

CHAPTER 11

The Problem of Meaning in Generative Grammar

MASSIMO PIATTELLI-PALMARINI and CARLO CECCHETTO

*Department of Cognitive Science,
San Raffaele University
Milan, Italy*

INTRODUCTION

We have been asked by the editors of this volume to write an overview of the problem of meaning in the generative tradition, that is, the one that goes back to the pioneering work of Noam Chomsky in the mid-Fifties and for which Chomsky is still the most important exponent. We have found this task entirely congenial, since we believe that the Chomskyan approach to linguistics opened the way to a highly successful research program, one that shed light on many important aspects of the language faculty, notably including our "knowledge of meaning" (Larson & Segal, 1996). However, the present review has simultaneously a broad and a more limited objective. On the one hand, we will try to present the reader with a broad geography of the issue, *as long as it can be reconstructed adopting a specific point of view*, that of the generative tradition. On the other hand, we do not intend to address ourselves uniquely, or even primarily, to generative scholars. We do believe, in fact, that adopting this particular approach to reconstruct the discussion on meaning in linguistics can offer interesting general insights. It is a fact that generative grammar, in the last 40

years of the history of linguistics, has been in competition with several other important approaches. From the analysis of this competition (and many ensuing interactions) we can learn more about these other approaches as well, notably about their weak and their strong points.

LEXICAL SEMANTICS AND COMPOSITIONAL SEMANTICS

A traditional claim about the generative tradition, and about Chomsky's work in particular, is that it focuses primarily on syntax and phonology, whereas semantics, another traditional sub-component of linguistics, is given a secondary, if not negligible, role.

Even if this claim contains a part of truth, it stands in need of various qualifications and a major restriction. Let us begin from the latter.

There are two distinct, though strictly related, aspects that a theory of meaning has to deal with: One is lexical semantics, the other the rules of composition of simple meanings into complex (sentential) ones. The central question of the former is: what is the meaning of a *single* word and how can the theory represent it? A typical controversy that arises at this level is the following: Is a word associated with a specific lexical entry in the mental lexicon, one that contains a specification of its semantic features, or is the meaning of the word to be captured by a list of meaning postulates (that is, restrictions on the domain of interpretation)?

The central question at the second (composite) level of semantic analysis is: How does the meaning of a single word contribute to the meaning of the sentence in which it occurs? A typical controversy that arises at this level concerns the primacy of syntax versus the primacy of semantics. Do the syntactic rules (those governing the composition of words into well-formed sequences) simply mirror the composition rules that operate on meanings, or is it the other way around? Or are syntax and semantics to a significant extent autonomous with respect to each other?

It may seem obvious that lexical semantics and compositional semantics must be strictly related, but this is not an innocuous claim.¹ There are in fact research traditions that consider only compositional semantics worth pursuing (arguably, the Montagovian tradition²) and other that consider only lexical semantics worth pursuing (arguably, Chomsky at some periods of his theorizing³). For the sake of clarity, we will consider the relation between Chomskyan linguistics and lexical semantics first, and then deal with compositional semantics.

LEXICAL SEMANTICS

Chomsky has always manifested a clear interest in lexical semantics and an optimistic attitude towards the success of such a theory. In this section we will first focus on some of Chomsky's general considerations on the topic and then offer an illustration of the role that a certain portion of lexical semantics plays in the generative theory of syntax.

General Considerations

Chomsky observes that the so-called argument of "the poverty of the stimulus" (in essence, given the richness, the complexity and the intricacy of linguistic competence, it is next to impossible that such competence results from the child's haphazard exposure to the typical occasional fragments of the local language), quite central to a re-conceptualization of the acquisition of syntax by the child, also applies to the acquisition of the lexicon: Children acquire lexical items at an extraordinary rate, even if, apart from a small group of technical terms, explicit definitions are rare, or even impossible. Moreover, as is the case with syntactic rules and principles, this acquisition takes place on the basis of occasional, unsystematic, and casual fragments of conversation. Not even the best dictionaries are able to offer noncircular definitions for most of the ordinary words we use. Nonetheless, children acquire, at an average, more than a dozen new words per day at the peak period of language growth (rightly called "lexical

explosion") (Miller, 1986; Miller & Gildea, 1987). The practically inevitable conclusion that is to be drawn from these data, analogous to the one drawn in the domain of syntax, is that the acquisition of the lexicon reflects, to a very substantial extent, internal causation rather than a simple absorption of the external inputs. The child possesses, owing to his/her internal biological constitution, a complex set of dispositions that are crucial to complete this task successfully, in a relatively short time, without the aid of definitions, and without the aid of elaborate explicit meaning-specifications by adults. Well-designed specific experiments, carried out in several languages, have consistently shown that powerful internal constraints drastically restrict the candidate meanings for each new word that the child encounters for the first time.⁴ So, the different lexical items do not simply form a list, but rather are integrated into a highly structured system. The nodes in the grid that gives structure to this system - for example notions like goal, source, agent, patient, cause, intention, event, +/- animate, +/- stative, +/- artefact and so on - must be already present in the mind of the child to be actively used in the acquisition process. Acquiring a new lexical item, as a consequence, is something like filling in an open slot in a pre-structured space of possible meanings or, more extremely, as proposed by Jerry Fodor, lexical acquisition is nothing else than acquiring labels for concept we already (innately) possess (Fodor's arguments for the strong innatism of "mentalese" are to be found in Fodor, 1975; and Fodor, 1980. Chomsky's thoughts on this topic can be found, among other places, in Chomsky, 1991, pp. 28-29).

