,500 derivational links.

a further requirement that the specific semantic relation between the base entry and the derived entry is available in one form or another for inference tasks. In other words, derivational semantics should be expressed within a formal semantic framework. Moreover, this formalism should be fully integrated in a general knowledge representation framework; for the process of language understanding draws on lexical knowledge as well as on world knowledge in general.

The shortcomings of the morphosemantic information in WordNet observed in Section 1.1 are mainly due to the limitations of the chosen representational means, which essentially consist of traditional sense relations like synonymy, hyponymy, etc. An approach worth mentioning here is the OntoWordNet project (Gangemi et al. 2003), whose objective is to axiomatize all types of conceptual information given in WordNet by semi-automatic methods.[2] In particular, this includes a formalization of WordNet *glosses*, which are textual definitions associated with word senses. Consider the WordNet 2.0 glosses for *examiner* and *examinee* presented in (4):

(4) a. examiner, tester: someone who administers a test to determine your qualifications

b. examinee, testee: someone who is tested (as by an intelligence test or an academic examination)

The gloss given in (4b) closely resembles the information that an examinee is the undergoer of an examination. Its logical form would thus provide a convenient formal representation of the derivational semantics of *examinee* and *testee*. However, WordNet glosses vary considerably in form and quality. To arrive at a corresponding logical explication of gloss (4a) in order to capture the derivational semantics of *examiner* and *tester* seems to be still possible but is surely more involved. All in all, it remains to be seen whether the OntoWordNet program is successful on a large scale. Another question is whether the employed logical formalism, which is closely related to so-called *description logics*, is capable of representing all facets of semantics necessary for natural language understanding.

### 1.3 Overview

In this article, we present an approach to derivational semantics that makes use of the so-called MultiNet paradigm, which is a general-

[2] A project with similar goals is described in Harabagiu et al. 1999.

purpose framework for natural language semantics and knowledge representation (Helbig 2001, Helbig and Gnörlich 2002). Specifically, we are concerned with the representation of derivational information in the computer lexicon HaGenLex, whose semantic component is based on the MultiNet formalism. HaGenLex is a core part of an NLP system that allows to transform natural language expressions into MultiNet representations (see e.g. Hartrumpf 2003). The HaGenLex-MultiNet system has already been successfully employed for building natural language interfaces to various types of databases (Helbig et al. 2000, Leveling and Helbig 2002) and is currently applied to text retrieval (Leveling 2004) and question answering.

The rest of this article is organized as follows: In Section 2, we briefly describe the structure of HaGenLex entries and introduce the lexically relevant features of MultiNet. Section 3 reviews various semantic classifications of deverbal nouns found in the literature and identifies some of their problems and deficiencies. Section 4 lists several insights of traditional grammar concerning the realization of arguments of deverbal nouns as postnominal attributes. In Section 5, we describe the current representation of deverbal nouns in HaGenLex, which takes the observations of Sections 3 and 4 into account.

With respect to maintaining and extending HaGenLex, we are interested in automatizing the generation of HaGenLex entries for deverbal nouns as far as possible. The remaining two sections focus on two problems arising in this context: the prediction of result nominalizations and the possible interpretations of the postnominal genitive attribute. In Section 6, we introduce an approach by Ehrich and Rapp that tries to explain this variation in the behavior of deverbal nouns in terms of the lexical semantics of the base verb. The final Section 7 investigates how the semantic descriptions proposed by Ehrich and Rapp can be represented within the HaGenLex-MultiNet framework, thereby improving the semantic representation in HaGenLex in general and the treatment of deverbal nouns in particular.

## 2 HaGenLex

HaGenLex (**Ha**gen **Ge**rma**n** **Lex**icon) is a domain independent lexical database for German, which has been developed since 1996 at the Department of Computer Science of the FernUniversität in Hagen. The lexicon currently comprises about 22,500 entries (11,250 nouns, 6,700 verbs,

3,200 adjectives, and 600 adverbs). The lexical material of HaGenLex has been compiled on the basis of frequency lists and publicly available dictionaries. The lexicographic work is supported by a workbench that provides an easy-to-use graphical user interface. See Hartrumpf et al. 2003 for more information on the technological environment of HaGenLex.

### 2.1 The MultiNet Formalism

The semantic information in HaGenLex is represented by means of the MultiNet formalism. MultiNet, short for *Multilayered Extended Semantic Networks*, is a domain independent semantic interlingua with broad coverage concerning the semantic phenomena of natural language. "Meaning" or "knowledge" is represented in form of semantic networks, which are labeled directed (hyper)graphs whose nodes represent concepts and whose edges represent semantic relations between concepts.

Figure 1 depicts a simple example of such a graph, which represents the semantics of the following sentence in terms of MultiNet:[3]

(5) Der Prüfling beantwortete jede Frage des Professors.
(The examinee answered every question of the professor.)

The example illustrates various core features of the MultiNet formalism: There is a distinction between generic concepts (*answer*, *examinee*, etc) and instantiated concepts ($c_1$, $c_2$ , etc). The instantiated concept $c_2$, for example, corresponds to the expression *der Prüfling*. Within MultiNet, the relation between instances and their generic concepts is referred to as *subordination*, in symbols, SUB (for conceptual objects) and SUBS (for situations). Moreover, the relations between a situation and its participants are represented by elements of a predefined set of semantic roles, here AGT (agent) and OBJ (neutral object). In addition, nodes are labeled with a symbol indicating the *ontological sort* of the represented concept. The concept $c_2$, for instance, is of the sort *d*, which is short for *discrete object*; see Appendix A.1 for all 45 ontological sorts used in MultiNet. Finally, concepts are specified with respect to several so-called *layer attributes* whose values indicate, e.g., the facticity of a concept (FACT), its determination of reference (REFER), and its quantificational content (QUANT). In the given Multi-

[3] The software environment of MultiNet includes a graphical editor which allows to create and modify MultiNet representations interactively as directed graphs; cf. Gnörlich 2002.

Net representation of sentence (5), the determiners and quantifiers are reflected by the values of the attributes REFER and QUANT.

For those who prefer some kind of quasi-logical form, the MultiNet representation shown in Figure 1 could be reformulated as follows:

(6) SUBS($c_1$, *answer*) & TEMP($c_1$, **past**) & SORT($c_1$) = *da* & FACT($c_1$) = real
& AGT($c_1$, $c_2$) & SUB($c_2$, *examinee*) & SORT($c_2$) = *d* & REFER($c_2$) = det
& OBJ($c_1$, $c_3$) & SUB($c_3$, *question*) & SORT($c_3$) = *io* & QUANT($c_3$) = every
& ATTCH($c_4$, $c_3$) & SUB($c_4$, *professor*) & SORT($c_4$) = *d* & REFER($c_4$) = det

From the perspective of standard predicate logic, a subordination term like SUB($c_4$, *professor*) corresponds to a predication, where the predicate *professor* is itself regarded as a concept – a technique that is known as *reification*. It should be noticed that (6) is to a certain degree related to what is sometimes called a "Neo-Davidsonian" approach to natural language semantics: The concept corresponding to the verb of the sentence is predicated of the described event, whereas the participants of the event, expressed by the grammatical subject and object, are characterized by their semantic roles (see e.g. Parsons 1990).

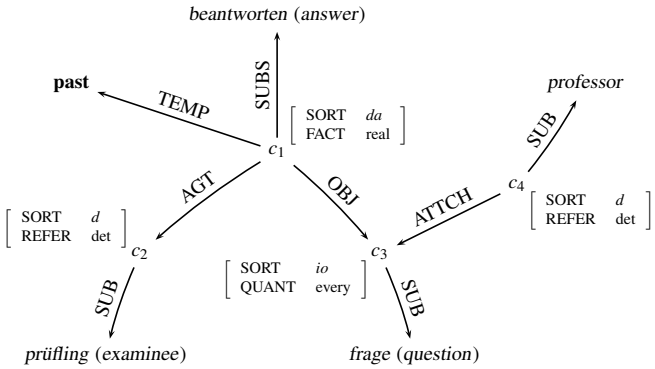



*Figure 1: Semantic representation for sentence (5)*

MultiNet predefines about 140 semantic relations and functions. For each of them there is a detailed explication in terms of a verbal explanation, typical paraphrases, formal axioms, and various comments and examples. Appendix A.2 contains a list of relations which are relevant to the subject matter of this article, together with brief descriptions and signature specifications. For a comprehensive exposition of the MultiNet paradigm, the reader is referred to Helbig 2001.

## 2.2 Lexical Structure and Content

Within the HaGenLex-MultiNet approach, each sense of a content word, i.e., each lexeme, uniquely corresponds to a *lexicalized concept*. So, each HaGenLex entry, when employed for the semantic analysis of a given natural language expression, gives rise to a concept node of the resulting MultiNet representation. In Figure 1, four of the nodes correspond to lexicalized concepts, viz. *beantworten*, *prüfling*, *frage*, and *professor*.

For reasons of simplicity we do not distinguish between different senses in this example (and at many other places in this article). It should be emphasized, however, that HaGenLex makes use of a system of two-place numerical indices for sense identification that allows a systematic distinction between homographs and polysemes.

Each HaGenLex entry is semantically classified by the ontological sort of the respective concept. For instance, the lexeme *professor* is characterized by the sort *d* (*discrete object*). An additional set of binary semantic features allows a more finely grained semantic classification, whose main purpose is to define *selectional restrictions* and thereby to effectively support the process of *disambiguation* during syntactico-semantic analysis. Examples of such features are HUMAN, ARTIFACT, and MOVABLE; see Helbig 2001 or Helbig et al. 2003 for a full list.

In HaGenLex, the *valency* or *case frame* of a lexeme is characterized both syntactically and semantically. In particular, each verb entry specifies the semantic roles corresponding to the complements of the verb. These roles appear in the MultiNet representation of a sentence containing the respective verb. The relations AGT and OBJ shown in Figure 1 are thus provided by the lexical entry of the verb, whose case frame is sketched in (7).

| (7) *beantworten* | AGT | ORNT | OBJ |
|---|---|---|---|
| | [HUMAN +] | [HUMAN +] | [INFO +] |
| | NP[nom] | NP[dat] | NP[acc] |
| | | optional | |

There is a set of standard sense relations defined in MultiNet, including SYNO (*synonymy*) and ANTO (*antonymy*). *Hyperonymy* corresponds to the subordination relation SUB (or SUBS) on the level of generic concepts. Furthermore, a systematic linking between HaGenLex entries and word senses of the lexical database GermaNet has been established, in order to project GermaNet sense relations onto HaGenLex (Osswald 2004).

