Meaning and Lexicalization of Word Formation

Workshop at the 14th International Morphology Meeting, Budapest, 13-16 May 2010

List of papers

Maria Bloch-Trojnar, John Paul II Catholic University of Lublin
Action nominalizations in English – an LMBM approach

Germana Olga Civilleri, University Roma Tre
Lexicalization of Deverbal Nouns: A Morpho-Semantic Analysis in Ancient Greek

Angeliki Efthymiou, Democritus University of Thrace, Greece
How many factors influence the meaning of denominal verbs? The case of Modern Greek verbs in -(i)ázo.

Jessica Forse, Swansea University
Analysing *en* and its Romance equivalents in Jackendoff’s Parallel Architecture

Pius ten Hacken, Swansea University
Semiproductivity and the Place of Word Formation in Grammar

Claire Hopkin, Swansea University
Lexicalization in morphology and in the Parallel Architecture

Renáta Panocová, P.J. Šafárik University, Košice, Slovakia
Semantics of diminutivization. Evidence from Russian

Kaarina Pitkänen-Heikkilä, University of Helsinki
Term-formation in a special language – How do the words specify the scientific concepts?

Maria Rosenberg, Stockholm University
Lexicalized Semantic Structures and Their Formal Expression in Word-Formation

Martin Schäfer, Friedrich-Schiller-Universität Jena, Germany
Semantic transparency and anaphoric islands

Carmen Scherer, University of Mainz
Lexicalization and lexical semantics – a diachronic case study of German noun derivation

Barbara Schlücker, Freie Universität Berlin
Combined concepts: Meaning and meaning relations in German A+N compounds

Alexandra Soares Rodrigues & Graça Rio-Torto, University of Coimbra, Portugal
Semantic coindexation: evidence from Portuguese derivation and compounding

Niina Zhang, National Chung Cheng University
Temporal opaqueness of root meanings
Action nominalizations in English – an LMBM approach
Maria Bloch-Trojnar
John Paul II Catholic University of Lublin
Poland

In the paper we shall review the systematic as well as idiosyncratic properties of de-verbal nominalizations (such as the productivity of formal markers, interaction of regular and lexicalized senses, syntactic properties) with a view to testing the predictive and explanatory potential of the framework of Lexeme Morpheme Base Morphology (LMBM) put forward and refined by Beard (1995). The presentation of an analysis in this model may be of interest to workshop participants on two grounds.

Firstly, LMBM is in line with Jackendoff’s Parallel Architecture since it advocates a strictly modular structure of grammar, which stems from the ubiquitous lack of isomorphism between grammatical modules. However, Jackendoff concentrates on the relationships between semantics, syntax and phonology leaving no room for morphological considerations. Beard (1995: 381) offers a fuller picture and argues that ‘the only place in language, where semantic, grammatical, and phonological representations are directly related to each other, is the lexicon. The direct relation of these representations defines the lexeme. Elsewhere, at the syntactic, semantic, and phonological levels, information from one domain must be translated into the representations of any other domain which employs that information. Morphology does all the translation’. Therefore, in LMBM there are strict boundaries between: the LEXICON (the storehouse of lexemes), GRAMMAR (structural relations of syntax and a set of morphological categories), SEMANTIC MODULE and the MORPHOLOGICAL SPELLING COMPONENT (the component mapping grammatical function to phonology).

Secondly, LMBM converges with the onomasiological approach to WF in that emphasis is placed on meaning/function of derivatives. However, there are crucial differences in the way derivational categories are marked formally. In the model of Štekauer the Morpheme-to-Seme-Assignment Principle matches onomasiological structure with lexical and affixal morphemes stored in the lexicon. In LMBM, there is a rigid distinction between lexemes and morphemes. Only the former are true linguistic signs. Furthermore, the Separation Hypothesis says that there is no direct connection between the side of morphology that deals with morphophonological operations and the side that specifies semantic-syntactic changes. Consequently, morphemes are not determinants of meaning but should be conceived of as clues which signal a relationship.

The analysis of the category of de-verbal nominalizations will span all derivational types without arbitrarily excluding one or the other, i.e. nominalizations marked with -ing, the so called Latinate suffixes: -(at)ion, -ment, -ance/-ence, -age, -al etc. and nominals formed by means of conversion. It will be proposed that de-verbal nominalizations result from two distinct lexical rules deriving process (uncountable) and event (countable) nominalizations, productively marked by -ing and conversion respectively. The action/result dichotomy has to be explained by a different class of rules – performative speech act rules – not controlled by the functional subcomponent of the lexicon (cf. Beard 1987). It will be demonstrated how the representation of regular nominalizations differs from that of lexicalized ones and how the morphological component, interwoven with other components, maps grammatical information to semantics and phonology.
References


Lexicalization of Deverbal Nouns: A Morpho-Semantic Analysis in Ancient Greek
Germana Olga Civilleri
University Roma Tre

Thanks to the rich articulation of its morphological level, Ancient Greek is a particularly interesting language to study word-derivation rules. We analyze the formation of the A.G. deverbal nouns (as morphological nominalizations), trying to associate to the morphological derivation patterns a semantic content.