Lexical semantics, therefore, plays an important role in the generative framework. Many researchers belonging to this tradition have actively tried to elaborate a theory of the lexicon. An example of such a theory is the theory of semantic markers (also called decompositionalist theory). In the generativist tradition (then called transformational grammar - a notion about which we shall have more to say in a moment) the central idea, originally presented by Katz and Fodor (1963), was that the meaning of a word consists of a structured set of

semantic markers into which it can be decomposed. These markers being a limited, innate and universal set of features, every lexical meaning was conceived as exactly decomposable into a specific choice for the (binary) values that each of these basic components can assume. Let us give an example. A simple lexical item such as *man* is literally identified with the amalgamation of three (innate) features: +HUMAN, +ADULT and +MALE. The word *girl*, by contrast, would be the amalgamation of +HUMAN, -ADULT and -MALE, and so on in a straightforward way. According to the theory, a specific distribution of binary values assigned to a proper subset of such lexical semantic "markers" corresponded to (and was in an important sense identical with) a certain lexical entry. Quite a number of semantic inferences and presuppositions were thus explained, in ingenious ways, lending some initial success to this theory (soon to be abandoned by its very proponents). Developments of this idea were also known under the name of "generative semantics", enjoying for a while a certain popularity in the profession (for a historical presentation see Newmeyer, 1986).

We will not go beyond this rather sketchy summary of the theory of semantic markers, because it soon encountered major problems, some of them unsolvable within the theory itself. It embodied, however, an important insight, that remains valid until this day: The idea that some sort of decomposition is involved in determining the meaning of a word. This is now considered common wisdom, in spite of the fact that hardly anyone today would subscribe to the strong version of the theory, according to which the semantic markers can *exhaust* the content of the lexical item. The main problems with this view can be succinctly described. For example a psycholinguistically oriented measurement of reaction times in comprehension (a technique that was not to be pursued much farther in the generative tradition proper) failed to confirm, *prima facie* at least, that comprehension stems from actual decomposition (see for example Fodor, Fodor & Garrett, 1975; and Johnson-Laird, 1983, especially chapter 10). Roughly, an illustration of the idea underlying these kinds of experiments follows: By

hypothesis, the meaning of the word *bachelor* is a set of semantic markers that contains, among other things, the negative feature -MARRIED.

If real-time comprehension requires decomposition, we would expect a subject asked to evaluate a sentence in which the word *bachelor* appears to show a longer processing delay (or at least not shorter) than he or she would for the same sentence in which *bachelor* has been replaced by the sequence *not married*. In fact, at bottom, genuine decomposition should entail reducing *bachelor* to *not married* (and then, eventually, to other more basic markers). What is actually observed is that people in relevant contexts are reliably faster in processing sentences where *bachelor* occurs, while the correspondent sentences where *not married* occurs require a longer time. This suggests that in order to understand *bachelor*, one is not obliged to reduce this word to the feature specification -MARRIED. In the absence of a full theory of how comprehension works, and of how elements pertaining to semantic competence map onto mechanisms of semantic performance (reaction times measure the latter, not the former), these findings do not say *per se* that the decompositionist theory is wrong, yet, somehow, they were almost unanimously judged to shed a negative light on it.

Moreover, also from a theoretical (especially, competence-oriented) point of view, the strong version of the theory, according to which the decomposition in semantic features *exhausts* the meaning of the word, was unsatisfactory, as observed by Chomsky himself. Take the word *assassinate*: it is highly plausible that it shares some semantic feature (say, the feature CAUSE TO DIE that itself might be decomposed in more primitive components) with other words like *kill*, *murder*, *massacre*, etc. It is not equally plausible, however, that it also contains an abstract additional feature to specify that the person who dies is in some way prominent, that the crime has socio-political implications, and so on. For these and other reasons, decompositionism in the strict version is not popular today. However, it is still widely believed that some decomposition is part of the story, even if it is not *all* the story. Because of space limitations, we cannot discuss the

work in lexical semantics within the Chomskyan tradition throughout (but see Pustejovsky, 1995; Hale & Keyser, 1993; and Fodor & Lepore, 1996b). We prefer to concentrate on an example, representative of how the current version of the generative theory of syntax contains at its core the idea that lexical items entering the derivation of the sentence are associated with semantic features.⁵ We will see that Theta Theory, an important part of lexical semantics, plays a key role at the very core of the Chomskyan program: the theory of syntax.

Theta Theory

An important module, among the different ones in which the generative theory of syntax is organised, deals with thematic relations: It is the so-called Theta Theory. Like the other modules, it constitutes a well-structured delimited realm of hypotheses and predictions in the theory itself, and at the same time a specific distinct component in the mental workings that the theory attributes to each individual speaker-hearer. The contents of Theta Theory are, thus, to be understood as being part and parcel of the hearer-speaker's tacit knowledge of language.