The internal representation of HaGenLex entries as typed feature structures is described in Helbig et al. 2003, where also the inheritance-based architecture of the lexicon is explained and supplementing resources and maintenance tools are presented.

| | |
|---|---|
| *nomen actionis* | refers to the action or process described by the base verb; examples: *Prüfung* (*examination*), *Besuch* (*visit*), *Explosion* (*explosion*), *Besäufnis* (*booze-up*) |
| *nomen agentis* | refers to the performer of an action; examples: *Tänzer* (*dancer*), *Prüfer* (*examiner*), *Erfinder* (*inventor*), *Träumer* (*dreamer*), *Ankömmling* (*arrival*) |
| *nomen patientis* | refers to the undergoer of an action; examples: *Prüfling* (*examinee*), *Entdeckung* (*discovery*) |
| *nomen acti* | refers to the result of an action; examples: *Schluchzer* (*sob*), *Übersetzung* (*translation*), *Fälschung* (*forgery*), *Verblüffung* (*amazement*) |
| *nomen instrumenti* | refers to the instrument used for an action; examples: *Öffner* (*opener*), *Rasierer* (*razor*, *shaver*), *Feile* (*file*) |
| *nomen loci* | refers to the place where an action or event happens; examples: *Bäckerei* (*bakery*), *Ausgabe* (*issue desk*) |

*Table 1: Traditional onomasiological classification of deverbal nouns*

## 3 Morphosemantic Classification of Deverbal Nouns

Concerning the treatment of derived nouns in the lexicon, the following two questions are of special interest: First, what is the *semantic representation* of derived nouns, especially with respect to their base entry? Second, to what extent needs this information to be explicitly stored in the lexicon and which parts can be automatically reconstructed by lexical rules from the base entry?

### 3.1 Traditional Morphosemantic Classification

In order to classify the possible meanings of deverbal nouns, the literature on word formation (e.g. Erben 1993, Fleischer and Barz 1995) typically employs the onomasiological categories listed in Table 1. The two classes at the bottom of Table 1, *nomen instrumenti* and *nomen loci*, are more or less unproblematic. A simple test pattern for *nomina instrumenti* can be constructed in German by an instrumental *mit*-clause (or a *with*-clause in English):

(8) a. Eine Feile ist etwas, mit dem man etwas feilt.
(A file is something with which you file things.)

b. Eine Abfindung ist etwas, mit dem man jemanden abfindet.
(A compensation is something with which you compensate people.)

Locative clauses give rise to a similar test pattern for *nomina loci*:

(9) a. Eine Bäckerei ist ein Ort, an dem etwas gebacken wird.
(A bakery is a place where you bake things).

b. Eine Ausgabe ist ein Ort, an dem etwas ausgegeben wird.
(An issue desk is a place where you issue things).

The categories *nomen actionis* and *nomen agentis* are more problematic since they are usually applied not only to actions but also to other events and processes. Examples in Table 1 are *Explosion* (for *nomina actionis*) and *Ankömmling* (for *nomina agentis*). We can eliminate this problem by extending the notion of *nomen actionis* to general events and processes. But

notice that reference to states is also possible, exemplified by *Besitz* (*possession*) and *Misstrauen* (*mistrust*, *distrust*):[4]

(10) a. Der Besitz dieser Waffe ist strafbar.
(The possession of this weapon is punishable.)

b. Das Misstrauen gegen die Regierung besteht weiterhin.
(The distrust of the government continues to exist.)

A similar observation holds for the *nomen agentis* category. For example, the deverbal nouns *Besitzer* (*owner*) and *Vorsteher* (*chairman*) refer to the subjects of the stative verbs *besitzen* (*own*) and *vorstehen* (*chair, preside*). It is thus not appropriate to speak of *agent nominalizations* in these cases. According to Eisenberg (1998, p. 264), it is more reasonable to understand *nomina agentis* as derived nouns referring to agentive entities, i.e., persons, institutions, animals, etc. In general, however, we seem to be forced to apply the *nomen agentis* category to any deverbal noun that refers to the subject of the base verb.

The most problematic categories are that of *nomina acti* and *nomina patientis*. Different conceptions can be found in the literature. Eisenberg (*op cit*, p. 265), for instance, only mentions *nomina acti* formed by *er*-affixation, which mostly describe the result of actions that are movements or sound utterances, witness (11).

(11) Dreher (turn), Huster (cough), Hopser (leap), Schluchzer (sob), Stupser (nudge)

For other authors, the category of *nomina acti* has a much broader scope. Fleischer and Barz (1995, p. 174) do not only count deverbal nouns referring to an effected object like *Fälschung* (*forgery*) as a *nomen acti* but also *Erwerbung* (*acquisition*), which is surely not effected and whether it should be regarded as affected appears to be a matter of taste. It is moreover somewhat irritating that Fleischer and Barz at the same time characterize *nomina acti* as referring to *states* that are consequences or results of actions (*op cit*, p. 86), for which nouns derived from psychological verbs like *Verblüffung* (*amazement*) provide typical examples. The nouns listed in (11), in contrast, are classified as process descriptions (*op cit*, p. 154), i.e. as *nomina actionis*. Finally, they restrict *nomina patientis* to persons being the object of an action, with *-ling* and *-and* as the only affixes taken into account.

[4] Alternatively, one could try to establish a more appropriate terminology, say, by speaking of *nomina situationis*.

Engel (1996, p. 505) also states that *nomina patientis* refer to persons which are the objects of an action. One of his examples, however, is *Aufkleber* (*sticker*), which is clearly not a person. Engel continues to characterize *nomina patientis* as nouns that take subject position in passive constructions and can be asked for using the corresponding relative clause:

(12) a. Ein Prüfling ist jemand, der geprüft wird.
(An examinee is somebody who is examined.)

b. Ein Aufkleber ist etwas, das aufgeklebt wird.
(A sticker is something which is sticked on.)

c. Eine Erfindung/Erwerbung ist etwas, das erfunden/erworben wird.
(An invention/acquisition is something which is invented/ acquired.)

As example (12c) illustrates, the criterion of taking subject position in passive sentences does not exclude effected and affected objects from being classified as *nomina patientis*.

### 3.2 Motsch's Classification

Motsch (1999, p. 343) does not identify a *nomen patientis* class at all. Instead he introduces a single semantic class for "themes" of an action or event which also includes resultatives, his *nomina acti*. His usage of the theme role is thus rather broad (*op cit*, pp. 36f): a theme can be the patient of an action, an affected or effected object, or a passive undergoer.[5] Motsch proposes a paraphrase test along the lines of (12), which works well for his examples *Lieferung* (*delivery*), *Spende* (*donation*), *Aufkleber* (*sticker*), and *Prüfling* (*examinee*). However, he also needs to account for resultatives like *Beschreibung* (*description*), to which the passivation test does not apply:

(13) *Eine Beschreibung ist etwas, das beschrieben wird.
(*A description is something which is described.)

[5] Whereas the semantic classifications of Eisenberg, Fleischer and Barz, and Engel are purely verbal, Motsch proposes a semantic representation consisting of predicate-argument structures that are constructed from certain elementary predicates like TUN (DO), AFFIZIER (AFFECT), etc. Semantic roles then correspond to certain argument positions of these predicates. Motsch emphasizes that his representations have only semi-formal character.

He nevertheless seems to subsume this type of resultative under his theme class as his examples *Bestuhlung* (*seating*) and *Strukturierung* (*structuring, structure*) show.

With respect to the semantic characterization of deverbal nouns, Motsch distinguishes three cases: first, pure nominalization, second, semantic shift and reinterpretation, and third, objects that are characterized by their semantic role; see Table 2. The first class essentially consists of the *nomina actionis*. The third class is subdivided into *nomina agentis*, *nomina instrumenti*, *nomina loci*, and the nouns referring to "themes" mentioned before, which include the traditional *nomina patientis*. Since Motsch describes *nomina agentis* as referring to animates, usually persons, that are agents of certain events, he is faced with the problems indicated in Section 3.1.

As for the second class, Motsch lists *iterative events*, *punctual events*, and *result states*. Deverbal nouns referring to iterative events are typically formed by so-called combinatorial derivation taking a prefix and a suffix: *Gefrage*, *Gesaufe*, or by *(er)ei*-suffixation: *Fragerei*, *Sauferei*. What Motsch considers as nouns referring to punctual events is precisely the class exemplified in (11) – the *nomina acti* of Eisenberg. Result state interpretations, finally, typically apply to nouns derived from verbs referring to a psychological or physical change of state; an example mentioned before is *Verblüffung* (*amazement*).

| | | |
|---|---|---|
| event/state | 1 | *nomen actionis* |
| | 2.1 | iterative reinterpretation |
| | 2.2 | punctual reinterpretation |
| | 2.3 | result state |
| object | 3.1 | *nomen agentis* |
| | 3.2 | *nomen instrumenti* |
| | 3.3 | theme, result, *nomen acti* |
| | 3.4 | *nomen loci* |

*Table 2: Motsch's (1999, Sect. 2.1) classification of deverbal nouns*

## 4 Valency of Deverbal Nouns

It is often seen as problematic (cf. e.g. Eisenberg 1999, Sect. 8.4, Zifonun 1997, Sect. G1-2.2) to assume that nouns have valency and arguments in the same way as verbs do because, first, nouns do not have fixed com-

plement positions and, second, there is considerable freedom in realizing the semantic roles syntactically. Nevertheless, there seems to be consensus that deverbal *nomina actionis* inherit the semantic roles of their base verbs. Consider e.g. the verb *zerstören* (*destroy*), whose participants can be described by the MultiNet roles AGT and AFF, respectively; cf. (14a).

(14) a. [Die Vandalen]$_{AGT}$ zerstörten [Rom]$_{AFF}$.
(The Vandals destroyed Rome.)

b. die Zerstörung [Roms/von Rom]$_{AFF}$ durch [die Vandalen]$_{AGT}$
(the destruction of Rome by the Vandals)

The event described by the corresponding deverbal noun *Zerstörung* has the same participants, which can be expressed as attributes of the noun; witness (14b).

