Ten Choices for Lexical Semantics

Sergei Nirenburg and Victor Raskin[1]
Computing Research Laboratory
New Mexico State University
Las Cruces, NM 88003
{sergei,raskin}@crl.nmsu.edu

Abstract. The modern computational lexical semantics reached a point in its development when it has become necessary to define the premises and goals of each of its several current trends. This paper proposes ten choices in terms of which these premises and goals can be discussed. It argues that the central questions include the use of lexical rules for generating word senses; the role of syntax, pragmatics, and formal semantics in the specification of lexical meaning; the use of a world model, or ontology, as the organizing principle for lexical-semantic descriptions; the use of rules with limited scope; the relation between static and dynamic resources; the commitment to descriptive coverage; the trade-off between generalization and idiosyncracy; and, finally, the adherence to the "supply-side" (method-oriented) or "demand-side" (task-oriented) ideology of research. The discussion is inspired by, but not limited to, the comparison between the generative lexicon approach and the ontological semantic approach to lexical semantics.

It is fair to say that lexical semantics, the study of word meaning and of its representation in the lexicon, experienced a powerful resurgence within the last decade.[2] Traditionally, linguistic semantics has concerned itself with two areas, word meaning and sentence meaning. After the ascendancy of logic in the 1930s and especially after "the Chomskian revolution" of the late 1950s-early 1960s, most semanticists, including those working in the transformational and post-transformational paradigm, focused almost exclusively on sentence meaning. The modern movement known as lexical semantics restores word meaning as the center of interest, without, however, reviving the theoretical perspective, i.e., reasoning about the nature of meaning, or methods, e.g., lexical fields, componential analysis of word meaning, etc., of their predecessors.[3]

The major theoretical thrust in lexical semantics has shifted toward generativity, possibly, with a view of extending the generative approach to grammar into the lexicon, that is, toward seeking ways of grounding lexical semantics in generative syntax. There is a thirty-year-old tradition in transformational syntax of "encroaching" into the domain of semantics, often without recognizing or admitting it, usually in order to make some lexical features work for syntax. Neither Chomsky's

[1] Also of the Interdepartmental Program in Linguistics and its Natural Language Processing Laboratory, Purdue University, West Lafayette, IN 47907-1356.

[2] See, for instance, Frawley and Steiner (1985), Benson *et al*. (1986), Cruse (1986), *Computational Linguistics* (1987), Boguraev and Briscoe (1989), B. Levin and Pinker (1991), Boguraev (1991), Zernik (1991), Pustejovsky (1991, 1993, 1995), Pustejovsky and Bergler (1992), Briscoe *et al.* (1993), Dorr (1993), Saint-Dizier and Viegas (1995), Dorr and Klavans (1994/1995, 1995), Viegas *et al*. (1996a).

[3] See, for instance, Dolgopol'skiy (1962), Garvin *et al*. (1967), Goodenough (1956), Greenberg (1949), Kroeber (1952), Lounsbury (1956), Masterman (1959), Vinogradova and Luria (1961), Shaykevich (1963), Trier (1931), Weisgerber (1951), Zholkovsky *et al*. (1961).

"strict subcategorization" of nouns, involving a handful of such features as ANIMATE, HUMAN, ABSTRACT, etc. (Chomsky 1965), nor Gruber's (1965/1976) nor Fillmore's (1968--see also 1971 and 1977) cases, nor the even more feeble involvement with lexical semantics on the part of lexical functional grammar (Bresnan 1982) amounted to or were intended to be significant contributions to lexical semantics. One difference between these earlier excursions from generative syntax into semantics and the current ones is that the latter do profess--and often show--an interest in at least some elements of word meaning *per se*.

A different approach, which may be referred to as 'ontological semantics,' stresses the unity of word, sentence and text semantics, striving to formulate a comprehensive semantic theory which, while relying on syntactic clues in its workings, is not closely coupled with a theory of syntax. This approach seeks to develop the semantic lexicon as a static knowledge source for the dynamic task of deriving text meaning representation, which, on this view, is the task of sentence and text semantics. While lexical semantics concentrates on distinguishing and relating word senses, ontological semantics stresses the task of fully describing those senses individually as well as relating them to each other, on the basis of an underlying model of the world, or ontology.

The lexical semantics of the 1920s-early 1960s concentrated on the notion of meaning itself, but it was largely marginalized by the then prevalent asemantic structuralist dogma and ghettoized within a largely atheoretical traditional framework. Towards the end of the structuralist era and within its more "liberated" offshoots, some theoretical developments in lexical semantics took place within systemic linguistics (see, for instance, Halliday 1961, 1983, 1985, Berry 1975, 1977, Winograd 1983: 272-356) and the Meaning-Text school (see, for instance, Zholkovsky *et al*. 1961, Apresyan *et al*. 1969, 1973, Mel'čuk 1974, 1979). Katz and Fodor (1963), though working in a predominantly syntactic generative paradigm, were also interested in the meaning of words, especially for the purposes of disambiguation and selection restriction in compositional semantics. The contributions triggered by the appearance of Katz, Fodor, and Postal's work (e.g., Weinreich 1966) contained discussions of important issues in lexical semantics, e.g., polysemy bounds, and this tradition is alive and well today (e.g., the work of Cruse 1986, Talmy 1983, Fillmore 1978, and others). Early work on language within artificial intelligence (Charniak and Wilks 1976, Schank 1975, Schank and Colby 1973, Schank and Riesbeck 1981, Wilks 1975a,b) also paid attention to word meaning. It is remarkable how little of this body of work is reflected or referenced in some of the present-day approaches to lexical semantics.

The lexical semantics community is quite heterogeneous. It includes former syntacticians, reformed and dyed-in-the-wool formal semanticists, and a considerable chunk of the NLP community, including both linguists and computer scientists, some with a strong statistical and probabilistic bent, others still pursuing the use of large corpora and MRDs, some actively engaged in developing NLP systems and some still working on designs for the systems of immediate future.

As communities tend to do, we argue about much the same examples and issues, get excited about promises of new methods and follow the ever changing scientific fashions and vicissitudes of funding. Nevertheless, we have different premises and goals; different views of theory, methodology, and applications; different limits and constraints as well as notions of feasibility; and different reasoning methods. We often use the same terms differently and different terms to refer to similar concepts. We also have very different outputs.

It is probably still fair to say that we are all capable of recognizing a good lexical entry in any guise and paradigm, though we have seen relatively few lexical entries, excellent or merely adequate. We

have seen many partial entries and concentrated much too often on some detail(s) in them and/or on their notation rather than their full content. Often, these entries are the results of almost accidental convergences among totally different approaches, and our terminological and substantive differences may follow from those differences, though we may be not always aware of that.

It seems, therefore, truly worthwhile, at least to us, to figure out similarities and differences among a variety of computational-linguistic lexicons in terms of the similarities and differences of the approaches that underlie those lexicons. This paper is a step in this direction. The task of comparison is difficult and we certainly do not hope that the results are in any way final or exhaustive.

This project was triggered in part by our reactions to Pustejovsky (1995), the first solely authored treatise on the new lexical semantics, but the paper promptly went well beyond the book and its merits and demerits. In spite of the frequent and convenient references to the book, which summarizes and represents well a popular approach to lexical semantics, it would be grossly unfair to read this paper as a review article of the book because we raise many issues that Pustejovsky did not concern himself with and we do not necessarily maintain that he should have done so. Instead, the result is an attempt to position our own, Pustejovsky's, and any other computational semantic approach within a grid of several major alternatives.

For this purpose, and fully aware of the losses of depth that this device brings about, we proceed by discussing a series of issues, expressed as dichotomies, each of which provides a possible opposition of views concerning lexical semantics, based on a single parameter. Unavoidably, we may have created a few strawmen on the way, but no position we will list is without well recorded support in the community. The dichotomies are as follows:

1. generative vs. non-generative
2. syntactic vs. semantic
3. pragmatic vs. non-pragmatic
4. semantic theory vs. formal semantics
5. ontological vs. non-ontological
6. scope-sensitive vs. non-scope-sensitive
7. static vs. dynamic
8. depth (and theory-dependence) vs. coverage of description
9. generalization vs. idiosyncraticity
10. supply-side vs. demand-side

We will see that the various workers in the field agree on some issues, disagree on others, and consider still others obvious non-issues. We will also see that some workers are tending to different fields and aspiring to different harvests, and that it is important for us to recognize that. Still, we believe that this exercise is useful. Though each of us could continue working within his or her chosen paradigm without looking outside (as this is easier, and there is no need to argue about tenets which are fully shared within one's own approach), putting one's work in the perspective of other work, at the very least, gives one an idea about the lacunae in one's program and suggests priorities for future development. It is also important to understand what can be realistically expected from an approach, one's own or anybody else's, depending on the choices the approach makes with regard to the ten issues mentioned above. We realize that we can only represent the opinions of our colleagues to the best of our ability and expect that some misunderstandings, especially of some minutiae, will inevitably creep in.

1. Generative vs. Non-Generative

1.1 Generative Lexicon: Main Idea

The main lexical-semantic idea behind the generative lexicon (Pustejovsky 1991, 1995) is basically true, if not novel. It is two-prong:

- senses of a polysemous lexical item can be related in a systematic way, with types of such relations recurring across various lexical items;
- by identifying these relations, it is possible to list fewer senses in a lexical entry and to derive all the other senses with the help of (lexical) rules based on these relations.

The paradigmatic (static, as opposed to syntagmatic, or context-determined) relations among word meanings have been explored and implemented in dictionaries of various sizes and for various languages by the members of the Meaning-Text school of thought since the mid-1960s (Zholkovsky *et al*. 1961, Apresyan et al. 1969, 1973, Mel'čuk 1974, 1974). These scholars vastly enriched the list of paradigmatic relations, similarly to the way it is done in generative lexicons, though the latter focuses only on those word meanings which are senses of the same lexical item. Even closer to the currently popular lexical rules, Givón (1967) and McCawley (1968: 130-132) came up with similar ideas earlier.

Our own experience in lexical semantics and particularly in large-scale lexical acquisition since the mid-1980s[4] also confirms that it is much more productive to derive as many entries as possible from others according to as many lexical rules as can be found: clearly, it is common sense that acquiring a whole new entry by a ready-made formula is a lot easier and can be done by less skilled acquirers than acquiring an entry from scratch. In fact, we have established a pretty reliable quantitative ratio: acquiring from scratch can be as slow as 1-2 entries an hour by the "master" lexicographer; acquiring via a lexical rule is up to 30 entries an hour by an inexperienced lexicographer (obviously, the former acquires lexical entry **types**, or templates, which are emulated by the latter--see, for instance, Raskin and Nirenburg 1995, 1996a,b, Viegas and Nirenburg 1996, and Viegas and Raskin 1996).

Of course, there are other motivating forces behind the generative lexicon theory, one of them being investigating the extension of the generative grammar methods and goals from the realm of syntax to the realm of lexical semantics. We will comment on this issue in Section 4.

Some claims made about the generative lexicon strike us as somewhat misleading and often spurious. These claims, however, do not seem essential to the concept of generative lexicon, and therefore, in what follows (Sections 1.2-1.4), we critically examine them, in the spirit of freeing a good idea of unnecessary ballast.

---

[4] See, for instance, Nirenburg *et al*. (1985, 1987, 1989, 1995), Nirenburg and Raskin (1986, 1987a,b), Raskin (1987a,b, 1990), Nirenburg and Carlson (1990), Meyer *et al*. (1990), Nirenburg and Goodman (1990), Nirenburg and Defrise (1991), Nirenburg and L. Levin (1992), Onyshkevych and Nirenburg (1992, 1994), Raskin et al. (1994a,b), Raskin and Nirenburg (1995, 1996a,b), Viegas *et al*. (1996b).

1.2 Generative vs. Enumerative?

The generative lexicon is motivated, in part, by the shortcomings of the entity it is juxtaposed against, the enumerative lexicon. The enumerative lexicon is criticized for:

- having the senses for each lexical item just listed without any relations established among them;
- the arbitrariness of (or, at least, a lack of a consistent criterion for) sense selection and coverage;
- failing to cover the complete range of usages for a lexical item;
- inability to cover novel, unattested senses.

Such enumerative lexicons are certainly real enough (most human-oriented dictionaries conform to the description to some extent), and there are quite a few of them around. However, there may be good enumerative lexicons, which cannot serve as appropriate foils for the generative lexicon.

Enumerative lexicons could, in fact, be acquired using a well thought-out and carefully planned procedure based on a sound and efficient methodology, underlain, in turn, by a theory. The acquisition environment would include a library of automatic and semi-automatic search means, concordance browsers and other acquisition tools to make the production of lexical entries easier, faster, and more uniform. The methodology would include the following steps:

- obtaining a large and representative set of corpora, complete with a fast look-up tool;
- examining a large set of existing lexicons and subjecting them to the semi-automatic polysemy reduction procedure (see Raskin and Nirenburg 1995: 41-45 and Beale *et al*. 1995);
- determining a small set of entry types which have to be created from scratch;
- identifying sets of large-scale lexical rules which derive lexical entries from entry types and from other lexical entries semi-automatically.

Such an enumerative lexicon will cover exactly the same senses as the generative lexicon, with the relations among these senses as clearly marked.[5] A good enumerative lexicon can be seen as at least weakly--and, in fact, strongly--equivalent (see Chomsky 1965: 60) to a generative lexicon **after all the lexical rules have fired**. Whether, in a computational application, lexical rules are triggered at acquisition or run time may have a computational significance, but their generative capacity, again in the sense of Chomsky (1965), i.e., their output, is not affected by that, one way or another (see Viegas *et al*. 1996b).

The corpora are used in such an enumerative approach for look-up purposes, not to limit the senses of lexical entries. Instead, all the applicable lexical rules are applied to all eligible lexical entries, thus creating virtual or actual entries for all the derived senses, many of them not attested in the corpora.

---

[5] It would be instructive to recall at this point that the term ‘generative,’ as introduced into linguistics by Chomsky (1957) and as inherited by Pustejovsky, means a way of enumeration through mathematical derivation, albeit a sophisticated and intelligent one.

1.3 Generative Lexicon and Novel Senses

In view of the equivalence of the outputs of the generative lexicon, on the one hand, and a high-quality enumerative lexicon, on the other, the claimed ability of the generative lexicon to generate novel, creative senses of lexical items needs to be examined more closely. What does this claim mean? What counts as a novel sense? Theoretically, it is a sense which is not previously attested to and which is a new, original usage. This, of course, is something that occurs rather rarely. Practically, it is a sense which does not occur in a corpus and in the lexicon based on this corpus. Neither the generative lexicon nor a good enumerative lexicon will--or should--list all the senses overtly. Many, if not actually most senses are derived through the application of lexical rules. But even if not listed, such a derived sense is present in the lexicon virtually, as it were, because it is fully determined by the pre-existing domain of a pre-existing lexical rule.

Does the claim of novelty mean that senses are novel and creative if they are not recorded in some given enumerative lexicon? If so, then the object chosen for comparison is low-quality (unless it was built based exclusively on a given corpus of texts) and therefore not the most appropriate one, as one should assume a similar quality of the lexicons under comparison. While the literature is not quite explicit on this point, several contributions (e.g., Johnston *et al*. 1995, Copestake 1995) seem to indicate the implicit existence of a given inferior lexicon or a non-representative corpus against which the comparison is made.

The other line of reasoning for justifying the claim of novelty involves the phenomena of type shifting and type coercion. A creative usage is one which arises from a rule that would overcome a sortal or other incongruity to avoid having to reject an input sentence as ill-formed. But there are rules that make type shifting and type coercion work. They are all pre-existing, not *post-hoc* rules, and, therefore, just as other lexical rules, fully determine, or enumerate (see below), their output in advance.

It is perhaps appropriate here to resort to simple formalism to clarify this point further, especially as the proponents of the generative lexicon approach[6] seem to treat formalism-based analogies and illustrations as the best devices for clinching arguments. Let L be the finite set of all lexical rules, l, used to derive senses from other senses; let T be the finite set of all type-shifting and coercion rules, t; let S be the (much smaller) set of the senses, s, of a lexical entry, e, in the generative lexicon G. Then,$G = \{e_1^G, e_2^G, ..., e_n^G\}$ and $S_e = \{s_1^e, s_2^e, ..., s_m^e\}$. If $l(s_e)$ is a sense of an entry derived with the help of lexical rule l and $t(s_e)$ is a sense of an entry derived with the help of type-shifting, or coercion, rule t, then let us define V_e as the set of all such derived senses of an entry: $V_e = \{v: \forall v \exists s \exists e \exists l \exists t\ v = l(s_e) \vee v = t(s_e)\}$. Let W^G be the set of all derived senses for all the entries in G: $W^{GLT} = \{w: \forall w \exists s \exists e \exists l \exists t\ w = l(s_e) \vee w = t(s_e)\}$. Finally, let U^{GLT} be the set of all senses, listed or derived in G: $U^{GLT} = =W^{GLT} \cup C^G$, where $C^G = \{c: \forall c \exists s \exists e\ c = s_e\}$. U^{GLT} represents the weak generative capacity of G, given the pre-defined sets L^G and T^G of lexical and type-shifting rules associated with the generative lexicon.