The two main strategies for the derivation of nouns from verbal stems in A.G. are suffixation and apophony:

- Suffixes have more likely an intrinsic clear semantics that modifies the value of the verbal stem. Anyway there is an internal gradation in the semantic potential of each suffix. We can represent such a condition by means of a continuum: at one extreme, we find the suffixes with an inherent lexical meaning (“semantically full suffixes”), at the other extreme, the suffixes with no inherent meaning (“semantically abstract suffixes”):

<table>
<thead>
<tr>
<th>Semantically full suffixes</th>
<th>Semantically abstract suffixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>-sis, -mos</td>
<td>-ia</td>
</tr>
<tr>
<td>more productive,</td>
<td>-mis</td>
</tr>
<tr>
<td>more verbal, (argument</td>
<td>-non</td>
</tr>
<tr>
<td>structure)</td>
<td>less productive,</td>
</tr>
<tr>
<td></td>
<td>resulting nouns are</td>
</tr>
<tr>
<td></td>
<td>nearer to bare nouns</td>
</tr>
</tbody>
</table>

- However, some elements lead us to suppose that the apophony (at least the o-grade) have a semantic function too.

When other “bothering factors” do not occur, the meaning of the derived noun normally results from the meaning of the base plus the derivation rule, in a compositional way. For example, the noun anábasis (derived from the stem ana-ban-) has an entirely compositional meaning, in the following way:

(1) \( aná-ba-si-s \)  
\[ \text{PREF. up-} \text{V go-} \text{SUFF.-tion-Nom.sing} \]  
“going up”

Here the suffix -si(s), which has a processual meaning (similar to the English suffix -tion), is added to the prefixed verbal stem, marking the categorial status of the deverbal noun and giving it a processual meaning.

But there are many cases in which the core rule is not productive and a loss of transparency of the compositional meaning takes place:

1) semantics can change depending on the phrasal context;
2) there are cases of polysemy: indeed, through metaphorric and metonymic shifts, lexemes and single suffixes can develop a new meaning which is connected to the original one;
3) many words are lexicalised, i.e. come into the lexicon as whole units!

We will explain the lexicalised lexemes as cases of polysemy – in which metaphorric and metonymic shifts from the original (compositional) meaning play a role – by analyzing the data elicited from the Homeric Greek corpus and we will show how the lexicalized nouns lie at different degrees of lexicalization. In particular, in order to
more clearly point out these facts, in this talk we will use those words that show “semantically full suffixes”, since they seem to be more productive and have a more evident core/original meaning from which the lexicalization process moves.

References


BENEDETTI, Marina (2002) “Radici, morfemi nominali e verbali: alla ricerca dell’inaccusatività indoeuropea” in Archivio Glottologico Italiano 80, 1 (pp. 20-45);


CHANTRAINE, P. (1968) La formation des noms en grec ancien, Klincksieck, Paris;


GAETA, Livio (2002) Quando i verbi compaiono come nomi, Franco Angeli, Pavia;


KURYLOWICZ, J. (1956) L’apophonie en indo-européen, Wrocław;


How many factors influence the meaning of denominal verbs?
The case of Modern Greek verbs in –(i)ázo

Angeliki Efthymiou
aefthym@eled.duth.gr
Democritus University of Thrace, Greece

Although generative morphology has concentrated on structural and phonological issues concerning word formation to the neglect of semantic issues for many years (cf. Lieber 2004, ten Hacken 2009), recent works assign an important role to the semantics of word formation processes (cf. for example Jackendoff 2009, Lieber 2004, 2009, Plag 1999, Štekauer P. 2005).

The aim of this paper is to examine the factors involved in Modern Greek verb forming processes. Our evidence comes from the Modern Greek suffix –(i)ázo. –(i)ázo usually attaches to nominal and adjectival [-learned] [+negative] bases and derives [-learned] verbs, which express a whole range of related concepts such as resultative, inchoative, ornative, locative, instrumental, similitative and performatival: e.g. komatçázo ‘brake/tear into pieces’, dropçázo ‘to disgrace’, tsuvaljázo ‘to bundle into a sack’, niçázo ‘scratch with one’s nails’, kuvendjázo ‘to bundle’. Given that the most robust semantic pattern of –(i)ázo derivatives is the inchoative meaning ‘be provided with many and usually unwanted endogenous entities’ (e.g. ritiðázo ‘to wrinkle, become wizened’), we address the following four questions.

1. What is the role of the meaning of the base? Are all bases suitable for the expression of the inchoative meaning? Is the evaluative (or cumulative) meaning assigned by the base of the derivative or by the suffix? For example, in the case of ritiðázo ‘to wrinkle, become wizened’ the negative meaning is already expressed in the base noun ritiða ‘wrinkle’.

2. What is the role of the word formation process in which –iázo participates in the creation of the meaning? How can we distinguish the meaning of these verbs from the meaning of the other Modern Greek verb forming suffixes (cf. for example laspóno ‘muddy’ vs laspçázo ‘become mash’ (láspi ‘mud’))? How can we distinguish the evaluative connotation of –iázo verbs from their underlying causative/resultative semantic structure?

3. Is the phonological shape of the suffix related to its evaluative and cumulative meaning? Is it a coincidence that the phonological sequence /iá/ is found also in other Modern Greek suffixes like –iá and –iáris, which form [-learned] derivatives that express pejorative or collective meanings (e.g. kokaljáris ‘skinny person’, zitçanjá ‘beggarhood, typical behavior of a beggar’) (cf. Anastassiadis-Symeonidis 1997, Efthymiou 1999)?

4. Does the evaluative/expressive meaning of the suffix and the [-learned] register of its derivatives affect its productivity? Does the rivalry with other suffixes influence the meaning or the productivity of –iázo (cf. for example ritiðázo ‘to wrinkle’ (intransitive) [-learned] vs. ritiðóno ‘to wrinkle’ (transitive and intransitive) [+/-learned])? How does this correlate with the fact that –iázo seems to be the prevailing default verb forming suffix in Modern Greek for the interpretation ‘become provided with many unwanted x’ in Modern Greek (cf. Efthymiou (to appear))? Elaborating on these questions we show that the computation of the meaning of a word formation process is a rather complex work, since it is influenced by various factors, such as the semantic, pragmatic and structural properties of the base, the
evaluative connotation of the suffix and its derivatives and the productivity of the word formation process.