A natural question now is, to what extent the syntactic realization of "arguments" of the deverbal noun can be predicted on the basis of the syntactic or semantic characteristics of the base verb. From a practical point of view, such a prediction of syntactic argument realization would be highly useful since it allows to automatically generate fully specified entries for deverbal nouns from their base verb entries.

According to standard textbooks on German grammar (e.g. Eisenberg 1999), there is a certain stock of regularities concerning the syntactic realization of the "arguments" of a deverbal noun. The most common pattern is that of (14), where the (direct) accusative object of a transitive verb becomes a genitive attribute of the deverbal noun or a postnominal *von*-PP, whereas the subject of the verb is turned into a postnominal *durch*-PP. (The parallel construction in English employs an *of*-PP and a *by*-PP, respectively.)

Prepositional complements of the verb reappear in identical form as attributes of the deverbal noun; e.g. *Teilnahme an* (*participation*), *Unterweisung in* (*instruction*). The same is true of clausal complements; witness (15):

(15) a. Der Anwalt behauptete, dass der Vertrag ungültig ist.
(The lawyer claimed that the contract is invalid.)

b. die Behauptung des Anwalts, dass der Vertrag ungültig ist
(the lawyer's claim that the contract is invalid)

Genitive and dative objects of the base verb, on the other hand, become postnominal PPs of the deverbal noun,[6] where the respective preposition is not predictable in a uniform way:

(16) a. Bedürfnis nach (need), Gedenken an (remembrance),

b. Mitteilung an (notification), Unterwerfung unter (subjugation, submission).

In case the genitive or dative object of the verb can be alternatively expressed by a prepositional object, this preposition is employed for the corresponding postnominal PP of the deverbal noun. Examples are *Erinnerung an* (*memory*, *recollection*, *remembrance*) and *Beitritt zu* (*joining*, *accession*). If there is no accusative object, the subject cannot be expressed by a *durch*-PP but only by a genitive NP, known as *genitivus subiectivus*, or by a *von*-PP.[7]

The *genitivus subiectivus* is also compatible with transitive verbs like *begrüßen* (*greet*), *belehren* (*teach*, *advise*), *durchsuchen* (*search (through)*), *unterstützen* (*support*), etc. But notice that for many transitive verbs the postnominal genitive of the *nomen actionis* does not allow a subject interpretation; examples are *befreien* (*free*, *liberate*) and *zerstören* (*destroy*).[8]

Let us now look at the valency of other nominalization types besides *nomina actionis*. *Nomina agentis* of transitive verbs trivially do not allow a *genitivus subiectivus* interpretation of a postnominal genitive attribute, whereas the direct object can be realized as a genitive NP or a *von*-PP (17a). For *nomina patientis*, things are the other way around: the postnominal genitive does not correspond to the direct object but to the subject of the base verb (17b). In the case of result state nominalizations, the postnominal genitive can refer to the carrier of that state, which corresponds to the direct object of the base verb if the latter is transitive (17c) and to the subject otherwise – e.g. *Verzweiflung* (*desperation*). (Notice that a postnominal *über*-PP can refer to the subject of transitive psychological verbs: *Verblüffung über*.) As mentioned before, the class of result object

---

[6] The reflexive verbs *bemächtigen* (*seize (hold)*) and *vergewissern* (*make sure*) seem to be potential exceptions to this claim:

a. *Die Boulevardpresse bemächtigte sich wiederholt dieses Themas.*

b. *(?)Die wiederholte Bemächtigung dieses Themas durch die Boulevardpresse (hatte zur Folge …)*

[7] A possible exception is *huldigen*: *die Huldigung der Operndiva durch ihre Fangemeinde*

[8] There are also *nomina actionis* with highly marked *genitivus obiectivus* interpretation, in which case the standard realization of the direct object is a PP; e.g. *Angriff auf* (*attack*).

nominalizations is not homogenous since some of its members behave like *nomina patientis*, that is, refer to the direct object of the base verb, e.g., *Erfindung* (*invention*), while others refer to implicit objects that are created during the described action or event, e.g., *Abdeckung* (*cover*). In the second case, the *genitivus obiectivus* is strongly preferred (17d).

(17) a. den Finder der Brieftasche belohnen
(to reward the finder of the wallet)

b. den Prüfling des Professors beruhigen
(to calm the examinee of the professor)

c. die Verblüffung des Publikums ignorieren
(to ignore the surprise of the audience)

d. die Abdeckung des Abflussrohres entfernen
(to remove the cover of the drain)

## 5 Deverbal Nouns in HaGenLex

In this section, we describe how the semantic relation between deverbal nouns and their base verbs can be represented within the HaGenLex-MultiNet framework. In addition, we address the question of how to determine the case frame of deverbal nouns. (The MultiNet sorts and relations used in the following are explained in the Appendix.)

### 5.1 Change of Sort Relations

As mentioned in Section 1.1, there is only one kind of morphosemantic relation in WordNet. MultiNet, in contrast, provides several so-called *change of sort* relations to distinguish the semantics of various forms of derivation. Table 3 gives an overview of these relations, with examples and brief explanations. Obviously, the inventory of change of sort relations does not cover all possible interpretations of deverbal nouns listed in Table 2. Indeed, there are only explicit change of sort relations for event and state nominalizations, i.e. for *nomina actionis*, namely CHEA and CHSA, respectively.

| Relation | Signature | Description | Example |
|---|---|---|---|
| CHEA | $dy \times ad$ | event nominalization | entdecken – Entdeckung<br>(discover – discovery) |
| CHPA | $ql \times at$ | property nominalization | breit – Breite<br>(broad/wide – breadth/width) |
| CHPE | $ql \times dy$ | inchoative deadjectival verbs | kalt – erkalten<br>(cold – cool down/go cold) |
| CHPS | $p \times [as \cup st]$ | state description by property | still – Stille<br>(quiet – quietness) |
| CHSA | $st \times as$ | state nominalization | langweilen – Langeweile<br>(bore – boredom) |
| CHSP1 | $si \times p$ | present participle | schlafen – schlafend<br>(sleep – sleeping) |
| CHSP2 | $si \times p$ | past participle | absperren – abgesperrt<br>(close off – closed off) |

*Table 3: Change of sort relations*

It should be noticed that within MultiNet the concepts corresponding to a verb and its *nomen actionis* differ with respect to their semantic sort: As indicated by the signature of the relation CHEA, the concept expressed by a verb referring to a dynamical situation is of sort *dy*, whereas the concept of the respective *nomen actionis* is of sort *ad*, i.e., a dynamical situational *object*. The underlying assumption is that concepts expressed by nouns are regarded as "objectified", in contrast to those expressed by verbs, and that this *intensional* distinction should be reflected on the level of conceptual modeling.

### 5.2 Participant Nominalization

By a participant of an event we mean an agent, patient, instrument, result object etc. The derivational semantics of a participant nominalization can be straightforwardly represented within the MultiNet paradigm by employing the thematic role of the respective participant, viz. AGT, OBJ, BENF, INSTR, RSLT, etc. For instance, if a deverbal noun (e.g. *Erfinder* (*inventor*)) refers to the agent of the base verb (here *erfinden* (*invent*)), then the semantic relation between the verb concept and the noun concept is just AGT.

There is one problem with this approach: *nomina agentis* built by *er*-suffixation often have a habitual or professional connotation. This is the case e.g. for *Erfinder* (*inventor*); an inventor is someone who invents things. The situation is different for examples like *Finder* (*finder*). A finder

is a finder only with respect to a certain object and the respective finding act. You cannot say (without strong reinterpretation) that John is a finder, but you can of course say that John is an inventor. Although this distinction may be of minor importance for text retrieval tasks, it becomes relevant if knowledge representation and inference is at issue. A possible solution within the MultiNet framework is to make use of the relation CTXT to "contextualize" a finder to a specific finding event.

### 5.3 Event Reinterpretations and Result States

The current version of HaGenLex does not account of iterative and punctual (i.e. semelfactive) event reinterpretations (the classes 2.1 and 2.2 of Table 2) but treats the respective deverbal nouns semantically in the same way as standard event nominalizations, i.e., by the relation CHEA (or CHSA). That is, the semantic relation that captures the derivation from *hopsen* to *Hopserei*, *Gehopse*, and *Hopser* is at the moment uniformly expressed by the relation CHEA.

This is not to say the MultiNet formalism is not capable of representing the iterative and semelfactive shift in a more fine grained way. Indeed, it is demonstrated in Helbig 2001, Sect. 7.4, how to express different Aktionsart phenomena within MultiNet. In the case of iterative or semelfactive reinterpretations, one needs to take the quantificational and mereological structure of events into account.

As to result state interpretations, we can use the same semantic relation RSLT as we did for result objects. The RSLT relation is then employed as a lexical semantic relation that holds between lexicalized concepts, like e.g. between *verblüffen* (*amaze*, *baffle*) and *Verblüffung* (*amazement*, *bafflement*).

### 5.4 Transfer of Case Frames

Since we are primarily interested in building a lexical database for NLP applications, the central requirement is to provide the syntactico-semantic parser with enough information to decide on the semantic structure of the input expression. So, in one way or another, the parser needs to know about the possible syntactic realizations of the semantic roles as attributes of the noun. A straightforward approach is to list them explicitly in the lexicon. The transfer of roles from a verb like *befreien* (*free*, *release*, *liberate*) to its event nominalization *Befreiung* with corresponding syntactic realiza-

tions is currently represented in HaGenLex as sketched in (18) (cf. Helbig 2001, Sect. 12.2).[9]

| (18) | *befreien* | AGT<br>NP[nom] | OBJ<br>NP[acc] | AVRT<br>*von*/*aus*-PP<br>optional |
|---|---|---|---|---|
| | *Befreiung* | OBJ<br>NP[gen]/*von*-PP<br>optional | AGT<br>*durch*-PP<br>optional | AVRT<br>*von*/*aus*-PP<br>optional |

The pattern underlying the transfer of the first two complement positions in (18) is the same as that mentioned first in Section 4 and exemplified by (14): the accusative object of the verb becomes a genitive attribute or a *von*-PP; the subject becomes a postnominal *durch*-PP. The third position of (18) illustrates the fact that prepositional complements of the verb turn into prepositional attributes of the deverbal noun.

