Piet Swanepoel
University of South Africa

Problems, Theories and Methodologies in Current Lexicographic Semantic Research*

"The basis of a sound and efficient lexicographic work is a good theory. (Zgusta 1971:17)"

Abstract

The papers on "word meaning/lexical semantics" reflect the multidisciplinary nature of current research on the analysis and description of lexical semantic information in dictionaries. The main aim of this paper is (i) to highlight a number of issues with regard to the relationship between theory and praxis that still plague traditional lexicography and influence its effective participation in this multidisciplinary enterprise, and (ii) to place the reported research within the context of the lexical semantic problems, theories, methodologies, techniques, and lexical data that underlie current research on lexicographic semantics.

1. Introduction

The papers on "word meaning/lexical semantics" have in common their focus on the problems experienced in the analysis, description and acquisition of lexical semantic information. With most other lexicographic research they share the ultimate aim of improving the lexical semantic information in lexicographic resources, be they dictionaries of the traditional kind that are meant for human users, or the lexica of natural language processing (NLP) systems.

These papers also attest, however, to the multidisciplinary input of current research on lexicographic semantics. The advent of the computer age has not only radically changed the praxis of the compilation of dictionaries, it has also given birth to new kinds of dictionaries, a variety of new lexicographic methodologies and techniques, and a range of new disciplines that concern themselves with the compilation of dictionaries (in the widest possible sense of the word) and the theoretical and practical problems they pose. Although there is much overlap and cross–fertilization, experts from the fields of (traditional) lexicography, theoretical linguistics, computational linguistics, computational lexicography and knowledge engineering all tackle the linguistic semantic problems experienced within their fields with their own kind of theoretical and practical expertise in accordance with the specific aims of their lexica.

In the main part of this paper I will approach the difficult task of placing the reported research on lexical meaning within this interdisciplinary framework of lexical semantic problems, theories, methodologies,

techniques, and data that underlies current research on lexicographic lexical semantics.

Before doing so, however, I would like to address a number of issues concerning the relationship between theory and praxis that still plague those involved in the compilation of dictionaries for human use (or so–called "traditional lexicography"; cf. in this regard the forum discussion in *Dictionaries* 14 (1992/1993)). Although none of the papers in this theme cluster address any of these problems, these matters need specific attention as they are hampering the general progress of lexicographic research. The discussion of these problems will set the scene for a few brief remarks on the changing face of lexicographic research in the nineties.

2. Setting the scene for the nineties

There are a number of popular myths and stereotypes in metalexicographic research that are hampering progress in the field. The most important of these myths is the so–called "theorylessness" of practical lexicography. The most damaging stereotype is that of the metalexicographer as the cynical critic of dictionaries and their creators who indulges in matters theoretical with no consideration for the "harsh" practicalities of dictionary making.

The semanticist–lexicographer Anna Wierzbicka has become notorious for her remarks that practical lexicography has no theoretical basis, that "even the best lexicographers, when pressed, can never explain what they are doing, or why" (Wierzbicka 1985:5) and that given the lack of help from semantic theory in particular, "it is the lexicographers' achievements, not their failures, which are truly remarkable" (Wierzbicka 1990:366).

Although theoretical linguistics is not the only source of the theory of lexicography, the impression of the lexicographer working in a theoretical vacuum is strengthened by the observation of authorities such as Atkins (cf. Atkins 1992/1993:5) that "Many people in contemporary lexicography deal with theoretical linguists by keeping their heads down below the barricades and getting on with writing dictionary entries." Even Ted Briscoe, a well–known computational linguist, says in an interview in the first issue of the *Cambridge Language Reference News* ((1) 1994) that lexicography (in contrast to linguistics) "is more atheoretical – a practical activity pursued to sell dictionaries".

Remarks like these should, however, be regarded with suspicion. No praxis is theory–free and lexicography, the art and craft of the compilation of dictionaries, has never lacked a theoretical basis. Atkins (1991:186;1992/1993:5–7)) aptly remarks that even Samuel Johnson had a *Plan of a Dictionary* in 1747 when he embarked on his Dictionary, and that the Preface to his great Dictionary in 1755 testifies to the fact that most of the lexical semantic problems facing lexicographers today were known to him.

The fact that theoretical linguists in their choice of data and the formulation of their theories pass over the empirical descriptive problems lexicographers face does not necessarily imply that lexicography has no theoretical base. Lexicographers are of necessity always meta–lexicographers: their practical descriptive activities have always been and will be informed and guided by principles or assumptions of a linguistic nature. These assumptions may not always be articulated, or if articulated, not strictly adhered to (which gives rise to accusations of the "gulf" between theory and practice), and they may be eclectically constituted, i.e. they may not form a systematic or strictly coherent body of hypotheses on lexical semantic matters, but that does not diminish their status as guiding principles for lexicographic praxis.

Lexicographers' linguistic theoretic intuitions have always been shaped by the experience gained by the analysis and description of "raw" linguistic data, "gut feelings", traditional lexicographical practices (even some for which no explicit motivation or theoretical backing exists), and what has been around by way of utilisable semantic theories for specific lexicographic tasks. All of this is, of course, subject to the dictates of the relevant kind of dictionary and the assumed or real information needs of various categories of dictionary users.