References
Efthymiou A. (to appear) “The Semantics of Verb forming Suffixes in Modern Greek”. In the proceedings of the 19th International Symposium of Theoretical and Applied Linguistics, 3-5 April 2009, School of English, Aristotle University of Thessaloniki.
Analysing *en-* and Its Romance Equivalents in Jackendoff’s Parallel Architecture

Jessica Forse
Swansea University
Department of Translation and Digital Communication
224671@swan.ac.uk

Much of the existing research into morphological processes focuses on the form of words, arguably at the expense of their meaning, so that many interesting aspects of morphology relating to semantics remain unexplored. This work explores one of these semantic aspects, concentrating on prefixation. This paper has developed out of a larger project which explores the extent to which Jackendoff’s formalism of lexical semantics, Conceptual Structure, can be applied to word formation (WF) processes in Romance languages and how it should be expanded and/or adapted to account for the processes and the differences between the individual languages.

Jackendoff (2009) offers some suggestions as to how to encode derivational WF processes, such as the formation of actor nouns from verbs, in his formalism of Conceptual Structure. This paper expands this line of investigation, focusing on one other WF process in particular, the formation of verbs using the prefix *en-* (and its allomorph *em-*) in English, and its counterparts in some Romance languages.

In a form-oriented approach, headedness is generally thought of as a matter of syntactic category: the head of a derived form is that which determines the category it belongs to. English is considered to be a predominantly “right-headed” language. However, the prefix *en-* is one of a small group of English “category-changing” prefixes, which violate the Right-hand Head Rule (RHHR) as proposed by Williams (1981). Whereas prefixes such as *counter-* and *un-* adhere to the RHHR, as in *counterattack* and *unhappy*, the exceptional left-hand head, *en-*, determines the category of the derived form, creating verbs, as in *enslave*.

When analysed in semantic terms, the relative contribution that the prefixes and the bases make to the meaning of the derived form varies. In *unhappy*, for example, the prefix *un-* simply contributes the meaning NOT, and would be formalised in Conceptual Structure as in (1b).

(1) WF Process: Prefixation with *un-*
   (a) Input: happy \( \rightarrow \) [Property HAPPY]
   (b) Output: unhappy \( \rightarrow \) [Property NOT [HAPPY]]

In Conceptual Structure, *conceptual constituents* are classified in terms of their *ontological categories*, such as Thing, Property and Situation. Crucially, the output of the WF process in (1), *unhappy*, belongs to the same ontological category as the input, *happy*, i.e. they are both Properties. This means that the input constitutes the main concept of the derived form; the prefix *un-* simply negates it.

The prefix *en-* contributes much more strongly to the meaning of derived forms than prefixes such as *un-*. It affects the concept of the input in a different way, modifying its ontological category, as in (2).
(2) WF Process: Prefixation with *en-*

(a) Input: slave $\rightarrow [\text{Thing SLAVE}]$

(b) Output: enslave (X enslaves Y) $\rightarrow$

$[\text{Action CAUSE} ([\text{Thing X}], [\text{Event GO}_{\text{Circ}} ([\text{Thing Y}], [\text{Path TO}_{\text{Circ}} ([\text{Property SLAVE}]]))])]]$

The conceptual structure in (2a) shows that the input, *slave*, is a Thing, whereas the basic conceptual structure in (2b) shows that the output, *enslave*, is an Action. Therefore, in this case, it is the prefix *en-* which has the stronger semantic contribution to the derived form.

The function-argument structure of an ontological category represents how a conceptual constituent belonging to this category can be decomposed in terms of its functions and arguments. The stronger semantic contribution of *en-* can be attributed to the fact that it determines the initial function of the conceptual structure, CAUSE, as in (2b). The semantic contribution of the base *slave* is more deeply embedded.

In this paper, a number of *en-* verbs are analysed in terms of Jackendoff’s formalism. The relation between the semantics of the input and that of the output of the WF process is examined in order to investigate the principal changes in meaning which this particular prefix engenders and the semantic contributions that *en-* makes to its derived forms. Marchand (1960: 113-5) classifies the semantics of *en-* verbs into the following types: *en-X* can mean (1) ‘put in X’, e.g. *encage*; (2) ‘make into X’ or ‘make X’, e.g. *enslave* or *enfeeble*; or (3) ‘X in’ or ‘X up’, e.g. *enwrap*. However, there are also cases which do not fit easily into such a classification; for example, *empower*, has a semantic schema of the type ‘give X’. Such counterexamples form the basis of a further investigation and a comprehensive list of semantic types for *en-* is developed.

Following on from this initial investigation into the principle changes in meaning which *en-* can bring about, some Romance equivalents of *en-*, i.e. French, Spanish and Portuguese prefixes which bring about similar semantic changes in verb formation, are examined in greater detail. The inputs and outputs of these WF processes are analysed in terms of Jackendoff’s formalism and the similarities and differences between the processes in each language are discussed. Given the obvious relationship between the prefixes in the four languages, I will also take etymological data into account in my analysis.

**References**


In this presentation I will address the problems of accounting for morphological productivity in Jackendoff’s (2002) Parallel Architecture. In a sense, productivity can be seen as the property underlying the entire generative enterprise. Chomsky (1957) takes the possibility of producing and understanding new sentences as the main motivation for a rule-based grammar.