If the *genitivus subiectivus* is possible, this is noted in the lexicon too. For instance, the case frame for *Unterstützung* (*support*) is roughly of the following form:

| (19) | *Unterstützung* | BENF<br>NP[gen]/*von*-PP | AGT<br>*durch*-PP<br>NP[gen]/*von*-PP |
|---|---|---|---|

Clearly, it would be extremely useful for the automatic generation of *nomina actionis* entries to have criteria that restrict the realization of the subject and the direct object as a postnominal genitive. Lexical information about the possibilty of *genitivus subiectivus* and *genitivus obiectivus* is of course relevant for semantic language processing because it helps to eliminate incorrect interpretation alternatives. For if the lexical entry of a certain deverbal noun specifies that only the semantic role corresponding to the direct object of the verb can be realized as a genitive NP, this reduces the number of possible interpretations of a postnominal NP considerably.

Schulz (1999, p. 138) proposes a first, tentative heuristics that solely rests on the MultiNet role of an argument. It is argued that if the role of an accusative object of the verb is AVRT, INIT, OPPOS, ORNT, or RSLT, then the postnominal genitive can only be interpreted as a *genitivus obiec-*

[9] The prenominal realization by the so-called Saxonian genitive (e.g. *Roms Zerstörung*, *Peters Unterstützung*) is not explicitly listed in HaGenLex.

*tivus*, whereas in the case of AFF, BENF, MCONT, or OBJ, the *subiectivus* as well as the *obiectivus* are possible. There are, however, many counter-examples to this hypothesis. Consider e.g. the verb *erschlagen* (*slay*, *strike dead*), whose direct object is characterized by the semantic role AFF. A genitive NP following *Erschlagung* never refers to the agent of the action but always to the affected object, witness *die Erschlagung des Feindes*. Another example is the verb *überreichen* (*hand over*, *present*). Within MultiNet, the accusative object of this verb is characterized by the semantic role OBJ, which indicates that the entity is only a "passive" participant of the event and does not undergo a substantial change. Postnominal genitive NPs of the event nominalization *Überreichung* can only refer to the entity handed over, and not to the agent.

In Section 6, we will discuss a different approach to explain the possible interpretations of the postnominal genitive, which takes aspectual properties of the verb into account. The distinction between verbs that express a change of state and those which do not turns out to be of central importance in this context. The thesis then is that *ung*-nominalizations based on verbs expressing a change of state, as *erschlagen* (*strike dead*) and *entdecken* (*discover*), only allow the *genitivus obiectivus*.

### 5.5 Current Coverage within HaGenLex

The current version of HaGenLex contains approximately 2,300 derived noun entries that are fully characterized with respect to the morphosyntactic and semantic relation to their respective base entries. Of these, almost 1,900 are deverbal; the rest consists of deadjectival *nomina qualitatis* formed by *(ig)keit*- and *heit*-suffixation. The deverbal noun entries split into *nomina actionis* (over 1,600), formed by *ung*- and *(at)ion*-suffixation, and *nomina agentis* with *er*-suffix.

For these entries, the characterization of their derivational morphology has been automatically generated by exploiting the CELEX morphological database (Baayen et al. 1995). This can be done in two ways: For every HaGenLex verb (or adjective, respectively), the corresponding deverbal (or deadjectival) nouns in CELEX can be extracted. Vice versa, each noun entry of HaGenLex can be checked whether it is deverbal (or deadjectival) according to CELEX.

The derivational information is then used to automatically generate new (or to modify existing) HaGenLex entries. In general, the resulting

entries need further editing.[10] For one thing, we have seen in Section 3 that derivational semantics is not completely determined by derivational morphology. For instance, the suffix *-er*, which is productive for *nomina agentis*, occurs with *nomina instrumenti* as well. Given that a noun is a *nomen agentis*, its case frame can be inferred automatically. Whether its preferred meaning is habitual, professional, or situational, however, has to be determined intellectually by the lexicographer. Another specification not fully determined by the verb entry is the choice of the preposition for arguments that are realized as genitive or dative objects of the verb (cf. Section 4).

For *ung*-nominalizations based on transitive verbs there are two central decisions which up to now resist an automatic treatment within HaGenLex, namely the existence of result interpretations and the possibility of the *genitivus subiectivus*. The remaining sections of this article are to a large part devoted to revealing semantic properties of the verb that are responsible for the different behavior shown by deverbal nouns in this respect.

We conclude this section by exemplifying the feature structure representation of deverbal nouns in HaGenLex. Figure 2 shows two such entries (in abbreviated form), which represent the event interpretation of *Erfindung* (*invention*) and the *nomen agentis Erfinder* (*inventor*). As illustrated by these examples, HaGenLex entries may contain additional MultiNet specifications besides those listed in the case frame. Technically, these specifications appear as values of a certain lexical feature (NET) in a quasi-logical notation. Consider the entry corresponding to the concept *Erfinder.1.1*. Notice that the case frame of this entry does not specify a semantic role at all. Indeed, the relation between the inventor and the invented object is mediated by an act of invention. It is precisely this indirect relation that is expressed by the MultiNet specification encoded in the NET value of the entry, repeated as (20) in the notation introduced in Section 2.1.[11]

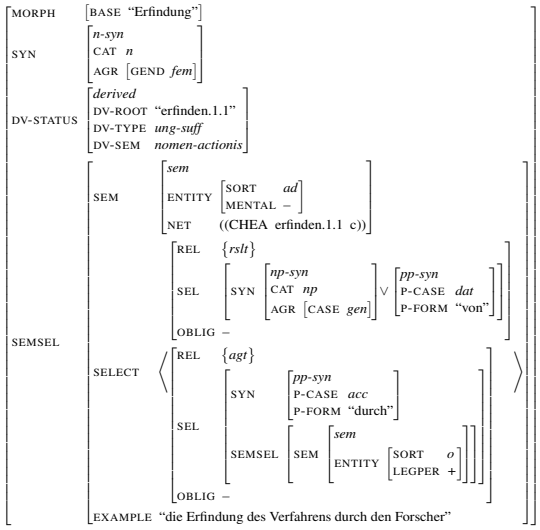
(20) AGT(*erfinden.1.1*, *c*) & RSLT(*erfinden.1.1*, $x_1$)

Here, *c* stands for the concept of the given entry and $x_1$, $x_2$, etc. correspond to the arguments listed in the case frame.

[10] Editing HaGenLex entries is supported by a powerful workbench for the lexicographer.

[11] The specific notation chosen for the lexicon implementation is suited to processing in the SCHEME programming language.

*Erfindung.1.1 (*event nominalization*, nomen actionis)*




*Erfinder.1.1 (*agent nominalization*, nomen agentis)*

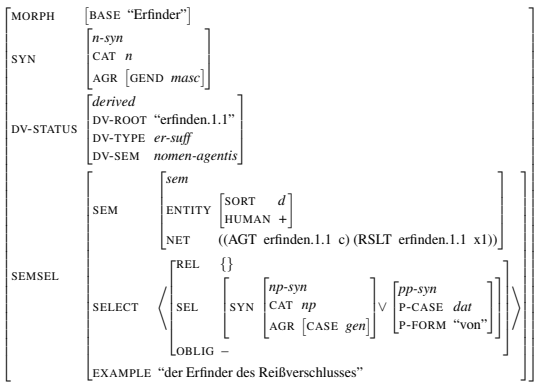



*Figure 2: HaGenLex entries for event and agent nominalization of* erfinden *(*invent*)*

## 6 The Approach of Ehrich and Rapp

The semantic classifications reviewed in Section 3 are primarily descriptive. This is also the case for Motsch's approach who associates with each semantic class a certain semi-formal semantic template. Although these classifications are useful in that they provide an overview of the possible phenomena associated with nominalization, we are more interested in an *explanatory theory* that predicts the possible senses of a deverbal noun as well as their valency on the basis of the semantic structure of the underlying verb.

### 6.1 Lexical Semantic Structure

Ehrich and Rapp (2000) present such an approach for *ung*-nominalizations, which cover about 80% of the deverbal nouns in German. They claim, first, that the syntactic realization of the arguments of a lexical entry is determined by the *Lexical Semantic Structure* of that entry and by category specific linking rules, and, second, that the semantic structure of the base verb indicates the possible senses of the corresponding *ung*-nominalization. Ehrich and Rapp consider *nomina actionis*, result state nominalizations, and result object nominalizations. Since the last class also includes nouns like *Entdeckung*, which satisfy the theme pattern of Motsch, the three classes under investigation can be identified with the classes 1, 2.3, and 3.3 of Table 2.[12]

A possible result interpretation of a deverbal noun by definition presupposes that it is part of the meaning of the base verb to bring about some sort of result. This characterization is closely related to the venerable aspectual distinction between *telic* and *atelic* verbs.[13] Telic verbs like *discover* are said to denote eventualities with an inherent culmination point, in contrast to atelic verbs like *explore*. Ehrich and Rapp refer to telic and atelic eventualities as *events* and *processes*, respectively. Since state nominalizations subsume non-resultive states arising as *nomina "actionis"* of stative verbs like *Bewunderung* (*admiration*) as well as result states of

[12] A mere terminological mismatch is that Ehrich and Rapp apply the terminus *nomen acti* to the type of nouns listed in (11), that is, to class 2.2; cf. *op cit*, p. 254.

[13] We are concerned here with *lexical aspect* or *Aktionsart*; see e.g. Rothstein 2004.

events, the possible semantic sorts relevant for *ung*-nominalizations are given by the leaves of the tree in Figure 3.