---

[6] The works that we consider as belonging to this approach, some more loosely than others, include Asher and Lascarides (1995), Atkins (1991), Briscoe (1993), Briscoe and Copestake (1991, 1996), Briscoe *et al*. (1990, 1993, 1995), Copestake (1990, 1992, 1995), Copestake and Briscoe (1992), Copestake *et al*. (1994/1995), Johnston *et al*. (1995), Lascarides (1995), Nunberg and Zaenen (1992), Ostler and Atkins (1992), Pustejovsky (1991, 1993, 1995), Pustejovsky and Boguraev (1993), Saint-Dizier (1995), Sanfilippo (1995), Sanfilippo *et al*. (1992).

U^{GLT} is also an **enumerable** set in the calculus, *I*, defined by the set of rules $L^G \cup T^G$applied to C^G in the sense that there is a finite procedure, *P*, of (typically, one-step) application of a rule to a listed (or, rarely, derived) sense, such that each element in U^{GLT} is generated by *P* (P includes zero, or non-application, of any rule, so as to include C^G in the calculus). In fact, U^{GLT} is also **decidable** in the sense that for each of its elements, *i*, there is an algorithm in *I*, which determines how it is generated, i.e., an algorithm, which identifies, typically, a listed entry and a rule applied to it to generate *i*. The set of all those identified pairs of listed entries and rules applied to them determines the strong generative capacity of *G*.

Then, the only way the lexicon may be able to generate, i.e., define, a sense *s* is if $s \in U^{GLT}$. In what way can such a sense, *h*, be novel or creative if it is already predetermined in *G* by *L* and *T*? This notion makes sense only if the existence of a proper subset *B* of U^{GLT} is implied, such that $h \in U^{GLT} \wedge h \notin B$. Then, a deficient enumerative lexicon, *M*, would list all the senses of *B* and not use any lexical or type-shifting rules: $E = \{e_1^e, e_2^e, ..., e_k^e\}$, $B = \{b: \forall b\, \exists s\, \exists e\, b{=}s_e\}$ and $L^E = T^E = \varnothing$.

Obviously, if a lexicon, *O*, does enumerate some senses and derives others in such a way that every sense in U^{GLT} is either listed or derived in *O* as well, so that both the weak and strong generative capacities of *O* equal--or exceed--those of U^{GLT}, then *G* does not generate any novel, creative senses with regard to *O*. It also follows that the generative lexicon approach must specify explicitly, about each sense claimed to be novel and creative, relative to what corpus or lexicon is it claimed to be novel and creative.

The above both clarifies the notion of a novel, creative sense as used in the generative lexicon approach and renders it rather trivial. It is not that there is anything special or wonderful about senses claimed to be such but rather that the corpus or lexicon, relative to which these senses are novel and creative, are incomplete. The claim of novelty is then reduced to a statement that it is better to have a high-quality corpus or lexicon than a lower-quality one, and, obviously, nobody will argue with that! What is less helpful is discussing the ability of the generative lexicon to deal with novel, creative senses without an explicit reference to an incomplete corpus or lexicon, in which these senses are not attested, and without a clear statement that it is precisely this non-attestation in an incomplete resource which makes these senses novel and creative, exclusively relative to that resource. The lack of clarity on this account, permeating the generative lexicon approach and unfortunately not remedied in Pustejovsky (1995), leads to an overstatement and a popular misunderstanding of the novel, creative sense as something nebulously attractive--or attractively nebulous.

We suggest that a truly novel and creative usage will not have a ready-made generative device, for which it is a possible output, and this is precisely what will make this sense novel and creative; in other words, it will not be an element of U^{GLT}. Such a usage will present a problem for a generative lexicon, just as it will for an enumerative one or, as a matter of fact, for a human trying to treat creative usage as metaphorical, allusive, ironic, or humorous.

To summarize, the rule formalism used in conjunction with the generative lexicon is based on a mathematical tautology: the rules produce what is put in them and nothing else. Anything else is a "miracle"--except that there are rules for those metaphorical, allusive, ironic, humorous, and other non-literal usages, which, when explored and listed, can be added as the third set of rules to *I*, and then even such senses--at least, the most "petrified," "stale," or clichéized of them--will no longer be novel and creative. The generative lexicon approach, however, has not yet addressed those rules.

1.4 Permeative Usage?

Another claimed advantage of the generative lexicon is that it "remembers" all the lexical rules that relate its senses. We submit, however, that, after all these rules have worked, the computational applications using the lexicon would have no use for them or any memory--or, to use a loaded term, trace--of them whatsoever; in other words, the decidability of the fully deployed set of all listed and derived sentences is of no computational consequence.

Pustejovsky (1995: 47-50) comes up with the notion of permeability of word senses to support this Lexical-rule remembrance claim. Comparing *John baked the potatoes* and *Mary baked a cake*, he wants both the change-of-state sense of *bake* in the former example and the creation sense in the latter to be present, to overlap, to permeate each other. His desire to see both of these meanings present is linked, of course, to his conviction that these two meanings of *bake* should not be both listed in the lexicon but rather that one of them should be derived from the other. His argument, then, runs as follows: see these two distinct senses? Well they co-occur in the same sentence, thus permeating each other. Therefore, they should not be listed as two distinct senses. Or, putting it more schematically: See these two senses? No, you don't!

Our counterargument below is simpler. Yes, there are perhaps two distinct senses--if we can justify the distinction. No, they do not co-occur in the same normal (not deliberately ambiguous) usage. Yes, we do think they may be listed as distinct senses, each closely related to the semantic properties of their themes. Yes, they can also be derived from each other, but what for and at what price?

We also think the permeative analysis of the data is open to debate because it seems to jeopardize what seems to us to be the most basic principle of language as practiced by its speakers, namely, that each felicitous speech act is unambiguous. It is known that native speakers find it very hard to detect ambiguity (see, for instance, Raskin 1977 and references there). It stands to reason that it would be equally difficult for them to register permeation, and we submit that they actually do not, and that the permeating senses are an artifact of the generative lexicon approach. This, we guess, is a cognitive argument against permeation.

The reason for the native speaker's unconscious blocking of ambiguity is that it is a complication for our communication and it raises the cognitive processing load (see, e.g., Gibson 1991). So the hearer settles on the one sense which happens to be obvious at the moment (see, again, Raskin 1977 and references there), and blocks the others. There are "non-bona-fide" modes of communication which are based on deliberate ambiguity, such as humor (see, for instance, Raskin 1985c: xiii, 115; cf. Raskin 1992), but functioning in these modes requires additional efforts and skills, and there are native speakers of languages who do not possess those skills without, arguably, being judged incompetent[7].

Encouraging permeative usage amounts to introducing something very similar to deliberate ambiguity, a kind of a sense-and-a-half situation, into semantic theory, both at the word-meaning level as permeability and at the sentence-meaning level as co-compositionality (see also Sections 3, 4, and 7 below). It seems especially redundant when an alternative analysis is possible. One of the senses of *cake* should and would indicate that it often is a result of baking--there are, however, cold,

[7] There is a popular view in most cultures that a person lacking a sense of humor also lacks in intelligence, but no research has succeeded in establishing a link between the sense of humor and intelligence (see, for instance, Mindess *et al.* 1987 for a report on the best known failed attempt to establish such a link).

uncooked dishes that are referred to as cakes as well. No sense of *potato* would indicate that--instead, *potato*, unlike *cake*, would be identified as a possible theme of *cook*, and *cook* will have *bake* and many other verbs as its hyponyms. This analysis takes good care of disambiguating the two senses of *bake* via the meaning of their respective themes, if a need for such disambiguation arises. In fact, it still needs to be demonstrated that it is at all necessary to disambiguate between these two senses for any practical or theoretical purpose, other than to support the claim of permeability of senses in the generative lexicon approach, which, in turn, is used only to support, against both cognitive and empirical evidence, the fixation of the generative lexicon approach on reducing the number of listed senses in the lexicon to a preferable minimum of one. This sounds like *ad hoc* circularity, a kind of *ad hoc* square.

1.5 Generative Vs. Enumerative "Yardage"

To summarize, some central claims associated with the generative lexicon seem to juxtapose it against low-quality or badly acquired enumerative lexicons and to ignore the fact that any reasonable acquisition procedure for a high-quality enumerative lexicon will subsume, and has subsumed in practice, the generative devices of the generative lexicon--without the ballast of some of the spurious and unnecessary claims made by the approach.

When those claims are blown up and the smoke clears away, it appears that the difference between the generative lexicon and the high-quality enumerative lexicon is only in the numbers, and unimportant numbers to boot. The former aspires to minimize the number of listed senses for each entry, reducing it ideally to one. The enumerative lexicon has no such ambitions, and the minimization of the number of listed entries in it is affected by the practical consideration of the minimization of the acquisition effort as mentioned in Section 1.1 above.

To reach the same generative capacity from a smaller range of listed senses, the generative lexicon will have to discover, or postulate, more lexical rules, and our practical experience shows that this effort may exceed, in many cases, the effort involved in listing more senses, even though each such sense may involve its creation from scratch.

In a pretty confused argument against Pustejovsky's treatment of *bake* and his efforts to reduce the two meanings to one (see Section 1.4 above),[8] Fodor and Lepore (1996) still manage to demonstrate that whatever gain can possibly be achieved by that reduction, more effort will have to be expended on dealing both with the process of achieving this goal and with the consequences of such treatment of polysemy. We cannot help agreeing with their conclusion, no matter how exactly achieved, that "the total yardage gained would appear to be negligible or nil" (*op. cit.*: 7).

2. Syntax vs. Semantics

The issue of the relations between syntax and semantics is perhaps the one with the most history of the ten dichotomies this paper is dealing with. At least ever since the interpretive-generative feud in the early transformational semantics of the 1960s, the issue informed, one way or another, most of the work in both fields, and the initial salvos of the war, namely, Katz and his interpretive coauthors' attempt to postulate deep structure as the impermeable boundary between syntax and seman-

[8] Coming from a very different disciplinary background, the authors largely follow our line of reasoning here at some times, but also take unnecessary detours and make some unnecessary claims of their own in the process of pursuing totally different goals--different not only from ours but also from Pustejovsky's. We will have a chance to return to this amazingly irrelevant review later on.

tics, on the one hand, and the repudiation of deep structure, primarily, as that boundary by their generative (and unaffiliated) critics, have been reiterated again and again for the subsequent three decades, thus shaping and reshaping the "linguistic wars."[9]

In computational linguistics, the issue has taken the form of encroaching into semantics from syntax (see Introduction above) in order to obtain as much meaning information as possible with the simpler syntactic tools and thus to avoid meaning analysis, which continues to intimidate many NLP scholars. We do not believe this can or should be done; we feel that the time and talent spent on the ingenious ways of bypassing semantics on the way to meaning should have been spent on computational semantics; we are convinced that meaning is essential for NLP and that it is unapproachable other than directly through computational semantics, assisted, of course, by syntax wherever and whenever necessary. The generative lexicon approach appears to be making the other choice, without much discussion of the issue, which seems to indicate that they assume the basic non-feasibility of computational semantics.

Early on in the book, Pustejovsky (1995: 5-6) makes the methodologically important statement that "there is no way in which meaning can be completely divorced from the structure that carries it. This is an important methodological point, since grammatical distinctions are a useful metric in evaluating competing semantic theories." He goes on to discuss the semantic import of the study of categorial alternation, as made popular in recent years by B. Levin (e.g., 1993), and states that "...the diversity of complement types that a verb or other category may take is in large part also determined by the semantics of the complements themselves... I will argue... that alternation classifications do not constitute [semantic?--SN&VR] theory." It is surprising, however, that throughout the rest of the book, it seems to be tacitly assumed that every syntactic distinction determines a significant semantic distinction and that every semantic distinction (or at least every one the generative lexicon is comfortable with) has a syntactic basis or at least a syntactic clue or a syntactic diagnostic.

This is a departure from the more moderate opinion quoted above, and this more radical stance is expressed repeatedly as the dependence of semantics on "basic lexical categories" (*op.cit.*: 1), on "syntactic patterns" and "grammatical alternations" (*op.cit.*: 8), as the search for "semantic discriminants leading to the distinct behavior of the transitive verbs" in the examples (*op.cit.*: 10), or as an "approach [that] would allow variation in complement selection to be represented as distinct senses" (*op.cit.*: 35). It is indeed in the analyses of examples (as well as examples used by other lexical semanticists subscribing to the idea of generative lexicon--see, for instance, Lascarides 1995: 75) that the apparently complete and unquestioned dependency on syntax comes through most clearly.

Thus, dealing with his own variations of Chomsky's (1957) famous examples of *John is eager to please* and *John is easy to please* in terms of *tough*-Movement and the availability or non-availability of alternating constructions (*op.cit*.: 21-22), Pustejovsky makes it clear that these different syntactic behaviors, essentially, constitute the semantic difference between adjectives like *eager* and adjectives like *easy*. We have demonstrated elsewhere (Raskin and Nirenburg 1995) that much more semantics is involved in the analysis of differences between these two adjectives and that these differences are not at all syntax dependent. *Easy* is a typical scalar, whose value is a range on the EASE/DIFFICULTY scale, and which modifies events; *eager* is an event-derived adjective modi-

[9] See Katz and Fodor (1963), Katz and Postal (1964), Weinreich (1966), Katz (1967, 1970, 1971), McCawley (1971), Lakoff (1968, 1970, 1971a,b); cf. also Montague (1974), Fodor (1978), Dowty (1979), Enç (1988), Harris (1993).

fying the agent of the event. This difference does explain the different syntactic behaviors of these adjectives but not the other way around.

One interesting offshoot of the earlier syntax vs. semantics debates has been a recent strong interest in "grammatical semantics," the subset of the semantics of natural languages which is overtly grammaticalized (see, for instance, Frawley 1992--cf. Raskin 1994; in computational-semantic literature, B. Levin 1993 and Nirenburg and L. Levin 1992---who call this field "syntax-driven lexical semantics"--- are noteworthy). This is a perfectly legitimate enterprise as long as one keeps in mind that semantics does not end there.

Wilks (1996) presents another example of an intelligent division of labor between syntax and semantics. He shows that up to 92% of homography recorded in LDOCE can be disambiguated based exclusively on the knowledge of the part of speech marker of a homograph. Homography is, of course, a form of polysemy and it is useful to know that the labor-intensive semantic methods are not necessary to process all of it. Thus, semantics can focus on the residual polysemy where syntax does not help. In a system not relying on LDOCE, a comparable result may be achieved if word senses are arranged in a hierarchy, with homography at top levels, and if disambiguation is required only down to some nonterminal node in it. Needless to say, the work of semantics is made easier by this to a very small extent, but every little bit counts!

It is also very important to understand that grammatical semantics does not assume that each syntactic distinction is reflected in semantic distinction--instead, it looks at those semantic distinctions which do have some morphological and syntactical phenomena associated with it. Consequently, grammatical semantics does not engage in recurring frustrating searches for a semantic distinction for various subclasses of lexical items conforming or not conforming to a certain rule that the generative lexicon approach constantly requires (see, for instance, Briscoe *et al*. 1995, Copestake 1995, or Briscoe and Copestake 1996).

The dependence on syntax in semantic analysis may lead to artificially constrained and misleading analyses. Thus, analyzing the sense of *fast* in *fast motorway* (see, for instance, Lascarides 1995: 75) as a new and creative sense of the adjective as opposed, say, to its sense in *fast runner,* ignores the important difference between syntactic and semantic modification precisely because of the implicit naive conviction that the use of the adjective with a different noun subcategory, which constitutes, since Chomsky (1965), a different syntactic environment for the adjective, automatically creates a different sense for *fast*. As established in Raskin and Nirenburg (1995), however, many adjectives do not modify semantically the nouns they modify syntactically, and this phenomenon covers many more examples than the well-known *occasional pizza* or *relentless miles*. Separating syntactic and semantic modification in the case of *fast* shows that it is, in fact, a modifier for an event, whose surface realization can be, at least in English, syntactically attached to the realizations of several semantic roles of, for instance, *run* or *drive*, namely, AGENT in *fast runner*, INSTRUMENT in *fast car*, and LOCATION (or PATH) in *fast motorway*. Throughout these examples, *fast* is used in exactly the same sense, and letting syntax drive semantics distorts the latter seriously.

Postulating a new sense for *fast* in *fast motorway* begs the notorious issue of the "plasticity" of the adjectival meaning (see Raskin and Nirenburg 1995: 21; specifically, on the plasticity of adjectival meaning, see also Marx 1977, 1983, Szalay and Deese 1978, and Lahav 1989), i.e., the tendency of many, if not all, adjectives to modify their meanings depending on that of the nouns they modify syntactically. The meaning of the adjective *good* is perhaps the most explored example of this phenomenon (see, for instance, Ziff 1960, Vendler 1963, Katz 1972, Pustejovsky 1995: 32, Fodor and

Lepore 1996: 11). Our own argument against the proliferation of different senses for *good* is two-fold: first, the adjective practically never modifies semantically the noun it modifies syntactically, expressing instead a general positive attitude to the concept evoked by the noun on the part of the speaker; secondly, we argue against the further detailing of the meaning on the grounds of granularity and practical effability, i.e., basically, that, in MT, for instance, an equally generalized notion of goodness will be expressed in another natural language by a similar adjective, which does appear to be a universal or near-universal--the fact that, by and of itself, would indicate the integrity of the vague concept of goodness in the human mind (Raskin and Nirenburg 1995: 28-29, 43-47, and 49-50). The upshot of this discussion is that no purported syntactic distinctions should lead automatically into the fracturing of one meaning into several.