In the context of morphology, two perspectives can be distinguished. One takes productivity to be a dichotomy. This view is advocated, for instance, by Schultink (1961). The other view takes it to be a gradual property. The problem with the latter view is that it is difficult to measure productivity, because, as recognized by Aronoff (1976), it is hard to determine the universe with respect to which the set of realized forms should be measured. Following earlier work, Bauer (2001) considers these two perspectives as separate properties.

Jackendoff (2002) introduces the Parallel Architecture as a model for linguistic competence that can account for the processing, acquisition, and evolution of language. The lexicon connects the three parallel structures. Whereas originally Jackendoff assumed that each structure had its own set of formation rules, the discussion of idioms and idiosyncratic constructions leads him to conclude that these formation rules are nothing else than lexical entries with highly underspecified structures in two of the three representations. In this way, the lexicon and the formation rules can be united in a single component.

Jackendoff (2009) then addresses the question of how morphology can be accommodated in such a system. As everything else can be seen as lexical, the default assumption is that morphology is also lexical. Jackendoff proposes a system in which productivity determines whether a rule is encoded as a lexical entry. If it is, it is fully productive and there is no difference to syntactic rules. If it is not, the rule does not exist and it is no more than an epiphenomenon, emergent from entries.

One problem associated with this approach to morphology is that the distinction between productive and non-productive cuts across the distinction between inflection and derivation. As is clear from Booij (2003) and other works, some morphologists do not consider this distinction to be important anyway. However, certain differences that are inherent in the distinction remain to be accounted for.

A more serious problem is the question of semiproducetive word formation rules. If word formation rules are of the same type as syntactic rules, they can only be fully productive. However, word formation is not rule-driven, but only rule-based. Lexicalization is a central aspect of word formation. This distinction can be seen if we compare the expressions (1) and (2).

(1) The library building is blocking my sea view.
(2) calendar converter

The syntactic rules active in (1) lead to a perfectly clear sentence. The result of the word formation rule producing (2), however, can only be interpreted on the basis of
the context of formation. We have to account for different things in (1) and (2). In (2) we have to state that the rule is available but not applied as automatically as syntactic rules are. Jackendoff (2009) proposes to treat such rules as epiphenomena. I challenge this position in ten Hacken (in press). In a reaction, Jackendoff (p.c.) proposes to let (2) be generated by rules with a diacritic feature to indicate that the result is not automatically available. Ten Hacken (in press) proposes instead to assume a separate word formation component with rules triggered by the pragmatic need of naming a new concept.

In my paper I will explore the consequences of the two different positions and argue that a separate word formation component has advantages at various levels of analysis.

References


There has been an increase in approaches to morphology which employ theories of lexical semantics, notably that of Jackendoff (1983; 1990; 2002), in order to describe the semantics of word-formation (c.f. Plag, 1998; Lieber, 2004). However, there is a communication problem between Generative Morphology and Jackendoff’s Parallel Architecture in that the term lexicalization is used for two different concepts. In Generative Morphology, the term lexicalization usually refers to the process whereby a lexical item formed by a word-formation process is stored permanently in the mental lexicon (Plag, 2003; Anderson, 1992). Lexicalization can thus be seen as an expected step in word-formation. Permanent storage of a lexical item allows for a shift or specialization in meaning, and this too has been called lexicalization (Huddleston and Pullum, 2002). In the latter view, the lexicalized meaning of a lexical item is contrasted to its compositional meaning.

In contrast, the view of lexicalization in Jackendoff’s Parallel Architecture concerns the verbalizing of elements of Conceptual Structure. Conceptual Structure concerns all meaning, not just linguistic meaning, and has interfaces to perception and action etc. It is therefore possible to examine which elements of pre-linguistic meaning are verbalised or lexicalized.

It is interesting to examine how these two different conceptualizations of lexicalization can be related. One possible link lies in Pustejovsky’s (1995) Generative Lexicon. In this theory, polysemy is accounted for through the idea of a dotted type, which unites the different sense of a word. We can then talk about the way in which clusters of different types are lexicalized, or put into words. The idea of a dotted type allows us to unite the compositional meaning and lexical meaning that have been identified in Generative Morphology and account for, for example the process-result alternation found in many nominalizations. Likewise, this view of lexicalization can be easily linked with Jackendoff’s view of verbalizing a concept.

Pustejovsky’s theory can then provide the tools to examine the semantics of a derived word more closely. For example, an alternation can be observed in many nominalizations between the senses of ‘process’ and ‘result’.

(1) a. The construction of the building took three years.
   b. The large construction blocked the view.

(2) a. Their spontaneous settlement among the natives.
   b. There was no settlement at Stamford at the time.

These alternations can be described using the idea of a dotted type:

(3) settlement_lcp = process, result, process.result

Taking for example the nominalization settlement, the way in which the type cluster is lexicalized can then be examined. The following examples indicate that there are certain restrictions on its usage.
(4)  

a. Clashes also occurred because of the settlement of peasants on the pastures.
b. ?? Clashes also occurred because of the settlement of the Joneses on the same street as the Smiths.
c. The settlement of Jamestown in Virginia.
d. ? The settlement of Guildford in Surrey.

We can therefore use Pustejovsky’s theory of split lexicalization to examine how the simple and complex types in this type cluster are lexicalized. This allows us to examine the correspondence between a morphologically derived nominalization, and the equivalent syntactically derived form, in this case settlement versus act of settling. Pustejovsky’s theory arguably allows better identification of the specialization of meaning identified within Generative Morphology, and it can be used to link this specialization of meaning with the elements of Conceptual Structure expressed by the different forms. In this way, the perspective on lexicalization in the Generative Lexicon provides a link between those in Generative Morphology and Conceptual Structure.