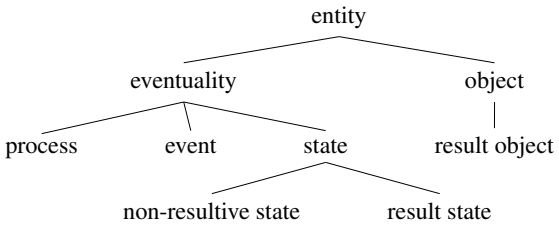



*Figure 3: Possible semantic sorts of* ung*-nominalization according to Erich and Rapp (2000)*

As said before, the assumption is that only deverbal nouns based on telic verbs may allow result interpretations (cf. *op cit*, p. 294). This condition is clearly not sufficient: *fertigstellen* (*complete*) is telic but *Fertigstellung* (*completion*) has only an interpretation as a *nomen actionis*. Ehrich and Rapp postulate specific constraints on the semantic structure of telic verbs that are correlated with possible interpretations of the *ung*-nominalization as a result object or result state, respectively. They adopt a decompositional approach to lexical semantics in order to explicate the change of state component inherent in telic verbs. Their semantic representation of verbs makes use of the basic predicates DO, BE, POSS, and APPL. Here DO is used for actions (with one or two arguments) and BE represents one-place states. The two-place predicates POSS and APPL stand for static relations; POSS indicates the relation of possession, both material (e.g. *own*) and psychological (e.g. *admire*); APPL is intended as generalized local relatedness. The telic aspect of a verb is then represented by means of a change of state predicate BECOME (BEC) that takes state predicates as arguments.

To give some examples, transitive atelic activity verbs like *schlagen* (*hit*) are represented by the two-place process predicate DO, while transitive verbs of possession are represented by POSS:

| | | |
|---|---|---|
| (21) | DO((*x*,*y*) *p*) | schlagen (hit), unterstützen (support), … |
| | POSS((*x*,*y*) *s*) | besitzen (own), kennen (know), bewundern (admire), … |
| | APPL((*x*,*y*) *s*) | umgeben (surround), … |

Telic verbs are represented by applying BECOME to the respective result state term:

(22) BEC((BE((*x*) *s*)) *e*) einschlafen (fall asleep), sterben (die), …
BEC((POSS((*x*,*y*) *s*)) *e*) erkennen (recognize), finden (find), …
BEC((APPL((*x*,*y*) *s*)) *e*) erreichen (reach), …

If it is part of the meaning of a telic verb that the change of state is the result of the action of one of the participants, then the semantic representation contains a DO term in addition to the BECOME term. The semantic template then has the form:

(23) DO((...) *p*) & BECOME((⟨State-Predicate⟩((...) *s*)) *e*)

Here, the conjunction of DO and BECOME is meant to express an implicit causation relation between the respective action (or process) and the change of state event.[14] Examples for such causative telic verbs are given in (24).

(24) a. DO((*x*,*y*) *p*) & BEC((BE((*y*) *s*)) *e*)
renovieren (renovate), fertigstellen (complete), …

b. DO((*x*) *p*) & BEC((BE((*y*) *s*)) *e*)
erfinden (invent), ausgraben (dig up), …

Notice that the semantic structure of verbs of *causative change by affection* (24a) involves the two-place DO predicate, whereas the semantic structure of *availability verbs* (24b) contains the one-place version of DO. Put differently, in the first case, the agent acts (more or less) directly on the object (patient, theme, undergoer, …) in order to give rise to the result state, whereas in the second case there is no such direct continual influence of the agent on the object.

Speaking of agents and the like points to a presumed inventory of semantic roles. In the approach of Ehrich and Rapp, each (non-situational) argument of a primitive predicate is associated with a specific thematic role as shown in Table 4. The second argument of DO is called the *affected theme* or *patient*, in symbols, $\text{THEME}_{\text{AFF}}$. Arguments within the scope of BECOME are referred to as *effected arguments* by Ehrich and Rapp. Since an argument can occur in more than one primitive predicate of a decom-

[14] Moreover, the event argument *e* occuring in the BECOME term is taken as identical to the event argument of the whole causative construction; cf. *op cit*, p. 258.

position, it may carry multiple thematic roles. For instance, in the semantic structure of (24a), which subsumes active change of state verbs, the affected theme is identical with the effected theme.

Another example for multiple roles is given by the so-called *treatment verbs*, which are subdivided into purely locative treatment verbs (25a) and treatment verbs with effected theme (25b).

(25) a. DO((*x*,*y*) *p*) & BEC((APPL((*z*,*y*) *s*)) *e*)
bemalen (paint), absperren (close off), …

b. DO((*x*,*y*) *p*) & BEC(((BE(*z*) & APPL(*z*,*y*)) *s*) *e*)
zusammenfassen (summarize), beurteilen (judge, assess), …

| | |
|---|---|
| DO((*x*) *p*) | *x* = AGENT |
| DO((*x*,*y*) *p*) | *x* = AGENT, *y* = $\text{THEME}_{\text{AFF}}$ |
| BE((*x*) *s*) | *x* = THEME |
| POSS((*x*,*y*) *s*) | *x* = EXPERIENCER, *y* = POSSESSUM |
| APPL((*x*,*y*) *s*) | *x* = APPLICATUM, *y* = RELATUM |

*Table 4: Primitive predicates of Ehrich and Rapp (2000) with associated thematic roles*

### 6.2 Interpretation of *ung*-Nominalizations

The general thesis of Ehrich and Rapp is that a deverbal noun and its base verb have the same Lexical Semantic Structure but differ with respect to their *argument structure*. The argument structure of a lexeme is essentially the list of those arguments of its Lexical Semantic Structure that are projected to the subcategorization frame, plus the *referential argument*. An illustrative example is provided by the treatment verbs with effected theme (25b). The argument structure of verbs of this type is simply the list of arguments preceding the Lexical Semantic Structure in (26):

(26) $\lambda y\, \lambda x\, \lambda e$ [DO((*x*,*y*) *p*) & BEC(((BE(*z*) & APPL(*z*,*y*)) *s*) *e*)]

The referential argument *e* represents the event referred to by the verb itself, whereas *x* and *y*, the agent and the affected theme, correspond to participants of the event which are typically denoted by the grammatical subject and the direct object, respectively. The effected theme *z*, in con-

trast, is an *implicit* argument, which is not projected to the subcategorization frame.[15]

Ehrich and Rapp propose a system of *linking rules*, which control how the elements of the argument structure are syntactically realized. A powerful linking theory would of course be extremely useful for a semantically based lexicon like HaGenLex. In this report, however, we are primarily concerned with the question of how the possible interpretations of a deverbal noun depend on the semantic structure of the base verb (but see Section 6.3 below).

According to Ehrich and Rapp, all possible interpretations of an *ung*-nominalization have the same Lexical Semantic Structure, which is identical with that of the base verb. The assumption is that different interpretations pick out different referential arguments. Since telicity is taken as a necessary condition for a verb to have an *ung*-nominalization with resultative interpretation, the semantic structure of such a verb must contain a BECOME term. The possible interpretations as an event, result state, or result object nominalization then correspond to different referential arguments as indicated by the last three lines of Table 5 (where the dots in front of BECOME indicate a possible DO term).

| | | |
|---|---|---|
| process | *p* | DO((…) *p*) |
| non-resultive state | *s* | ⟨State-Predicate⟩((…) *s*) |
| event | *e* | … BECOME((⟨State-Predicate⟩((…) *s*)) *e*) |
| result state | *s* | … BECOME((⟨State-Predicate⟩((…) *s*)) *e*) |
| result object | *z* | … BECOME((⟨State-Predicate⟩((*z*,…) *s*)) *e*) |

*Table 5: Semantic templates for nominalization sorts, with referential argument*

Let us now turn to the specific constraints Ehrich and Rapp assume for result interpretations. They observe that the verbs in (24b), (25a), and (25b) allow a result object interpretation of the *ung*-nominalization while those in (24a) do not. In order to explain this behavior, Ehrich and Rapp claim that a result object interpretation is possible if the *effected argument*

[15] For verbs of class (25a), Ehrich and Rapp postulate the same argument structure as for those of class (25b). In contrast to (25b), however, the effected theme of (25a) can be realized syntactically as a so-called *argument adjunct* by a *mit*-phrase (*with*-phrase); cf. Section 6.3 below. It is argued (*op cit*, pp. 266f) that this option is blocked for template (25b) because the argument of BE can only be realized in a certain structural syntactic position, which in the case of (25b) is already occupied by the affected theme.

*with highest rank is referentially independent*, in which case that argument becomes the referential argument of the deverbal noun. Here the *rank* of an argument is determined with respect to the position in the primitive predicates, where the first argument of a two-place predicate has higher rank than the second one. So in (25), the effected argument with highest rank is *z*, whereas in (24), it is *y*. What is special about (24a) is that *y* is *referentially dependent* in the sense that it is also the second argument of DO, i.e., an affected theme.

As for result state interpretations, the entity being in such a state is necessarily given by an effected argument, i.e., by an argument of a state predicate within the BECOME term. According to Ehrich and Rapp, this argument is always the *effected argument with lowest rank*. In addition, the argument must be an *affected theme, which is neither created nor destroyed*. All in all, verbs subsumed by (25a) are assumed to give rise to a result state interpretations of their *ung*-nominalization, those satisfying (25b) and (24b) do not, and for (24a) it depends on whether the effected theme is created or only modified.

### 6.3 Linking and the Interpretation of the Postnominal Genitive

A key assumption of Ehrich and Rapp is that the subcategorization frame of a deverbal noun is not derived from the subcategorization frame of the base verb but that both frames are derived from one and the same Lexical Semantic Structure by category specific linking rules. Due to lack of space we only sum up the consequences of Ehrich and Rapp's linking theory with respect to the interpretation of the postnominal genitive of deverbal nouns; for an introduction to their theory of verbal linking the reader is referred to *op cit*, Sect. 2.3.

According to Ehrich and Rapp, the argument structure of an *ung*-nominalization consists of all thematic arguments of its Lexical Semantic Structure if the latter does *not* contain a BECOME term; otherwise, the argument structure consists solely of *the effected argument with lowest rank* (besides the referential argument). Consider e.g. the deverbal nouns *Unterstützung* (*support*) and *Beurteilung* (*judgment*, *assessment*), whose Lexical Semantic Structure coincides with that of the respective base verbs *unterstützen* (21) and *beurteilen* (25b). The corresponding argument structures are presented in (27), where parentheses indicate optionality:

(27) a. Unterstützung (support)
$(\lambda y)\ (\lambda x)\ \lambda p\ [\mathrm{DO}((x,y)\ p)]$

b. Beurteilung (judgment, assessment)
(λ*y*) λ*e* [DO((*x*,*y*) *p*) & BEC(((BE(*z*) & APPL(*z*,*y*)) *s*) *e*)]

In the Lexical Semantic Structure (27b) associated with *Beurteilung*, the effected argument with lowest rank is *y*.