In the background of this theory lies the traditional idea, in part borrowed from much older theories of grammar, and in part from formal logic, that lexical items come either as predicates, or as arguments, or else as adjuncts. Let us pause a moment to specify what these terms mean. Predicates are elements that typically require "saturation": One-place predicates (or properties) such as *run* or *collapse* require one "agent" (*Bill*, or *the house*, or *Mary*, etc.). Expressions like "*Bill runs*" or "*The house collapsed*" are well-formed, because the open space of the predicate is properly saturated. The same applies, *mutatis mutandis*, to two-place predicates or relations such as *feed* (*Bill feeds the dog*), to three-place-predicates such as *give* (*John gave a book to Bill*), and so on. The arguments are the elements that can saturate the open spaces of predicates and they are typically (though not exclusively) nominal expressions (like *the farmer*, *a rabbit*, *John*, etc.). Adjuncts

are elements, like adverbs, that contribute with further qualification, but are not strictly necessary to the well-formedness of the sentence, (for instance, *regularly* in "the farmer regularly feeds a rabbit"). Theta Theory focuses on the relation between arguments and predicates and identifies some fixed semantic roles that arguments play with respect to predicates. The idea is very simple. Take a three place predicate like *give*. Whatever its three arguments are, one of them will have the role of AGENT of the event of giving, the other will have the role of GOAL and the third will be the THEME. So, in a sentence like 1 below, *John* is the agent, *Bill* is the goal and *the book* is the theme:

John gave a book to Bill

(1)

Examining different kinds of predicates, in a variety of well-formed sentences, and in many languages, it is possible to identify a list of Thematic Roles that must be assigned to arguments: apart from those already cited, further potential candidates are PATIENT, EXPERIENCER, SOURCE, BENEFACTIVE and some others. It is, moreover, theoretically convenient to order the thematic roles in a kind of universal hierarchy, such that predicates (most notably complex verbal forms) are accordingly saturated or "discharged" by the arguments present in the sentence in a certain fixed sequence, from top to bottom (Grimshaw, 1990; Hale & Keyser, 1993).

We do not need to go further into the details of Theta Theory (for a more complete presentation of Theta Theory and its relations with the other modules of the theory of syntax see Haegeman, 1994). We think that even this summary presentation is enough to clarify the point we are interested in: Theta Roles are among the core syntactico-semantic features that are associated with a lexical item. The idea is that a predicate like *kill* is stored in the mental lexicon together with a so-called "theta grid": in addition to the information that it is a two-place predicate, the thematic roles of the two arguments, and their respective syntactic

valencies are also specified. One of them must be an AGENT and the other a PATIENT. Simplifying a bit, we may add that these roles, in a given language, will map precisely onto certain strategically decisive hierarchical positions in the syntactic structure of the sentence. This universally fixed co-mapping of semantic and syntactic specifications (given certain uniform differences between different languages) contributes to the determination of the meaning of *kill* in any sentence in which it appears in that language. It is interesting to add that our innate knowledge of this complex universal mapping does indeed offer an explanation for the remarkable facts about acquisition that we have alluded to earlier, notably for the child's highly efficient guesses about the meaning of novel lexical items made on the basis of the specific syntactic forms in which they appear upon first presentation.

This may already suffice to show that thematic roles are absolutely central to the theory of syntax within the generative tradition. To better see why it is so, one must also be reminded of the role that syntactic transformations (a notion we will introduce in a moment) play in this research program. Virtually every sentence is the result of a transformational operation: Its derivation is a process during which certain lexical units, inserted from the lexicon and placed in certain positions, move to other positions in the syntactic tree. These movements are traditionally called syntactic transformations, and they gave the initial name to the theory as a whole. As of the mid-Fifties, Chomsky was able to demonstrate, by means of justly celebrated arguments, that a transformational grammar is an intrinsically more efficient and psychologically more plausible formal device than any of the preceding "Phrase Structure Grammars" postulated in the structuralist tradition. A very simple example involves interrogative sentences. Take:

Who(m) are you looking for?

(2)

You are looking for him

(3)

It is now assumed by many that the interrogative phrase in sentence 2 has moved from the position that follows the proposition *for* (the position occupied by *him* in sentence 3. This explains many interesting properties of interrogative. For example, if the movement of *whom* has taken place, there is a certain level of representation (the one that precedes the movement) in which the preposition *for* governs the interrogative noun phrase assigning it the accusative case. So, the accusative case of *whom* in sentence 2 can be related to the presence of the preposition *for* (this seems correct, since the object pronoun that follows *for* in sentence 3 is also accusative, as expected). The core intuition here is that sentence 2 is derived from a structure like the one exemplified in sentence 3, and not built from scratch as a separate and independent expression. In particular, sentence 2 is the result of having "moved", the interrogative element correspondent to "him" in sentence 3 to the very beginning of the sentence, obtaining an interrogative, though some key syntactic dependencies have been left intact. In English, this involves the morphological device of prefixing a *wh*- and (optionally in current speech, mandatorily in good written style) preserving the manifest accusative suffix *-m*. In other languages, in general, other morpho-syntactic devices will be mobilized in carrying out this kind of "transformation". Further evidence that something has, indeed, moved is also offered by the impossibility, in sentence 2, of preserving any noun or pronoun in the final (or "base") position (sentences like: "Who(m) are you looking for him?" or "Who(m) are you looking for John?" are totally impermissible). This is a rather trivial example of transformation but, it should make the idea clear enough.

Let us now come to the next crucial point, one that brings together the notion of transformation and that of thematic positions. We can ask, more generally, how to justify the presence of transformations in the first place. Why do constituents move? Why is the sequence "are you looking for whom?", under normal circumstances, i.e., barring special emphasis, as an "echo-question," ungrammatical?