Distinguishing word senses on the basis of differences in syntactic behavior does not seem to be a very promising practice (cf. the Dorr *et al*. 1995 attempt to develop B. Levin's approach into doing precisely this) also because such an endeavor can only be based on the implicit assumption of isomorphism between the set of syntactic constructions and the set of lexical meanings. But it seems obvious that there are more lexical meanings than syntactic distinctions, orders of magnitude more. That means that syntactic distinctions can at best define classes of lexical meanings, and indeed that is precisely what the earlier incursions from syntax into semantics achieved: just very crude, coarse-grained taxonomies of meanings in terms of preciously few features. On top of that, the case of *good* above further weakens the isomorphism assumption by demonstrating that it does not hold also because there are cases when several purported syntactic distinctions still correspond to the same meaning.

3. Pragmatic Vs. Non-Pragmatic

Throughout the generative lexicon approach, the bogeyman of pragmatics, variously characterized as "commonsense knowledge," "world knowledge," or context, is used as something that is separate from the enterprise itself, complex, and, alternatively, not worth doing or not possible to do, at least for now (see, for instance, Pustejovsky 1995: 4, Copestake 1995). Occasionally--and increasingly so lately--brief ventures into this dangerous unknown are undertaken in the framework of syntax-driven lexical semantics (see, for instance, Asher and Lascarides 1995, Lascarides 1995). It is noteworthy that semantics as such seems to be excluded from these ventures which, basically, jump from lexically augmented syntax into pragmatics. Even curiouser, the exclusion of semantics is combined with a prevalent assumption of and an occasional reference to its being done truth-conditionally; one cannot help wondering whether this situation constitutes a tacit recognition of the fact that truth-conditional semantics is of no use to lexical semantics (see Section 4 below).

Pustejovsky again starts out with a very careful stance of "abstracting the notion of lexical meaning away from other semantic influences," which "might suggest that discourse and pragmatic factors should be handled differently or separately from the semantic contribution of lexical items in composition" (Pustejovsky 1995: 6), but then adopts the position that pragmatics, under various names, is what is locked out of the lexicon itself. In fact, chronologically, this is the position that many lexical semanticists adopted from the start (see Section 2 above), and the more cautious formulation of the principle appeared much later to impart greater theoretical respectability for the approach

The generative lexicon is in good company with respect to this technique. In setting up a "waste-basket" for the phenomena it will not be held accountable for, the approach upholds a great tradition in American linguistics. Bloomfield (1933) dubbed the vast set of phenomena which segmentation and distribution---the only techniques he allowed---could not handle "semantics"

and excluded it from linguistics. Chomsky started on the same path early on but later, after a reluctant incorporation of Katz and Fodor's semantics into standard theory, changed the name of the dustbin to "performance" and then to "pragmatics."

Pustejovsky prudently admits that the separation of lexical semantics from pragmatics "is not a necessary assumption and may in fact be wrong, [but] it will help narrow our focus on what is important for lexical semantic descriptions" (1995: 6). This would sound like a legitimate methodological principle, unless it was claimed that "the issue is clearly an empirical one" (*op.cit.*: 4). The issue can be thought of as empirical if one can find some empirical evidence in natural language, through native speakers' behavior, for or against the existence of the distinction between lexical and pragmatic knowledge, often inaccurately equated with "world knowledge" or "encyclopedic knowledge." Claims to this effect have been made in post-transformational semantic theory of the 1980s, and they should not be ignored in the current debates (see Raskin 1985a,b, 1985c: 134-135; for more discussion of the issue--from both sides--see Hobbs 1987, Wilensky 1986, 1991; cf. Wilks 1975a, 1975b: 343).

In NLP, however, it is perfectly clear, and definitely empirically so, that both types of information are essential for successful text analysis and generation. As Wilks and Fass (1992: 1183) put it, "knowledge of language and the world are not separable, just as they are not separable into databases called, respectively, dictionaries and encyclopedias" (see also Nirenburg 1986). Where each type of knowledge should come from depends on one's theoretical basis (see Section 4) and its foundations (Section 5), and, related to that, one's clarity on the difference between the static and dynamic knowledge resources (Section 7).

4. Semantic Theory Vs. Formal Semantics etc.

4.1 Semantic Theory in the Generative Lexicon Approach

Pustejovsky (1995) clearly leads the way in laying out a theoretical foundation of sorts for the generative lexicon approach. Many other lexical semanticists defer to his semantic theory (see, for instance, Copestake 1995: 22, Saint-Dizier 1995: 153, Sanfilippo 1995: 158), primarily by adopting his qualia structure, i.e., the original and most essential part of the notation offered by the theory. Pustejovsky (1995) makes an attempt to see more in his theoretical approach than just the use of qualia notation.

Discussion of semantic theory in Pustejovsky (1995) is couched primarily in terms of *desiderata.* On this view, semantic theory consists of lexical semantic theory and treatment of 'compositionality,' which appears to cover what is known in transformational and post-transformational semantic theory as sentence meaning--in other words, whatever it takes to interpret the meaning of the sentence. The description of compositionality in the Generative Lexicon is rather general, as are indeed the other parts pertaining to semantic theory.[10] One important issue which we discuss later (see Section 7) is that composition of sentence meaning from the meanings of its components is

[10] It is also not entirely clear why the common terms 'meaningfulness' and 'ambiguity' should have been replaced with 'semanticality' (*op.cit.*: 40) and 'polymorphism' (*op.cit.*: 57). The one innovative feature of semanticality compared to meaningfulness concerns the variability of the former (*op.cit.*: 42), at the level of wishful thinking, which has been an element in our own *desiderata* for semantic theory as well for quite some time (see Nirenburg and Raskin 1986), except that we consider the variability in the depth of semantic representation and not just the degree of meaningfulness.

only a part of the processes needed for text meaning representation.

The boundary between lexical semantic theory and the realm of compositionality is described in the Generative Lexicon in a rather fluid manner: compositionality keeps popping up in the discussion of the lexicon. Unlike Katz and Fodor (1963), with their clearly defined (though syntactically determined) "amalgamation rules," Pustejovsky does not introduce any mechanism for combining word meanings into sentence meaning, though some such mechanism is implicit in the discussion and plays a very important role in the "evaluation of lexical semantic models" (1995: 1), a declared central concern of Pustejovsky's book.

The aim of the lexical semantic theory is to describe "a core set of word senses, typically with greater internal structure than is assumed in previous theories" (*op.cit.*: 2). This, however, does not mean for Pustejovsky additional depth of semantic description but rather the relations between the underdescribed[11] core senses (see Section 1 above). Four issues are identified as "the most pressing problems of lexical semantics," none of which seems to be a lexical semantic issue *per se*. In standard terms, these issues are ambiguity and meaningfulness of natural language utterances (two items straight out of Katz and Fodor's interpretive theory of sentence meaning), the creative use of word senses in sentences, and sense permeability (see also Section 1 above).

In our understanding, the crucial issues in lexical semantic theory are word meaning (in the approach taken by the Mikrokosmos project -- see, e.g., Onyshkevych and Nirenburg 1994, Mahesh and Nirenburg 1995, Nirenburg *et al*. 1995 -- it is realized either as the relationship between a word sense and an element or a group of elements of the underlying model of the world or as a procedure for assignment of a particular set of values to appropriate elements of a text meaning representation) and the other static (that is, context-independent) knowledge (no matter whether of syntactic, semantic, pragmatic, prosodic, stylistic or other nature) which helps a) to disambiguate a word during analysis and b) to select the most appropriate word from a set of synonyms during generation. As described above, lexical semantics is a necessary but not sufficient component of the general theory of computational semantics. Knowledge of context-dependent rules, often difficult to index through a lexicon (cf. Fodor and Lepore 1996) and knowledge about complex events and objects in the world and relationships among them (a static but non-lexical source of clues for disambiguation and synonymy resolution), as well as a specification of the procedures for deriving meanings from texts and *vice versa*, are also needed.

4.2 Semantic Theory and Formalism

Pustejovsky rejects formal semantics at the outset both as falling short of achieving the goals he considers desirable (1995: 1) and as "an application of ready-to-wear formalism to a new body of data" (*op.cit.*: 2).[12] Nevertheless, formalisms permeate the book and similar lexical-semantic writings. Truth values of model-theoretical semantics tend to creep in too, for reasons that are not immediately clear (*op.cit.*: 23, Lascarides 1995: 78), and are viewed as objectionable by formal semanticists (Fodor and Lepore, 1996: 3). Most of the time, the formalism provides a (redundant,

[11] Thus, when a lexical rule is described in the generative lexicon approach, typically the description focuses on just one distinguishing feature of one lexical item which changes its value in the other. No actual definition of the meaning of either item is otherwise treated.

[12] In an extreme but not uncommon manifestation, formal semantics is happy to denote the meaning of *dog* as DOG and consider its job done (Fodor and Lepore, 1996:1). This job is, however, not that of computational semantics, which requires much more informative and substantive lexical entries--see more below.

in our estimation) translation of what has already been expressed otherwise in the text, possibly, as an insider *lingua franca*. Occasionally, it is evoked only to be rejected as useless (see, for instance, Johnston *et al*. 1995: 69, 70, Lascarides 1995: 77, 78). The above may be no more than a quibble. However, a word of caution is in order.

First, coming unintroduced and unjustified independently, contrary to the traditions of formal linguistics, the formalism exercises unclear authority, often producing formal statements that demand interpretation, thus leading analyses in a certain direction, not necessarily dictated by the material being formalized. This is why it must be modified, constrained, and augmented (see, e.g., Copestake's 1995: 25 analysis of *apple juice chair*) to fit what the user wants it to do. For believers, the formal statements are licensed; for others, including those who can readily use the formalism, this license is lacking.

SecondlyXS, in the absence of an explicitly stated axiomatic theory, which it is supposed to serve, the formalism can be adjusted to express any rule the user wants it to express, thus depriving the tool of authority. Apparently, the people exchanging messages in the formalism subscribe to some mutually accepted, though unexplained, rules of usage in an apparent consensus--at least, we are definitely willing to give them the benefit of the doubt for that. Our point is that the appropriate place to introduce and justify the formalism is in the pages where a semantic theory is formulated, and the appropriate use of formalism should be consistent with the clearly stated principles of such a theory.

The generative lexicon approach is, of course, interested, as is all computational semantics, in representing both lexical meaning and sentence meaning. It also prefers to do things on the basis of linguistic theory, which is precisely how computational semantics should be done. In spite of Pustejovsky's (1995: 1) initial and fully justified rejection of formal semantics as a basis of achieving the generative lexicon goals, the approach did not find anything more in contemporary linguistic semantics for dealing with sentence meaning than rather shallow analyses of quantifiers and other closed-class phenomena. Formal semantics currently holds a monopoly on the former enterprise[13] and extends at least into the quantifier part of the latter.[14] A trend that can be called 'grammatical semantics' deals with the closed classes and with overtly grammaticalized meanings of the open classes (Frawley 1992 is the most extensive recent enterprise in this strand).

This creates a real problem for the generative lexicon approach: there is no ready-made semantic theory they can use for the task of meaning representation of a sufficiently fine granularity that NLP requires. Their failure to recognize the situation results in their automatic and fruitless acceptance of formal semantics as the legitimate and monopolous "owner" of linguistic semantics and in the excessively soft underbelly of semantic theory in Pustejovsky (1995). It also leaves them, essentially, without a framework for syntagmatic relations between lexical meanings to parallel

---

[13] The very term 'compositionality' comes from formal semantics, where it is equated with sentence meaning, and from the philosophy of language, into which it was introduced by Frege (1892) and Russell (1905). It is also customary to use it in linguistic semantics, certainly after Katz and Fodor (1963), to denote the representation of the meaning of a sentence exclusively on the basis of what the sentence itself contains, i.e., the meanings of the words and of their combinations. See, for instance, Lyons (1995: 204-209) for a succinct account of the formal meaning of compositionality. Pustejovsky (1995) seems to vacillate between the formal and linguistic-semantic meanings of the term.

[14] See, for instance, Lewis (1972), Parsons (1972, 1980, 1985, 1990), Stalnaker and Thomason (1973), Montague (1974), Dowty (1979), Barwise and Perry (1983), Keenan and Faltz (1985), Partee *et al*. (1990), Chierchia and McConnell-Ginet (1990), Cann (1991), Chierchia (1995), Hornstein (1995).

paradigmatic relations established by lexical rules, resulting in Pustejovsky's inconclusive explorations of "compositionality" (see Section 4.1 above; cf. Section 7 below).

An unexpected and rather comic effect of this theoretical muddle is that Fodor and Lepore (1996) mistake Pustejovsky's (1995) lip service to formalisms for a commitment to their brand of the philosophy of language and heap tons of totally undeserved and irrelevant criticism on him. It is Fodor and Lepore (1996), again, who provide a tellingly extreme example of the way formal semantics defines itself out of any usefulness for linguistic or NLP semantics, by insisting that "the lexical entry for *dog* says that it refers to dogs"[15] and contains absolutely nothing else. This, purely extensional concept of meaning--that of Russell's (1905) and early Wittgenstein's (1921)--may make sense in some paradigms and for some goals, even if it has been largely rejected both in the philosophy of language and even if the authors' main reason for rejecting any further complexity in lexical meaning is that they cannot think of a reliable way to proceed with defining that complexity and choose--and can afford to do so because of the ultimate non-descriptivity of their goals--to settle on an ultimately atomistic approach to meaning (with no atom related to any other and no atoms making up a molecule). But the ensuing inability to relate meanings to each other or to use them for disambiguation of the sentences they appear in flies in the face of the goals of the generative lexicon approach, of lexical and computational semantics, and indeed of linguistic semantics. The goals of the last as modeling the semantic competence of the native speakers with regard to their ability to disambiguate sentences, to "amalgamate" their meanings out of lexical meanings, to paraphrase, and to detect anomalies, defined by the same Fodor and Katz (1963), have never been contested since, even if they seem to be largely abandoned by most linguistic semanticists for now.[16]

While the linguistic semanticists of the 1990s can, apparently, shelve the uncontested goals of linguistic semantics for now, electing to do other things, no such choice is available to computational semantics. It is in the business of representing meaning explicitly, both at the lexical and sentential levels and at an appropriate granularity, to get the job of knowledge-based NLP done. And no amount of "prime" semantics (as in the apocryphal question and answer: "What is the meaning of life?" "Oh, *life'*"), with the meaning of ϕ being set up as ϕ'--or, for that matter, the meaning of *dog* defined as dogs, will be of any help to the enterprise undertaken both by us and by the adherents of the generative lexicon approach.

4.3 The Qualia and Some Comments on Notation

Rather like the concept of lexical functions in the Meaning-Text theory, the Qualia structure of the generative lexicon theory is perceived by many to be the central part of the theory and is often used or referred to separately from the rest of the theory components. It consists of a prescribed set of four roles with an unspecified open-ended set of values. The enterprise carries an unintended re-

---

[15] In the original, our italicized *dog* is underlined and our underlined dogs is double-underlined.

[16] Having turned away from these goals and abandoned the field of linguistic semantics, Fodor has kept faith, however, in this review, in the impossibility of meaning in context, postulated explicitly in Katz and Fodor (1963) on the ground of theoretical unattainability. Is it legitimate at all, one may wish to ask, to pursue such a "minimalist," *dog*/dogs view of meaning? The answer has to be that it is indeed, for instance, in a formal theory which values simplicity above all, which chooses to ignore the well-established limitations and failures of purely extensional, denotational semantics, and which has no descriptive goals whatsoever. For the generative lexicon approach, however, such a formal semantics will deny any real possibility of establishing relations among different senses. Fodor and Lepore's (1996) failure to appreciate this is rather astonishing.

semblance to the type of work fashionable in AI NLP in the late 1960s and 1970s: proposing sets of properties (notably, semantic cases or case roles) for characterizing the semantic dependency behavior of argument-taking lexical units (see, e.g., Bruce 1975). That tradition also involved proposals for systems of semantic atoms, primitives, used for describing actual meanings of lexical units. This latter issue is outside the sphere of interest of the generative lexicon, though not, in our opinion, of lexical semantic theory (see Section 4.2 above).

The definitions of the four qualia roles are in terms of meaning and carry all the difficult problems of circumscribing the meaning of case roles. Assignment of values to roles is not discussed by Pustejovsky in any detail, and some of the assignments are problematic, as, for instance, the value "narrative" for the constitutive role (which is defined as "the relation between an object and its constitutive parts" (1995: 76)), for the lexicon entry of *novel* (*op.cit.*: 78). The usage of 'telic' has been made quite plastic as well *(op.cit.*: 99-100), by introducing 'direct' and 'purpose' telicity, without specifying a rule about how to understand whether a particular value is direct or purpose.

It seems that the main distinction between this proposal and earlier AI proposals, outside of a concern for connectivity with syntax in the generative lexicon, is a particular notation.There is nothing wrong with a notation to emerge as the main consequence of a theoretical approach. After all, theory defines the format of descriptions. However, a discussion of the justification of the introduced qualia (why are these and not others selected? what does the theory have to say about the process of value assignment?) would have been very welcome.

