1 The Basic Stance of I-Semantics

Chomsky (1986b) distinguishes between two ways to treat language as an object of inquiry. The study of “externalized language” or E-language treats language as an external artifact used by human beings, and seeks to characterize its properties as part of the external world with which humans interact. By contrast, the study of “internalized language” or I-language treats language as a body of knowledge within the minds/brains of speakers, and seeks to characterize its properties within the context of a more general theory of psychology. Since at least the time of Aspects of the Theory of Syntax (1965), the goal of generative grammar has explicitly been an account of I-language.

Clearly, an account of how humans actually use language in the external world eventually requires an account of their grasp of the principles of language, i.e. a theory of I-language. Thus a theory of I-language is ultimately necessary, whether or not one desires an account of E-language as well. (Chomsky has in fact argued further that E-language is not a rewarding subject of scientific inquiry at all.)

Semantics, the theory of the relation between language and the world, suffers from the same duality: one can examine it either as an abstract relation external to language users (E-semantics) or as a body of knowledge within the brain/mind of the language user (I-semantics). Again, an account of how humans use language requires an account of how they grasp the relationship between language and the world, so ultimately a theory of I-semantics is desirable.

More specifically, if the theories of syntax and phonology are situated within the stance of I-language, semantic theory may be concerned either with the relation of language users’ knowledge of language to the external world (I’ll call this “semi-E-semantics”), or with the relation of language users’ knowledge of language to their grasp or understanding of the world. In the latter
case, a full theory then requires as well an account of the relation between the world and the language user's grasp of it, which falls under the standard topic of "perception and cognition." Figure 20.1 schematizes the relationship of these different approaches.

Most research in semantics either is inexplicit about its stance or else professes E-semantics. However, a growing body of work in several somewhat independent traditions has explicitly adopted the stance of I-semantics, aspiring to study the relation between human language and human concepts, and to use language as a tool for exploring the structure of human cognition. The present chapter discusses the consequences of adopting such a stance and some of the more prominent results and disputes that have emerged.

2 Basic Issues for I-Semantics

Some basic questions for a theory of I-semantics are the following:

(1) **The nature of meaning**
   a. What are the terms in which humans grasp the world? That is, what are the formal and substantive properties of human concepts/thoughts/ideas?
   b. How may these properties be formalized so as to develop a fully explicit and predictive theory?

(2) **Correspondence to language**
   a. How are the terms in which humans grasp the world related systematically to linguistic expressions?
   b. In what respects, if any, are these terms not related systematically to linguistic expressions, being under the influence of nonlinguistic knowledge, context, etc.? Is there a natural division between linguistic and nonlinguistic factors?

(3) **Correspondence to the world**
   a. How are the terms in which humans grasp the world related systematically to the world itself through the perceptual systems?
b. In what respects, if any, are these terms not related systematically to the physical world? That is, to what extent are human concepts/thoughts/ideas abstract?

These parallel issues familiar from E-semantics. (1) corresponds to the formal problem of creating an intensional language adequate for expressing all the sentences of natural language, such that semantic properties such as sameness and difference of meaning, inference, and presupposition can be computed as formal relations among sentences. (2) corresponds to the problem of mapping syntactic structure into this intensional language. (3) corresponds roughly to the problem of reference (mapping the intensional language onto a model) and the problem of the intentionality of linguistic expressions, though, as will be seen presently, the similarity is only skin-deep.

But in addition, since in I-semantics the formal theories investigated in (1)–(3) are supposed to be instantiated in a human brain, there are fundamental issues that are not shared with E-semantics:

(4) **Brain instantiation**
   a. How are the systems in (1)–(3) wired in the brain, in terms of both broad localization and local neural connectivity?
   b. To what extent are these systems representative of brain systems as a whole?

(5) **Developmental questions**
   a. How does a person acquire the terms in which (s)he grasps the world and the systems that relate these terms to linguistic expression and perceptual input?
   b. To what extent is concept acquisition free to adapt to the environment, and to what extent is it guided by genetically determined constraints, at either the individual or species level?

(6) **Evolutionary questions**
   a. What aspects of the human way of grasping the world can be traced back to nonhuman antecedents, and which represent evolutionary innovations?
   b. To what extent are these innovations representative of evolutionary innovation as a whole?

That is, I-semantics, unlike E-semantics, in principle connects with more general issues of psychology, neuroscience, and biology. Even if such connections are in their infancy, the aspiration to establish them sets boundary conditions on the enterprise that, as will be seen, propel inquiry in directions quite different from traditional semantic theory.