The proposed argument structures give indeed correct interpretations of the postnominal genitive for the examples presented in (27): *die Unterstützung des Kollegen* can mean both, supporting the colleague or being supported by the colleague; *die Beurteilung des Kollegen*, in contrast, shows a strongly preferred *genitivus obiectivus* interpretation. One has to keep in mind that this behavior is confined to the *nomen actionis* reading of *Beurteilung*, that is, *Beurteilung* here refers to the act of judgment.

*Beurteilung* also allows a result object interpretation. In Ehrich and Rapp's theory, the corresponding argument structure is that of (27b) with referential argument λ*z* instead of λ*e*. Now observe that under the result interpretation, *die Beurteilung des Kollegen* can refer to the judgment given by the colleague. Ehrich and Rapp explain this seeming violation of the proposed argument structure by pointing out that *des Kollegen* should not be analyzed as a *genitivus subiectus* in this case but as a *genitivus auctoris*, which indicates the originator and also occurs with non-derived nouns as in *die Idee des Kollegen* (*the idea of the colleague*).

A further point to notice is that Ehrich and Rapp's theory of argument structure and linking is only concerned with "structural" argument positions, where within the nominal phrase only the postnominal genitive is considered as structural. For this reason, the syntactic realization of the agent by a *durch*-PP (*by*-PP) is analyzed as a so-called "argument adjunct", which is not controlled by the argument structure of the deverbal noun. Although there is currently no such distinction between arguments and argument adjuncts in HaGenLex, it should be emphasized that we can nevertheless make use of Ehrich and Rapp's predictions concerning the interpretation of the postnominal genitive (as far as they turn out to be reliable).

### 6.4 Critical Remarks

Within the present article, the main purpose of reviewing Ehrich and Rapp's approach is to motivate possible extensions and modifications of the semantic representation in HaGenLex in order to allow an automatic prediction of result interpretations. Nevertheless, some critical remarks

on Ehrich and Rapp's theory may be in order (with no completeness intended):

*(i)* Despite first appearance it is far from clear how the semantic formalism of Ehrich and Rapp fits into the formal framework of standard predicate logic. Leaving aside the idiosyncratic argument parentheses as in POSS ((*x*,*y*) *s*), the main obstacle is that such a state expression cannot serve as an argument of another predicate. So BECOME is not a predicate in the logical sense but more something like an operator, whose syntactic behavior and semantic denotation remains to be explicated. Moreover, whatever the '&' in 'DO (…) & BECOME (…)' means, it is obviously not logical conjunction.

*(ii)* Change of possession verbs like *schenken* (*give as present*) and *liefern* (*deliver*) seem to be a problem for Ehrich and Rapp's theory. The proposed semantic template is DO ((*x*,*y*) *p*) & BEC ((POSS ((*z*,*y*) *s*)) *e*). According to the theory, since the first argument of POSS, i.e. the EXPERIENCER, is the effected argument with highest rank and in addition is referentially independent, there should be a result object interpretation of *Schenkung* and *Lieferung*. Such an interpretation indeed exists, but it refers to the POSSESSUM and not the EXPERIENCER.

*(iii)* Locative treatment verbs (25a) are assumed to give rise to result object nominalizations whose referential argument is the APPLICATUM, i.e. the first argument of the embedded APPL predicate. On semantic grounds, however, the APPLICATUM should *not* be identified with the result object since the latter is typically a newly created object that materially consists of the APPLICATUM (which is often a substance).

*(iv)* The affective change of state template (24a) subsumes verbs of creation and modification. On the other hand, Ehrich and Rapp notice that within this classes there is a difference with respect to the existence of result object nominalizations. The semantic representation is thus not fine grained enough to allow a correct prediction in this case. From the viewpoint of knowledge representation and reasoning it is of course important to know whether an entity is created or just modified.

## 7 Lexical Semantic Structure in HaGenLex

The goal of this section is to sketch how lexical semantics as employed in approaches like that of Ehrich and Rapp can be expressed by means of the MultiNet paradigm. The gain of such an undertaking is twofold: First, the lexical semantic information in HaGenLex becomes more detailed, espe-

cially with respect to the classification of verbs and deverbal nouns by their event structure. Second, we are able to employ Ehrich and Rapp's criteria for predicting whether a given verb gives rise to an *ung*-nominalization with a result object or result state interpretation and whether the *nomen actionis* permits the *genitivus subiectivus*.

The MultiNet sort hierarchy does not distinguish between telic and atelic situations, as the reader may convince himself by scanning the subsorts of *si* (and *abs*) in the table presented in Appendix A.1. In particular, there is no correlate of Ehrich and Rapp's sortal distinction between processes and events. Similarly, the MultiNet ontology does not distinguish between non-resultive states and result states (cf. Figure 3). However, this lack of sortal distinction should be of minor importance because it is one of the main goals of the decompositional approach to give a *structural* characterization of telicity and atelicity in terms of event structures. In other words, Ehrich and Rapp's sortal distinction between processes and events and between non-resultive states and result states is redundant since the sort of an argument can be recovered from the structural position of that argument in the Lexical Semantic Structure.

The central question then is how to represent Lexical Semantic Structure within MultiNet. Let us check how far we get with the semantic case frame approach of HaGenLex described in Sections 2 and 5. There are two semantic roles expressing a change of state, namely AFF and RSLT. The relation AFF holds between a dynamical situation and one of its participants if the latter is changed by the situation; RSLT holds if the participant is created by the situation. If the semantic role RSLT occurs in the case frame of a verb, then the verb can be clearly counted as telic. This is different for the role AFF, which in HaGenLex also serves to characterize the direct object of atelic verbs like *foltern* (*torture*).

Moreover, the intended usage of the relation AFF is too restrictive to capture every type of change of state. Indeed, the paraphrase test given in the MultiNet explication of AFF requires that *the object is changed* by the event, which implies more than just a change of state (see Helbig 2001, Sect. 18.2.1). For instance, in HaGenLex, the direct object of the verb *leeren* (*empty*) is characterized by the role OBJ and not by AFF, because if something, e.g., a bottle, is emptied, that object itself is not regarded as changed – in accordance with the everyday usage of the word *change*. Hence the required aspectual information cannot be expressed by MultiNet roles alone.

Despite of the fact that MultiNet lacks a BECOME operator for turning states into result states of events, it is nevertheless possible to represent

result states by means of the RSLT relation. To this end, the lexicalized concept corresponding to a telic verb is assumed to stand in a categorical RSLT relation to a generic concept representing the result state of the respective event. In order to pin down that this generic state concept represents the result state inherent in the semantics of the given verb, one can employ the change of sort relations CHSP2 and CHPS (see Table 3), whose combination applied to the verb concept gives us a concept which (in German) is best verbalized by the so-called *stative passive* (*Zustandspassiv*).

Figure 4 illustrates this reconstruction of Lexical Semantic Structure within MultiNet for the verb *erfinden* (*invent*). The given semantic representation says that the lexicalized concept *erfinden* is of sort action with an agent and a resulting object; moreover, it is inherent in the *erfinden* concept, that the action has a result state, whose carrier (theme, …) is the result object. If we compare the MultiNet representation of Figure 4 with the Lexical Semantic Structure (24b) proposed by Ehrich and Rapp for verbs like *erfinden*, we can concede a considerable degree of correspondence. The primitive predicate DO is represented by the ontological sort *da* for actions and the agent role AGT; the BECOME((BE(…))…) construct corresponds to the result state construction and the role SCAR. A more difficult issue is to recover the distinction between the one-place and two-place versions of DO from the MultiNet representation. Recall that in the theory of Ehrich and Rapp it makes a difference with respect to the prediction of result object nominalizations whether an effected theme is also an affected theme, i.e. the second argument of DO. We come back to this problem below.

Notice that, in general, the representation of lexical semantics in Mul-

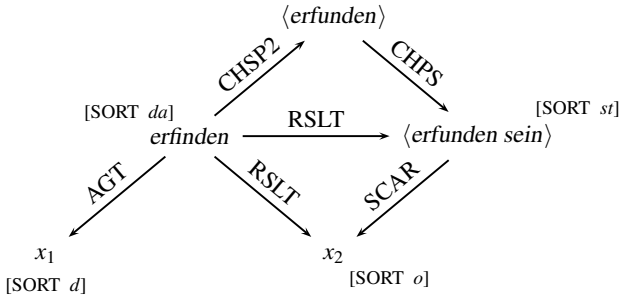



*Figure 4: MultiNet reconstruction of event structure for* erfinden *(*invent*)*

tiNet works on the level of generic concepts; hence the absence of SUB and SUBS relations in Figure 4. (Instantiation is only required if reference to specific objects and events comes into play.) Notice further that the concept names in angle brackets are only included for the sake of convenience; they do not introduce any semantic aspect not already covered by the lexical concept *erfinden* and the given MultiNet structure.

Let us turn briefly to the representation of such structures in HaGenLex. The semantic roles spanning the HaGenLex case frame of the example verb *erfinden* are AGT and RSLT. In the feature structure of the entry, these roles occur in the subcategorization list together with other semantic and syntactic specifications of the arguments. Recall from Section 5.5 that HaGenLex allows additional MultiNet specifications in terms of the quasi-logical representation introduced in Section 2.1. In the case of the example presented in Figure 4, the HaGenLex entry contains the term (28b), where, by convention, $c$ stands for concept of the entry in question (here, for *erfinden*) and $x_1$ and $x_2$ correspond to the arguments listed in the case frame. The term (28a) represents the part of the MultiNet semantics that is incorporated in the subcategorization list of the entry.

(28) a. $\mathrm{AGT}(c_1, x_1)$ & $\mathrm{RSLT}(c, x_2)$ & $\mathrm{SORT}(x_1) = d$ & $\mathrm{SORT}(x_2) = o$

b. $\mathrm{RSLT}(c, n_1)$ & $\mathrm{SCAR}(n_1, x_2)$ & $\mathrm{SORT}(n_1) = st$ & $\mathrm{CHSP2}(c, n_2)$ & $\mathrm{CHPS}(n_2, n_1)$

The previous discussion shows that result states (as well as their arguments) can be easily represented by means of MultiNet. Hence, provided that HaGenLex entries systematically contain this kind of representation, we can make use of the following conclusions Ehrich and Rapp draw from the *absence* of the BECOME term: First, for atelic verbs, i.e., those verbs whose lexical semantics does not contain a result state specification, the *ung*-nominalization has no result interpretation. Second, the *nomen actionis* of atelic verbs allows the *genitivus subiectivus*.