The answer in the generative tradition can seem very clear (but it is not, as we will see): constituents move because the positions where they are engaged in thematic relations with other constituents (that is, the place where they receive theta roles) are different from those in which they check grammatical features (like agreement, case etc.). Everyone knows that certain thematic relations typically "go with" certain grammatical features. For instance, an agent typically is a subject and takes the nominative case, a theme typically is a complement object and takes the accusative case, etcetera (in English this is particularly clear with personal pronouns *He loves her*, *she loves him* etc.).

This raw pre-theoretical notion of "going with" does not work, as one can easily understand, reflecting on passive sentences in which the subject is not an agent. However, it has been formalized and substantiated in the theory in a sophisticated form that does not suffer from the problems we have just alluded to: first, it is assumed that there is a sort of split between the thematic relations *as such*, and the grammatical features *as such*, each corresponding to a certain fixed position in the syntactic structure. Subsequently, they are recomposed, as a result of movement of the lexical item from one position to the other. So, constituents are first directly inserted from the lexicon into the thematic positions (typically, positions close enough to the constituent from which they receive a theta role: for example, a theme is inserted close enough to the verb for which it saturates an empty position). Then, later on, they move to the position where they receive case, agree with other constituents and fulfil other grammatical tasks (an interrogative phrase, for instance, must move to a sentence-initial position for reasons related to the explicit interrogative morphology it carries - the *wh*-morpheme in English: a theme in a passive sentence moves to the subject position to receive the nominative case, as is clear from languages like Latin where the case is visible).

In a certain sense, then, the answer to the general question "Why is movement needed?" rests on the theory of thematic relations. Of course, such an

answer is not really a satisfactory one, since one could also ask: why cannot constituents check grammatical features directly in the initial (or base), thematic position? In a sense, generative grammar does not really answer this deeper question. Maybe human natural languages just are like that, and it is conceivable, in the abstract, that other intelligent creatures could speak languages in which checking or case-assigning takes place directly in the base position. We can conceive of such a possibility in the abstract, but maybe for humans such languages would be un-learnable and un-usable. Or possibly there is some deeper conceptual necessity for it to be so. In the generative tradition, so far, the fact that grammatical features are checked in places different from those in which thematic roles are assigned comes close to being a primitive, something that cannot be in turn deduced from something else (but Chomsky's recent works contain some speculations on this topic (see for example, Chomsky, 1995b paragraph 4.7.1 and Moro, 1995 who explicitly addresses this problem and proposes an alternative view). However, given the Theta Theory, and the other components of the theory that deal with Case, Agreement, etc., the all-pervasive transformational property of natural languages can be reduced to this primitive. Keeping in mind the importance of transformations (virtually every sentence is the result of some transformations!), it should be clear why the role of thematic theory is absolutely crucial. So, even the theory of syntax *strictu sensu* presupposes a theory of the lexicon.

Meaning Postulates

What about other theories of lexical meaning that have been proposed? Are they in line, or at least compatible, with the general characters of generative grammar? Without any claim of completeness, we must mention at least the theory of semantic networks. Simply stated, in its most radical version, it identifies the meaning of a word with its position in a semantic network (this position, moreover, is supposed to mirror faithfully the way in which the informa-

tion conveyed by the word has been originally acquired). This theory, originally proposed in the old associationist psychological framework, has recently been revamped in the connectionist literature. It is hardly compatible with the strong innatist hypothesis that characterises the generative tradition. However, in a much attenuated form, and restricting it to localized chunks of the lexicon, most notably to technically or at any rate specialized terms, few doubt that it has some merit. Even in the generativist framework, it can be safely conceded that the meanings of *map* (restricted to land geography) as opposed to *chart* (restricted to nautical use, or astronomy) are mutually inter-defined. The same applies to a host of scientific terms that are introduced via mutual differences (such as *prophase*, *metaphase*, and *telophase*). Given the limited aim of this paper, and the formidable problems encountered by the semantic networks theory as a general theory of the lexicon (this is at least implicit in all the consideration we have made already, as well as in those we are about to make now), it will not be developed any further.⁶

Instead, some attention will be devoted to another theory: the idea that the meaning of a lexical item can be captured by a set of "meaning postulates". A meaning postulate is a device originally used by logicians to avoid some problems that arise in the framework of possible worlds semantics (cf. Carnap, 1952). However, Montague explicitly uses it in his semantic analysis of English and Fodor once considered it as the basis for a theory of lexical meaning. Again, the idea is pretty simple, even if the technical implementation can be very complex: a meaning postulate is nothing other than a restriction on the domain of interpretation. For the reader who is not at all familiar with this approach, we first present some very basic information on model theoretic semantics (for an introduction to model theoretic semantics, and indeed to the formal approach to semantics in general, see Chierchia & McConnell-Ginet, 1990; and Larson & Segal, 1995). After that, we will specify what a meaning postulate is.

A language, be it a natural language like English or an artificial one like propositional calculus, must be interpreted; that is, there must be some model that represents the world referred to by the language under question. For instance, if a certain artificial language can be used to prove theorems in arithmetic, then the domain of interpretation of that language will include the set of natural numbers. Colloquial English can be used to talk about a certain portion of reality; thus, the model of colloquial English will include those individuals and sets that are part of the portion of reality we want to describe.