It should also be noticed that the questions in (4)–(6) strongly parallel well-known issues in generative linguistics, because of its focus on I-language. An
interesting difference is that evolutionary questions about language tend to be rather obscure, because language per se is not attested other than in humans. By contrast, many aspects of the human grasp of the world have obvious animal parallels, so that evolutionary issues can play an interesting role in theory construction.

3 Some References

Some early work on semantics within generative grammar recognized these goals to some degree. Katz and Fodor (1963), Katz (1972), and Weinreich (1966) situate semantic theory in the issue of human competence, but they do not address issues such as (4)–(6). Bierwisch (1967, 1969) is unusual for explicitly linking the search for semantic primitives and semantic universals to perceptual psychology.

In the 1970s a wide range of work developed attempting to link semantics and cognition. Fodor, Bever, and Garrett (1974) and especially Fodor (1975) place traditional philosophical issues about meaning in a psychological framework. Miller and Johnson-Laird (1976) develop a massive account of lexical meaning in a framework that aspires to both computational explicitness and psychological plausibility. Jackendoff (1976, 1978) discusses the relation between grammatical patterns and the cognitive structures they express. Within the AI tradition, Schank’s (1973) “conceptual dependencies,” Minsky’s (1975) “frames,” and Schank and Abelson’s (1975) “scripts” are all explicitly claimed to be theories of the connection between language and human concepts.

Psychological experiments showing the character of word meanings, calling into question classical theories, were pioneered by Rosch (Rosch 1973, 1975, 1978; Rosch and Mervis 1975; Rosch et al. 1976). Berlin and Kay (1969), Clark and Chase (1972), and Levelt (1984) are other noteworthy pieces of work of this period.


More recently, two main schools within linguistics can be loosely distinguished whose central tenet is to identify meaning with human concepts. One is Cognitive Linguistics or Cognitive Grammar, which is (in many cases deliberately) independent of generative grammar. The main figures in this movement are Lakoff (1987, 1990), Lakoff and Johnson (1980), Lakoff and Turner (1989), Langacker (1987, 1991), and Talmy (1978, 1980, 1983, 1985); other important works are Brugman (1981), Fillmore (1982), Fauconnier (1985), Herskovits (1986), Levinson (1992), and Vandeloise (1986). The other school seeks to integrate

The discussion in the rest of this chapter, rather than elaborating the ideas of these trends one by one, concentrates on the issues and results of research in I-semantics as a whole; concurrences and differences among approaches will be mentioned as they arise.

4 The Nature of Truth and Reference in I-Semantics – Preliminaries

In order to treat semantics as an issue about the structure of the human organism, it is necessary to be careful about basic goals of the enterprise. In particular, the traditional preoccupation with explicating the notion of the truth or falsity of a sentence must be re-evaluated. For there is no longer a direct relation between an utterance and the world that renders the utterance true or false; there is instead the sequence of three relations diagrammed in the upper line of Figure 20.1.

As a consequence, a definition of truth for I-semantics parallels the definition of grammaticality for I-linguistics. In I-linguistics, the statement (7a), which appears to be about sentences abstracted away from speakers, is always taken as an abbreviation for statement (7b), which puts grammaticality squarely in the mind of the language user.

\[(7) \quad \begin{align*}
   &a. \text{String } S \text{ is a grammatical sentence of language } L. \\
   &b. \text{A speaker of language } L \text{ judges string } S \text{ grammatical (subject to limitations of memory and processing, and under an idealization of uniformity among speakers).}
\end{align*}\]

Similarly, the traditional Tarskian definition of truth (8a) must be reinterpreted in I-semantics as an issue of judgment, as in (8b).

\[(8) \quad \begin{align*}
   &a. \text{Sentence } S \text{ of language } L \text{ is true iff conditions } C_1, \ldots, C_n \text{ obtain in the world.} \\
   &b. \text{A speaker of language } L \text{ judges sentence } S \text{ true iff conditions } C_1, \ldots, C_n \text{ obtain in his or her construal of the world (subject to limitations of memory and processing, and under an idealization of uniformity among speakers).}
\end{align*}\]
That is, truth is no longer regarded as a relation between a sentence and the
world, but rather as a relation between a sentence and a speaker's construal
of the world. Parallel reinterpretations must be adopted for logical relations
such as entailment, presupposition, and so forth.