It is accepted today that theoretical linguistics, and for the purpose of this paper, theoretical semantics, offers the lexicographer different ways of looking at language and word meaning, that it not only serves as a means of "consciousness raising" but indeed as a source for the formulation of guiding principles that can assist in the formulation of editorial policies and to guide the lexicographer in the analysis and synthesis of linguistic data for a specific dictionary (cf.Atkins 1992/1993:29; Cruse 1992/1993:88).

Theoretical lexicography/metalexicography consists, however, of more than merely the selected input of theoretical linguistics (cf. Wiegand 1983). As I have already remarked, the compilation of dictionaries has become a multidisciplinary activity, and metalexicography, as Hausmann et al. (1986: 272) define it, is the sum of the scientific activities that endeavour to collect, describe, and evaluate *all* (emphasis mine – PS) theoretical questions and problems of method and procedure arising in the compilation of dictionaries, and consequently aim to improve lexicographic practice.

According to Frawley (1992/93:1) there still is, however, a manifest tension among so–called "practicing and practical" lexicographers ("who make dictionaries for a living") and the theoretical lexicographers or meta–lexicographers, i. e. "those concerned with what dictionaries should or could be").

As an empirical observation, this is an overgeneralization as it simply does not provide for the many cases of cooperation and mutual support amongst researchers in various fields of expertise that are currently doing research on and practically constructing dictionaries of various kinds. This remark merely reinforces the stereotype of the metalexicographer as the cynical

critic of dictionaries and their creators and who in his/her self- indulgence in matters theoretical has no sense at all of the practicalities of dictionary making. Such stereotyping does an injustice to both the contribution of the metalexicographer and the so-called "practicing" (or "working" lexicographer) to the field of dictionary making.

Their experience of practical lexicography often makes lexicographers the best people in the field to tackle the problems of improving dictionaries, but that is only possible if they are indeed prepared to keep abreast of theoretical developments in the fields that are relevant to lexicography.

Given the time constraints the lexicographer works under, however, the metalexicographer is by definition the one to fill in for the practical lexicographer in this regard. It is, after all, the task of the metalexicographer to not merely act as a kind of broker on matters theoretical, but also to formulate sound theoretical solutions to the practical problems that the lexicographer experiences. Discussions of theoretical or metatheoretical issues form part and parcel of such research, and are not merely manifestations of the "self-indulgence" of a batch of theorists (cf. Atkins 1992/1993:24). Most researchers in the field are by now acutely aware of the "harsh" world of practical dictionary-making (Atkins 1992/1993:25), of the tight constraints placed on the lexicographer by production procedures and schedules, limitations of dictionary space, lack of time and personpower, and the dictates of the market place (cf. Atkins 1992/1993:5).

Researchers in the field have demonstrated the need, and indeed the immense opportunities and results of cooperation between specialists within all the fields that are of immediate importance to practical lexicography. The kind of research problems in lexicography has itself drawn various other experts to the field, giving birth to the Janus-faced type of researcher in whose body the linguist, lexicographer and/or computer expert lives in symbiosis (cf. Zgusta 1992/1993:133). Researchers like Beryl Atkins, Nicoletta Calzolari, Charles Fillmore, Dirk Geeraerts, and James Pustejovsky are typical examples.

A number of the researchers have been instrumental in the development of tools for lexicographers to automatically and semi-automatically select lexicographical evidence from large corpora (cf. Boguraev & Brisco (Eds.) 1989), or in the development of systems in which the required information has been or will be prepackaged for them, as in the proposed semantically coded lexicon of the Cambridge Language System (cf. Proctor 1994).

This cooperation is, however, not limited to the devising of solutions to lexicographic problems, it also entails highlighting the problems themselves. An example of this is the exposure by computational lexicographers (in their efforts to re-use and revamp existing dictionaries in machine-readable form for NLP-applications) of the lack of systematicity and incompleteness in a number of popular dictionaries, and of the many ways in which the lexicographer falls back on the user's ability to fill in on inadequate

dictionary entries (especially in sense descriptions) (cf. especially Boguraev & Briscoe (Eds.) 1989).

3. Popular semantic theories, lexical semantic problems, and methodologies

After this general introduction, I would like to take a closer look at some of the lexical semantic problems that are tackled in the papers in this theme cluster, some of the more popular lexical semantic theories that researchers fall back on and the way this links up with the research reported on in some recent published literature.

The past few years have seen the development and entrenchment of a research methodology within the field of lexicography that comprises any two or more of the following phases:

(i) identification of a lexicographical (semantic) problem;
(ii) analysis of the problem within the framework of one or more linguistic frameworks (hypotheses, methods, data); and/or
(iii) a lexicographic perspective on the problem, i.e. an analysis of current and past approaches (assumptions, methods, procedures, techniques) to the problem within lexicography itself;
(iv) presentation of a solution to the research problem from (ii) and/or (iii) and
(v) presentation of a motivation for the proposed solution either in terms of the theoretic assumptions, methods and data presented in (ii) and/or (iii) or in terms of any other relevant considerations such as the dictionary type, constraints on space, user's information needs or research findings within other relevant fields (e.g. information science or cognitive psychology).