References


The morphological process of diminutivization in Slavic languages has been a frequently studied field in linguistic theory. It is assumed that the nature of diminutives in Slavic languages is derivational (cf. Manova 2005). This paper will concentrate on the process how diminutive nouns in Russian receive their meaning.

The theoretical framework of Štekauer’s (2005) onomasiological model of word-formation and Horecký’s multilevel model of word-formation (1994) will be a starting point of our research. Both models take into account the reference to extralinguistic reality and include conceptual, semantic, onomasiological, onomatological and phonological level. In Horecký’s model, the semantic level plays the most prominent role. Horecký (1994: 19) claims that semantic features form the semantic level of the language and characterize words, certain sets of words and larger systems of words; a specific system of words is formed by sets of words or the so-called word-formation fields that are derived and may be defined by the word class of the WF base and that of a resulting naming unit.

A complete semantic definition includes not only semantic features, but also specifies relations between the features and their hierarchical arrangement in the form of a tree-diagram. For instance, according to Horecký (1994: 31) the most abstract and hierarchically highest semantic feature of deadjectival adjectives is intensity (INT) and adjectives denoting a higher degree of quality are characterized as +INT. Adjectives having a low degree of intensity are assigned the semantic feature –INT, e.g. sladký : sweet, počerný : slight dark. Qualitative adjectives marked +INT are further divided based on the fact whether they denote simply a higher measure or emphasize a higher measure of quality. The most appropriate semantic feature expressing this difference is the semantic feature of gradus (GRAD). Consequently, adjectives of the type novučičký : very new and malilinký : the very smallest are assigned the semantic feature +GRAD (Horecký, 1994: 31). Deverbal substantives, desubstantival substantives, deadjectival substantives, deverbal adjectives, desubstantival adjectives, deadjectival verbs, desubstantival verbs and deverbal verbs in Slovak were described in a similar way by the same author.

Clearly, diminutive nouns are a subset of desubstantival substantives. It is assumed that derivation of diminutives merely modifies the meaning of basic words, which is in line with Dokulil’s (1962: 47) understanding of the so-called modificational onomasiological category. To our knowledge, diminutives in Russian have not been examined within the above framework. Russian displays a rich inventory of diminutives. As Bratus (1969: 2) claims, out of 25,000 of the most commonly used Russian words, more than a thousand nouns and adjectives have or can have diminutive forms. On the other hand, not all nouns in Russian can be diminutivized, e.g. мужество : courage, учитель : teacher, растение : plant, луна : moon, рис : rice etc. Thus, the aim of this paper is to describe a complete semantic definition of diminutive nouns in Russian including semantic features, relations between them and their hierarchy.
References


Term formation in a special language –
How do the words specify the scientific concepts?

Kaarina Pitkänen-Heikkilä
University of Helsinki
Department of Finnish, Finno-Ugrian and Scandinavian Studies

Term formation in special languages usually requires a strong contribution from experts of special fields. It is necessary to understand the scientific concepts thoroughly and to also know the tradition of the discipline’s term formation as well as the language’s word formation methods and norms.

This paper deals with the Finnish-language botanical terminology that was created by Elias Lönnrot in the middle of the 19th century (Pitkänen 2008). The terminology consists of 1500 plant morphology terms, e.g. emi ‘carpel, pistill’ and hede ‘stamen’. I am interested in why a certain word (formed by word formation) has been chosen to represent a particular concept. How do the words specify the scientific concepts, and finally, how do terms get their meaning?

The study integrates lexicological and terminological methods. In lexicology, the word and its various meanings serve as the focus, whereas the theory of terminology focuses on the concept and concept systems (Sager 1990: 55–56).

Finnish botanical terminology can be divided into three groups depending on the origin: terms have been either 1) accepted from the existing vocabulary and used in their original meanings (e.g. juuri ‘root’), 2) chosen from the existing vocabulary and used for the new, specific botanical meanings (e.g. purje ‘sail’ → ‘the large upright petal of a sweet pea or related flower’ [vexillum]), or 3) created on the basis of the existing vocabulary and used for the new, specific botanical meanings (e.g. kärhi [a retrogressive derivation from kärhys ‘big, branchy tree’] ‘a slender thread-like appendage of a climbing plant’ [cirrus, tendril]). The created terms are new derivations and compounds, and constitute 70% of all the terms. These will be discussed in this paper.

Two thirds of the created new words have been formed either loosely or precisely according to either Latin or Swedish terms, and one third has been formed completely differently from its equivalents in these foreign languages. It is worth noting that many loan translations contain rare vocabulary from Finnish dialects used as equivalents to foreign parts of terms. Lönnrot’s aim was that the scientific terminology would be inspired by the indigenous language, in order that the Finnish agricultural population would also be able to understand scientific texts.

Characteristic of Lönnrot’s botanical terms is the utilisation of the vocabulary of various Finnish dialects, using compounds reflecting the generic and partitive relations of concepts and exploiting particular repeating structures for the coordinate concepts (e.g. certain affixes). These kinds of structures form term systems that reflect the scientific concept systems.