One immediate result of the qualia being formulated the way they are is, from our point of view, a simplification (not to say, impoverishment) of the semantic zone of the lexical entry, as compared, for instance, to lexical entries in the Mikrokosmos project (see Viegas *et al*. 1996a). NLP analyzers usually demand ever more information in the entries to drive inferences (ironically, though, not at all unexpectedly, Fodor and Lepore 1996 take Pustejovsky to task for making his lexical entries too complex while we accuse him of impoverishing them!). The relative paucity of semantic information in the generative lexicon approach, seen from the standpoint of processing needs, seems to invite the strict demarcation of the realm of lexical semantics and the rest of the knowledge needed for processing. As a result, in some versions of this approach, whatever is not included in the lexicon is assigned to pragmatics, which is so hard to investigate in this paradigm that some lexical semanticists explicitly refuse to do it and, significantly, refer to it as "the pragmatic dustbin" (see, for instance, Copestake 1995: 24).

Why be content with so little semantics in the qualia? One plausible explanation sends us back to Section 2. The semantics in the generative lexicon approach is a venture from syntax into an uncharted territory, and such ventures are kept short in distance and brief in duration. Semantic depths which cannot be directly tested on syntax are left untouched. As we already noted, the enterprise sounds very similar to Chomsky's own incursion into semantics in his lexicon (1965), which he ever persisted in referring to as a part of syntax, not semantics. The generative lexicon approach makes a step forward 30 years later and does call this venture "semantics" but it remains as attached to the syntactic leash as standard theory and as committed to minimizing its semantic exposure: "Some of these [qualia --- SN&VR] roles are reminiscent of descriptors used by various computational researchers, such as Wilks (1975[a]), Hayes (1979) and Hobbs et al. (1987). Within the theory outlined here, these roles determine a *minimal* semantic description of a word *which has both semantic and grammatical consequences*." (Pustejovsky 1991:418, fn 7) (Emphasis added--SN&VR; note that this statement squarely puts the generative lexicon into the supply-side paradigm in lexical semantics--see Section 10 below--as well as equating the lexicon enterprise to

grammatical semantics--see Section 2 above).

But coming back to the notational elements in the approach, one would expect to have all such elements as the four qualia specified explicitly with regard to their scope, and this is, in fact, what theories are for. After all, Chomsky's theory was ridiculed for introducing and meticulously defining five different types of brackets. What is the conceptual space, from which the qualia and other notational elements of the approach emerge? Why does Pustejovsky's theory miss an opportunity to define that space explicitly in such a way that the necessity and sufficiency of the notational concepts he does introduce become clear to us--including, of course, an opportunity to falsify its conclusions on the basis of its own explicitly stated rules?[17] To suggest our own explanation of this missed opportunity, we have to raise the issue of ontology.

5. Ontological Vs. Non-Ontological

Ontological modeling is taken by many as an essential part of lexical and computational semantics: "the meanings of words should somehow reflect the deeper conceptual structures in the cognitive system, and the domain it operates in" (Pustejovsky, 1995: 6).[18]. We basically agree with Pustejovsky that "this is tantamount to stating that the semantics of natural language should be the image of nonlinguistic conceptual organizing principles, whatever their structure" (*ibid*), even though we believe that a discussion of the exact nature of this "imaging" or, as we would prefer to see it, "anchoring," should be included in lexical semantic theory.

We believe that the notational elements that are treated as theory within the generative lexicon approach can, in fact, be legitimately considered as elements of semantic theory if they are anchored in a well-designed conceptual ontology. Then and only then, can one justify the postulation of a certain number of theoretical concepts, a certain set of roles and features, and a prescribed range of values. The alternative is a lack of certainty about the status of these notions and an osmosis- or emulation-based usage of them: a new feature and certainly a new value for a feature can always be expected to be produced if needed, the *ad hoc* way.

Proposals have been made in the generative lexicon paradigm for generalizing meaning descriptions using the concept of *lexical conceptual paradigms* (e.g., Pustejovsky and Boguraev 1993, Pustejovsky and Anick 1988, Pustejovsky *et al*. 1993). These paradigms "encode basic lexical knowledge that is not associated with individual entries but with sets of entries or concepts" (Bergler 1995: 169). Such "meta-lexical" paradigms combine with linking information through an associated syntactic schema to supply each lexical entry with information necessary for processing. While it is possible to view this as simply a convenience device which allows the lexicographer to specify a set of constraints for a group of lexical entries at once (as was, for instance, done in the KBMT-89 project (Nirenburg *et al*. 1992), this approach can be seen as a step toward incorporating an ontology.

Bergler (1995) extends the amount of these "meta-lexical" structures recognized by the generative

---

[17] An examination of the Aristotelian roots of the qualia theory fails to fill the vacuum either.

[18] For similar theoretical views, see also Lewis (1972), Miller and Johnson-Laird (1976), Miller (1977), Kay (1979), Hayes (1979), Jackendoff (1983), Johnson-Laird (1983), Barwise and Perry (198)3, Fauconnier (1985), Hobbs and Moore (1985), Hobbs (1987), Hobbs *et al*. (1987), Lakoff (1987), Langacker (1987), Knight (1993). The largest and most popularly cited implementation of an ontology is CYC--see, for instance, Lenat *et al*. (1990). On the limited applicability of CYC to NLP, due largely to its insufficient and non-uniform depth, see Mahesh *et al*. (1996).

lexicon to include many elements that are required for actual text understanding. She, for instance, incorporates a set of properties she calls a "style sheet," whose genesis can be traced to the "pragmatic factors" of PAULINE (Hovy 1988) or TAMERLAN (e.g., Nirenburg and Defrise 1993). She stops short, however, of incorporating a full-fledged ontology and instead introduces nine features, in terms of which she describes reporting verbs in English. A similar approach to semantic analysis with a set number of disjoint semantic features playing the role of the underlying meaning model was used in the Panglyzer analyzer (see, for instance, The Pangloss Mark III 1994)).

There is, of course, a great deal of apprehension and, we think, miscomprehension about the nature of ontology in the literature, and we addressed some of these and related issues in Nirenburg *et al*. (1995). One recurring trend in the writings of scholars from the AI tradition is toward erasing the boundaries between ontologies and taxonomies of natural language concepts. This can be found in Hirst (1995), who acknowledges the insights of Kay (1971). Both papers treat ontology as the lexicon of a natural (though invented) language, and Hirst objects to it, basically, along the lines of the redundancy and awkwardness of treating one natural language in terms of another. Similarly, Wilks *et al*.(1996: 59) see ontological efforts as adding another natural language (see also Johnston *et al*. 1995: 72), albeit artificially concocted, to the existing ones, while somehow claiming its priority.

By contrast, in the Mikrokosmos approach, an ontology for NLP purposes is seen not at all as a natural language but rather as a language-neutral "body of knowledge about the world (or a domain) that a) is a repository of primitive symbols used in meaning representation; b) organizes these symbols in a tangled subsumption hierarchy; and c) further interconnects these symbols using a rich system of semantic and discourse-pragmatic relations defined among the concepts" (Mahesh and Nirenburg 1995: 1). The names of concepts in the ontology may look like English words or phrases but their semantics is quite different and is defined in terms of explicitly stated interrelationships among these concepts. The function of the ontology is to supply "world knowledge to lexical, syntactic, and semantic processes" (*ibid*), and, in fact, we use exactly the same ontology for supporting multilingual machine translation.

An ontology like that comes at a considerable cost--it requires a deep commitment in time, effort, and intellectual engagement. It requires a well-developed methodology based on a clear theoretical foundation (see Mahesh 1996). The rewards, however, are also huge: a powerful base of primitives, with a rich content and rich inheritance, that is made available for the lexical entries, assuring their consistency and non-arbitrariness.[19] We address and reject as inapplicable (Nirenburg *et al*. 1995) the standard charge of irreproducibility for ontologies: on the one hand, we accept as expected that different groups and individuals will come up with different ontologies, even for a limited domain; on the other hand, we believe--and, actually, know for a well-established fact--that groups and individuals with similar training and facing an identical task would, indeed, come up with very similar ontologies. Note that no two grammars or lexicons of particular languages, even in a given theoretical paradigm, are expected by anybody to be identical!

---

[19] The MikroKosmos lexicons fit Fillmore and Atkins' (1992: 75) vision of an ideal dictionary of the future: "...we imagine, for some distant future, an online lexical resource, which we can refer to as a "frame-based" dictionary, which will be adequate to our aims. In such a dictionary (housed on a workstation with multiple windowing capabilities), individual word senses, relationships among the senses of the polysemous words, and relationships between (senses of) semantically related words will be linked with the cognitive structures (or 'frames'), knowledge of which is presupposed by the concepts encoded by the words."

To enhance the uniformity of ontology acquisition (for instance, by different acquirers), we have developed weak semi-automatic methods of acquisition, supported by semi-automatic acquisition tools. We have also discovered heuristic techniques and recurring patterns of acquisition. Again, this adds to the cost of lexical semantic work. This cost, however, buys a very desirable (at least, to us) result--the much enhanced depth and complexity of lexical entries not resulting from lexical acquisition but rather contributed "free of charge" by the inheritance, properties, and constraints on pre-acquired ontological concepts on which the lexical entries are based.[20]

6. Scope-Sensitive Vs. Non-Scope-Sensitive

This awkwardly named issue deals with a simple question: are we interested in any generalization or only in a generalization of a large enough scope? The answer is often determined by external reasons, usually due to the requirements of an application. It may be summarized as an answer to the following question: when is a rule or generalization worth pursuing, i.e., spending efforts on its discovery, definition, and application? A whole range of opinions is possible, on a case by case basis, when language field work is pursued. However, for many, the answer is the extreme "always" and is justified by a general methodological view of scientific endeavor. This tradition dates back to the Young Grammarians of the late 19th century (see, for instance, Osthoff and Brugman 1878, Paul 1886, Delbrück 1919; for a succinct summary, see also Zvegintzev 1964), who saw language as fully rule-governed and the minimum number of hopeless exceptions explained away on the basis of a wastebasket rule of analogy). This determinist position was, basically, endorsed by Bloomfield (1933), who added a physical (actually, neurological, in current terms) aspect to it.

The Chomskian paradigm sees language as a rule- or principle-governed activity; this position is justified by the belief, or hope, that every newly discovered rule may give us an additional glimpse into the mental mechanisms underlying language, a worthy scientific concern. The nasty question about this approach is this: what if a rule covers, say, only three lexical units? Is it not more efficient just to list them without generalizing about them? Is not that perhaps what our mind does as well?

The firmly negative answer ("never generalize") is not common in NLP applications these days--after all, some generalizations are very easy to make and exceptions to some rules do not faze too many people; morphology rules are a good example. A skeptical position on generalization, i.e., "generalize only when it is beneficial," is usually taken by developers of large-scale applications, having to deal with deadlines and deliverables. Only rules with respectable-sized scopes are typically worth pursuing according to this position (see Viegas *et al*. 1996b). The "nasty" question here is: are you ready then to substitute "a bag of tricks" for the actual rules of language? Of course, the jury is still out on the issue of whether language can be fully explained or modeled--short of really knowing what is going on in the mind of the native speaker--with anything which is not, at least to some extent, a bag of tricks.

Rules and generalizations can be not only expensive but also in need of corrective work due to overgeneralization; and this has been a legitimate recent concern (see, for instance, Copestake 1995, Briscoe *et al*. 1995). Indeed, a rule for forming the plurals of English nouns, though certainly justified in that its domain (scope) is vast, will produce, if not corrected, forms like *gooses* and *childs*. For this particular rule, providing a "stop list" of (around 200) irregular forms is relatively cheap and therefore acceptable on the grounds of overall economy. The rule for forming mass

[20] It is notable, again, that Fodor and Lepore (1996: 1) attack Pustejovsky (1995) for too much complexity in his lexical entries while we advocate here much more complex entries. See also fn. 10 above.

nouns determining meat (or fur) of an animal from count nouns denoting animals (as in *He doesn't like camel*), discussed in Copestake and Briscoe (1992) as the "grinding" rule is an altogether different story. The delineation of the domain of the rule is rather difficult (e.g., one has to deal with its applicability to *shrimp* but not to *mussel*; possibly to *ox* but certainly not to *heifer* or *effer*; and, if one generalizes to non-animal food, its applicability to *cabbage* but not *carrot*). Some mechanisms were suggested for dealing with the issue, such as, for instance, the device of 'blocking' (see Briscoe *et al*. 1995), which prevents the application of a rule to a noun for which there is already a specific word in the language (e.g., *beef* for *cow*). Blocking can only work, of course, if the general lexicon is sufficiently complete, and even then a special connection between the appropriate senses of *cow* and *beef* must be overtly made, manually.

Other corrective measures may become necessary as well, such as constraints on the rules, counter-rules, etc. They need to be discovered. At a certain point, the specification of the domains of the rules loses its semantic validity, and complaints to this effect are made within the approach (see, for instance, Briscoe and Copestake 1996 about such deficiencies in Pinker 1989 and B. Levin 1993; Pustejovsky 1995: 10 about B. Levin's 1993 classes).

Criticism of the generative lexicon approach becomes, at this point, similar methodologically to Weinreich's (1966) charge of infinite polysemy against Katz and Fodor (1963): if a theory does not have a criterion stipulating when a meaning should not be subdivided any further, then any superentry may be infinitely polysemous, and in Katz and Fodor's interpretive semantics, the all-important boundary between semantic markers, to which the theory is sensitive, and semantic distinguishers, ignored by the theory, is forever moving, depending on the grain size of description Ultimately, it is easy to show that semantic distinguishers may remain empty as one would need to include everything in the description into the theory.

Similarly, a semantic lexicon which stresses generalization faces the problem of having to deal with rules whose scope becomes progressively smaller, that is, the rule becomes applicable to fewer and fewer lexical units as the fight against overgeneration (including blocking and other means) is gradually won. At some point, it becomes methodologically silly to continue to formulate rules for creation of just a handful of new senses. It becomes easier to define these senses extensionally, simply by enumerating the domain of the rule and writing the corresponding lexical entries overtly.

Even if the need to treat exceptions did not reduce the scope of the rules postulated to do that, the relatively small size of the original scope of a rule, such as the grinding rule (see also Atkins 1991, Briscoe and Copestake 1991, Ostler and Atkins 1992), should cause some apprehension. The inapplicability of the grinding rule to meanings outside it should raise a methodological question about the nature of one's interest in this rule. Is it of interest to one *per se*, just because "it is there," one thinks, or is it representative of a certain *type* of rules? Unless one claims and demonstrates the latter, one runs a serious risk of ending up where the early enthusiasts of componential analysis found themselves, after years and years of perfecting the application of their tool to terms of kinship (see, for instance, Goodenough 1956, Greenberg 1949, Kroeber 1952, Lounsbury 1956): the semantic field of kinship was eventually found to be unique in rendering itself applicable to the method; other semantic fields quickly ran the technique into the ground through the runaway proliferation of semantic features which had to be postulated to ensure adequate coverage. In all the excellent articles on grinding and the blocking of grinding, we have found no explicit claims that the rule addresses a property applicable to other classes of words, thus leading to rules similar to the grinding rule. In other words, the concern for maximum generalization with one narrow word class is, inexplicably, not coupled with a concern for the portability of the methodology outside

that class.

We believe that the postulation and use of any small rule, without an explicit concern for its generalizability and portability, is not only bad methodology but also bad theory because a theory should not be littered with generalizations which are not overtly useful. The greater the number of rules and the smaller the classes which are their scopes, the less manageable--and elegant--the theory becomes. Even more importantly, the smaller the scope and the size of the class, the less likely it is that a formal syntactic criterion (test) can be found for delineating such a class (the use of such a criterion for each rule seems to be a requirement in the generative lexicon paradigm).This means that other criteria must be introduced, those not based on surface syntax observations. These criteria are, then, semantic in nature (unless they are observations of frequency of occurrence in corpora). We suspect that if the enterprise of delineating classes of scopes for rules is taken in a consistent manner, the result will be the creation of an ontology. As there are no syntactic reasons for determining these classes, new criteria will have to be derived, specifically, the criteria used to justify ontological decisions in our approach.

This conclusion is further reinforced by the fact that the small classes set up in the battle against overgeneralization are extremely unlikely to be independently justifiable elsewhere within the approach, which goes against the principle of independent justification which has guided linguistic theory since Chomsky (1965), which established the still reigning and, we believe, valid paradigm for the introduction of new categories, rules, and notational devices into a theory. Now, failure to justify a class independently opens it to the charge of *adhoc*-ness, which is indefensible in the paradigm. The only imaginable way out lies, again, in an independently motivated ontology.

7. Dynamic Vs. Static Resources

An important part of a computational theory, besides its format, the notation it introduces, and its foundation, is the architecture of the processing environment associated with the theory. We commented in Section 4 above, that the boundary between lexical semantics and the compositional module of semantic theory is not firm enough either for Pustejovsky or for the other scholars sharing his premises. We believe that much criticism of the approach on theoretical grounds is due to this lack of crispness on the matter. This explains a great deal of confusion about where the lexicon stops and the *terra incognita* of pragmatics begins: in fact, it makes the boundary between the two completely movable--the less you choose to put in the lexicon, the more pragmatics is out there. We believe that the semantic reality of language needs to be somewhat better defined than that, that a procedural component describing the generation of text meaning representations is an integral part of any semantic theory beyond the pure lexical semantics, as is a clear division of labor among the various modules of semantic theory. This does not exclude multidirectional dependencies and interactions but it does set up our expectations for each resource more accurately.