The language in question will contain at least singular terms (roughly, names for individuals) and names for properties (like *red*, or *prime*). The interpretation of a singular term in the model (its denotation) will be an individual corresponding to it: the interpretation of *John* will be the individual John. What about names for properties? They will be interpreted as subsets of the domain. For example, the denotation of *prime* in a model for arithmetic will be the (infinite) set of all the prime numbers, the denotation of *red* in a model for English will be the set of all and only the things that are red, and so on. If the language contains names for binary relations (like *see* or *bigger than*), these will be interpreted in the domain as sets of ordered pairs. For example, the denotation of *bigger than* will be the set of the ordered pair $\{x, y\}$ such that x is bigger than y . Analogously, the denotation of *see* will be the set of the ordered pairs $\{x, y\}$, such that x sees y ; The generalization to names of ternary relations like *give* is straightforward. Adopting this strategy, we can easily see how the truth value of a sentence can be calculated. Imagine that someone wants to know if the sentence "John sees Mary" is true in a certain model. What he has to do is *i)* check which individuals are the denotations of *John* and *Mary*, *ii)* pick out the set of ordered pairs that is the denotation of *see*, and finally, *iii)* verify that this set contains a pair whose first member is the denotation of *John* and whose second member is the denotation of *Mary*. If such a pair exists in the model, then the sentence is true (in the model), otherwise it is false (in the model).

Bearing in mind this very sketchy characterization of such standard semantic theory, consider a relation of synonymy, for example the one that holds between *bachelor* and *unmarried*, and ask how a theory of the lexicon can represent this kind of relation. A very natural answer comes to mind: it can be said that a model for the language that contains the two words *bachelor* and *unmarried* must be such that if some individual in the model is a bachelor, then he is also unmarried. This might seem a mere trick, but it is not. If nothing is said, the words *bachelor* and *unmarried*, in (some of) the possible models of the language that contains them, can have different denotations, that is, they can be associated with different sets (after all, they are two different linguistic expressions). What we need in order to express the fact that the two expressions are synonymous is a condition like sentence 4 (that all the possible models for a language containing *bachelor* and *unmarried* must satisfy):

for any x , x is a bachelor if and only if x is unmarried (4)

So, technically, a meaning postulate is a device used to reduce the number of the admissible models for a certain language L .

Coming now to our main point here, it has been proposed that the meaning of a lexical item A is the set of meaning postulates necessary and sufficient to build a model adequate for the language L in which the expression A occurs. The idea is clear: we do not consider a model in which an arbitrary x can be a bachelor without being unmarried as adequate. Arguably, the proponent of the meaning postulates theory goes on to say, analogous constraints can be proposed for each one of the lexical items. So, lexical semantics can be reduced to a list of sets of meaning postulates.⁷

This idea, at least in its pure form, today is no longer very popular (but see Fodor & LePore, in press, for a recent revamping). There are several problems

with it, but we will mention only a couple of the most salient ones (for a criticism of the meaning postulates theory based on different considerations see Johnson-Laird, 1987; and also Kripke's, 1980; and Putnam's, 1975, notorious arguments for a direct reference theory). The reason for doing this is twofold: First a treatment based on meaning postulates is a particularly perspicuous version of what has been called "conceptual role semantics". In a nutshell, this theory identifies the meaning of an expression with the role it plays in the totality of the language system it belongs to, or, in other words, with the set of links that hold between the expression and the rest of the lexicon. This is an hypothesis that, even today, receives much attention.⁸ Many people would be ready to subscribe to some version of "conceptual role semantics", even if this idea does not always receive enough specification. Meaning postulates, then, are a nice way to implement what is called "conceptual role semantics, because they allow for a formal (that is to say, precise) implementation. An additional reason for presenting meaning postulates here is that they will lead us to a topic -the controversy on externalism- that will be treated in some detail in the second part of this paper (the one devoted to compositional semantics).

So, let us come to the salient problems. What is wrong, if anything, with meaning postulates? A major problem the theory has to face is that not all the words are like *bachelor*. As a matter of fact, it is now widely believed that the set of words that admit explicit definitions is rather restricted. Trying to give an exhaustive list of the meaning postulates for a natural-kind term like *dog* or *water* is likely to lead to frustration. Dictionaries do attempt to give definitions for these terms, but they are patently insufficient. It comes as no surprise that, in the end, *dog* is defined allusively, in terms of some typical canine property (possibly supplemented with some drawing or picture). That can be acceptable for the purpose of a dictionary, but it goes without saying, it is not a genuine definition, nor something out of which we can build a list of meaning postulates for *dog*.

The source of the problem with meaning postulates seems to be the fact

that, in order to identify something as a particular instance of a certain kind, we do not need to be able to define exhaustively the kind itself. Rather, our concept of the kind largely rests on a presentation (in one way or another) of some of its typical instances; that is, the meaning of a natural-kind word seems to be encoded as a schema or a prototype. We can identify Fido as a dog because it is close to the prototypical dog we have in mind, not because we can give the definition of doghood (maybe in terms of DNA sequences). Several experiments based on reaction times, and on a factorial decomposition of similarity traits within a kind, have shown that, for instance, a canary is more quickly identified as an instance of *bird* than a penguin, lettuce as a better instance of *vegetable* than an olive, and so on. The closer an actual exemplar is to the prototype, in a mental similarity space that is remarkably uniform across different subjects, the quicker it will be identified. Specimens that are closer to the prototype in that space are perceived and judged as "better exemplars" of the kind than specimens that lie far from the prototype.