The new terms’ meanings are predictable (see Štekauer 2005) via the old, indigenous vocabulary and the logical system of terms that comes from botanical Latin. In the
interpretation of a term it is important to know the meaning (or etymology) of the used word and structures, the history of the special field and its vocabulary, and the history of the scientific concept and concept system. Thus a new, consciously developed terminology can be understood through the old, familiar vocabulary and structures as well as through the new, logical system of terms. This paper deals with these concepts and considers the balance between the different factors that shape the meaning of the botanical terms formed by word formation in 19th century Finnish.

References


Lexicalized Semantic Structures and Their Formal Expression in Word-Formation

Maria Rosenberg, PhD
maria.rosenberg@fraita.su.se
Department of French, Italian and Classical Languages
Stockholm University
SE-106 91 Stockholm, Sweden

This study addresses word-formation and its lexical representation. The objective is to examine the relation between semantic and formal structure. Semantic structures are supposed to be lexicalized and more or less productive. We adopt an onomasiological perspective, going from meaning to form (cf. Marchand 1969, Štekauer 2005). Given this perspective, the method is empirical. The data draws from a parallel corpus (cf. Cysouw & Wälchli forthcoming), consisting of Swedish NN-compounds and their French counterparts, limited to NN-compounds (kafferast ‘coffee break’ vs. pause-café) and constructional idioms (skidstav ‘ski stick’ vs. bâton de ski). The analysis aims at systematic generalizations, and is both quantitative and qualitative with respect to formal and semantic structure.

We adopt a constructional approach to morphology: “[w]ord-formation patterns can be seen as abstract schemas that generalize over sets of existing complex words with a systematic correlation between form and meaning” (Booij 2009:201). The architecture of grammar is tripartite and parallel (cf. Jackendoff 2002). In the hierarchical lexicon, abstract schemas coexist with individual instantiations: outputs of a productive rule can be listed (Booij 2009). This approach complies with Ryder’s (1994) claim that through knowledge about productive semantic patterns, new compounds are created and interpreted, and with Jackendoff (2009) who assumes productivity of compounds to involve a set of principles, which enable the interpretation of new compounds. Lexicalized compounds conform mostly to these principles. Odd interpretations of compounds are rare (cf. Isabelle 1984). Lapata (2002) mentions three problems regarding compounds and their interpretation: (i) high productivity engenders previously unseen formations; (ii) implicit internal semantic relation; (iii) contextual and pragmatic impact. In Jackendoff’s (2009) conceptual semantics, a compound’s meaning is a function of the meaning of its constituents, and there are several basic functions for the semantic relation within compounds. This view, with which we agree, differs from Lieber’s (2004:49) opinion that a general characterization of the semantic relations within root compounds is probably impossible (cf. also Selkirk 1983).

NN-compounds are right-headed in Swedish, but left-headed in French (cf. Williams 1981, Lieber 1992, Booij 2009). The internal relation of Swedish NN-compounds is often implicit, whereas it can be partly specified by prepositions in the corresponding French constructions. Despite formal differences, we aim at a unified semantic account, by using Jackendoff’s (2009) lexical representation of compounds. More precisely, we examine how basic functions are expressed formally. Swedish NN-compounding is highly productive, and NN-compounds are easily formed for new concepts. French prefers lexicalized NPs, although NN-compounding is productive (cf. Fradin 2009). Hence, our study attempts to examine under which circumstances French allows NN-compounds and in what way these are more restricted compared to Swedish. Possible formations and neologisms may shed light upon restrictions governing certain constructions and their productivity.
In conclusion, the set of possible interpretations provided by our study can serve to disambiguate constructions in context. In a broader perspective, our study intends to have relevance for language processing tasks, i.e. machine translation, with respect to Romance and Germanic languages (cf. Johnston & Busa 1999).

References
Compounds may or may not constitute anaphoric islands, and compounds may or may not be semantically transparent. Focusing on A N compounds in German and Mandarin Chinese, this paper argues that anaphoric islandhood can be used as a measure for semantic transparency.

Compounds can be more or less semantically transparent, but explicit criteria, excepting hyponymy, do not exist. Thus, *blackbird* is more transparent than *commonwealth*, because the former is a hyponym to bird, whereas the latter is not a hyponym to wealth. But even within hyponymic compounds there is an intuitively felt difference in semantic transparency between compounds like *Grünglas* ‘green.glass’ or *Kleinstadt* ‘small.town’ on the one hand and others like *Grünspecht* ‘green.woodpecker’ or *Schwerkraft* heavy.force ‘gravity’ on the other hand. Testing for anaphoric islandhood confirms this intuition: only the former allow anaphoric reference to the respective head nouns, in consequence licensing head noun deletion, as shown in (1) for the pair *Grünglas/Weißglas* ‘green.glass/white.glass’.

(1) *Ich* bin das Grünglas losgeworden, das weiße Glas liegt noch im Auto.

*I got rid of the green glass, the white glass is still in the car.*

This is not possible for the latter, which constitute anaphoric islands and do not allow anaphoric reference to their head, making head noun deletion impossible, cf. (2) for the pair *Grünspecht/Schwarzspecht* ‘green.woodpecker/black.woodpecker’.

(2) *Mein* Vater hat in seinem Garten schon mal einen Grünspecht gesehen, aber noch nie einen schwarzen Specht.

*My father has in his garden once a green.woodpecker seen, but never ever a black woodpecker.*

Intended: ‘My father once saw a green woodpecker in his garden, but he has never seen a black woodpecker.’

My talk assesses the usefulness of this test by looking closely at a broader selection of examples from German. In particular, I will discuss (a) possible confounds to the test, most notably the prerequisite of contrastive A N pairs (cf. e.g. the case of *Schwerkraft* ‘gravity’) (b) the relationship of the semantically transparent A N compounds to their phrasal counterparts and (c) the relationship between transparent A N compounds and lexicalization/conventionalization.