This reinterpretation places a crucial burden on I-semantic theory. It is no
longer possible simply to characterize “the world” logically or set-theoreti-
cally, as is frequent in formal semantics. Rather, it is an empirical problem to
determine what sorts of entities inhabit the world as humans construe it.
These entities may or may not be characterizable in standard logical or set-
theoretic terms, and in fact they prove not to be, as will be seen below. Fur-
thermore, truth-conditions must be stated in the vocabulary of human construal
of the world.

The claim of this approach is that when people communicate linguistically,
they do not communicate about the world plain and simple, but about the
world as humanly understood. The entities to which speakers refer are not
entities in “the world” plain and simple, but rather entities available in the
human construal of the world. These include physical objects and events,
illusory objects such as virtual contours, fictional objects such as Santa Claus,
social constructs such as marriages and university degrees, mental constructs
such as intentions and beliefs, and theoretical constructs such as numbers and
logical operators. From the point of view of I- semantics, they all have equally
robust status. (This position is worked out in detail by Jackendoff 1983 and
Lakoff 1987.)

It is a further empirical question how these different sorts of entities are
related to the “real” real world, or alternatively how “meanings in the head”
come to be “meaningful.” We return to these issues, which are still controver-
sial, at the end of the chapter.

5 Compositionality and Universality of Semantic Vocabulary

Like traditional formal semantics, all approaches to I- semantics are committed
to the composition of phrase and sentence meanings from the meanings of
constituent lexical items. With the exception of some of the AI-based approaches,
all take it for granted that syntactic structure acts as an important guide in the
construction of phrase meanings from constituents.

However, nearly all approaches to I- semantics go beyond most work in
formal semantics in making a serious commitment to semantic decomposition
of lexical items into smaller units.2 This practice has borne considerable fruit
in explaining patterns of inference and patterns of syntactic behavior, some of
which will be seen below; this alone is sufficient justification for the undertaking.

But in addition the assumptions of I- semantics lead to foundational questions
of lexical decomposition. Parallel to the basic problem in I-linguistics of how a child acquires the grammar of a language, I-semantics faces the question of how a child acquires vocabulary, in particular the meanings of words. Where do the concepts come from that are associated with sounds and syntactic structures?

On the view that the mind can be characterized in terms of formal systems, learning has to be treated as the combining of existing formal units into larger constructs. Moreover, the “existing formal units” must start somewhere: there must be a vocabulary of primitives and principles of combination (i.e. axioms) to get the system off the ground. This system, by logical necessity, must be instantiated in the brain innately, since by definition it cannot be learned. (This argument is due to Fodor 1975.)

In fact, all theories of learning presuppose some innate basis. It may be something as simple as principles of association, or something as rich as Fodor’s (1975) repertoire sufficient to encompass the meanings of all monomorphemic words. The interesting empirical question thus is not whether there is an innate basis behind word meanings, but exactly what this innate basis is. In practice, most approaches to I-semantics assume a fairly rich combinatorial basis for word meanings, with a sizable repertoire of substantive conceptual primitives.

It is sometimes objected (e.g. Partee 1993b) that the innateness of semantic primitives should imply a wider degree of universality in the lexical semantics of languages of the world than actually exists. There are two answers to this objection. The first takes a cue from generative grammar. The innateness of Universal Grammar in phonology and syntax does not imply overwhelming uniformity in these aspects of language: languages can pick and choose among a repertoire of possibilities, some of which are even mutually inconsistent with each other (i.e. accusative vs. ergative case marking). The same could easily be true of semantics: of the concepts made available in the innate basis, languages may choose to grammaticalize or lexicalize different selections, so that different languages may appear semantically incommensurate.

Second, it is not necessarily the case that all (or any) semantic/conceptual primitives are independently expressible as words. Just as the smallest isolable speech sounds (phonemes) are composites of distinctive features that cannot appear independently, so it appears to be the case that all word meanings are composite, made up of semantic/conceptual constituents that cannot appear in isolation. That is, word meanings are “molecular” entities in the “chemistry of concepts,” while semantic/conceptual primitives are subatomic or even quarklike. This being the case, the ultimate decomposition of a lexical item cannot be expressed in terms of a linguistic paraphrase.

The impossibility of paraphrase definitions is often taken as an argument against lexical decomposition (e.g. Fodor, Garrett, Walker, and Parkes 1980), but in fact paraphrase definitions can be expected only if every semantic primitive can itself be expressed by a word, and if every principle of semantic composition can be paralleled by a principle of syntactic composition. Otherwise
(as is actually the case), the absence of paraphrase definitions is altogether natural.