The article "Anatomy of a verb entry" by Atkins, Kegl and Levin (1988) is, to my mind, a prime example of this methodology in as much as the authors clearly illustrate the lack of systematicity in the treatment of certain systematic relationships in the semantics and syntax of the English verb system (illustrated specifically by *bake*); they back up this claim with a detailed analysis of current dictionary practice the inadequacies of which (e.g., omissions, lack of systematicity, and problematic encodings) become apparent against the backdrop of the linguistic facts as is illustrated and evidenced by corpus data. Their own suggestions for the improvement of the lexicographic description of verbs involves (i) the description of all verbs according to a "dictionary neutral" template of the semantic–syntactic relationships of verbs on which lexicographers can make principled decisions about the content and presentation for specific dictionaries, and (ii) suggestions for the encoding of these systematic relationships in dictionary entries, with due consideration of the restrictions of selection and

encoding set by dictionaries in their overall design policies.

3.1 Theories

Although the lexicon has come to occupy a central role in theoretical linguistics (cf., for example, Tomaszczyk & Lewandowska–Tomaszczyk 1990 and Harras, Hass & Strauss 1991), we have as yet

(i) no generally accepted grand theory of lexical (semantic) competence that makes it possible to account for the complexities, generalizations, idiosyncrasies of all linguistic, especially semantic, properties of lexical items,

and

(ii) no globally coherent theory of the ideal semantic representations for all classes of word category types for the various kinds of lexica (cf. Pustejovsky 1991:414).

We do have, however, a number of lexical (semantic) theories around, of which most concentrate on some, but not all lexicographically relevant aspects of lexical semantics. As will be discussed later, a large amount of the lexicographic research is aimed at devising the ideal kinds of lexical entries for different dictionary types.

In the papers in this theme cluster the authors rely on the assumptions and analytical tools of a number of theoretical frameworks, such as the frame–based semantics of Fillmore and Atkins (1992), the generative lexicon of Pustejovsky (1991), the conceptual semantics of Jackendoff (1991), the prototype semantics of Lakoff (1987), Dik's Functional grammar (cf. Dik 1989), the set of Lexical Functions developed in the work of Mel'chuk (cf. Mel'chuk & Zholkovsky 1988), and a number of relational models for the representation of semantic knowledge to enable NLP systems to carry out some limited language processing application .

Some of these theories are tailored to the analysis of some types of words, some specific aspect of the meaning of lexical items or some aspect of the overall semantic structure of the vocabulary of a language, but they are simply inadequate for the analysis of all word types and the semantic structure of the vocabulary as a whole. In a number of cases, therefore, researchers make use of a variety of theoretical and analytical tools to account for the semantics of lexical items (cf. also Lehrer 1992). Lexicographers often use the analytical tools at their disposal eclectically, without necessarily committing themselves to the basic tenets of the theoretical frameworks in which these tools are embedded. In most cases,

empirical adequacy, i.e. the ability of the theory to cover the empirically relevant semantic data/facts, is of primary concern.

In a number of cases lexicographers, linguists and computational experts have had to develop, supplement or bring some refinement to the existing analytical tools/semantic metalanguage. Examples of this are the development of an inventory of frame specific role categories that go far beyond the familiar types of Agent and Patient, (cf. for example Fillmore and Atkins 1992), the development of complex semantic features sets, like the qualia structure proposed by Pustejovsky (1991), the fine-grained analysis of lexical semantic relations (cf. Beckwith et al. 1991:212 and *International Journal of Lexicography*, 3(4) 1990) and inheritance principals (cf. Chaffin 1992; Cruse 1992; Evens 1988; Mel'cuk & Zholkovsky 1988; Proctor 1994; Pustejovsky & Anick 1993), and the enrichment of semantic field theories with the frames, schemata, or idealized cognitive models that structure such fields (cf. Croft 1993; Lakoff 1987,and especially the contributions in Lehrer and Kittay (Eds.) 1992).

Theoretical pluralism in the approach to specific semantic problems has much to recommend it if such an approach aids one in getting a clearer and fuller understanding, analysis and description of a specific semantic phenomenon. Higashimori ("Cognition, synonymy and definitions"; this volume) illustrates this in his analysis of the synonymy of particles (e.g. *huh*, *eh*) within the framework of the Relevance Theory of Sperber and Wilson (1986), the synonymy of prepositions within the prototype theory of Lakoff (1987) and that of verbs within Langacker's variant of Cognitive grammar (cf. Langacker 1990).

3.2 Lexical semantic problems

At the macrostructural level semantic theory has to assist the (computational) lexicographer or computational linguist in making decisions on which lexical items to handle as cases of monosemy, polysemy or homonymy. But semantic theory has the most to contribute to dictionary making at the level of the microstructural analysis and description of lexical meaning where it has to offer solutions to a number of vexing questions. Some of these seem to be eternal debating points (cf. Zgusta & Farina 1988 for an overview), which underscores the fact that despite the number of lexical semantic theories around there are still no straight and simple answers to a number of complex questions, and that the lexicographer often has to rely on his/her own judgement in the absence of simple guidelines.