The Mikrokosmos environment, for example, contains the following **static** resources:

- an ontology, which includes script-like prototype complex events;
- a lexicon which has already subsumed the ontology (by being anchored in it) and the results of the application of lexical rules;
- the output of a syntactic parser; and
- the design of the text-meaning representation language (see Onyshkevych and Nirenburg 1994).

Besides these, ontological computational semantics needs at least the following **dynamic** resources:

- context information, including the analysis of sentences before and after the current one and extralinguistic context involving knowledge about a particular speech situation, its participants, etc.;
- rules of the (crucial) microtheory of semantic dependency construction, which is, among other things, connected with the microtheory of modification, including discrepancies between syntactic and semantic modification, for a system which does not give syntax a privileged status among the clues for semantic dependency;
- large family of analysis (dynamic) microtheories, such as those of aspect, time. modality, etc.

It is the dynamic resources which define our view on compositionality, and we fail to see how compositionality can be viewed without identifying such modules. We also think that 'compositionality' is an unfortunate term for the agglomerate of dynamic resources within semantic theory. After all, the traditional use of the term covers the combinatorics of assembling the meaning of the sentence out of the meanings of the words which make up the sentence, the combinatorics which involve selection restrictions in weeding out incompatible combinations of the senses of the polysemous words.

But semantic theory also requires resources which manipulate meanings of whole sentences. Thus, the context dynamic resource is not compositional at all. If everything non- or supracompositional is assigned by the generative lexicon approach to pragmatics, this will mean that this dynamic resource is taken out of semantics. We do not think it would be a good decision: not only would it assign something knowable to the land of the unknowable or only very partially and selectively knowable, but it will also declare each and every sentence uninterpretable within semantics.

Another blow against compositionality as a term and as an approach is the issue of the tension between the meaning of the text and word meaning. The compositional approach assumes the latter as a given, but one has to be mindful of the fact that word meaning is, for many linguists, only a definitional construct for semantic theory, "an artifact of theory and training" (Wilks 1996). Throughout the millennia, there have been views in linguistic thought that only sentences are real and basic, and words acquire their meanings only in sentences (see, for instance, Gardiner 1951, who traces this tradition back to the earliest Indian thinkers; Firth 1957, Zvegintzev 1968, and Raskin 1971 treat word meaning as a function of the usage of a word with other words in sentences). Of course, this opinion effectively denies the existence of lexical semantics as a separate field.

Any approach deriving text meaning centrally from lexical meaning is, basically, a **prototype-oriented theory.** The generative lexicon paradigm also attempts to account for the novel, creative senses (see Section 1.2). We do not necessarily want to argue against this approach, but we would like to stress that word meaning is secondary to text meaning as a theoretical construct. After Frege and Montague, it has become commonplace in formal semantics to take compositionality more or less uncritically as the sole basis of text meaning creation, while a more empirically driven view seeks to limit compositionality and strongly supplement it with supracompositional elements of meaning (cf. the construction grammar of Fillmore (1988) and Fillmore *et al*. (1988) and the findings of discourse researchers, as they go beyond sentence meaning and, therefore, still further away from building semantic representations in lockstep with syntactic processing).

Supracompositional elements are, of course, based in context, that is, in factoring contextual information into text meaning. Part of this context is obviously extralinguistic, which strengthens even further the necessity of combining world knowledge and linguistic (including lexical) knowledge into the computation of lexical meaning. The practical reality of NLP and computational linguistics makes it obvious for the community that world knowledge and linguistic knowledge cannot be separated. We believe strongly that the same is true of theoretical semantics, thus adding credence to Pustejovsky's thesis that computational work enriches linguistic theory: "I believe that we have reached an interesting turning point in research, where linguistic studies can be informed by computational tools for lexicology as well as appreciation of large lexical databases." (Pustejovsky 1995: 5). This view would have been even more appropriate if "lexicology" had been replaced by "semantics." The generative lexicon approach would have gained further credibility if it had acquired a principled view of what is there in semantic theory besides lexical semantics.

8. Depth (and Theory-Dependence) vs. Coverage of Description

The generative lexicon approach shares with the rest of theoretical linguistics the practice of high selectivity with regard to its material. This makes such works great fun to read: interesting phenomena are selected; borderline cases are examined. In the generative lexicon approach, new and relatively unexplored lexical rules have been focused upon, at the expense of large-scope rules that may be more obvious. In theoretical linguistics, borderline cases are dwelt upon as testing grounds for certain rules that may prop up or expose vulnerability in a paradigm. In both cases, an assumption is tacitly made that the ordinary cases are easy to account for, and so they are not processed. As we mentioned elsewhere (Raskin and Nirenburg 1995), in the whole of transformational and post-transformational semantic theory, only a handful of examples has ever been actually described, with no emphasis on coverage.

Contrary to that, large-scale applications require the description of every lexical-semantic phenomenon (and a finer-grained description than that provided by a handful of features, often conveniently borrowed from syntax), and the task is to develop a theory for such applications underlying a principled methodology for complete descriptive coverage of the material. The implementation of any such project would clearly demonstrate that the proverbial common case is not so common: there are many nontrivial decisions and choices to make, many of them widely applicable and extrapolable to large classes of data. In practice, the descriptive task is always inseparable from pure theory. This was well understood by the American descriptivists of the first half of the century, but it has not been part of the linguistic experience--or education--for several decades now (see also Section 4.2 above).

Good theorists carry out descriptive work in full expectation that a close scrutiny of data will lead to, often significant, modifications of their *a priori* notions. Thus, the sizable theoretical-linguistic scholarship on the lexical category of adjective barely touches on the concept of scale (see Raskin and Nirenburg 1995: 4-21), while even a cursory look at the data shows that it is very natural to represent the meaning a prevalent semantic subclass of adjectives[21] using scales: e.g., *big* (scale: SIZE), *good* (scale: QUALITY), or *beautiful* (scale: APPEARANCE). Consequently, the discovery of a few dozen scale properties underlying and determining the meaning of the statistically dominant subclass of English adjectives becomes an important descriptive subtask, basically completely unanticipated by the preceding theory.

[21] It is especially true of English, where the grammatical realization of non-scalar, denominal modifiers is usually nominal.

Conversely, the much touted borderline case of the relative/qualitative adjective, such as *old* used relatively in *old boyfriend*, in the sense of 'former,' on the one hand, and qualitatively in *old man*, on the other, the case which has been presumed to be central to the study of the adjective (*op. cit.*: 16-17), has very little descriptive significance or challenge to it, fading and merging with other numerous cases of multiple senses of the same word within the adjective superentry and receiving the same, standard treatment.

9. Generalization vs. Idiosyncraticity

There are many reasons to attempt to write language descriptions in the most general manner --- the more generally applicable the rules, the fewer rules need to be written; the smaller the set of rules (of a given complexity) can be found to be sufficient for a particular task, the more elegant the solution, etc. In the area of the lexicon, for example, the ideal of generalizability and productivity is to devise simple entries which, when used as data by a set of syntactic and semantic analysis operations, regularly yield predictable results in a compositional manner. To be maximally general, much of the information in lexical entries should be inherited, based on class membership or should be predictable from general principles.

However, experience with NLP applications shows that the pursuit of generalization promises only limited success. In a multitude of routine cases, it becomes difficult to use general rules--Briscoe and Copestake (1996) is an attempt to alleviate this problem through nonlinguistic means. The enterprise of building a language description maximizing the role of generalizations is neatly encapsulated by Sparck Jones: "We may have a formalism with axioms, rules of inference, and so forth which is quite kosher as far as the manifest criteria for logics go, but which is a logic only in the letter, not the spirit. This is because, to do its job, it has to absorb the *ad hoc* miscellaneity that makes language only approximately systematic" (1989, p.137).

This state of affairs, all too familiar to anybody who has attempted even a medium-scale description of an actual language beyond the stages of morphology and syntax, leads to the necessity of directly representing, usually in the lexicon, information about how to process small classes of phenomena which could not be covered by general rules. An important goal for developers of NLP systems is, thus, to find the correct balance between what can be processed on general principles and what is idiosyncratic in language, what we can calculate and what we must know literally, what is compositional and what is conventional. In other words, the decision as to what to put into a set of general rules and what to store in a static knowledge base such as the lexicon becomes a crucial early decision in designing computational-linguistic theories and applications.

The trade-off between generality and idiosyncraticity is complicated by the possibility of the following compromise (cf. L.Levin and Nirenburg 1994). If direct generalizations cannot be made, some researchers say that there may still be a possibility that the apparent variability in grammar rules and lexicon data can be accounted for by parameterization: there may exist a set of universal parameters that would explain the differences among various phenomena in terms of the difference in particular parameter settings. For example, in HPSG, idiosyncratic information could appear in the sort hierarchy of sentence types and would be subject to the principles of grammar (such as the head feature principle and subcategorization principle) along with other sentence types. This is better than dealing with completely ungeneralized material. But the search for a set of universal parameters, however important, does not, in our opinion, hold a very bright promise from the standpoint of coverage. More important, therefore, is the development of a principled methodology of dealing with idiosyncratic information in the lexicon; in other words, instead of explaining away

every idiosyncracy in terms of a non-*ad hoc* general rule, that needs to be discovered or postulated and independently justified, developing a standard methodology for dealing with the exceptions qua exceptions in a systematic and efficient way. This is a different route for generalization, interesting theoretically and efficient in NLP practice.

The extent to which ungeneralizable--or at least not easily generalizable--idiosyncraticity needs to be reflected in the description of the lexicon is defined largely by the computational task at hand and the coverage/quality ratio that determines the optimal grain size of the description. Once determined, however, everything falling within that grain size has to be accounted for independently of whether there is a nice general rule for it or not. This pressure works against a theory, where the methodology calls for each category and each rule to be maximally extended. And it is no wonder that the simple, more obvious, and easily formalizable rules of syntax tempt lexical and compositional semanticists into using them, often resulting in diverse semantic phenomena being lumpted together. Thus, it must bother a theorical linguist working on adjectives to discover that the distinction between the attributive and predicative use of adjectives, a near-universal across languages, does not have much semantic significance.

The generality/idiosyncraticity balance in computational semantics conforms to the principle of optimal efficiency of analysis, and there is much more to the theoretical basis of this principle than just ergonomic or commonsense-related considerations. As our work on the microtheory of adjectival semantics (Raskin and Nirenburg 1995) has abundantly demonstrated, computational semantics must take full advantage of the related linguistic theory and transform its knowledge, in a principled, systematic, documented way, to work best in computational semantic descriptions. There are two ways this systematic transformation of knowledge can be conceptualized. On the one hand, it is an applied theory, which mediates between linguistic theory and computational analysis, thus serving as the theoretical basis for an application of linguistics to natural language processing (see, for instance, Raskin 1971, 1987a,b). On the other hand, it is a methodology because the applied theory "induces" a set of methods and techniques for doing so.[22]

---

[22] There is a great deal of confusion in linguistic and computational linguistic literature, as indeed in science and even philosophy of science in general, about the usage of the terms 'theory' and 'methodology.' We approach this issue in the Chomsky (1965) tradition (albeit not followed up or strictly adhered to by himself) when we see theory, *T*, as a set of statements defining the format of linguistic descriptions, *D*, for the data set, *S*, within the range of the theory, up to the shape of various kinds of brackets, so and methodology, *M* (see Raskin and Nirenburg 1995: 3) as a set of licensed techniques for implementing those description on the basis of the theory, so that $M(T, S) = D$. (It is precisely because methodology is so completely based on theory that the above confusion may occur. Another contributing factor is, of course, the non-terminological use of the word *theory* to denote any generalization or systematicity. The still available dictionary definition of *methodology* as a "science of methods"is not helpful either because, in reality, there is no such science that would deal with the development of methods for any discipline or field--instead, in every discipline, the methodology is indeed a function, a derivative, as it were, of the theory and the data.). A specific example of this conceptual relationship, with regard to the by now familiar case of adjectival semantics, can be presented as follows: a (good) linguistic **theory** will establish the role of scale-type properties in adjectival meaning, and, accordingly, the adjective entries in the lexicon will be based on scales, along the lines of Raskin and Nirenburg (1995); an **applied theory** will anchor the adjectival scales on ontological properties and determine the list of those scales which will appear in the actual lexical entries; a **methodology** will determine both the development of the appropriate part of the ontology and the list of scale properties. An important and controversial part of methodology is the provenance of heuristics (*op. cit.*: 55-59; see also Attardo 1996).

10. Supply-Side Vs. Demand-Side

Two distinct methodological positions can be detected in lexical semantics today. The generative lexicon approach belongs firmly to what could be called the supply-side school of thought in contemporary lexical semantics (see Nirenburg 1996), while the Mikrokosmos project belongs to the demand-side school. Let us try and define the difference as well as the commonalities between the two approaches.

Traditions of research in linguistic semantics can be distinguished in the following way suggested by Paul Kay: "Students concerned with lexical fields and lexical domains ('lexical semanticists') have interested themselves in the paradigmatic relations of contrast that obtain among related lexical items and the substantive detail of how particular lexical items map to the nonlinguistic objects they stand for. 'Formal semanticists' (those who study combinatorial properties of word meanings) have been mostly unconcerned with these issues, concentrating rather on how the meanings of individual words, whatever their internal structure may be and however they may be paradigmatically related to one another, combine into the meanings of phrases and sentences (and recently, to some extent, texts). Combinatorial semanticists have naturally been more concerned with syntax, especially as the leading idea of formal semantics has been the specific combinatorial hypothesis of Fregean compositionality" (1992: 309).

Formal semantics, as defined by Kay, is rejected by both sides in lexical semantics (see, however, Section 4.2 for further clarification). While there are many distinctions among the approaches to lexical semantics, we would like to focus here on one dimension of differences. Some researchers are concentrating on describing the paradigmatic relations of contrast and, more importantly, semantic derivation relations among lexical meanings without a reference to knowledge about the world (Kay's "nonlinguistic objects") and, consequently, deemphasizing the definition of core lexical meaning. Some other researchers stress the mapping between lexical items and nonlinguistic objects they stand for and, consequently, make the definition of core lexical meaning a central goal. This task is declared by the former group to be outside their purview: "Undoubtedly, the inferential processes involved in language comprehension extend beyond the limited mechanisms provided within unification-based formalisms; however, it is not clear yet whether lexical operations *per se* require them" (Briscoe 1993: 11; see also Section 2).

The former group also takes the task of describing lexical meaning to be almost seamlessly connected to lexicalized syntactic theory: "[T]he role of the lexicon in capturing linguistic generalizations[:] more and more of the rules of grammar are coming to be seen as formal devices which manipulate (aspects of) lexical entries, and in the sense that many of these rules are lexically governed and must, therefore, be restricted to more finely specified classes of lexical items than can be obtained from traditional part-of-speech classifications" (Briscoe 1993: 2), whereas the other group does not treat syntactic information as privileged but just as one of many clues helping to determine meaning.

Methodologically, the first group is pursuing the formulation of lexical meaning theories as algebraic entities in which the maximizing factor is formal elegance, descriptive power (attained, for instance, through generalization of rules), economy of descriptive means, and absence of exceptions. As a result of this, difficult issues which cannot at present be subject to such discipline are not treated--they are either ignored or declared to be outside the purview of the theory. Thus, for instance, the issue of basic lexical meaning is rarely discussed in this tradition, while attention centers on regularity of meaning **shifts** (and, most recently, **differentiae**--see, for instance, Hirst 1995 or Johnston et al. 1995) under the influence of paradigmatic morphosyntactic transformations and

of unexpected syntagmatic cooccurrence of lexical forms in texts. The first group of researchers is much more dependent, therefore, than the second one on the availability of such ready-made or easily obtainable artifact resources as machine-readable dictionaries, tagged corpora, frequency lists, etc.[23] It is the methodology of the first group that can be reasonably and inoffensively referred to as a supply-side approach: it is based only on what can be offered by the state of the art (or, we should perhaps say, science) in the formal description of linguistic theory divorced from world knowledge.

For supply-siders, the main issues include (among others):

- lexical semantics as outgrowth of lexical grammar, or grammatical semantics;
- lexical semantics as a counterpoint to formal semantics, i.e. an emphasis on lexical meaning rather than on compositional sentence meaning;
- formalisms for representing lexical knowledge (e.g., the LRL of the ACQUILEX project): feature structures, typed feature structures and rules for their symbolic manipulation, default inheritance hierarchies, etc.;
- establishing lexical rules for relating word senses;
- constraining the power of lexical rules so that they do not overgenerate;
- capturing valuable generalizations about applicability of lexical rules.

A strong temptation for a supply-sider is provided by the availability or an easy accessibility of a computational procedure that seems to yield results that, when a real application is contemplated, may prove to be usable. How about, say, a procedure which runs over a very large corpus and furnishes a word list? Can one assume that this procedure should be immediately adopted as part of an NLP project of the future? It is fair to say that, for a supply sider, the answer is yes. In fact, a supply-sider's arsenal, consists of any such available or easily developable procedures. [Church's exploratory data analysis].