How then does one justify a proposed semantic/conceptual primitive? First notice that it cannot be justified in isolation; it makes sense only in the context of the rest of the axiomatic system in which it is situated. In turn, a particular proposal for an axiomatic system is justified not only logically but empirically, in terms of how well it explains the semantic phenomena of language and cognition. Sometimes the issue is raised as how one can ever tell whether a set of proposed primitives is "really" primitive. Suffice it to say that this question seems never to have discouraged physicists, who just take it as a spur to see if deeper decomposition leads to deeper explanation. In practice this is what happens in lexical semantics as well.

6 Relation of Human Concepts to Natural Language Semantics

The last section used the locution "semantic/conceptual" as though the two are one and the same within an I-semantic theory. Most practitioners of I-semantics in fact take for granted the identification of natural language meanings and human concepts, but two alternative positions deserve discussion.

The basic hypothesis in I-semantics is that the syntactic structures of natural language must be related to the concepts they express by a set of rules of translation or correspondence, where "concepts" include all the richness and interconnection of human knowledge (sometimes called "encyclopedic" knowledge and sometimes "pragmatics"). However, the question arises as to whether the relation is direct (Figure 20.2a), or whether there is an independent identifiable level that can be called "semantics proper" intervening between syntax and concepts (Figure 20.2b).

For those advocating a two-stage mapping like Figure 20.2b, the issue is what role the independent level of semantic structure plays. In early work of Katz (1972, for instance), semantic structure is conceived of as a level of representation which captures only semantic relations among sentences such as entailment, synonymy, ambiguity, and anomaly, but which does not represent

---

Figure 20.2

a. syntactic structures semantic/conceptual structures
   \[ \xrightarrow{\text{mapping}} \]

b. syntactic structures semantic structures conceptual structures
   \[ \xrightarrow{\text{mapping}} \xrightarrow{\text{mapping}} \]
“world knowledge” in general. Such a level of semantics is sometimes assumed by more recent formal semanticists as well. In fact it has proven impossible to separate out purely “semantic” relations from world knowledge and metaphoric construal without massively losing generalizations (Jackendoff 1981, 1983; Lakoff and Johnson 1980; Lakoff 1987), so Katz’s conception of semantic structure cannot be maintained.

A second conception of semantic structure has arisen independently in work of, for instance, Pinker (1989) and Bierwisch (1982, 1992), in which it is the level of representation which encodes just those aspects of meaning that make a difference in syntactic structure. For instance, many languages make syntactic distinctions based on number, gender, and animacy, but none make distinctions based on color or the difference between dogs and cats. The idea, then, is that semantic structure encodes number, gender, and animacy but not color or species. Similarly, languages make syntactic distinctions between eventive and stative verbs, but not between verbs of running and verbs of walking; again only the former distinction would appear in semantic structure. Sometimes the implication is that while conceptual structure is universal, semantic structure varies from language to language, depending on what conceptual distinctions are grammaticalized. (See e.g. Talmy 1980, Choi and Bowerman 1991 for discussion of crosslinguistic differences.)

It is possible, however, to reproduce this configuration of facts within a theory structured as in Figure 20.2a, by situating the language-particular aspects of semantics directly in the mapping between syntax and conceptual structure. The idea is that this mapping is only sensitive to a restricted repertoire of conceptual distinctions, so that the rest are necessarily opaque to syntax. The variability among languages in this respect is then attributed to variability in the conceptual factors to which the mapping is sensitive. As a result, no extra level of strictly linguistic semantics is necessary.

Notice how this parallels the behavior of the syntax-phonology interface. There are many phonological distinctions which play no role in syntax; but this does not lead to the postulation of a separate level of representation that contains only the phonological distinctions relevant to syntax. Rather, the idea is that the mapping between phonology and syntax is simply not sensitive to segmental content (say the difference between the phoneme sequences *dog* and *cat*), to number of syllables (*camel* vs. *dromedary*), and so forth. On the other hand, the mapping is sensitive to word and phrase boundaries, to affixal structure, and so forth. That is, the mapping between levels of representation is only a partial homology, not a one-to-one correspondence of every element of the structures.