A number of papers in this theme cluster addresses aspects of the problem of polysemy. Some of the general questions concerning polysemy for which lexicographers are looking for answers, are the following:

What should count as a dictionary sense?
On what basis should the senses of a lexical item be differentiated?

How many and which senses should be distinguished for a specific lexical item?
In which cases should a sense be assigned to another existing sense and in which cases should it be considered a sense on its own? (Or: When do you lump together into one general sense and when do you split off into more specific meanings?)
What are the possible relationships between the senses of polysemous lexical items, and should these relationships be described in dictionaries?
How should predictable meaning shifts ("regular polysemy") be handled in dictionaries?
How do you map the multidimensionality of polysemy relationships onto the two–dimensional structure of the traditional dictionary entry. Or: What is the most appropriate sense structure for the dictionary type: a "flat" structuring in which all senses have equal status or a hierarchical structuring with the embedding of related senses. If the choice falls on a hierarchical structuring, the question is: How can we decide what depth of "nesting" is appropriate? (cf. Atkins 1992/1993:18,20–21 and Geeraerts 1990 in which some of these issues are discussed).

De Regt ("The description of multiple meaning in some Biblical Hebrew lexicographical projects"; this volume) contributes in his paper to the lumping x splitting problem. The author shows that in a number of Biblical Hebrew dictionaries no distinction is made between so–called "linguistic" and "contextual" senses of words and that contextual usages are recorded as separate senses. He proposes that contextual senses that belong to the same semantic domain be lumped together and that the syntagmatic characteristics of a word in each of its attested senses be supplied more fully.

In her paper "Ideological polysemy in Afrikaans dictionaries" (this volume) Carstens explores the nature of "ideological polysemy". Ideological senses manifest themselves as peripheral sense variations that are associated with specific ideological frames. Descriptions of ideological sense distinctions can be analyzed in terms of their peripheral descriptive values, valuations and group foci.

A number of researchers tackle the problem of so–called "regular" or predictable polysemy. According to Atkins (Atkins 1991:178; cf. also Ostler & Atkins 1992) there is as yet

(i) no comprehensive description of the phenomenon of regular polysemy (e.g. the tr. and intr. use of some verbs, and classes of metonymic extensions of the type container–contained), or
(ii) any set of criteria for its inclusion and description in dictionaries which could be systematically applied during dictionary compilation.

Rozina ("On verbs with completely affected goal"; this volume) contributes to the first of these issues with an analysis of the pattern of

regular polysemy "action–state–process" of the Russian equivalents of the verb "to cover/to fill". As in most research on this topic, her primary aim is to pinpoint the meaning components of this class of verbs that license the polysemy of the mentioned type (cf. also Ostler & Atkins 1992).

Bouillon and Viegas ("A semi–polymorphic approach to the interpretation of adjectival constructions: a cross–linguistic perspective"; this volume) tackle the problem of the polysemy of adjectival constructions within the framework of Pustejovsky's model of the Generative Lexicon (cf. Pustejovsky 1991). Within this model part of the polysemy of the adjective is spread to the accompanying nouns of the adjective so that the polysemous senses of Adj.+ N constructions are a function of the way the semantics of the adjectives interact with the qualia structure of the nouns. The authors' analysis underscores Pustejovsky's theory (cf. Pustejovsky 1991:422–424) that systematic ambiguity, of which regular polysemy or logical polysemy is an instantiation, can be accounted for in terms of principles of semantic composition, such as that of co–specification, co–compositionality and type coercion, rather than in terms of the semantics of individual lexical items themselves. (Cf. in this regard also Taylor 1992(a) and 1992(b) for an analysis of these phenomena within the framework of cognitive grammar.)

Both these contributions tie in with the search for general principles according to which the senses of polysemous items can be predicted (cf., however, Lehrer 1989). In the design of adequate lexicons for NLP systems such principles are of course of primary importance as they facilitate the economic representation of lexical semantic knowledge with the aid of inheritance relations that link the semantics of individual lexical items with multidimensional semantic networks (cf. Proctor 1994, Pustejovsky & Anick 1993).

In the case of traditional dictionaries neither the systematic nor the comprehensive encoding of regular polysemy seems to be a prerequisite. Ilson (1990:130) notes that the principle of consistency would require that each predictable polysemous sense of the lexical items of a class be given in the dictionary entry for each member of the set. But doing so wherever possible "would swell real dictionaries to an unmanageable size, with material much of which would be of doubtful practical use". Ilson notes further that it is not "necessary to deal even partially with *all* types of semantic regularity just because you have decided to deal with some" (1990:131). More specifically, he suggests (i) that predictable senses only be given if they occur frequently (frequency based on the lexicographer's instinct, backed up by corpus evidence on frequency of usage), especially in EFL–dictionaries where more attention is given to the variety of uses of the most common words, or if the dictionary user is likely to want to produce them; (ii) that they be omitted from learner's dictionaries if they are of a universal kind; but that they be treated with language specific regularities in dictionaries for the mother tongue speaker (if they are not treated elsewhere).