One is entitled to ask, then, what is a prototype? One proposal is that it is a list of default values for the main features of the kind. A prototypical bird, for example, flies, so a default value will be "able to fly". Of course, there are birds that do not fly (penguins, for instance). Accordingly, they are judged "poor" exemplars of *bird*. This is why what counts are *default* values, and not rigid ones: it is part of our ability to master the meaning of the term *bird*, that we feel free to change the relevant value from the default one (flies) to a non-default one (does not fly) when presented with a penguin. It is not, thus, a binary judgement of all-or-nothing (belongs to the kind/ does not belong to the kind), but rather a graded judgement of more-or-less, ranging continuously from prototypical exemplars, to peripheral ones, to non-exemplars. (for an early discussion, see Rosch, 1978; recent papers on the topic include Fodor & Lepore, 1996a; Kamp & Partee, 1995; and Osherson & Smith, in press).

Therefore, it is now fairly widely believed that those psychological theories

based on meaning postulates are not appealing, at least if they pretend to be theories of our lexical knowledge in general. On the other hand, they can still be adequate theories for a limited part of the lexicon: the one composed by words like *bachelor* that admit explicit definitions.

Finally, let us come to the relation that holds between the Chomskyan approach to language and one which uses meaning postulates. This is the bridge towards the second part of this paper.

Note that to introduce a meaning postulate is to operate at the denotation (language-external) level. As we stressed before, you have a class of possible models for a certain language *L* and you exclude some of them from the list of those acceptable, given the vocabulary (or lexicon) of the language in question. This may seem innocuous, even unavoidable. After all, a language, it can be argued, is used to talk about an external reality and so what can be wrong with operating at the level of denotation? A model is a portion of the reality referred to by the language. Admittedly, the model can be more or less adequate, but there should be no problem in principle with using model-theoretic devices to capture lexical meanings. It is however exactly this *prima facie* innocuous claim that Chomsky elects as a major polemical objective in his latest philosophical papers. Chomsky's opposition to externalism, it seems to us, undermines, among other things, the possibility of meaning postulates theories. So if Chomsky's latest work is right, meaning postulates theories are wrong. This leads us to compositional semantics.

COMPOSITIONAL SEMANTICS

Syntax and Semantics in Montague

The best thing is to begin by defining the problem for this part of the theory of meaning. Frege was the first to formulate it (see Geach & Black, 1952):⁹ A speaker can understand a potentially infinite number of sentences, even if his/her ability to do so necessarily depends on the use of finite means (as shown by the

fact that the brain is finite). How can that happen? The answer given by Frege lies in the so-called "Principle of Compositionality": the meaning of a sentence can be built up compositionally from the meanings of its constituents in a way that depends *entirely* on the grammatical modes of combination. In addition, the grammatical modes of combination (the rules of syntax, as we might say) are explicitly recursive and it is the presence of this recursive procedure that guarantees to the speaker the possibility of understanding a potentially infinite number of sentences.¹⁰ Speaking of compositionality is not a mere metaphor. Frege and, after him, Montague and Montagovian scholars built up a system of rigid compositional rules, according to which constituents possessing the same syntactic role are treated as having the same basic semantics. The result is a semantics that is strictly isomorphic with syntax, each compositional rule on meanings corresponding to a syntactic rule, and vice versa.

We will now give an example, to give the flavour of the extent to which such an isomorphism holds in Montague. Take two sentences like the following:

John smiles (5)

Every man smiles (6)

It seems that the two noun phrases *John* and *every man* play the same syntactic role in the two sentences: they are the subject of a simple intransitive verb. This actually corresponds to a naive concept of syntax, as we will see in a moment, but, for the time being, let us go on with our example. The "Principle of Compositionality", then, forces the semanticist to assign the same semantics to the two noun phrases. This is the source of a non trivial problem: Given what we said earlier, the denotation of the verb *smile* is the set that includes the entities of the domain that smile (for simplicity, we abstract here from many important factors, like the tense specification). The denotation of *John* is John. The semantics of sentence 5, then, would be roughly the following: if the denotation of *smile*

contains the denotation of *John*, the sentence is true, otherwise it is false. Assuming the "Principle of Compositionality" and the hypothesis that *John* and *every man* play the same syntactic role, we are in trouble. In fact, we seem to be committed to assert that the denotation of *every man*, like the denotation of *John*, is an individual of the domain, but this is clearly problematic: *every man* is a quantificational expression, it does not have the function of naming a singular individual!¹¹ Montague (renewing the treatment for quantificational expressions proposed by Frege) makes a characteristic move at this point, turning, in a sense, the problem upside down: he asks what an adequate semantic account of the quantificational expressions is and, then, extends this semantic account to proper names as a special case. The desired result is obtained: an identical semantic treatment corresponds to the same syntactic role. We will now sketch Montague's solution: counterintuitive as it might seem at first blush, it is formally elegant and attains the desired goal.