It is unclear which of these views of semantic structure will eventually prevail. The rest of this chapter will assume the latter view without further argument, accepting the identification of semantic representations with (possibly a subset of) human conceptual structures, in accordance with the majority of work in the field.
7 The Noncategorical Nature of Conceptual Categories

The emphasis in I-semantics on lexical meaning has given rise to considerable investigation into the phenomenon of categorization: how humans place individuals into categories, and how systems of categories are constructed mentally. As a semantic issue, this problem can be framed in terms of stating the truth-conditions for the sentence This is an X, where X is the category in question. By contrast with the main tradition of formal semantics, which adopts rather uncritically the Tarskian criterion of necessary and sufficient truth-conditions, much I-semantic research (in particular Jackendoff 1983 and Lakoff 1987) stresses the insufficiency of Tarskian conditions to characterize the richness of human categories. This section presents examples of some of the problems that have arisen.

7.1 Categories with graded boundaries

Consider the category red. This cannot be identified with a particular perceived hue, since a broad range of hues are all called red. However, as hues shade imperceptibly from red toward orange, there comes a point where observers are no longer clear about their judgments. There may be hesitation and interobserver disagreement. In addition, the judgment of a particular hue may depend on what hues have immediately preceded it in presentation (if presented after focal red, it is judged orange, but if after focal orange, it is judged red). That is, there is a focal range in which judgments are secure and consistent, but this shades into a borderline range in which there is conflict with a neighboring category, and in which judgments become less secure and more context-dependent.

This “fuzziness” in the boundary of a category is not a matter of speakers not “knowing the meaning of red”; rather it is inherent in the structure of the concept itself. One can make red more “Tarskian” by stipulating rigid boundaries with orange, pink, purple, and brown, but then one is not dealing with the ordinary meaning of the word. Similarly, one can create a new category red–orange at the boundary, but then the same sort of fuzziness occurs at the boundary of red and red–orange. (The existence of such phenomena is noted by Putnam 1975; Berlin and Kay 1969 is the classic work on color judgments.)

Similar boundary problems arise with words like hot and tall that express significant deviation from a norm. What is the lower bound of temperature for, say, hot soup, or the lower bound of height for tall woman? It is inherent in the structure of these concepts that the boundaries are not classically sharp, in the sense that any particular woman can be definitively said to be either in or out of the set of tall women.
Incidentally, these adjectives also present evidence for the interdependence of linguistic meaning and encyclopedic knowledge in judgments of truth. For, as has often been observed, a small elephant is bigger than a big mouse: the norm to which the adjective is applied depends on one’s knowledge of the standard size of the animals in question. (Bierwisch and Lang 1989 discuss in detail adjectives that relate an instance to a norm.) Similarly (as Talmy 1978 points out), what counts as nearness between stars is metrically quite different from nearness between pieces of furniture; nearness too is defined in terms of the normal distance expected among individuals of the category in question. (See Pustejovsky 1991 for more highly elaborated examples of this sort, having to do with what Millikan 1984 calls the proper function of objects; Katz’s 1966 treatment of good is an early, somewhat clumsily executed example.)

7.2 “Cluster” concepts

The classic example of this phenomenon is Wittgenstein’s (1953) analysis of game, in which he demonstrates that it is impossible to find a necessary condition that distinguishes games from other activities. He suggests that the word is understood in terms of “family resemblance”: there is a cluster of conditions that define games; but no games satisfy all of them, and none of them is shared by all games. That is, none of the conditions is necessary, but various suitable combinations of them are sufficient to permit an individual to be judged a member of the category. Such categories are now called cluster concepts.

This analysis is amplified by Rosch (1973, 1975, 1978; Rosch and Mervis 1975), who shows experimentally that categorization judgments may contain a cline of “typicality,” ranging from typical exemplars of the category (e.g. a robin is a “typical” bird) to atypical exemplars (e.g. a penguin is an “atypical” bird). Among the various sources for typicality effects, one is a set of conditions that form a cluster concept: examples that satisfy fewer of the conditions are generally regarded as less typical than examples that maximally satisfy the conditions.

The effects of cluster conditions can be observed in concepts involving as few as two conditions. A case first discussed by Fillmore (1982) is the verb climb. Consider the following examples:

(9) a. Bill climbed (up) the mountain.
b. Bill climbed down the mountain.
c. The snake climbed (up) the tree.
d. ?*The snake climbed down the tree.