Once one has decided what any dictionary is going to be used for, the second problem is always that of acquiring the necessary data for the compilation of such a dictionary (cf. Evens 1988:9). Computational lexicographers/linguists (cf., for example, Boguraev & Briscoe (Eds.) 1989) have devised a number of computational tools for the acquisition of lexicographic data for this purpose from existing dictionaries in machine readable form.

Peters, Federici, Montemagni and Colzolari ("From machine readable dictionaries to lexicons for NLP; the COBUILD dictionary: a different approach"; this volume) address the acquisition problem. They discuss the development of a syntactic-semantic-parser to generate representations of the semantic attributes of lexical items from their definitions in the COBUILD dictionary. These representations contain, amongst others, information on the genus terms of lexical items and on the semantic features of their preferred collocates. Eck, Skuce and Meyer ("Definitions in the document-production process: exploring the potential of knowledge engineering technology"; this volume), on the other hand, come up with a number of tools that can be used to capture the conceptual characteristics of specialized lexical items (terms) from the documents that are normally produced by technical writers.

The use of machine readable dictionaries for the construction of large scale lexical knowledge bases for NLP systems has uncovered two major problem areas, viz.

(i) the lack of comprehensiveness of description, i.e. the empirical inadequacy of dictionary entries with respect to both the selection of types of semantic data and the comprehensive description of each type;
(ii) the lack of systematicity in lexical semantic description, i.e. the lack of equal treatment of individual members of a specific syntactic and/or semantic class with regard to the kind of semantic information provided, depth of coverage and consistency in encoding techniques.

With regard to the latter, systematicity would demand that the entries for members of the same class are treated uniformly with respect to syntactic coding, type of definition wording, and patterning of examples (cf. Atkins, Kegl and Levin 1988:110).

Ensuring a systematic (consistent) and exhaustive presentation of the linguistic facts of dictionaries is a task that is tackled on various fronts. The most important of these at present, as is also reflected in the reported research in this theme cluster, are

(i) the identification of the various lexical semantic features of various word categories that motivate their lexical syntactic properties

and

(ii) devising ideal templates for the semantic description of semantic–syntactic vocabulary types or lexicographic types (cf. Apresjan 1992/1993) (e.g. words denoting emotions, natural kinds, verbs of perception, verbs of sound, change of state verbs, creation verbs etc.)

The research being done in these two areas often overlap, as most of the proposed templates contain a slot for a description of the mapping of a lexical item's semantic characteristics onto its syntactic environment (cf., e.g., Atkins, Kegl & Levin 1988:104; Jackendoff 1991; Levin 1991; Levin & Rappaport 1991; Nirenburg & Levin 1992; Pustejovsky 1991).

Research has shown that there is a systematic relationship between the semantics and syntax of word categories like those mentioned above in that the semantic characteristics make it possible to predict their syntactic properties, i.e. the syntactic realization of its arguments and its selectional restrictions, but also some of their extended uses (cf. the discussion of regular polysemy above). Atkins and Levin (1991) show, for example, that the knowledge that *bake* in one of its uses is a change of state verb allows for the prediction that it will participate in the causative/inchoative alternation; being a verb of creation leads to the predictability of its participation in the benefactive alternation.

The article by Fillmore & Atkins (1992) on the semantics and syntax of the *risk*–group is illustrative of this kind of research. The syntactically relevant semantic characteristics/roles are derived from the analysis of the frame, i.e. a structured body of knowledge, to which all the words in a set are linked. This analysis provides a set of structured semantic categories in terms of which the meanings of individual words can be explicated and in terms of which it can be shown how different words related to a frame highlight or activate different parts of it. In their proposed dictionary the mapping between the semantics and the syntax of a lexical item of the set is captured in a "valence description", i.e a description that specifies, in both semantic and syntactic terms, what a lexical item requires of its constituents and its context, and what it contributes to the structures that contain it.

As Fillmore & Atkins' analysis of the *risk*–frame illustrates, identifying the members of vocabulary sets, determining the content and structure of the identified frames and identifying the syntactically relevant semantic characteristics requires a large amount of ingenuity and hermeneutical skills of the analyst. Alonge ("Motion verbs: data on meaning components in dictionaries and identification of syntactic properties"; this volume) shows, however, that the relevant semantic characteristics are, as in the case of

motion verbs in Italian, the genus terms (MOTION) and the categories of differentiating meaning characteristics (e.g. MANNER, DIRECTION, PATH, MEDIUM of MOTION) that are supplied in dictionary definitions of this class. Dini ("*Più* and dictionary syntax: a case study"; this volume) gives special attention to the way dictionary definitions miss the generalizations on certain categories of words and simply ignore the constraints placed on the use of these words. This fact stresses the point that lexicographers should be alert to the kinds of generalizations on linguistic data that theoretical linguistics has to offer them in their efforts to give a comprehensive description of lexical items.