The starting point is that *smile* denotes a set and that we want sentence 6 to have a truth value computable on the basis of its constituents. Montague's solution is that *every man* denotes a set of sets and that the sentence is true if, and only if, the set denoted by *smile* is a member of that set of sets. The inclusion order is now reversed: in order for the sentence to be true, the predicate must be a member of the set denoted by the subject (it is not the subject that must belong to the set denoted by the predicate). This may appear awkward, or, at best, to be the result of a merely formal trick, but it is not. There is an important intuition underlying this treatment. Remember that the denotation of a property is a set. So, a set of sets can be seen as a set of properties. If we now ask what set of properties corresponds to the noun phrase *every man*, a natural answer comes to mind: that set will be the one that includes all the properties that hold of *every man*. For example, if all the men of the domain of interpretation are tall, the denotation of *every man* will include the set of the tall men. But, if all the men but one are tall, the denotation of *every man* will not include the set of the tall

men. Similarly, if all the men of the domain of interpretation smile, the denotation of *every man* will include the set of the men who smile. This ensures that sentence 6 is true exactly in the situation in which our intuition indicates that it is true: when all the men smile.

Before moving to proper names, let us take another moment to observe that this treatment applies straightforwardly to other quantificational expressions. Take the following sentence:

No man smiles

(7)

If our semantics is to be strictly compositional, here as well the quantificational expression must denote a set of sets, and the sentence must be true just in case the set denoted by *smile* is an element of that set of sets. What is the set of sets denoted by *no man*? The answer is refreshingly easy: it is the set of the properties that hold of no man. We leave to the reader the simple task of checking that this choice gives the desired truth conditions for the sentence 7.

Finally, let us consider the case of the proper name *John* in sentence 5: once again, as required, given the strict compositionality of our approach, it has to denote a set of sets (or, putting it differently, a set of properties). Which one? The answer should, by now, sound perfectly predictable: The noun phrase *John* will denote the set of properties that hold of the individual John. An individual is then identified with all the properties it possesses.

Our example, we hope, will have shown that compositionality is not a vague metaphor in the Montagovian tradition. Rather, it is a fundamental and demanding constraint that the theory of meaning must obey.

In conclusion, we want to stress the point which is crucial for us. Compositionality, as it is conceived in the Montagovian tradition at least, imposes a very strict parallelism between syntax and semantics. This strict isomorphism leaves no room for an autonomous syntax, or *prima facie* symmetrically, for an autono-

mous semantics. But that holds only *prima facie*, because, given the fundamental semantic character of the Principle of Compositionality, it is the syntax that becomes ancillary, if not to a specific semantic theory directly, at least to the general requirements imposed by the theory of meaning.

It is against this ancillary view of syntax that Chomsky raised his strong objections.

The Classical Argument for the Autonomy of Syntax

Since Chomsky's arguments supporting the thesis of the autonomy of syntax from semantics are notorious, we will remind the reader of them only very briefly. For our present purposes, it suffices to recall one example, the most celebrated one. Let us examine the two sequences 8 and 9 below, and ask ourselves in what way their meanings differ:

Colorless green ideas sleep furiously

Furiously sleep ideas green colorless

*1-4/22
app. of 1.2/25*

(8)

(9)

Chomsky observes that, while we have no intuition about differences in meaning between sequence 8 and sequence 9, we have a very clear intuition on the grammaticality of the two sequences: (8) is a good English sentence, whereas (9) is not. Trivial as it might seem, this observation alone constitutes a serious obstacle for a theory that postulates a too strict isomorphism between syntax and semantics. That theory does not predict, nor can it accommodate, such a striking difference between grammaticality judgements and judgements on meaning.

However, this example, and the related discussions, constitute a well-known chapter of the history of recent linguistics, dispensing us from needless elaborations. Rather, we think that, today, after 40 years of fruitful empirical investigations, something else can better justify the autonomy of syntax from semantics, namely the growth of knowledge that that hypothesis has made

possible. Even more telling is the fact that the research program based on the autonomy of syntax has proved to be fruitful in the field of semantics proper as well, or at least in the territory that lies at the boundary between syntax and semantics. In line with the limited scope of this review, making, once more, no attempt at systematicity and completeness of coverage, we think that a representative example may be sufficient.

Quantificational Expressions in Government and Binding Theory

Capitalizing on what we have just seen above, the example we want to consider involves quantificational expressions. For reasons that are both historical and expository, it has the additional merit of allowing us to draw a comparison between the Montagovian program and the generative one. It is eminently appropriate, as we will see in a moment, to start with an apparently unrelated class of expressions, the interrogative ones. Consider a standard interrogative sentence in English:

Which man have you seen?

(10)

Interrogative sentences raise a host of intriguing problems, but the main aspect we want to focus on here is the following: There seems to be a link between the interpretation that the interrogative phrase (*Which man*) receives and the sentence-initial position it occupies.¹² Let us be reminded that in these sentences the interrogative constituent is presumed to move from the position where it receives a Theta Role (the one following the verb *seen*) to the left peripheral position where its interrogative value becomes manifest. Suppose that the interrogative phrase leaves behind a sort of empty slot, or gap, in the original position. This supposition, daring as it may sound at first blush, has an intuitive justification (after being prompted to do so, upon a moment's reflection, many of us come to develop a sort of "direct" intuition that there is, indeed, such a gap).