Climbing involves two independent conceptual conditions: (1) an individual is traveling upward, and (2) the individual is moving with characteristic effortful grasping motions (clambering). On the most likely interpretation of (9a), both
conditions are met. (9b) violates the first condition, and, since snakes cannot clamber, (9c) violates the second; yet both examples are acceptable instances of climbing. However, if both conditions are violated, as in (9d), the action cannot be characterized as climbing. Thus neither of the two conditions is necessary, but either is sufficient. Moreover, the default interpretation of (9a), in which both conditions are satisfied, is judged more stereotypical climbing; (9b) and (9c) are judged somewhat more marginal but still perfectly legitimate instances.

Parallel analyses have been proposed for the verbs *lie* ("tell a lie") (Coleman and Kay 1981) and *see* (Jackendoff 1983, based in part on Miller and Johnson-Laird 1976). Similar phenomena arise in lexical entries for nouns as well. For instance, a stereotypical *chair* has a stereotypical form and a standard function. Objects that have the proper function but the wrong form – say beanbag chairs – are more marginal instances of the category; and so are objects that have the right form but cannot fulfill the function – say chairs made of newspaper or giant chairs. An object that violates both conditions – say a pile of crumpled newspaper – is by no stretch of the imagination a chair. Lakoff (1987, chapter 4) applies such an analysis to the concept *mother*, which includes the woman who contributes genetic material, the woman who bears the child, and the woman who raises the child. In today's arrangements of adoption and genetic engineering, not all three of these always coincide, and so the term is not always used stereotypically.

Jackendoff (1983) calls a system of conditions of this sort a *preference rule system*; Lakoff (1987) calls it (one aspect of) an *idealized cognitive model*; the *frames* of Minsky (1975) have similar effects. Such concepts, like Tarskian concepts, are combinations of conditions. They differ from Tarskian concepts in that the conditions are combined differently. The combination is not a logical conjunction, since satisfaction of less than the full set of conditions is sufficient to categorize an object. But it is not a logical disjunction either, because a logical disjunction does not make the proper distinction between central and borderline examples. That is, a proper theory of word meanings must be richer than standard logical models. (Default logic comes closer to capturing the effect of such conditions, which function as default values where there is no evidence to the contrary, for instance in (9a).)

Appealing to the psychological goals of I-semantics, such enrichment of the theory proves to be plausible on broader psychological grounds. The manner in which conditions interact in cluster concepts is central in visual perception (Wertheimer 1923, Marr 1982), in phonetic perception (Liberman and Studdert-Kennedy 1977), in musical perception (Lerdahl and Jackendoff 1983), and in Gricean implicature (Bach and Harnish 1979). Moreover, such an interaction is plausibly instantiated in the brain, where the firing of a neuron is normally not a rigid function of other neural firings (like a logical conjunction), but rather a composite of many excitatory and inhibitory synapses of varying strengths. Thus cluster concepts, even though unusual in a logical framework, are quite natural in a psychological framework.
7.3 Image-like components of meaning

Putnam (1975) suggests that for many categories, speakers carry in their heads images of stereotypical instances (though, because he is committed to E-semantics, he does not consider such images part of meanings). Within I-semantics, appeal to image-like representations as part of concepts is widespread, though there are few concrete proposals. Jackendoff (1987a) and Landau and Jackendoff (1993) suggest that entities that are understood spatially (e.g. physical objects and actions) are partially encoded in terms of a spatial representation that is an enriched version of Marr’s (1982) 3D model structure. It is widely assumed that other sensory phenomena such as sounds, smells, tastes, and bodily sensations are also encoded partially in modality-specific imagistic terms.

Again, although parsimony might argue against introducing such further kinds of representations into meanings, psychological considerations abundantly support it. Part of the process of identifying a physical object is comparing its appearance to known examples or to a schema that summarizes known examples. Given that the visual system must encode such schemata independently – even in nonlinguistic organisms – there is no reason that the conceptual system should not be able to make use of this information. If it did not, information about appearance would have to be encoded in the conceptual system anyway in order for categorization to take place, resulting in a needless duplication of information.

There is a further reason to accept imagistic encoding as part of word meaning. Many detailed characteristics of shape (as well as sound, smell, taste, and body sensation) do not lend themselves to any sort of algebraic (or propositional) feature decomposition. For instance, the difference between the appearance of a duck and that of a goose can be expressed algebraically only in terms of ad hoc features such as [+long neck]. However, such a difference is altogether natural in a system like Marr’s 3D model structure, which is adapted to detecting subtle differences of shape and contour.