To enhance the comprehensive and systematic description of the semantics of lexical items a number of researchers has focused on the semantic characteristics of specific lexicographic types/words and on devising templates for the adequate description of each member of such sets. The main aim, is of course to identify the semantic characteristics that define the class, in terms of which the sense(s) of individual members of the set can be adequately described, and in terms of which the members can be differentiated from each other. Such sets of characteristics are also used, however, for setting up semantic hierarchies which function as multiple inheritance networks in NLP–systems.

Atkins and Levin (1991:244–249), for example, suggest the following template of semantic characteristics for the description of a verb of sound from which information for a dictionary entry can be selected:

(i) genus: "to sound"/ "to produce a sound" (indication of meaning shared by hyponyms)
(ii) dif(ferentiae) (to differentiate co–hyponyms):
  (a) the type of sound produced (loud, soft, low, high, etc)
  (b) the manner in which the sound is produced (by blowing, impact, vibration, etc)
  (c) person or device producing the sound
  (d) instrument used or not (+/– instrument)
  (e) selectional restrictions on subjects along the dimensions of animate, human and concrete)
  (f) subcategorization (complements to the verbs)

In this kind of research, attention is also given to the values associated with each semantic characteristic and the way these characteristics are structured relative to each other (as for example, in a script or frame) (cf. Apresjan 1992/1993; Swanepoel 1992(b); Fillmore & Atkins 1992).

These templates are meant to serve as a basis for the comprehensive and systematic descriptions of the semantics of individual members of lexical sets. The systematicity of such descriptions is of course enhanced if explicit indications are also given of how these differentiating characteristics (and their values) should/could map onto dictionary definitions).

A number of the papers in this theme cluster (cf. this volume) focus on the crucial (language–specific) semantic characteristics of particular vocabulary classes: Viberg ("Differentiation and polysemy in the Swedish verbal lexicon") analyses the verbs of physical contact, Braasch ("There's no accounting for taste – except in dictionaries...") the class of audition verbs, Monachini and Roventini ("Italian audition verbs: a corpus– and frame–based analysis") some verbs of perception, and Kirchmeier–Andersen, Pederson and Schosler ("Combining semantics and syntax in monolingual dictionaries") the set of motion verbs.

A number of researchers propose a multi–leveled representation format or template for the description of the semantics of classes of lexical items (cf. also Pustejovsky 1991). Delor, Antonín and Masalles ("Towards a "VROS" representation"; this volume), for example, present a proposal for a description of the semantics of verbs (for a multilingual lexical knowledge base) on the following levels: argument structure, event structure, selectional restrictions and compositional semantics. Some, though not all of this information, is also captured in Oppentocht's proposal ("Towards a lexical semantic model for the creation of NLP and human–friendly definitions"; this volume) for the description of verbs in a definition format that is suitable for use in both NLP–systems and for the construction of dictionaries for human users. Van der Vliet ("Conceptual semantics for nouns"; this volume) proposes a multileveled semantic representation for nouns and focuses in his paper on aspects of the contents and structure of the conceptual level. A representational format for terminological knowledge that meets the requirements of NLP–systems is put forward in the paper by Schütz and Ripplinger ("Controlling NLP through terminological information"; this volume).

Geeraerts's paper ("Varieties of lexical variation"; this volume) complements the research on lexical fields in as much as he highlights the various kinds of lexical variation that can occur in a semantic field and the factors that influence lexical choices.

Heyvaert ("The edges of definition"; this volume) approaches the question of the adequate description of lexical items from another perspective, viz. the influence that the choice of a specific theory has on what would count as an adequate definition/lexicographic description. In particular, he explores what kind of information descriptions would have to contain if one bases one's description on the "meaning is use"–theory.

Comprehensiveness and systematicity of descriptions are, of course, scalar concepts of which the specific value is determined, amongst others, by the aim or purpose of the semantic description. Kilgarriff ("The myth of completeness and some problems with consistency"; this volume) argues that in the case of dictionaries for human users, the importance of lexical items, measured in terms of their frequency in corpora, is another determining factor – more important words require more extensive

treatment (cf. also the discussion about the treatment of regular polysemy above).

4. Conclusion

As should be obvious from the foregoing, the papers on lexical semantics each make a very specific contribution to current metalexicographic research aimed at the development of adequate dictionaries, be they for NLP–systems or for human use. The need for a multidisciplinary effort to achieve these aims speaks for itself.

The computer systems and tools that are becoming available both to the researcher, the practical lexicographer and the human user are opening up a myriad of new possibilities for the presentation and utilization of masses of lexicographic information. The proposed "frame–based" dictionary of Fillmore & Atkins (1992) with its multiple windowing capabilities seems an important step forward, especially in the design of learner's dictionaries, as it will supply the user with the normal kind of semantic information and the frames, i.e. the cognitive structures, against which individual concepts are defined.

Developments such as these underline the need for research on new formats and models for the representation and retrieval of the vast amounts of knowledge associated with lexical items.