In order to transfer Ehrich and Rapp's constraints for result interpretations listed at the close of Section 6.2, we need to take the ranking of arguments and their referential independence into account. Recall that the rank of an argument is determined by its position in the primitive predicates, that is, by the thematic roles of the argument. It therefore suffices to define a corresponding ranking of MultiNet roles. In effect, this task comes down to choosing appropriate MultiNet roles for the arguments of Ehrich and Rapp's basic two-place state predicates POSS and APPL. Since POSS as well as APPL generalize over semantically rather different sorts

of states – viz. possession and psychological relations in the first case and local relationships and other kinds of "associations" in the second – we cannot expect to get a unique representation by MultiNet roles.

MultiNet provides a primitive semantic relation POSS for possession in the usual, material sense of the word. In such cases, the MultiNet roles of the participants are semantically empty because the semantics is determined by a MultiNet relation between the participants. For this purpose, MultiNet provides the roles ARG1 and ARG2. As to psychological states, mental experiencers are characterized by the MultiNet role MEXP, whereas the content of the state is described by the role MCONT; if the mental state is oriented towards an existing object then the respective argument carries both the MCONT and the OBJ role.[16] So, the role with highest rank for states of (material) possession is ARG1, and for mental states, it is MEXP.[17]

Representing the arguments of APPL is slightly more intricate – partly because of the rather vague semantics of APPL. For the present purposes, we can try to be content with a minimalistic MultiNet representation in terms of an unspecified state. That is, the two MultiNet roles at disposal are SCAR (state carrier, theme) and SSPE (state specifier). Since the APPLICATUM and the RELATUM correspond to the state specifier and the carrier, respectively, the SSPE role is assigned a higher rank than SCAR. A further problem, listed as point (iii) of Section 6.4, is that Ehrich and Rapp identify the APPLICATUM of concrete treatment verbs like *bemalen* (*paint*) with the result of such actions, which is questionable from a conceptual viewpoint.

Figure 5 depicts a possible MultiNet template for concrete treatment verbs that resolves the aforementioned deficiency. For ease of presentation we spare the canonical change of sort construction by CHSP2 and CHPS (cf. Figure 4). The shaded node stands for the lexical concept of the verb; symbols in brackets denote ontological sorts. The object treated by the action carries the role AFF. The result of the act is a discrete object that consists of the material employed for treating the affected object; this is expressed by the MultiNet relation MERO (mereology). Finally, the resulting object is attached to the affected object in some unspecified sense (ATTCH). Notice that the given structure lacks a direct role for the third

[16] Cf. Helbig 2001, Sect. 18.2.50.

[17] As mentioned in point (ii) of Section 6.4, Ehrich and Rapp's theory of result object interpretations unfortunately does not allow correct predictions for change of possession verbs.

argument $x_3$, which represents the material employed for the treatment. Alternatively, one could assign the instrument role (INSTR) to this argument. On the other hand, the primary instrument of a painting act is not the paint but the brush.

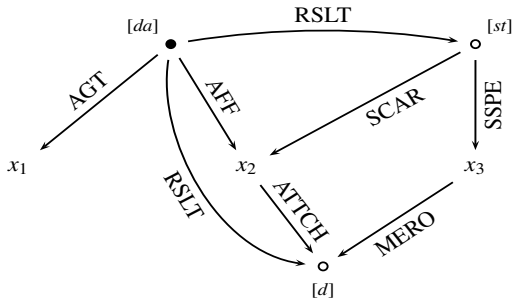



*Figure 5: MultiNet template for concrete treatment verbs like* bemalen *(*paint*)*

An interesting observation is that the verb *absperren* (*close off*), which is subsumed in Ehrich and Rapp's classification by the same template (25a) as *bemalen*, gives rise to a slightly different semantic representation in HaGenLex, with AFF replaced by OBJ. Indeed, there are reasons to subdivide the class of concrete treatment verbs into the class consisting of the verbs *bemalen* (*paint*), *beschmieren* (*smear*), *besprühen* (*spray*), etc. and the class consisting of *abdecken* (*cover*), *absperren* (*close off*), etc. Moreover, contrary to what Ehrich and Rapp claim, a result state interpretation of the *ung*-nominalization is acceptable for verbs of the second class but questionable for those of the first. This observation is supported by the fact that the verbs of the second class, in contrast to those of the first, are subject to an argument alternation called *locatum-subject alternation* in Levin 1993, Sect. 3.5:

(29) a. (Eine Barriere aus) Stacheldraht sperrte das Gelände ab.
((A barrier of) barbed wire closed off the area.)
b. *(Eine Schicht aus) Farbe bemalte die Wand.
(*(A coat of) paint painted the wall.)

A result state interpretation in the case of treatment verbs thus seems to be related to a stative reading of the verb. We leave the further investigation of this question to another occasion.

As noted before, applying the criterion of referential independence requires to decide whether a given action verb is to be described by the one-

place or by the two-place version of DO . A straightforward solution to this problem is to mimic the arity of DO by the number of direct roles. The MultiNet semantics of verbs belonging to class (24b), like *ausgraben* (*dig up*), would then have the form shown on the left of Figure 6. Within MultiNet, however, the most natural representation for these verbs is as depicted on the right of Figure 6. For it is precisely the purpose of the MultiNet role OBJ to characterize unaffected participants. But then we are faced again with the problem to distinguish verbs like *ausgraben* from those like *leeren* (*empty*) by their lexical MultiNet structure. Maybe a more detailed semantic description is required in this case which takes into account that if you empty something you remove something from it whereas if you dig up something you remove it from something else. So, the existence of result object nominalizations may fall out from a suitable semantic representation without recourse to the notion of "referential independence". We will not pursue this issue further here.


*Figure 6: Alternative representations of neutral objects*

In contrast to OBJ, the MultiNet role AFF immediately allows to decide on the arity of DO because AFF can be regarded as a special case of Ehrich and Rapp's $\text{THEME}_{\text{AFF}}$. The RSLT role, on the other hand, is much less helpful in this respect, for HaGenLex employs this role to characterize the direct object of verbs including *erfinden* (*invent*) as well as *fertigstellen* (*complete*). Recall that Ehrich and Rapp take these two verbs to have different Lexical Semantic Structure, namely (24a) and (24b), respectively, which differ with respect to the arity of the DO predicate. According to Ehrich and Rapp's theory, this difference explains the fact that *Erfindung* (*invention*) has a result object interpretation whereas *Fertigstellung* (*completion*) has not. In particular, the RSLT role is *not* an indicator for the existence of result object nominalizations. In order to capture the "affected-

ness" of the effected theme for verbs like *fertigstellen* (*complete*) one could think of assigning the roles RSLT and AFF simultaneously.[18]

The verbs listed in Figure 7 are assumed by Ehrich and Rapp to belong to class (24a) without exception. Within HaGenLex, there is a natural subclassification into verbs of creation and modification by means of the roles RSLT and AFF, respectively. This distinction matters to Ehrich and Rapp's prediction of result state nominalizations, although it is not modeled by their Lexical Semantic Structure; cf. point (iv) of Section 6.4. Since in template (24a), the effected theme is also an affected theme, Ehrich and Rapp's theory predicts that none of the verbs listed in Figure 7 gives rise to a result object interpretation of its *ung*-nominalization. However, *Beschädigung* (*damage*) and *Verletzung* (*injury*) are apparent counterexamples to this claim. Although their existence depends on the existence of the objects they are attached to, damages and injuries are discrete objects on a par with stains and holes. Furthermore, Ehrich and Rapp's theory not only denies result object interpretations for *Beschädigung* and *Verletzung* but predicts result state interpretations, which seems inappropriate as well because an injury is no more a state than a hole.[19]


beschädigen (damage)
renovieren (renovate)
verändern (change)
verletzen (injure)

abfassen (write)
errichten (erect)
fertigstellen (complete)
herstellen (produce)

*Figure 7: Causative telic verbs with affected objects and result objects*

Figure 8 presents a possible MultiNet representation for verbs like *beschädigen* and *verletzen* which takes the foregoing discussion into account. The result state inherent in the semantics of these verbs is charac-

[18] A remaining problem is that the role AFF, as currently specified within MultiNet, presumes the participant to exist while the event occurs, which is not the case for acts of creation.

[19] The state in question is of course to *have* an injury or a hole.

terized by the MultiNet structure to the effect that a result object (a damage or injury) is attached to the affected object, which in turn is the carrier of the result state. Notice that the semantic representation makes use of a feature of MultiNet not mentioned before in this article: the relation EQU, which indicates the equivalence of concepts, holds between a concept represented by a node of the MultiNet graph and a concept represented by an edge between nodes. MultiNet thus allows a certain form of higher order representation.

To sum up, the task of representing Lexical Semantic Structure à la Ehrich and Rapp by means of MultiNet has led us to the following results: The event structure operator BECOME can be easily reconstructed within MultiNet by using the RSLT relation. Reconstructing the thematic structure is more difficult since there is no direct correspondence with MultiNet roles. Nevertheless, a careful analysis of these difficulties can lead to significant improvements of the semantic representation in HaGenLex.

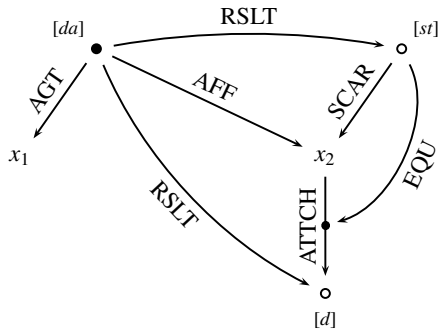



*Figure 8: Lexical MultiNet template for* beschädigen *(*damage*),* verletzen *(*injure*)*

Finally, although the predictions of Ehrich and Rapp concerning the existence of result interpretations have to be taken with care, we think it worthwhile to pursue the program further to automatically infer possible interpretations of deverbal nouns on the basis of the lexical semantics of the corresponding verbs. The MultiNet formalism seems to be well suited for this purpose.