This is not to say that all categorization can be reduced to imagistic representations. For instance, imagistic representation cannot encode the distinction between types and tokens or the place of ducks in the taxonomy of waterfowl, birds, and animals. Nor can it encode the quantificational aspects of meaning. Rather, it seems appropriate to think of algebraic and imagistic representations as sharing between them the burden of labor of encoding human knowledge of categories – if you will, Putnam’s “division of linguistic labor” applied to the modules of the mind.

The linkage thus established between sensory/imagistic encodings and propositional/algebraic encodings also helps explain how it is possible for us to talk about what we see. In order for people to visually identify an entity to which language makes reference, the semantic/conceptual structures derived from language cannot be compared directly against the world. Rather, they
must be placed in correspondence with representations arising from the visual modality, so that they may be checked against visual input. This is the only plausible causal mechanism by which visual input can have an influence on linguistic expression. Thus an insistence on the psychological principles of I-semantics has important consequences for the theory of reference. We return to this issue in section 10.

8 Spatial Expressions

In contrast with formal semantics, research in I-semantics has consistently been preoccupied with the interpretation of expressions of spatial location and motion. There are a number of natural reasons for this emphasis, all emerging from the fundamental goals of I-semantics. First, the field of spatial location and motion is very rich. Second, because sentences of spatial motion and location are perceptually verifiable (see the previous subsection), judgments of truth, entailment, and ambiguity are quite sharp and highly structured. Third, as will be seen in section 9, many other fields of concepts are structured algebraically along similar lines, so an understanding of spatial expressions potentially yields dividends in understanding language more broadly.

Fourth, again drawing on the biological aspect of semantics, it is to be expected that psychological encoding of spatial location and motion evolved long before language, since all our mammalian antecedents show the capacity to navigate through the world. Fifth, as Piaget (1970) stresses, children show an understanding of spatial concepts in their sensorimotor behavior well before they can speak. In studying spatial concepts, then, one is tapping into a kind of human knowledge that is biologically and developmentally prior to linguistic knowledge, which makes it of exceptional interest. (Many of these reasons are stressed by Miller and Johnson-Laird 1976 as well as later writers.)

There is room here only to touch on one basic phenomenon within this domain: figure-ground organization. However, this phenomenon is central, as it forms the foundation of the theories of Gruber (1965), Talmy (1978, 1983), Langacker (1987), and Jackendoff (1983), and underlies much detailed work on prepositional meanings such as Brugman (1981), Herskovits (1986), Vandelooise (1986), and Landau and Jackendoff (1993).

Consider a simple sentence like The cat sat on the mat. The two objects being described, the cat and the mat, are expressed asymmetrically: the cat is subject of the sentence, near the top of the syntactic tree, while the mat is embedded down inside the prepositional phrase. In principle the spatial relation between the cat and the mat could be encoded conceptually as (10), in which on is a two-place relation between the two objects; proposals along these lines are not infrequent.

(10) sit (the cat) and on (the cat, the mat)
However, the research cited above claims that the meaning of the sentence is asymmetrical in just the way that the syntax is:

1. The mat is conceptualized as a reference object or landmark.
2. The reference object helps to define a region of space. This region of space is expressed by the PP on the mat: it is the region in contact with the upper surface of the mat. Different prepositions define different regions based on the same reference object, for instance under the mat, beyond the mat, and off the mat.
3. The verb uses this region of space to locate a figural object, namely the cat. In the present example, the verb sat also indicates the posture of the figural object and the past time of the situation in question.

So an informal analysis of The cat sat on the mat is something like (11).

(11)  
    the cat  
    figural object  
    sat  
    location of  
    figural object  
    in region  
    +  
    posture of  
    figural object  
    +  
    past time  
    on the mat  
    reference object  
    region  

To see the asymmetry more clearly, we may reverse the roles, producing the sentence The mat lay under the cat. This sentence sounds distinctly odd, even though the physical situation it depicts is exactly the same; and certainly the grammar is impeccable. In order to see why this should be so, consider a situation in which the roles of the figural and reference objects can be satisfactorily exchanged, for instance (12).

(12)  
    a. The star is inside the circle.  
    b. The circle lies around the star.

Even here there is an asymmetry in the concepts expressed by these two sentences, as can be seen by using them as answers to questions. (This argument is due to Gruber 1965.)

(13)  
    a. Where is the star?  
       It is inside the circle.  
       ??The circle lies around it.  
    b. Where is the circle?  
       It lies around the star.  
       ??The star is inside it.