These developments also stress the need for adequate theories of human lexical knowledge. Prototype semantics is seen by a number of researchers as a viable theoretical framework for the exploration of such innovations as it incorporates fine–grained subtheories on a number of lexical semantic phenomena like cognitive models (cf. Lakoff 1987), polysemy (cf.Geeraerts 1990), syntagmatic modulation (cf. Taylor 1992(a) and 1992(b)), synchronic semantic chaining and the mechanisms of lexical semantic change (cf. Lakoff 1987; Langacker 1987 and 1990; Moerdijk 1989; Norvig and Lakoff 1987; Goossens 1990), semantic perspectivization (cf. Langacker 1987; Croft 1993), semantic motivation (cf. Swanepoel 1992(a)), and onomasiological salience in a conceptual domain/lexical field (cf. Geeraerts 1993 and this volume).

The specific problems of NLP–systems do, however, require theoretical developments of their own. One of the main points made by the research on the mapping of lexical semantics onto syntax is that the relationship between meaning and syntactic form is not arbitrary, nor always predictably systematic, but motivated. Making these motivating factors explicit is going to be a major task. Preliminary research on the disambiguation problem for NLP systems (cf., e.g., McRoy 1992) shows that lexical meaning is, for example, motivated by a number of other linguistic and non–linguistic knowledge structures. The exact nature of these mappings and their motivated character still has to be determined and the results translated into dictionary content and structure.

Note

* I would like to thank the following colleagues fot theircomments on the first draft of this paper : Willy Martin, Rosandre Hendriksen, August Cluver, Hilton Hubbard and Rufus Gouws. The usual disclaimer applies.

References

Apresjan, Jurij, D. 1992/1993. "Systemic Lexicography as a Basis of Dictionary–making". *Dictionaries* 14:79–87.
Atkins, B.T.S. 1991. "Building a Lexicon: The Contribution of Lexicography". *International Journal of Lexicography* 4.3:167–204.
Atkins, B.T.S. 1992/1993. "Theoretical Lexicography and Its Relation to Dictionary–making". *Dictionaries* 14:4–43.
Atkins, B., J. Kegl & B. Levin. 1988. "Anatomy of a verb entry". *International Journal of Lexicography* 1:84–126.
Atkins, B. & B. Levin. 1991. "Admitting Impediments". In: Zernik (Ed.) 1991:233–262.
Beckwith, R. et al. 1991." WordNet: A Lexical Database Organized on Psycholinguistic Principles". In: Zernik (Ed.) 1991: 211–232.
Boguraev, B. & T. Briscoe. (Eds.) 1989. *Computational Lexicography for Natural Language Processing.* London: Longman.
Chaffin, R. 1992. "The Concept of a Semantic Relation". In: Lehrer & Kittay (Eds.) 1992:253–288.
Croft, W. 1993. "The role of domains in the interpretation of metaphors and metonymies". *Cognitive Linguistics* 4–4:335–370.
Cruse, D. A. 1992. "Antonymy Revisited: Some Thoughts on the Relationship Between Words and Concepts". In: Lehrer & Kittay (Eds.) 1992:289–306.
Cruse, D.A. 1992/1993."On Polylexy". *Dictionaries* 14:88–89.
Dik, S.C. 1989. *The theory of Functional Grammar Part 1: The structure of the clause.* Dordrecht: Foris.
Evens, M.A.(Ed.) 1988. *Relational models of the lexicon.* Cambridge: Cambridge University Press.
Fillmore, C.J. & B.T.S. Atkins. 1992. "Toward a Frame–Based Lexicon: The Semantics of RISK and its Neighbours". In: Lehrer & Kittay (Eds.) 1992:75 –102.
Frawley, William. 1992/1993."Introduction". *Dictionaries* 14: 1–3.
Geeraerts, D. 1989."Introduction: Prospects and Problems of Prototype Theory". *Linguistics* 27:587–612.
Geeraerts, D. 1990."The Lexicographic Treatment of Prototypical Polysemy". In: Tsohatzidis (Ed.) 1990:195–210.
Geeraerts D. 1993."Generalized onomasiological salience". *Belgian Journal of Linguistics* 8:43–56.
Goossens, L. 1990. "Metaphtonymy: The Interaction of Metaphor and Metonymy in Expressions for Linguistic Action". *Cognitive Linguistics* 1–3:323–340.
Harras, G., U. Hass & G. Strauss. 1991. *Wortbedeutung und ihre Darstellung im Wörterbuch.* Berlin: de Gruyter.
Hausmann, F.J. et al. 1986."The International Encyclopedia of Lexicography. An Outline of the Project". *Lexicographica* 2:271–283.
Hyldgaard–Jensen, K. & A. Zettersten (Eds.) 1983. *Symposium zur Lexicographie.* Hildesheim: George Olms.
Ilson, R. 1990."Semantic Regularities in Dictionaries". In: Tomaszczyk & Lewandowska–Tomaszczyk (Eds.) 1990:123–132.
Jackendoff, R. 1991. *Semantic Structures.* Cambridge: MIT Press.
Lakoff, G. 1987. *Woman, Fire and Dangerous Things: What Categories Reveal about the Mind.* Chicago: University of Chicago Press.
Langacker, R..W. 1987. *Foundations of Cognitive Grammar. Volume 1: Theoretical Prerequisites.* Stanford: Stanford University Press.
Langacker, R.W. 1990. *Concept, Image, and Symbol: The Cognitive Basis of Grammar.* Berlin: De Gruyter.
Lehrer, A. 1989. *Polysemy, Conventionality and the Structure of the Lexicon.* Duisberg. LAUD Series A240.