## References

R. Harald Baayen, Richard Piepenbrock, and Leon Gulikers. *The CELEX Lexical Database. Release 2 (CD-ROM)*. Linguistic Data Consortium, University of Pennsylvania, Philadelphia, Pennsylvania, 1995.

Veronika Ehrich and Irene Rapp. Sortale Bedeutung und Argumentstruktur: ung-Nominalisierung im Deutschen. *Zeitschrift für Sprachwissenschaft*, 19(2):245–303, 2000.

Peter Eisenberg. *Grundriß der Deutschen Grammatik: Das Wort*. J. B. Metzler, Stuttgart, 1998.

Peter Eisenberg. *Grundriß der Deutschen Grammatik: Der Satz*. J. B. Metzler, Stuttgart, 1999.

Ulrich Engel. *Deutsche Grammatik*. Julius Groos, Heidelberg, 3rd edition, 1996.

Johannes Erben. *Einführung in die deutsche Wortbildungslehre*. Erich Schmidt Verlag, Berlin, 3rd edition, 1993.

Christiane Fellbaum and George A. Miller. Morphosemantic links in WordNet. *Traitement automatique des langues*, 44(2):69–80, 2003.

Wolfgang Fleischer and Irmhild Barz. *Wortbildung der deutschen Gegenwartssprache*. Niemeyer, Tübingen, 2nd edition, 1995.

Aldo Gangemi, Roberto Navigli, and Paola Velardi. The OntoWordNet project: Extension and axiomatization of conceptual relations in WordNet. In Robert Meersman, Zahir Tari, and Douglas C. Schmidt, editors, *On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE – OTM Confederated International Conferences*, Lecture Notes in Computer Science 2888, pages 820–838. Springer, Berlin, 2003.

Carsten Gnörlich. *Technologische Grundlagen der Wissensverwaltung für die automatische Sprachverarbeitung*. PhD thesis, FernUniversität in Hagen, Fachbereich Informatik, Hagen, Germany, 2002.

Sanda M. Harabagiu, George A. Miller, and Dan I. Moldovan. WordNet 2 – a morphologically and semantically enhanced resource. In *Proceedings of ACL-SIGLEX99: Standardizing Lexical Resources*, pages 1–8, Maryland, 1999.

Sven Hartrumpf. *Hybrid Disambiguation in Natural Language Analysis*. Der Andere Verlag, Osnabrück, 2003.

Sven Hartrumpf, Hermann Helbig, and Rainer Osswald. The semantically based computer lexicon HaGenLex – Structure and technological environment. *Traitement automatique des langues*, 44(2):81–105, 2003.

Hermann Helbig. *Die semantische Struktur natürlicher Sprache: Wissensrepräsentation mit MultiNet*. Springer, Berlin, 2001.

Hermann Helbig and Carsten Gnörlich. Multilayered Extended Semantic Networks as a language for meaning representation in NLP systems. In Alexander Gelbukh, editor, *Computational Linguistics and Intelligent Text Processing (CICLing 2002)*, Lecture Notes in Computer Science 2276, pages 69–85, Berlin, 2002. Springer.

Hermann Helbig, Carsten Gnörlich, and Johannes Leveling. Natürlichsprachlicher Zugang zu Informationsanbietern im Internet und zu lokalen Datenbanken. In Klaus-Dirk Schmitz, editor, *Sprachtechnologie für eine dynamische Wirtschaft im Medienzeitalter*, pages 79–94, Wien, Austria, 2000. TermNet.

Claudia Kunze and Andreas Wagner. Anwendungsperspektiven des GermaNet, eines lexikalisch-semantischen Netzes für das Deutsche. In Ingrid Lemberg, Bernhard Schröder, and Angelika Storrer, editors, *Chancen und Perspektiven computergestützter Lexikographie*, pages 229–246. Niemeyer, Tübingen, Germany, 2001.

Johannes Leveling. University of Hagen at CLEF 2003: Natural language access to the GIRT4 data. In C. Peters, M. Braschler, J. Gonzalo, and M. Kluck, editors, *Evaluation of Cross-Language Information Retrieval Systems, Fourth Workshop of the Cross-Language Evaluation Forum, CLEF 2003, Revised Papers*, Lecture Notes in Computer Science, Berlin, 2004. Springer. To appear.

Johannes Leveling and Hermann Helbig. A robust natural language interface for access to bibliographic databases. In Nagib Callaos, Maurice Margenstern, and Belkis Sanchez, editors, *Proceedings of the 6th World Multiconference on Systemics, Cybernetics and Informatics (SCI 2002)*, volume XI, pages 133–138, Orlando, Florida, July 2002. International Institute of Informatics and Systemics (IIIS).

Beth Levin. *English Verb Classes and Alternations*. University of Chicago Press, Chicago, 1993.

Wolfgang Motsch. *Deutsche Wortbildung in Grundzügen*. Schriften des Instituts für Deutsche Sprache, Band 8. de Gruyter, Berlin, 1999.

Rainer Osswald. Die Verwendung von GermaNet zur Pflege und Erweiterung des Computerlexikons HaGenLex. *LDV-Forum*: 19 (1/2): pages 43-51, 2004.

Terence Parsons. *Events in the Semantics of English*. MIT Press, Cambridge, MA, 1990.

Susan Rothstein. *Structuring Events*. Blackwell, Oxford, 2004.

Marion Schulz. *Eine Werkbank zur interaktiven Erstellung semantikbasierter Computerlexika*. PhD thesis, FernUniversität Hagen, 1999.

Gisela Zifonun, Ludger Hoffmann, and Bruno Srecker. *Grammatik der deutschen Sprache*. Schriften des Instituts für Deutsche Sprache, Band 7.3. de Gruyter, Berlin, 1997.

# A Elements of the MultiNet Formalism

## A.1 Ontological Sorts

| | |
|---|---|
| entity [*ent*] | |
| object [*o*] | |
| concrete object [*co*] | |
| discrete object [*d*] | *house, apple, tiger* |
| substance [*s*] | *milk, honey, iron* |
| abstract object [*ab*] | |
| attribute [*at*] | |
| measurable attribute [*oa*] | *height, weight, length* |
| non-measurable attribute [*na*] | *form, trait, charm* |
| relationship [*re*] | *causality, similarity, synonymy* |
| ideal object [*io*] | *question, justice, criterion* |
| abstract temoporal object [*ta*] | *Renaissance, Easter, holiday* |
| modality [*mo*] | *necessity, intention, permission* |
| situational object [*abs*] | |
| dynamic situational object [*ad*] | *race, robbery, movement* |
| static situational object [*as*] | *equilibrium, sleep* |
| situation [*si*] | |
| dynamic situation [*dy*] | |
| action [*da*] | *write, sing, sell, drive* |
| happening [*dn*] | *rain, decay, explode* |
| static situation [*st*] | *stand, be ill* |
| situational descriptor [*sd*] | |
| time [*t*] | *yesterday, Monday, tomorrow* |
| location [*l*] | *here, there* |
| modal situational descriptor [*md*] | *impossible, necessary, desirable* |
| quality [*ql*] | |
| property [*p*] | |
| total quality [*tq*] | *dead, empty, green* |
| gradable quality [*gq*] | |
| measurable quality [*mq*] | *small, expensive* |
| non-measurable quality [*nq*] | *friendly, tired* |
| relational quality [*rq*] | *inverse, equivalent, similar* |
| functional quality [*fq*] | |
| operational quality [*oq*] | *forth, last, next* |
| associative quality [*aq*] | *chemical, philosophical* |
| quantity [*qn*] | |
| quantificator [*qf*] | |
| numerical quantificator [*nu*] | *one, two, five, hundred* |
| non-numerical quantificator [*nn*] | *all, many, several* |
| unit of measurement [*me*] | *kg, meter, mile* |
| measurement [*m*] | *three miles, two hours* |
| graduator [*gr*] | |
| qualitative graduator [*lg*] | *very, especially, rather* |
| quantitative graduator [*ng*] | *almost, nearly, approximately* |
| formal entitity[*fe*] | (meta level entities like figures and tables) |

## A.2 Subset of Semantic Relations

(The change of sort relations are listed in Table 3, Section 5.1.)

| **Relation** | **Signature** | **Short description** |
|---|---|---|
| AFF | [*si* ∪*abs*]×[*si* ∪*o*] | Affected object |
| AGT | [*si* ∪*abs*]×*o* | Agent |
| ANTO | (identical sorts) | Antonymy |
| ARG1/2 | [*si* ∪*abs*]×*ent* | Argument specification for metalevel concepts |
| ATTCH | [*o* \ *at*]×[*o* \ *at*] | Attachment |
| AVRT | [*dy* ∪*ad*]×*o* | Averting/Turning away from an object |
| BENF | [*si* ∪*abs*]×[*o* \ *abs*] | Benefactee |
| CTXT | [*si* ∪*abs*]×[*o* ∪*si*] | Restricting context |
| EQU | (identical sorts) | Equivalence |
| EXP | [*si* ∪*abs*]×*o* | Experiencer |
| INSTR | [*si* ∪*abs*]×*co* | Instrument |
| MCONT | [*si* ∪*o*]×[*si* ∪*o*] | Mental content |
| MERO | *o*×*o* | Generalized part-whole relation |
| METH | [*si* ∪*abs*]×[*dy* ∪*ad* ∪*io*] | Method |
| MEXP | [*si* ∪*abs*]×*d* | Mental carrier of a state |
| OBJ | [*si* ∪*o*]×[*si* ∪*o*] | Neutral object |
| OPPOS | [*si* ∪*o*]×[*si* ∪*o*] | Entity being opposed by a situation |
| ORNT | [*si* ∪*abs*]×*o* | Orientation towards something |
| POSS | [*co* ∪*io*]×[*co* ∪*io*] | Possession |
| RSLT | [*si* ∪*abs*]×[*si* ∪*o*] | Result |
| SCAR | [*st* ∪*as*]×*o* | Carrier of state |
| SSPE | [*st*∪*as*]×*ent* | State specifier |
| SUB | [*o* \ *abs*]×[*o* \ *abs*] | Conceptual subordination |
| SUBS | [*si* ∪*abs*]×[*si* ∪*abs*] | Conceptual subordination (of situations) |
| SYNO | (identical sorts) | Synonmy |