This last point is especially interesting, because A N compounding is not available as a free ad-hoc option (as e.g. N N compounding) in German, but is heavily constrained in that the resulting word must refer to some well-established category of things.

Throughout, I will draw on data from Mandarin, where a similar contrast between two types of A N compounds holds (cf. Schäfer:2009).

**References**

Lexicalization is a gradual process that is quite common in word-formation. However, it is not whole word-formation patterns, such as A+N compounding in English, that undergo lexicalization but single words. Whereas some words formed by a particular word-formation pattern may be completely lexicalized and opaque others may be perfectly transparent and analyzable, compare *black board*, *black mail* and *black market* to *gray board*, *electronic mail* and *domestic market*. Other A+N compounds may be more or less lexicalized and compositional in meaning.

In this paper, the composition of meaning and lexicalization processes will be studied in the framework of lexical semantics. Lieber (2004) assumes that the lexical representation of words and affixes consists of two parts: The first part, called *skeleton*, consists of semantic-grammatic information that is relevant to the syntax. The skeleton is decomposable and formalizable. The second part, called *body*, contains semantic-pragmatic information. It is encyclopaedic and holistic.

The highly polysemous German -er-derivation (cf. 1) will serve as the starting point for discussing a couple of questions related to the constitution and the transparency of meaning:

1. How can the vast range of meanings in (1) be explained? Is it possible to derive this variety in terms of a theory of lexical semantics or are most – or at least some – of the meanings the result of lexicalization processes?
2. What changes appear in the semantic-syntactic skeleton and the semantic-pragmatic body of a nominal -er-derivative in the process of lexicalization?
3. Is there an interrelation between changes in frequency and the degree of lexicalization? Are derivatives with a higher token frequency more likely to be lexicalized? Do lexicalized and nonlexicalized -er-derivatives show particular frequency patterns?

To answer the third question, data will be drawn from the diachronic Mainz Newspaper Corpus (1609-2000, 9 measuring points, 1 million word forms, cf. Scherer 2005). Two groups of derivatives will be studied regarding their development within the last four centuries: first, a group of 23 high frequent -er-derivatives with a total token frequency of \( N>50 \) and, second, a group of 48 low frequent derivatives with a total token frequency of \( N=5 \).
References
Combined concepts:
Meaning and meaning relations in German A+N compounds

Barbara Schlücker
Freie Universität Berlin
barbara.schluecker@fu-berlin.de

A notorious problem of semantics in word formation is the meaning of nominal compounds. Nominal compounds are quite productive in Germanic languages, in particular N+N compounds. It seems reasonable to assume that one reason for their productivity lies in the fact that compounding combines two concepts but leaves open the exact nature of the relation between them. From this point of view, the meaning of compounds is inherently underspecified and therefore very flexible. The diversity of possible meaning relations is reflected on the one hand by approaches which try to spell out the numerous possible relations and on the other by those that propose a single underspecified relation, denying the feasibility of specifying all meaning classes.

This paper deals with a class of nominal compounds that has received much less attention in the literature so far, namely A+N compounds, thereby focussing on German. Although German A+N compounds do not display the same diversity of meaning relations as N+N compounds do, there are obviously several subclasses: apart from the predominant attributive relation with the adjective specifying the noun referent, as in (1), the adjective may also specify an implicit event, a resultant state or an implicit referent, see (2)-(4), and others, cf. Simoska (1999).

(1) Altbau ‘old building’ → building that is old
(2) Schnellgericht ‘instant meal’ → meal than can be prepared quickly
(3) Jungbrunnen ‘fountain of youth’ → fountain that makes the bathing person young
(4) Feinbäckerei ‘confectionery’ → bakery that produces fine pastries

The paper draws on insights on the meaning and the process of interpretation of both lexicalized and novel compounds from morphological, semantic as well as psycholinguistic theory. As for the latter, it is the relation between the constituent parts in N+N compounds (rather than the meaning of the constituents) that has been shown to be crucial for the processing of familiar and novel compounds (Gagné & Shoben 1997, Gagné 2001, Gagné & Spalding 2006, 2009). Semantic (or rather: conceptual) analyses of those relations in a fine-grained and yet generative and unrestricted system have been proposed by Lieber (2004, 2009) and Jackendoff (2009), among others.

The paper discusses to what extent such approaches can be transferred to the analysis of German A+N compounds. Based on semantic analyses of German A+N compounds such as proposed by Motsch (2004), Bücking (2009), and others, it discusses the interplay of the constituent meaning, the relational meaning, and the context, for the meaning constitution of A+N compounds; aiming at an analysis that can account for the diversity of meaning relations of German A+N compounds that is at the same time plausible from a psycholinguistic point of view with regard to processing, conceptual combination and the structure of the lexicon.
References


This paper tries to shed light on the questions: How do words formed by word formation get their meaning? What are the factors involved and what is the balance between them?

According to Jackendoff’s Parallel Architecture (2002), a word is an interface between Phonology, Syntax and Semantics, and formation of word meaning is independent of syntax. The PA permits us to solve some problems that come to light by adopting other perspectives (syntactically guided: Rappaport Hovav & Levin 1992; Beard 1995; Lieber 2004; or cognitively guided: Barker 1998; Ryder 1999; Panther & Thornburg 2002).

We propose to analyse Portuguese word formation of deverbal nouns and adjectives and of compounds, in order to test if these two genolexical mechanisms behave similarly respecting the semantic structure of their constituents.