Lehrer, A. 1992."A theory of vocabulary structure: Retrospectives and prospectives". In: Pütz (Ed.) 1992:243–256.
Lehrer, A. & Eva Feder Kittay (Eds.) 1992. *Frames, Fields and Contrasts. New Essays in Semantic and Lexical Organization.* Hillsdale: Lawrence Erlbaum.
Levin, Beth. 1991. "Building a Lexicon: The Contribution of Linguistics". *International Journal of Lexicography* 4.3:205–226.
Levin, Beth & Malka Rappaport Hovav. 1991."Wiping the slate clean: A lexical semantic exploration". *Cognition* 41:123–151.
McRoy, S.W. 1992. "Using Multiple Knowledge Sources for Word Sense Discrimination". *Computational Linguistics* 18.1:1–30.
Mel'cuk, I. & A. Zholkovsky. 1988."The explanatory combinatorial dictionary". In: Evens (Ed.) 1988:41–74.
Moerdijk, A. 1989."Benaderingen van metonymie". *Forum der Letteren* 30.2:115–134.
Nirenberg, S. & L. Levin 1992."Syntax-Driven and Ontology-Driven Lexical Semantics". In: Pustejovsky & Bergler (Eds.) 1992:5–20.
Norvig, P. & G. Lakoff. 1987."Taking: A Study in Lexical Network Theory". *Proceedings of the 13th Annual Meeting of the Berkeley Linguistic Society*: 195–212.
Ostler, N. & B.T.S. Atkins. 1992."Predictable Meaning Shift: Some Linguistic Properties of Lexical Implication Rules". In: Pustejovsky & Bergler (Eds.) 1992:87–100.
Proctor, M. 1994. "Semantic Coding. The First Scientific Support System for the Lexicography of the Future". *Cambridge Language Reference News* 1:4.
Pustejovsky, J. 1991."The Generative Lexicon". *Computational Linguistics* 17:409–441.
Pustejovsky, J. & P. P. Anick. 1993."Lexical Semantic Techniques for Corpus Analysis". *Computational Linguistics* 19.2:331–358.
Pustejovsky, J. & S. Bergler (Eds.) 1992. *Lexical Semantics and Knowledge Representation. First SIGLEX Workshop. Berkely, CA, USA, June 1991. Proceedings.* Berlin: Springer-Verlag.
Pütz, M. (Ed.) 1992. *Thirty Years of Linguistic Evolution.* Philadephia/Amsterdam: John Benjamins.
Sperber, D. & D. Wilson. 1986. *Relevance: Communication and Cognition.* Cambridge, Mass.: Harvard Press.
Swanepoel, P.H. 1992a. "Linguistic Motivation and its Lexicographical Application". *South African Journal of Linguistics* 10(2):49–60.
Swanepoel, P. H. 1992b. "Getting a Grip on Emotions. Defining Lexical Items that Denote Emotions". In: Tommola et al. (Eds.) 1992:419–432.
Taylor, J. 1992a."How many meanings does a word have?" *Spil* 25:133–168.
Taylor, J. 1992b."Old problems: Adjectives in Cognitive Grammar". *Cognitive Linguistics* 3–1:1–35.
Tommola, H. et al. (Eds.) 1992. *EURALEX '92. Proceedings 1–11. Papers submitted to the 5th EURALEX International Congress on Lexicography in Tampere, Finland.* Tampere: University of Tampere.
Tomaszczyk, J. & B. Lewandowska–Tomaszczyk (Eds.) 1990. *Meaning and Lexicography.* Amsterdam: John Benjamins.
Tsohatzidis, S.L. (Ed.) 1990. *Meanings and Prototypes. Studies on Linguistic Categorization.* London: Routledge.
Wiegand, H.E. 1983."Überlegungen zu einer Theorie der lexicographischen Sprachbeschreibung". In: Hyldgaard-Jensen & Zettersten (Eds.) 1983:35–72.
Wierzbicka, Anna. 1985. *Lexicography and Conceptual Analysis.* Ann Arbor: Karoma.
Wierzbicka, A, 1990. "'Prototypes save': On the Uses and Abuses of the Notion 'prototype' in Linguistics and Related Fields". In: Tsohatzidiz (Ed.) 1990:347–368.
Zernik, U. (Ed.) 1991. *Lexical Acquisition: Exploiting On-Line Resources to Build a Lexicon.* Hillsdale: Lawrence Erlbaum.
Zgusta, L. 1971. *Manual of Lexicography* . The Hague: Mouton.
Zgusta, L. 1992/1993. Lexicography, Its Theory, and Linguistics. *Dictionaries* 14:130–138.
Zgusta, L. & D.M.T.Cr. Farina. 1988. *Lexicography Today: An Annotated Bibliography of the Theory of Lexicography.* Tübingen: Max Niemeyer.