Moreover, it gives an immediate explanation for the fact that, as we have noted earlier, whereas in a declarative sentence the verb *seen* can be followed by an element like a pronoun (cf. *you have seen him*), this is clearly impossible in interrogative sentences (cf. the ungrammaticality of *which man have you seen him?*). Assuming that the interrogative phrase leaves a gap (or, as it is technically called, a trace) after the verb, one can say that the pronoun is excluded because its position is already occupied by the trace. But, most of all, in hindsight, the existence of these gaps has been so well corroborated by many interesting theoretical and empirical consequences, ranging over many languages, that it has gained an aura of obviousness within the generative tradition. Now, our main point here is that the gap can be interpreted as a variable in the logicians' sense. For example, a sentence like (10) can be paraphrased by saying: "For which x , such that x is a man, have you seen x ?". The moved interrogative constituent fixes the range of variation (in our case, the set of the x s that are men) of the variable left in the thematic position. So far, so good. Let us now move to quantificational expressions. First observe that logicians represent a quantifier as an operator that binds a variable from a position at the left periphery of a logical formula. So, typical logical representations involving universal and existential quantifiers are the following:

$$\forall x (x \text{ is even or } x \text{ is odd}) \quad (11)$$

$$\exists x (x \text{ is even and } x \text{ is prime}) \quad (12)$$

From this observation it is natural to speculate that, at some abstract level, the quantificational expressions of natural languages also move to the sentence-initial position, just like the interrogative expressions. An important difference is that the latter do it overtly. Yet, at a higher level of abstraction, this difference becomes inessential. In this way, at a certain higher level, natural

languages would have a uniform logical representation for interrogative and quantificational sentences, moreover, this representation is, in relevant respects, similar to the one assigned to quantificational formulas by logicians.

This all-important abstract level is called by generative linguists *Logical Form*. At this level (or, as it is called for short "at LF") sentence 13 would have the representation in sentence 14:

$$\text{I saw every man} \quad (13)$$

$$\text{Every man } (x) \text{ I saw } x \quad (14)$$

This conjecture, if right, is interesting, because it allows for a unitary analysis of two apparently unrelated empirical domains: quantificational sentences and interrogative ones. In addition, it makes visible a nontrivial correspondence between formal and natural languages. So let us briefly review the main arguments and data supporting this conjecture.

Extensive crosslinguistic research has given ample evidence for this hypothesis. We give one piece of evidence in the following and we refer to the works cited in note 12 for more extensive discussion.

An interesting interpretative fact about interrogatives can be illustrated by the following example:

$$\text{Who does his mother love?} \quad (15)$$

This sentence, interestingly, lacks (or at least strongly discourages) a reading that is, a priori, perfectly intelligible: Which is the person such that his own mother loves this person? This is a perfectly plausible meaning, yet sentence 15 does *not* have it (or at any rate the use of sentence 15 to express it would be exceedingly contrived and ineffectual). This is the reading that would admit as an answer: John (meaning that the mother of John loves John). We all

unproblematically understand sentence 15 as having a different meaning, namely: Who is the person x such that the mother of y (y being a male person *different* from x , and antecedently specified in the discourse) loves x ? This apparently marginal fact about straightforward versus impossible interpretations of sentences of type (15) has received much attention, also because it holds crosslinguistically to an astonishing point.¹³ If we characterize the missing reading saying that in it the pronoun *his* and the interrogative element *who* would have the same referential index - and if we go on assuming that the interrogative has moved from the theta position after the verb to the sentence initial position - there must be a constraint that says something like the following: in its leftward movement, an element cannot cross a pronoun which has the same referential index.

Now, consider the conjecture about the quantificational expressions illustrated by sentences 13 and 14. It was stated that a leftward movement of the quantifier takes place at an abstract level (at LF). This conjecture, if correct, makes an interesting prediction in the case of sentences like the following:

His mother likes everyone (16)

We said that an element cannot cross a coindexed pronoun in its leftward movement. Then, if the quantifier *everyone* actually moves (non-overtly), the following reading should not be allowed: "everyone is loved by his own mother." As is easily verified, this reading is in fact missing. This strongly supports the hypothesis that quantifiers, like interrogatives, leftward-move, leaving a variable in the base position. Note now that nothing like that happens with proper names: sentence 17 below can indeed mean that the mother of John loves John:

His mother loves John (17)

Let us pause a moment to reflect on what we have accomplished. Interestingly enough, the Montagovian approach and the generative one arrive at very different conclusions in a field typically investigated by semanticists: The status of quantificational expressions. While in the former approach their semantics is identified with that of proper names, in the latter, quantifiers are assimilated to a significant extent to interrogative phrases (since they both display an operator-variable structure).

Both analyses are well motivated, even if the kind of motivation behind each of them is very different, and it would be beside our point here to make a definite stand.¹⁴ What we wanted to show was merely that a treatment based on a non-ancillary (that is, autonomous) syntax proved to be theoretically and empirically productive even in the treatment of a typically semantic topic: the theory of quantification. This seems to us to add to the legitimacy of Chomsky's defence, initiated some 40 years ago, of the autonomy of syntax.

Bearing this in mind, one could yet observe that the problem that led Frege to introduce the "Principle of Compositionality" is still very much with us. And, as a matter of fact, a most challenging (and most exciting) task that the present day research in semantics is confronted with is exactly reconciling a system of compositional rules with the background assumption of a sophisticated theory of syntax.

Externalism and Internalism

We have seen that model-theoretic semantics identifies the meaning of a lexical entry with the extension it has in the domain of interpretation.¹⁵ To simplify things, let us say that intransitive verbs and adjectives denote sets, while transitive verbs denote sets of ordered pairs, and singular terms denote individuals (or, if one prefers, sets of sets, as in the Montague account). The second main characteristic of model-theoretic semantics is the fact that the rules of meaning composition are, to a great extent at least, set-theoretical notions (one