Empirical data show that word formation mechanisms align semantic structuring independently of syntactic structuring. Only semantic features are responsible for the meaning of the coined word. This is shown by:

i) agent deverbal adjectives can be formed by verbal bases that in their argument structure lack an argument (Agent) corresponding to the meaning of the deverbal adjective. E.g.: chovedor ‘that makes rain’, suador ‘that makes sweat’.

ii) event deverbal nouns, although sharing a verbal base and a general meaning ‘event of Vb’, display different semantics according to the suffix. E.g.: [[anda]² Vnç]N ‘adventure, journey’, [[anda]² mento]N ‘speed or way of something going’, [[anda]² dura]N ‘physical way of moving’. The difference between their meanings comes from the coindexation of semantic features of the affix and semantic features of the base.

iii) concrete meanings (e.g. ‘portion’, ‘residue’, ‘amounts’) of some deverbal nouns (e.g. [[serra]² dura]N ‘sawdust’, [[ceva]² dura]N ‘rests of the bird that a bird of prey has been feed on’). These concrete meanings vary across affixes. Deverbal nouns constructed on the same base but with other affixes don’t have the same concrete meanings ([[serra]² ção]N ‘the action of sawing’, [[ceva]² gem]N ‘the action of feeding’ (causative construction)).

We propose that variation in the meanings across deverbal nouns is due to the semantic feature of each affix, besides other information domains — subsumed in Semantic structure — such as pragmatic-referential fields.

Semantic features that are the object of coindexation come from the lexical units involved, and also from other semantic structures.


Despite their holistic meaning and their idiomaticity, these lexical constructions respect the major grammatical relations (conjunction, subordination, modification) underlying compounding. Their semantic structure reflects the ‘maximal semantic
frame’ associated to each of the constituents, as well as the plausible semantic relations relying them. The distance between possible and lexicalized meanings they convey requires additional meaning computation, involving unexpressed features of the constituents (profile and/or proper functions of denotata, extralinguistic and cultural features), and coercion functions (Jackendoff 2009).

References


It is generally realized that the meaning of word roots can be either transparent or opaque, consistently. In Chinese, the former can be illustrated by the meaning of bei ‘cup’ in cha-bei ‘tea-cup’, and the latter can be illustrated by the meaning of hong ‘red’ in yan-hong ‘eye-red => envy’. See Arad (2003) for formalization of these situations. This talk presents a case in which the meaning of a root in the same word is opaque in one construction, but transparent in another construction.

In Chinese, a noun (N) and a classifier (CL) may form an N-CL cluster, a complex word, as in (1). In such a word, a specific semantic relation, UNIT relation, is expressed. For instance, in (1a), the classifier pi is a unit for counting horses.

(1) a. ma-pi  b. zhi-zhang  c. chuan-zhi  d. jiang-xiang
    horse-CL  paper-CL  boat-CL  award-CL
    ‘horse’  ‘paper’  ‘boat’  ‘award’

Adopting Jackendoff’s (2009) method of analyzing the interpretation of complex words, I formalize the reading of an N-CL cluster as in (2), where the first element, $X_1$, is a noun and the second element, $Y_2$, is a classifier:

(2) $\text{UNIT} (X_1, Y_2)$, ‘$Y_2$ is a unit of $X_1$ for counting’.

Thus, classifiers are meaningful when they are merged with nouns (contra Watanabe 2009: Sec.4). However, if the cluster is not in construal with a numeral, the UNIT feature of the classifier is not active. The presence of a numeral encodes counting, and counting needs a unit. If no counting occurs, the UNIT feature plays no role. In data like (3), where no numeral occurs, the interpretation of the nominal remains the same if the classifier is absent:

(3) a. Zhexie ma-pi hen zhongyao.  b. Zhexie ma hen zhongyao.
    DEM horse-CL very important  DEM horse very important
    ‘The horses are important.’  ‘The horses are important.’

But when the cluster is combined with a numeral, integrated by the classifier ge, as in (4a) and (5a), the UNIT feature of the classifier of the cluster is obligatory. Their absence may cause either unacceptability, as in (4b) and (5b), or different interpretation, as in (6b).

(4) a. san ge shui-di
    three CL water-CL
    ‘three drops of water’
  b. *san ge shui
    three CL water
(5) a. san ge shu-ben
    three CL book-CL
    ‘three books’
  b. *san ge shu
    three CL book
(6) a. san ge hua-shu
    three CL flower-CL
    ‘3 bouquets of flowers’
  b. san ge hua
    three CL flower
  c. san shu hua
    three CL flower
    ‘3 bouquets of flowers’
The contrast between the a-forms and b-forms in (4) through (6) indicates that the ge in the former group, is different from the ge in the latter group. The contrast in (4) shows that the former ge may occur with a mass noun, whereas the latter ge may not (Chao 1968:508). The contrast in (5) shows that the former ge may occur with any count noun, whereas the latter ge has certain restrictions, although it may occur with many count nouns (Loke 1994). In (6a), the classifier shu is used for entities in bundles. (6a) has the same interpretation as (6c). The contrast between (6a) and (6b) simply shows that the former ge plays no semantic role. The contrast clearly shows that it is the classifier of the N-CL that contributes its UNIT feature to the whole nominal, and agrees with the numeral. All of these contrasts suggest that the former ge can be a place-holder, whereas the latter ge is a substantial classifier.

The inconsistency of the role of the UNIT feature of the classifier in the cluster can be explained, if ge and the classifier form a chain and there is only one UNIT feature in the chain. This finding is theoretical interesting, considering the different view of the traditional Lexical Integrity.

References