The second faction tends to be much more cautious with respect to such procedures. Before incorporating one or committing to develop another such procedure, they want to make sure that there is a legitimate place for this procedure in the overall architecture of an NLP system they are developing. Thus, is it a given that one needs a word list at any stage of analysis or generation? Or, is it possible that the words will appear at the input and need to be analyzed morphologically and lexically as part of the system analyzer, which will never actually use the word list as a useful resource?[24]

This "killsport" group's theoretical work is different in kind in other ways as well. Wilks (1994: 586) illustrates this difference well: "There is a great difference between linguistic theory in Chomsky's sense, as motivated entirely by the need to explain, and theories, whether linguistic, AI or whatever, as the basis of procedural, application-orientated accounts of language. The latter stress testability, procedures, coverage, recovery from error, non-standard language, metaphor, textual content, and the interface to general knowledge structures." Thus, the methodology of the second

[23] See, for instance, Ageno *et al*. (1992), Ahlswede *et al*. (1985), Amsler (1982, 1984a,b), Boguraev and Briscoe (1987, 1989), Boguraev *et al*. (1987), Calzolari (1984), Chodorow *et al*. (1985), Copestake (1990, 1992), Copestake *et al*. (1994/1995), Dorr *et al*. (1994/1995), Farwell *et al*. (1993), Lonsdale *et al*. (1994/1995), Markowitz *et al*. (1986), Pustejovsky *et al*. (1993), Sanfilippo and Poznanski (1992), Slator and Wilks (1987), Slocum and Morgan (1986), Voss and Dorr (1995), Walker (1984, 1987), Walker and Amsler (1986), Wilks *et al*. (1987), Wu and Xia (1994/1995), Zernik (1991).

faction can be characterized as **demand-side**, as it pursues theories which are capable of supporting practical applications.[25]

Note that the theories produced by the supply-side group can also strive to support practical applications, but they are clearly a side effect of theoretical pursuits. In practice, a lot of additional work is always needed to implement a supply-side theory as a computer program, because such theories are usually not formulated with such processing in mind and often do not easily lend themselves to such applications. The problems with demand-side theories include difficulties in algebraic definition and testability and falsifiability exclusively through experimentation.

Burning issues for the demand-siders include (among others):

- determining the number of lexemes in a lexicon (breadth);
- establishing criteria for sense specification and delimitation;
- granularity issue I: determining the threshold of synonymy (beyond which two word senses would share a meaning);
- granularity issue II: determining the threshold of ambiguity (that is, the appropriate number of senses for a lexeme, whether listed or derived with the help of lexical rules);
- tuning the depth and breadth of lexical description to the needs of a particular application;
- enhancing the levels of automaticity in lexical acquisition.

Specific issues which have, to a greater or lesser degree, been proven important to both supply- and demand-siders include:

- the theoretical and applicational status of the lexical rules;
- the typology and inventory of the lexical rules;

[24] One may be tempted to think here that the two approaches being opposed correspond to the dichotomy between the linguistic engineering (LE) approach, for supply-side, and NLP, for the other side (see also Sheremetyeva 1997 for more discussion of LE and its relation to NLP). In such a dichotomy, the distinctions run, apparently, along these lines: NLP is more theoretical while LE is more practical; NLP is interested both in developing working systems for particular tasks and--probably even more so--in using computer implementations as confirmations or falsifications of hypotheses about human processing of language, while LE is interested only in the former; accordingly, NLP is likely to pay less interest than LE to the available resources and ready-made and easily developable procedures to set and achieve the realistic goals of today, favoring instead more theoretical research for developing superior systems in the future. The reality is, however, different. Many supply-siders would welcome LE as just sketched; it is they, however, who typically work towards systems of the future, warehousing all the easily available online resources and procedures as automatically usable in those systems. It is the other faction which has to be cautious about such a general mobilization of all available techniques because the demand-side paradigm is always, by definition, interested in actually developing a system and has to concern itself with the compatibility of such a resource or procedure and the general system architecture. In fact, to put it rather bluntly, the automatic approval of each online resource or procedure for potential use is bad engineering; highly principled selectivity is good engineering. Leaving bad LE aside and without ascribing this defenseless position to anybody in the field, we can conjecture that the theoretical difference between NLP and good LE lies in that the former applies the entire theory of language to the development of language processing systems while LE applies a reasonable and justified subset of such a theory, with more limited goals. What follows from this position is that both approaches share criteria of quality and use only those resources which fit together and promote the optimal results as opposed to those resources, whose **only** merit is their easy availability.

• the place of lexical rules in the control structure of an NLP system.

Obviously, both sides are interested in lexical meaning, and while the one side is driven by the availability of acceptable tools and the other by the actual practical goals, both also share a commitment to a common theoretical paradigm, which is definitely post-Chomskian.

11. Conclusion

Throughout the paper, we have emphasized the importance for lexical semanticists, independently of the choices they make on the ten issues, to be explicit about their premises and goals. Frequently, when lexical semanticists meet, they politely perpetuate the illusion that they all work towards the same goal of creating and optimizing large lexicons for non-toy meaning-based NLP systems, such as MT systems. It is clear that, for some approaches, this is a more or less remote goal, and they assume that every development in their own approach automatically brings us closer to it. For others, this goal is the next deadline and the subject of their grant report.

These different scientific and sociological realities often determine the crucial distinctions in priorities and values as well as in the choice of theoretical and methodological support. A misunderstanding of these distinctions leads to attempts to judge one approach by the rules of another, and there is not much value in that. A grossly misguided review of Pustejovsky (1995) in Fodor and Lepore (1996) is a prime, if exaggerated example of such an attempt: the authors, coming from far outside of lexical and linguistic semantics, seem to chide Pustejovsky for not being a good philosopher of language!

We hope that our review of some of the major theoretical and methodological choices one can make within lexical semantics will contribute to a better understanding of what each of us is doing. It is in this spirit and from this perspective that we have felt quite unabashed in promulgating our own choices as corresponding to our goals and in questioning the appropriateness of the choices made in other approaches with regard to their professed or implied goals.

Acknowledgements.

The research reported in this paper was supported by Contract MDA904-92-C-5189 with the U.S. Department of Defense. Both authors feel indebted to the other members of the MikroKosmos team. The authors are grateful to James Pustejovsky and Yorick Wilks for constructive criticism.

References

Ageno, A., I. Castellon, G. Rigau, H. Rodriguez, M. F. Verdejo, M. A. Marti, and M. Taule 1992. SEISD: An environment for extraction of semantic information from online dictionaries. In: **Proceedings of the 3rd Conference on Applied Natural Language Processing--ANLP '92**. Trento, Italy, pp. 253-255.

[25] This methodology should not be squarely equated with language engineering, either. Indeed, the demand-side tradition does not, as is often believed, presuppose using a pure "bag of tricks" approach. Exponents of this approach, we have demonstrated our overriding concern for theoretical foundations for everything we do throughout this paper. The approach is based on the notion of a society of microtheories describing language as well as language processing phenomena (e.g., meaning assignment heuristics), and it grounds the lexical meaning in an artificial model of the world, as seen by speaker/hearer; the theory also includes the specification of a formalism (for representing not only lexical meaning and world knowledge but also the meaning of texts) and of a computational system for deriving text meaning from text input as well as producing text output from text meaning representation.

Ahlswede, Thomas E., Martha Evans, K. Rossi, and Judith Markowitz 1985. Building a lexical database by parsing Webster's Seventh Collegiate Dictionary. In: **Advances in Lexicology. Proceedings**. Waterloo, Ontario, Canada, pp. 65-78.

Amsler, Robert A. 1982. Computational lexicology: A Research Program. In: **AFIPS. Proceedings of the National Computer Conference**, Vol. V, pp. 657-663.

Amsler, Robert A. 1984a. Lexical knowledge bases. In: **Proceedings of COLING '84**. Stanford, CA: Stanford University, pp. 458-459.

Amsler, Robert A. 1984b. Machine-readable dictionaries. In: M. E. Williams (ed.), **Annual Review of Information Science and Technology**, Vol. 19. White Plains, N.Y.: Knowledge Industry Publications, pp. 161-209.

Apresyan Yury D., Igor' A. Mel'čuk, and Alexander K. Zholkovsky 1969. Semantics and lexicography: Towards a new type of unilingual dictionary. In: F. Kiefer (ed.), **Studies in Syntax and Semantics**. Dordrecht: Reidel, pp. 1-33.

Apresyan Yury D., Igor' A. Mel'čuk, and Alexander K. Zholkovsky 1973. Materials for an explanatory combinatory dictionary of Modern Russian. In: F. Kiefer (ed.), **Trends in Soviet Theoretical Linguistics**. Dordrecht: D. Reidel, pp. 411-438.

Asher, Nicholas, and Alex Lascarides 1995. Metaphor in discourse. In: Klavans *et al*., pp. 3-7.

Atkins, B. T. S. 1991. Building a lexicon: The contribution of lexicography. In: Boguraev, pp. 167-204

Attardo, Donalee Hughes 1996. Lexicographic Acquisition: A Theoretical and Computational Study of Linguistic Heuristics. An unpublished Ph.D. thesis, Graduate Interdepartmental Program in Linguistics, Purdue University, West Lafayette, Indiana.

Bartsch, Renate, and Theo Venemann 1972. **Semantic Structures**. Frankfort: Athenaum.

Barwise, Jon, and John Perry 1983. **Situations and Attitudes**. Cambridge, MA: M.I.T. Press.

Beale, Stephen, Sergei Nirenburg, and Kavi Mahesh 1995. Semantic Analysis in the Mikrokosmos Machine Translation Project. In **Proceedings of the Second Symposium on Natural Language Processing (SNLP-95)**, August 2-4. Bangkok, Thailand.

Bendix, Edward H. 1966. **Componential Analysis of General Vocabulary**. Indiana University Research Center in Anthropology, Folklore, and Linguistics, Publication 41. Bloomington, IN: Indiana University.

Benson, Morton, Evelyn Benson, and Robert Ilson 1986. **Lexicographic Description of English**. Amsterdam: John Benjamins.

Bergler, Sabine 1995. Generative lexicon principles for machine translation: A case for meta-lexical structure. In: Dorr and Klavans, pp. 155-182.

Berry, Margaret 1975. **Introduction to Systemic Linguistics 1: Structures and Systems**. New York: Saint Martin's Press.

Berry, Margaret 1977. **Introduction to Systemic Linguistics 2**: Levels and Links. London: Batsford.

Bloomfield, Leonard 1933. **Language**. New York: Holt.

Boguraev, Branimir (ed.) 1991. Special Issue: **Building a Lexicon**. *International Journal of Lexicography* 4:3.

Boguraev, Bran, and Ted Briscoe 1987. Large lexicons for natural language processing: Exploring the grammar coding system for LDOCE. *Computational Linguistics* 13, pp. 203-218.

Boguraev, Bran, and Edward J. Briscoe 1989. **Computational Lexicography for Natural Language Processing**. London: Longman.

Boguraev, Bran, Ted Briscoe, J. Carroll, D. Carter, and C. Grover 1987. The derivation of a grammatically indexed lexicon from the Longman Dictionary of Contemporary English. In: **Proceedings of ACL '87**. Stanford, CA: Stanford University, pp. 193-200.

Bolinger, Dwight 1965. Atomization of meaning. *Language* 41:4, pp. 555-573.

Bresnan, Joan (ed.) 1982. The Mental Representation of Grammatical Relations. Cambridge, MA: MIT Press.

Briscoe, T. 1993. Introduction. In: Briscoe *et al*., pp.1-12.

Briscoe. Edward J., and Ann Copestake 1991. Sense extensions as lexical rules. In: **Proceedings of the IJCAI Workshop on Computational Approaches to Non-Literal Language**, Sydney, Australia, pp. 12-20.

Briscoe, Ted, and Ann Copestake 1996. Controlling the application of lexical rules. In: Viegas *et al*., pp. 7-19.

Briscoe, Edward J., Ann Copestake, and Bran Boguraev 1990. Enjoy the paper: Lexical semantics via lexicology. In: **Proceedings of COLING '90**. Helsinki, Finland, pp. 42-47.

Briscoe, Edward J., Ann Copestake, and Alex Lascarides 1995. Blocking. In: Patrick Saint-Dizier and Evelyne Viegas (eds.), **Computational Lexical Semantics**. Cambridge: Cambridge University Press, pp. 273-302.

Briscoe, Ted, Valeria de Paiva, and Ann Copestake (eds.) 1993. **Inheritance, Defaults, and the Lexicon**. Cambridge: Cambridge University Press.

Bruce, B. 1975. Case system for natural language. *Artificial Intelligence* 6, pp. 327-360.

Calzolari, Nicoletta 1984. Machine-readable dictionaries, lexical database and the lexical system. In: **Proceedings of COLING '84**. Stanford, CA: Stanford University, p. 460.

Cann, Ronnie 1991. **Formal Semantics: An Introduction**. Cambridge: Cambridge University Press.

Carlson, Lynn, and Sergei Nirenburg 1990. World modeling for NLP. Technical Report CMU-CMT-90-121, Center for Machine Translation, Carnegie Mellon University, Pittsburgh, PA.

Charniak, Eugene, and Yorick Wilks (eds.) 1975. **Computational Semantics: An Introduction to Artificial Intelligence and Natural Language Comprehension**. Amsterdam: North-Holland.

Chierchia, Gennaro 1995. **Dynamics of Meaning: Anaphora, Presupposition, and the Theory of Grammar**. Chicago-London: University of Chicago Press.

Chierchia, Gennaro, and Sally McConnell-Ginet 1990. **Meaning and Grammar: An Introduction to Semantics**. Cambridge, MA-London: M.I.T. Press.

Chodorow, Martin S., Roy J. Byrd, and George E. Heidorn 1985. Extracting semantic hierarchies from a large on-line dictionary. In: **Proceedings of ACL '85**. Chicago: University of Chicago, pp. 299-304.

Chomsky, Noam 1957. **Syntactic Structures**. The Hague: Mouton.

Chomsky, Noam 1965. **Aspects of the Theory of Syntax**. Cambridge, MA: M.I.T. Press.

Computational Linguistics 1987. Special Issue on the Lexicon. 13: 3-4.

Copestake, Ann 1990. An approach to building the hierarchical element of a lexical knowledge base from a machine readable dictionary. In: **Proceedings of the First International Workshop on Inheritance in Natural Language Processing**. Toulouse, France, pp. 19-29.

Copestake, Ann 1992. The ACQUILEX LKB: Representation issues in semi-automatic acquisition of large lexicons. In: **Proceedings of the 3rd Conference on Applied Natural Language Processing--ANLP '92**. Trento, Italy, pp. 88-96.

Copestake, Ann 1995. Representing lexical polysemy. In: Klavans et al., pp. 21-26.

Copestake, Ann, and Ted Briscoe 1992. Lexical operations in a unification-based framework. In: Pustejovsky and Bergler, pp. 101-119.

Copestake, Ann, Ted Briscoe, Piek Vossen, Alicia Ageno, Irene Castellon, Francesc Ribas, German Rigau, Horacio Rodríguez, and Anna Samiotou 1994/1995. Acquisition of lexical translation relations from MRD's. In: Dorr and Klavans pp. 183-219.

Cruse, D. A. 1986. **Lexical Semantics**. Cambridge: Cambridge University Press.

Delbrück, B. 1919. **Einleitung in das Studium der indogermanischen Sprachen**, 6th ed. Leipzig.

Dillon, George L. 1977. **Introduction to Contemporary Linguistic Semantics**. Englewood Cliffs, N.J.: Prentice-Hall.

Dolgopol'skiy, Aron B. 1962. Izuchenie leksiki s tochki zreniya transformatsionnogo analiza plana soderzhaniya yazyka /A study of lexics from the point of view of the transformational analysis of the content plane of language. In: **Leksikograficheskiy sbornik** 5. Moscow: Nauka.

Dorr, Bonnie J. 1993. **Machine Translation: A View from the Lexicon**. Cambridge, MA: M.I.T. Press.

Dorr, Bonnie J., Joseph Garman, and Amy Weinberg 1994/1995. From syntactic encodings to thematic roles: Building lexical entries for interlingual MT. In: Dorr and Klavans, pp. 221-250.

Dorr, Bonnie J., and Judith Klavans (eds.) 1994/1995. Special Issue: **Building Lexicons for Machine Translation I**. *Machine Translation* 9: 3-4.

Dorr, Bonnie J., and Judith Klavans (eds.) 1995. Special Issue: **Building Lexicons for Machine Translation II**. *Machine Translation* 10: 1-2.

Dowty, David 1979. **Word Meaning and Montague Grammar**. Dordrecht: D. Reidel.

Enç, Mürvet 1988. The syntax-semantics interface. In: Frederick J. Newmeyer (ed.), **Linguistics: The Cambridge Survey. Vol. I. Linguistic Theory: Foundations**. Cambridge: Cambridge University Press, pp. 239-254.

Farwell, David, Louise Guthrie, and Yorick A. Wilks 1993. Automatically creating lexical entries for ULTRA, a multilingual MT system. *Machine Translation* 8:3, pp.127-146.

Fauconnier, Gilles 1985. **Mental Spaces**. Cambridge, MA: M.I.T. Press.

Fillmore, Charles J. 1968. The case for case. In: Emmon Bach and Robert T. Harms (eds.), **Universals in Linguistic Theory**. New York: Holt, Rinehart, and Winston, pp. 1-88.

Fillmore, Charles J. 1971. Types of lexical information, In: Steinberg and Jakobovits, pp. 370-392.

Fillmore, Charles J. 1977. The case for case reopened. In: Peter Cole and Jerry M. Sadock (eds.), **Syntax and Semantics** 8. New York: Academic Press, pp. 59-81.

Fillmore, Charles J. 1978. On the organization of semantic information in the lexicon. In: Donna Farkas *et al*. (eds.), **Papers from the Parasession on the Lexicon**. Chicago: Chicago Linguistic Society, pp.148-173.

Fillmore, Charles J. 1988. The mechanisms of "construction grammar." Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society. Berkeley, CA: University of California, pp. 35-55.

Fillmore, Charles J., and Beryl T. Atkins 1992. Toward a frame-based lexicon: The semantics of RISK and its neighbors. In: Lehrer and Kittay, pp. 75-102.

Fillmore, Charles J., Paul Kay, and M. C. Connor 1988. Regularity and idiomaticity in grammar: The case of *let alone*. *Language* 64: 3, pp. 501-538.

Firth, J. R. 1957. Modes of meaning. In: J. R. Firth, **Papers in Linguistics**. London-Oxford: Oxford University Press.

Fodor, Janet Dean 1977. **Semantics: Theories of Meaning in Generative Grammar**. Hassocks, Sussex: Harvester.

Fodor, Jerry, and Ernest Lepore 1996. The emptiness of the lexicon: Critical reflections on J. Pustejovsky's The Generative Lexicon. *Linguistic Inquiry* (in press).

Frawley, William 1992. **Linguistic Semantics**. Hillsdale, N.J.: Erlbaum.

Frawley, William, and R. Steiner (eds.) 1985. Special Issue: **Advances in Lexicography.** *Papers in Linguistics* 18.

Frege, Gottlob 1892. Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik* 100, pp. 25-50. English translation in: Peter T. Geach and Max Black (eds.), **Translations from the Philosophical Writings of Gottlob Frege**, 2nd ed. Oxford: Blackwell, 1966, pp. 56-78.

Gardiner, Alan 1951. **The Theory of Speech and Language**, 2nd ed. Oxford: Clarendon Press.

Garvin, Paul L., J. Brewer, and Madeleine Mathiot 1967. **Predication-typing: A pilot study in Semantic Analysis**. Supplement to *Language* 43: 2, Part. 2.

Gibson, Edward. 1991. **A Computational Theory Of Human Linguistic Processing: Memory Limitations And Processing Breakdown.** Ph.D. Dissertation. Computational Linguistics Program. Carnegie Mellon University.

Givón, Talmy 1967. **Transformations of Ellipsis, Sense Development and Rules of Lexical Derivation**. SP-2896. Santa Monica, CA: Systems Development Corporation.

Goodenough, W. 1956. Componential analysis and the study of meaning. *Language* 32:1.

Greenberg, Joseph 1949. The logical analysis of kinship. *Philosophy of Science* 6: 1.

Gruber, Jeffrey 1965/1976. Studies in Lexical Relations. An unpublished Ph.D. thesis, MIT, Cambridge, MA, 1965. A revised and extended version published as: **Lexical Structures in Syntax and Semantics**. New York: North-Holland, 1976.

Halliday, M. A. K. 1961. Categories of the theory of grammar. *Word* 17, pp. 241-92.

Halliday, M. A. K. (ed.) 1983. **Readings in Systemic Linguistics**. London: Batsford.

Halliday, M. A. K. 1985. **An Introduction to Functional Grammar**. London-Baltimore: E. Arnold.

Hayes, Patrick 1979. The naive physics manifesto. In: D. Mitchie (ed.), **Expert Systems in the Microelectronic Age**. Edinburgh: Edinburgh University Press.

Harris, Randy Allen 1993. **The Linguistics Wars**. New York-Oxford: Oxford University Press.

Hirst, Graeme 1995. Near-synonymy and the structure of lexical knowledge. In: Klavans et al., pp. 51-56.

Hobbs, Jerry 1987. World knowledge and word meaning. In: Yorick Wilks (ed.), **TINLAP-3. Theoretical Issues in Natural Language Processing-3**. Las Cruces, N.M.: Computational Research Laboratory, New Mexico State University, pp. 20-25.

Hobbs, Jerry, William Croft, Todd Davies, Douglas Edwards. and Kenneth Laws 1987. Commonsense metaphysics and lexical semantics. *Computational Linguistics* 13: 3-4, pp. 241-250.

Hobbs, Jerry, and Robert C. Moore 1985. **Formal Theories of the Commonsense World**. Norwood, N.J.: Ablex.

Hovy, Edward 1988. Generating Natural Language Under Pragmatic Constraints. An unpublished Ph.D. thesis, Department of Computer Science, Yale University, New Haven, CT.

Hornstein, Norbert 1995. **Logical Form: From GB to Minimalism**. Oxford-Cambridge, MA: Blackwell.

Jackendoff, Ray 1983. Semantics and Cognition. Cambridge, MA-London: M.I.T. Press.

Johnston, Michael, Branimir Boguraev, and James Pustejovsky 1995. The acquisition and interpretation of complex nominals. In: Klavans *et al*., pp. 69-74.

Katz, Jerrold J. 1967. Recent issues in semantic theory. *Foundations of Language* 3, pp. 124-194.

Katz, Jerrold J. 1970. Interpretive semantics vs. generative semantics. *Foundations of Language* 6, pp. 220-259.

Katz, Jerrold J. 1971. Generative semantics is interpretive semantics. *Linguistic Inquiry* 2, pp. 313-331.

Katz, Jerrold 1972. Semantic theory and the meaning of good. *Journal of Philosophy* 61, pp. 736-760.

Katz, Jerrold J. 1972/1974. **Semantic Theory**. New York: Harper and Row.

Katz, Jerrold J. 1978. Effability and translation. In: F. Guenthner and M. Guenthner-Reutter (eds.), **Meaning and Translation: Philosophical and Linguistic Approaches**. London: Duckworth, pp. 191-234.

Katz, Jerrold J., and Jerry A. Fodor 1963. The structure of a semantic theory. *Language* 39:1, pp. 170-210. Reprinted in: Jerrold J. Katz and Jerry A. Fodor (eds.), **The Structure of Language: Readings in the Philosophy of Language**. Englewood Cliffs: Prentice-Hall, 1964, pp. 479-518.

Katz, Jerrold J., and Paul M. Postal 1964. **An Integrated Theory of Linguistic Descriptions**. Cambridge, MA: M.I.T. Press.

Kay, Paul 1971. Taxonomy and semantic contrast. *Language* 47: 4, pp. 866-887.

Kay, Paul 1979. The role of cognitive schemata in word meaning: Hedges revisited. Berkeley, CA: Department of Linguistics, University of California.

Kay, P. 1992. At Least. In: Lehrer and Kittay, pp.309-332.

Keenan, Edward L., and L. M. Faltz 1985. **Boolean Semantics for Natural Language**. Dordrecht-Boston: D. Reidel.

Kempson, Ruth M. 1977. **Semantic Theory**. Cambridge: Cambridge University Press.

Klavans, Judith, Bran Boguraev, Lori Levin, and James Pustejovsky (eds.) 1995. **Symposium: Representation and Acquisition of Lexical Knowledge: Polysemy, Ambiguity, and Generativity. Working Notes**. AAAI Spring Symposium Series. Stanford, CA: Stanford University.

Knight, Kevin 1993. Building a large ontology for machine translation. In: **Proceedings of the ARPA Human Language Technology Workshop**. Princeton, N.J.

Kroeber, A. 1952. Classificatory systems of relationship. In: A. Kroeber, **The Nature of Culture**. Chicago: University of Chicago Press.

Lahav, Ran 1989. Against compositionality: The case of adjectives. *Philosophical Studies* 57, pp. 261-279.

Lakoff, George 1968. Instrumental adverbs and the concept of deep structure. *Foundations of Language* 4, pp. 4-29.

Lakoff, George 1970. Global rules. *Language* 46: 4, pp. 627-639.

Lakoff, George 1971a. On generative semantics. In: Steinberg and Jakobovits, pp. 232-296.

Lakoff, George 1971b. Presupposition and relative well-formedness. In: Steinberg and Jakobovits, pp. 329-340.

Lakoff, George 1987. **Women, Fire, and Dangerous Things: What Categories Reveal About the Mind**. Chicago-London: UNiversity of Chicago Press.

Langacker, Ronald W. 1987. **Foundations of Cognitive Grammar**, Vol. 1. Stanford: Stanford University Press.

Lascarides, Alex 1995. The pragmatics of word meaning. In: Klavans *et al*., pp. 75-80.

Lenat, Douglas B., R. V. Guha, K. Pittman, D., Pratt, and M. Shepherd 1990. Cyc: Toward programs with common sense. Communications of ACM 33, no. 8 (August 1990).

Lehrer, Adrienne, and Eva Feder Kittay (eds.) 1992. **Frames, Fields, and Contrasts: New Essays in Semantic and Lexical Organization**. Hillsdale, N.J.: Erlbaum.

Levin, Beth 1991. Building a lexicon: The contribution of linguistics. *International Journal of Lexicography* 4:3, pp. 205-226.

Levin, Beth 1993. **Towards a Lexical Organization of English Verbs**. Chicago: University of Chicago Press.

Levin, Beth, and Steven Pinker (eds.) 1991. Special Issue: **Lexical and Conceptual Semantics**. *Cognition* 41

Levin, Lori and Sergei Nirenburg 1994. Construction-based MT lexicons. In: Antonio Zampolli, Nicoletta Calzolari, and Martha Palmer (eds.), **Current Issues in Computational Linguistics: In Honour of Don Walker**. Dordrecht-Piza: Kluwer-Giardini, pp. 321-338.

Lewis, David 1972. General semantics. In: Donald Davidson and Gilbert Harman (eds.), **Semantics of Natural Language**. Dordrecht: D. Reidel, pp. 169-218. Reprinted in: Barbara H. Partee (ed.), **Montague Grammar**. New York: Academic Press, 1976, pp. 1-50.

Lonsdale, Deryle, Teruko Mitamura, and Eric Nyberg 1994/1995. Acquisition of large lexicons for practical knowledge-based MT. In: Dorr and Klavans, pp. 251-283.

Lounsbury, F. 1956. A semantic analysis of the Pawnee kinship usage. *Language* 32: 1.

Lyons, John 1977. **Semantics**. Cambridge: Cambridge University Press.

Lyons, John 1995. **Linguistic Semantics: An Introduction**. Cambridge: Cambridge University Press.

Mahesh, Kavi 1996. Ontology Development for Machine Translation: Ideology and Methodology. Memoranda in Computer and Cognitive Science MCCS-96-292. Las Cruces, N.M.: New Mexico State University

Mahesh, Kavi, and Sergei Nirenburg 1995. A situated ontology for practical NLP. A paper presented at the IJCAI ‘95 Workshop on Basic Ontological Issues in Knowledge Sharing. Montreal, August 19-21.

Mahesh, Kavi, Sergei Nirenburg, Jim Bowie, and David Farwell 1996. An Assessment of Cyc for Natural Language. Memoranda in Computer and Cognitive Science MCCS-96-302. Las Cruces, N.M.: New Mexico State University.

Markowitz, Judith, Thomas Ahlswede, and Martha Evans 1986. Semantically significant patterns in dictionary definitions. In: **Proceedings of ACL ‘86**. New York, pp. 112-119.

Marx, Wolfgang 1977. Die Kontextabhängigkeit der Assoziativen Bedeutung. *Zeitschrift für experimentelle und angewandte Psychologie* 24, pp. 455-462.

Marx, Wolfgang 1983. The meaning-confining function of the adjective. In: Gert Rickheit and Michael Bock (eds.), **Psycholinguistic studies in language processing**. Berlin: Walter de Gruyter, pp. 70-81.

Masterman, M. 1959. What is a thesaurus? In: M. Masterman (ed.), **Essays on and in Machine Translation by the Cambridge Language Research Unit**. Cambridge: Cambridge University Press.

McCawley, James D. 1968. The role of semantics in a grammar. In: Emmon Bach and Robert T. Harms (eds.), **Universals in Linguistic Theory**. New York: Holt, Rinehart, and Winston, pp. 124-169.

McCawley, James D. 1971. Interpretive semantics meets Frankenstein. Foundations of Language 7, pp. 285-296.

McCawley, James D. 1986. What linguists might contribute to dictionary making if they could get their act together. In: Peter C. Bjarkman and Victor Raskin (eds.), **The Real-World Linguist: Linguistic Applications in the 1980s**. Norwood, N.J.: Ablex, pp. 1-18.

Mel'čuk, Igor' A. 1974. **Opyt teorii lingvisticheskikh modeley "Smysl <--> Tekst"** /An Essay on a Theory of Linguistic Models "Sense <--> Text"/. Moscow: Nauka.

Mel'čuk, Igor' A. 197 9. **Studies in Dependency Syntax**. Ann Arbor, MI: Karoma.

Meyer, Ingrid, Boyan Onyshkevych, and Lynn Carlson 1990. Lexicographic principles and design for knowledge-based machine translation. Technical Report CMU-CMT-90-118, Center for Machine Translation, Carnegie Melon University, Pittsburgh, PA.

Miller, George A. 1977. Practical and lexical knowledge. In: Philip N. Johnson-Laird and P. C. Wason (eds.), **Thinking: Readings in Cognitive Science**. Cambridge: Cambridge University Press.

Miller, George A., and Christiane Fellbaum 1991. Semantic networks of English. *Cognition* 41, pp. 197-229.

Miller, George A., Christiane Fellbaum, J. Kegl, and K. Miller 1988. Wordnet: An electronic lexical reference system based on theories of lexical memory. *Révue Québecoise de Linguistique* 17, pp. 181-211.

Miller, George A., and Philip N. Johnson-Laird 1976. **Language and Perception**. Cambridge: Cambridge University Press.

Mindess, Harvey, Carolyn Miller, Joy Turek, Amanda Bender, and Suzanne Corbin 1985. **The Antioch Sense of Humor Test: Making Sense of Humor**. New York: Avon Books.

Montague, Richard 1974. **Formal Philosophy. Selected Papers of Richard Montague**. Ed. by Richmond Thomason. New Haven, CT: Yale University Press.

Nirenburg, Sergei 1986. Linguistics and artificial intelligence. In: Peter C. Bjarkman and Victor Raskin (eds.), The **Real-World Linguist: Linguistic Applications in the 1980s**. Norwood, N.J.: Ablex, pp. 116-144.

Nirenburg, Sergei 1996. Supply-side and demand-side lexical semantics. Introduction to the Workshop on Breadth and Depth of Semantic Lexicons at the 34th Annual Meeting of the Association for Computational Linguistics, University of California, Santa Cruz, CA.

Nirenburg, Sergei, Jaime Carbonell, Masaru Tomita, and Kenneth Goodman 1992. **Machine Translation: A Knowledge-Based Approach**. San Mateo, CA: Morgan Kaufmann.

Nirenburg, Sergei, and Christine Defrise 1991. Practical computational linguistics. In: Rod Johnson and Michael Rosner (eds.), **Computational Linguistics and Formal Semantics**. Cambridge: Cambridge University Press.

Nirenburg, Sergei, and Kenneth Goodman 1990. Treatment of meaning in MT systems. In: **Proceedings of the Third International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Language**. Austin, TX: Linguistic Research Center, University of Texas.

Nirenburg, Sergei, and Lori Levin 1992. Syntax-driven and ontology-driven lexical semantics. In: Pustejovsky and Bergler, pp. 5-20.

Nirenburg, Sergei, and Victor Raskin 1986. A metric for computational analysis of meaning: Toward an applied theory of linguistic semantics. **Proceedings of COLING '86**. Bonn, F.R.G.: University of Bonn, pp. 338-340. See the full version in: Nirenburg *et al*. 1985.

Nirenburg, Sergei, and Victor Raskin 1987a. The analysis lexicon and the lexicon management system. *Computers and Translation* 2, pp. 177-88.

Nirenburg, Sergei, and Victor Raskin 1987b. The subworld concept lexicon and the lexicon management system. *Computational Linguistics* 13:3-4, pp. 276-289.

Nirenburg, Sergei, Victor Raskin, and Rita McCardell 1989. Ontology-based lexical acquisition. In: Uri Zernik Ed.), **Proceedings of the First International Lexical Acquisition Workshop**, IJCAI '89, Detroit, Paper 8.

Nirenburg, Sergei, Victor Raskin, and Boyan Onyshkevych 1995. *Apologiae ontologiae*. Memoranda in Computer and Cognitive Science MCCS-95-281. New Mexico State University: Computing Research Laboratory. Reprinted in: Klavans *et al*. 1995, pp. 95-107. Reprinted in a shortened version in: **TMI 95: Proceedings of the Sixth International Conference on Theoretical and Methodological Issues in Machine Translation**. Centre for Computational Linguistics, Catholic Universities Leuven Belgium, 1995, pp. 106-114.

Nirenburg, Sergei, Victor Raskin, and Allen B. Tucker (eds.) 1985. **The TRANSLATOR Project**. Hamilton, N.Y.: Colgate University.

Nirenburg, Sergei, Victor Raskin, and Allen B. Tucker 1987. The structure of interlingua in TRANSLATOR. In: Sergei Nirenburg (ed.), **Machine Translation: Theoretical and Methodological Issues**. Cambridge: Cambridge University Press, pp. 90-113.

Nunberg, Geoffrey, and Annie Zaenen 1992. Systematic polysemy in lexicology and lexicography. In: **Proceedings of the EURALEX**. Tampere, Finland.

Onyshkevych, Boyan, and Sergei Nirenburg 1992. Lexicon, ontology, and text meaning. In: Pustejovsky and Bergler, pp. 289-303.

Onyshkevych, Boyan, and Sergei Nirenburg 1994. The lexicon in the scheme of KBMT things. Memoranda in Computer and Cognitive Science MCCS-94-277. Las Cruces, N.M.: New Mexico State University. Reprinted as: A lexicon

for knowledge-based machine translation, in: Dorr and Klavans 1995, pp. 5-57.

Osthof, H., and K. Brugman 1878. **Morphologische Untersuchungen**, Vol. I. Leipzig.

Ostler, Nicholas, and B. T. S. Atkins 1992. Predictable meaning shift: Some linguistic properties of lexical implication rules. In: Pustejovsky and Bergler, pp. 87-100.

Pangloss Mark III, The 1994. The Pangloss Mark III Machine Translation System. A joint Technical Report by NMCU CRL, USC ISI, and CMU CMT. Center for Machine Translation, Carnegie Mellon University, Pittsburgh.

Parsons, Terence 1972. Some problems concerning the logic of grammatical modifiers. *Synthèse* 21:3-4, pp. 320-334. Reprinted in: Donald Davidson and Gilbert Harman (eds.), **Semantics of Natural Language**. Dordrecht: Reidel, pp. 127-141.

Parsons, Terence 1980. Modifiers and quantifiers in natural language. *Canadian Journal of Philosophy* 6, Supplement, pp. 29-60.

Parsons, Terence 1985. Underlying events in the logical analysis of English. In: Ernest LePore and Brian P. McLaughlin (eds.), **Actions and Events: Perspectives on the Philosophy of Donald Davidson**, Oxford: Blackwell, pp. 235-267.

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

Partee, Barbara, Alice ter Meulen, and Robert E. Wall 1990. **Mathematical Methods in Linguistics**. Dordrecht-Boston-London: Kluwer.

Paul, H. 1886. **Prinzipien der Sprachgeschichte**, 2nd ed. Halle.

Pinker, Steven 1989. **Learnability and Cognition: The Acquisition of Argument Structure**. Cambridge, MA: M.I.T. Press.

Pustejovsky, James 1991. The generative lexicon. *Computational Linguistics* 17:4, pp. 409-441.

Pustejovsky, James 1995. **The Generative Lexicon**. Cambridge, MA: MIT Press.

Pustejovsky, James (ed.) 1993. **Semantics and the Lexicon**. Dordrecht-Boston-London: Kluwer.

Pustejovsky, James, and P. Anick 1988. On the semantic interpretation of nominals. In: **Proceedings of COLING '88**. Budapest, Hungary.

Pustejovsky, James, and Sabine Bergler (eds.) 1992. **Lexical Semantics and Knowledge Representation**. Berlin: Springer-Verlag.

Pustejovsky, James, Sabine Bergler, and P. Anick 1993. Lexical semantic techniques for corpus analysis. *Computational Linguistics* 19: 2, pp. 331-358.

Pustejovsky, James, and Branimir Boguraev 1993. Lexical knowledge representation and natural language processing. *Artificial Intelligence* 63, pp. 193-223.

Raskin, Victor 1971. **K teorii yazukovykh podsistem** /Towards a Theory of Linguistic Subsystems/. Moscow: Moscow State University Press.

Raskin, Victor 1977. Literal meaning and speech acts. Theoretical Linguistics 4: 3, pp. 209-225.

Raskin, Victor 1985a. Linguistic and encyclopedic knowledge in text processing. In: Mario Alinei (ed.), Round Table Discussion on Text/Discourse I, *Quaderni di Semantica* 6:1, pp. 92-102.

Raskin, Victor 1985b, Once again on linguistic and encyclopedic knowledge. In: Mario Alinei (ed.), Round Table Discussion on Text/Discourse II, *Quaderni di Semantica* 6:2, pp. 377-83.

Raskin, Victor 1985c. **Semantic Mechanisms of Humor**. Dordrecht-Boston-Lancaster: D. Reidel.

Raskin, Victor 1986. Script-based semantic theory. In: Donald G. Ellis and William A. Donohue (eds.), **Contemporary Issues in Language and Discourse Processes**. Hillsdale, N.J.: Erlbaum, pp. 23-61.

Raskin, Victor 1987a. Linguistics and natural language processing. In: Sergei Nirenburg (ed.), **Machine Translation: Theoretical and Methodological Issues**. Cambridge: Cambridge University Press, pp. 42-58.

Raskin, V. 1987b. What is there in linguistic semantics for natural language processing? In: Sergei Nirenburg (ed.), **Proceedings of Natural Language Planning Workshop**. RADC: Blue Mountain Lake, N.Y., pp. 78-96.

Raskin, Victor 1990. Ontology, sublanguage, and semantic networks in natural language processing. In: Martin C. Golumbic (ed.), **Advances in Artificial Intelligence: Natural Language and Knowledge-Based Systems**. New York-Berlin: Springer-Verlag, pp. 114-128.

Raskin, Victor 1992. Using the powers of language: Non-casual language in advertising, politics, relationships, humor, and lying. In E. L. Pedersen (ed.), **Proceedings of the 1992 Annual Meeting of the Deseret Language and Linguistic Society**. Provo, UT: Brigham Young University, pp. 17-30

Raskin, Victor 1994. Frawley: Linguistic Semantics. A review article. *Language* 70: 3, pp. 552-556.

Raskin, Victor, Donalee H. Attardo, and Salvatore Attardo 1994a. The SMEARR semantic database: An intelligent and versatile resource for the humanities. In: Don Ross and Dan Brink (eds.), **Research in Humanities Computing 3**. Series ed. by Susan Hockey and Nancy Ide. Oxford: Clarendon Press, pp. 109-124.

Raskin, Victor, Salvatore Attardo, and Donalee H. Attardo 1994b. Augmenting formal semantic representation for NLP: The story of SMEARR. *Machine Translation* 9:2, pp. 81-98.

Raskin, Victor, and Sergei Nirenburg 1995. Lexical semantics of adjectives: A microtheory of adjectival meaning. Memoranda in Computer and Cognitive Science MCCS-95-288. Las Cruces, N.M.: New Mexico State University.

Raskin, Victor, and Sergei Nirenburg 1996a. Adjectival modification in text meaning representation. **Proceedings of COLING '96**. Copenhagen, pp. 842-847.

Raskin, Victor, and Sergei Nirenburg 1996b. Lexical rules for deverbal adjectives. In: Viegas et al. 1996b, pp. 89-104.

Russell Bertrand 1905. On denoting. *Mind* 14, pp. 479-493.

Saint-Dizier, Patrick 1995. Generativity, type coercion and verb semantic classes. In: Klavans *et al*., pp. 152-157.

Saint-Dizier, Patrick, and Evelyne Viegas 1995. **Computational Lexical Semantics**. Cambridge: Cambridge University Press.

Sanfilippo, Antonio 1995. Lexical polymorphism and word disambiguation. In: Klavans *et al*., pp. 158-162.

Sanfilippo, A., and V. Poznanski 1992. The acquisition of lexical knowledge from combined machine-readable sources. In: **Proceedings of the 3rd Conference on Applied Natural Language Processing--ANLP '92**. Trento, Italy, pp. 80-88.

Sanfilippo, A., E. J. Briscoe, A. Copestake, M. A. Marti, and A. Alonge 1992. Translation equivalence and lexicalization in the ACQUILEX LKB. In: **Proceedings of the 4th International Conference on Theoretical and Methodological Issues in Machine Translation--TMI '92**. Montreal, Canada.

Schank, Roger C. (ed.) 1975. **Conceptual Information Processing**. Amsterdam: North-Holland.

Schank, Roger C., and Kenneth M. Colby (eds.) 1973. **Computer Models of Thought and Language**. San Francisco: Freeman.

Schank, Roger C., and Christopher K. Riesbeck (eds.) 1981. **Inside Computer Understanding: Five Programs plus Miniatures**. Hillsdale, N.J.: Erlbaum.

Shaykevich, Anatoliy Ya. 1963. Distributsiya slov v tekste i vydeleniye semanticheskikh poley v yazyke /Word distribution in text and distinguishing semantic fields in language. *Inostrannye yazyki v shkole* 2.

Sheremetyeva, Svetlana Olegovna 1997. Metodologiya minimizatsii usiliy v inzhenernoy lingvistike. /A Methodology of Minimizing Efforts in Linguistic Engineering/. An unpublished Doctor of Science (post-Ph.D., Dr. Hab.) dissertation. Department of Linguistics, St. Peterbourgh State University, St. Peterbourgh, Russia (forthcoming).

Slator, Brian M., and Yorick A. Wilks 1987. Towards semantic structures from dictionary entries. In: **Proceedings of the Second Annual Rocky Mountain Conference on Artificial Intelligence**. Boulder, CO, pp. 85-96.

Slocum, Jonathan, and Martha G. Morgan 1986. The role of dictionaries and machine-readable lexicons in translation. Unpublished Ms.

Stalnaker, Robert, and Richard Thomason 1973. A semantic theory of adverbs. *Linguistic Inquiry* 4, pp. 195-220.

Steinberg, Danny D., and Leon A. Jakobovits (eds.) 1971. **Semantics: An Interdisciplinary Reader in Philosophy, Linguistics and Psychology**. Cambridge: Cambridge University Press.

Szalay, L. B., and J. Deese 1978. **Subjective Meaning and Culture: An Assessment Through Word Association**. Hillsdale, N.J.: Erlbaum.

Talmy, Leonard 1983. How language structures space. In: Herbert Pick and Linda Acredolo (eds.), **Spatial Orientation: Theory, Research, and Application**. New York: Plenum Press.

Trier, Jost 1931. **Der Deutsche Wortschatz im Sinnbezirk des Verstandes**. Heidelberg.

Vendler, Zeno 1963. The grammar of goodness. *The Philosophical Review* 72:4, pp. 446-465.

Viegas, Evelyne, and Sergei Nirenburg 1996. The Ecology of Lexical Acquisition: Computational Lexicon Making Process. In: **Proceedings of Euralex '96**, Göteborg University, Sweden.

Viegas, Evelyne, Sergei Nirenburg, Boyan Onyshkevych, Nicholas Ostler, Victor Raskin, and Antonio Sanfilippo (eds.) 1996a. **Breadth and Depth of Semantic Lexicons. Proceedings of a Workshop Sponsored by the Special Interest Group on the Lexicon of the Association for Computational Linguistics**. Supplement to **34th Annual Meeting of the Association for Computational Linguistics. Proceedings of the Conference**. Santa Cruz, CA: University of California.

Viegas, Evelyne, Boyan Onyshkevych, Victor Raskin, and Sergei Nirenburg 1996b. From *submit* to *submitted* via *submission*: On lexical rules in large-scale lexicon acquisition. In: **34th Annual Meeting of the Association for Computational Linguistics. Proceedings of the Conference**. Santa Cruz, CA: University of California, pp. 32-39.

Viegas, Evelyne, and Victor Raskin 1996. Lexical Acquisition: Guidelines and Methodology. Memoranda in Computer and Cognitive Science MCCS-96-2xx. Las Cruces, N.M.: New Mexico State University (in preparation).

Vinogradova, Olga S., and Alexandre R. Luria 1961. Ob"ektivnoe issledovanie smyslovykh otnosheniy /An objective investigation of sense relations. In: **Tezisy dokladov mezhvuzovskoy konferentsii po primeneniyu strukturnykh i statisticheskikh metodov issledovaniya slovarnogo sostava yazyka**. Moscow.

Voss, Clare, and Bonnie J. Dorr 1995. Toward a lexicalized grammar for interlinguas. In: Dorr and Klavans, pp. 143-184.

Walker, Donald E. 1984. Panel session: Machine-readable dictionaries. In: **Proceedings of COLING '84**. Stanford, CA: Stanford University, p. 457.

Walker, Donald E. 1987. Knowledge resource tools for accessing large text files. In: Sergei Nirenburg (ed.), **Machine**

Translation: Theoretical and Methodological Issues. Cambridge: Cambridge University Press, pp. 247-261.

Walker, Donald E. and Robert A. Amsler 1986. The use of machine-readable dictionaries in sublanguage analysis. In: Ralph Grishman and Richard Kittredge (eds.), **Analyzing Language in Restricted Domains**. Hillsdale, N.J.: Erlbaum, pp. 69-83.

Weinreich, Uriel 1962. Lexicographic definition in descriptive semantics. In: Fred W. Householder and Sol Saporta, (eds.), **Problems in Lexicography**. Bloomington, IN: Indiana University, pp. 25-43. Reprinted in: Weinreich 1980, pp. 295-314.

Weinreich, Uriel 1966. Explorations in semantic theory. In: T. A. Sebeok (ed.), **Current Trends in Linguistics**, Vol. III. The Hague: Mouton, pp. 395-477.

Weinreich, Uriel 1968. Lexicography. In: Thomas E. Sebeok (ed.), **Current Trends in Linguistics**, Vol. I. The Hague: Mouton, pp. 60-93. Reprinted in: Weinreich 1980, pp. 315-357.

Weinreich, Uriel 1980. **On Semantics**. Ed. by William Labov and Beatrice S. Weinreich. Philadelphia: University of Pennsylvania Press.

Weisgerber, Leo 1951. **Das Gesetz der Sprache**. Heidelberg.

Wierzbicka, Anna 1985. **Lexicography and conceptual analysis**. Ann Arbor, MI: Karoma.

Wilensky, Robert 1986. Some problems and proposals for knowledge representation. Cognitive Science Report 40, University of California, Berkeley, CA.

Wilensky, Robert 1991. Extending the lexicon by exploiting subregularities. Report UCB/CSD 91/616, Computer Science Department, University of California, Berkeley, CA.

Wilks, Yorick 1975a. An intelligent analyzer and understander of English. *CACM* 18, pp. 264-274.

Wilks, Yorick 1975b. Preference Semantics. In: Edward L. Keenan (ed.), **The Formal Semantics of Natural Language**. Cambridge: Cambridge University Press.

Wilks, Y. 1994. Stone Soup and the French Room. In: A. Zampolli, N. Calzolari and M. Palmer (eds.), **Current Issues in Computational Linguistics: In Honor of Don Walker**. Pisa, Italy: Giardini and Dordrecht, The Netherlands: Kluwer, pp.585-595.

Wilks, Yorick 1996. Homography and part-of-speech tagging. Presentation at the MikroKosmos Workshop. Las Cruces, N.M.: Computing Research Laboratory, August 21.

Wilks, Yorick, and Dan Fass 1992. Preference semantics. In: Stuart C. Shapiro (ed.), **Encyclopedia of Artificial Intelligence**, 2nd ed. New York: Wiley, pp. 1182-1194.

Wilks, Yorick A., Dan Fass, Cheng-Ming Guo, James E. M. McDonald, Tony Plate and Brian M. Slator 1987. A tractable machine dictionary as a resource for computational semantics. In: Sergei Nirenburg (ed.), **Proceedings of Natural Language Planning Workshop**. RADC: Blue Mountain Lake, N.Y., pp.95-123.

Wilks, Yorick A., Brian M. Slator, and Louise M. Guthrie 1996. **Electric Words: Dictionaries, Computers, and Meanings**. Cambridge, MA: M.I.T. Press.

Winograd, Terry 1983. **Language as a Cognitive Process. Volume I: Syntax**. Reading, MA: Addison-Wesley.

Wittgenstein, Ludwig 1921. **Tractatus Logico-Philosophicus**. English translation: London: Routledge and Kegan Paul, 1922.

Wu, Dekai, and Xuanyin Xia 1994/1995. Large-scale automatic extraction of an English-Chinese translation lexicon. In: Dorr and Klavans, pp. 285-313.

Zernik, Uri (ed.) 1991. **Lexical Acquisition: Exploiting On-line Resources to Build a Lexicon**. Hillsdale, N.J.: Erlbaum.

Ziff, Paul 1960. **Semantic Analysis**. Ithaca, N.Y.: Cornell University Press.

Zholkovsky, Alexandre, Nina N. Leont'eva, and Yuriy S. Martem'yanov 1961. O printsipial'nom ispol'zovanii smysla pri mashinnom perevode /On a principled use of sense in machine translation/. In: **Mashinnyy perevod** 2. Moscow: USSR Academy of Sciences.

Zvegintzev, Vladimir Andreevich 1964. Mladogrammaticheskoe napravlenie /The Young Grammarian School/. In: V. A. Zvegintzev (ed.), **Istoriya yazykoznaniya XIX-XX vekov v ocherkakh i izvlecheniyakh**. Moscow: Prosveshchenie, pp. 184-186.

Zvegintzev, Vladimir Andreevich 1968. **Teoreticheskaya i prikladnaya lingvistika** /Theoretical and Computational Linguistics/. Moscow: Prosveshchenie.