That is, sentence (12a) is a much better way of saying where the circle is, and sentence (12b) is a much better way of saying where the star is. This difference
is explained by the analysis in (11), which says that the reference object (the object of the preposition) is being used to help locate the figural object (the subject of the sentence), not the other way around. In these terms, it is clear why The mat lay under the cat sounds odd: it is telling how to find the mat by using the cat as a point of reference; but it is hard to imagine circumstances in which one would want to do this.

For an even more striking case pointed out by Talmy (1983), next to appears to be totally symmetrical: if X is next to Y, it stands to reason that Y is next to X. Yet The bike is next to the house is fine, but The house is next to the bike is distinctly odd. The reason is that relatively large immovable objects are normally used as landmarks for locating small movable objects, but not the other way round. (Notice that if the house in question happens to be a toy “Monopoly” house, so it is relatively small and movable, The house is next to the bike is much more acceptable.)

The asymmetries in these sentences have nothing to do with the syntax of English, which says only that the subject NP comes first, the object of the preposition follows the preposition, and so forth. And, as (12)–(13) make clear, the asymmetries have nothing to do with the physical world per se either. Rather, they have to do with the way the human mind conceptually organizes the perceived world: how we use objects to find other objects.

This conceptual organization, and the grammatical organization that goes with it, is replicated in language after language of the world. Thus we have good reason to believe it is part of the innate basis of human concepts.

These examples also show that language expresses the location of objects not in terms of absolute space, but always in terms of figures placed against a background. This background is a region of space whose organization is determined by reference objects (among which may be the speaker and hearer). In terms of a formal logic, this may or may not be a natural way for language to work. But in a psychological context, it is altogether natural, as it accords entirely with the figure-ground organization found in the perceptual systems, and therefore provides a natural mapping to procedures of perceptual verification of the sort discussed in section 7.3.

9 The Same Organization in Abstract Domains of Thought

A second foundational result in I-semantics also originates with Gruber (1965), who showed that many grammatical patterns used to describe physical objects in space also appear in expressions that describe non-spatial domains.

The groups of sentences in (14) through (17) illustrate this result. The first example, (14a), exhibits the figure-ground organization just discussed; the whole set illustrates the larger grammatical and conceptual patterns within which
this particular pattern is situated. Notice especially the parallels indicated by
the italicized words.

(14) **Spatial location and motion**
    a. The messenger *is* in Istanbul.  
      (Location)
    b. The messenger *went from* Paris to Istanbul.  
      (Change of location)
    c. The gang *kept* the messenger in Istanbul.  
      (Caused stasis)

(15) **Possession**
    a. The money *is* Fred’s.  
      (Possession)
    b. The inheritance *finally went to* Fred.  
      (Change of possession)
    c. Fred *kept* the money.  
      (Caused stasis)

(16) **Ascription of properties**
    a. The light *is* red.  
      (Simple property)
    b. The light *went/changed from green to* red.  
      (Change of property)
    c. The cop *kept* the light red.  
      (Caused stasis)

(17) **Scheduling activities**
    a. The meeting *is on* Monday.  
      (Simple schedule)
    b. The meeting was *changed from Tuesday to* Monday.  
      (Change of schedule)
    c. The chairman *kept* the meeting on Monday.  
      (Caused stasis)

Each of these groups contains one sentence with the verb *be*, one with *go* or *change*, and one with *keep*. When *be* appears with a preposition (as in (14a) and (17a)), the same preposition can appear with *keep*; if *be* appears without a preposition (as in (15a) and (16a)), so does *keep*. On the other hand, *go* and *change* characteristically appear along with the prepositions *from* and *to*.

There is no reason for these patterns that derives from the physical nature of the situations expressed. Changing possession does not necessarily entail changing location: the sale of a house or stocks does not involve motion at all. An object’s color has nothing to do with where it is or who owns it. Setting the appointed time for a meeting or trip bears no apparent relationship at all to the other three.

On a more abstract level, however, the meanings of the four groups of sentences are parallel.

1. The “*be*” sentences all describe some state of affairs in which some characteristic is ascribed to the subject of the sentence: location in a
region in (14), belonging to someone in (15), having a property in (16),
and having an appointed time in (17).
2. The “go/change” sentences all describe a change involving the subject
of the sentence, in which it comes to have the characteristic ascribed
by the corresponding “be” sentence. The subject’s characteristic at
the beginning of the change is described by the phrase following “from,”
and at the end of the change by the phrase following “to.”
3. The “keep” sentences all describe the subject of the sentence causing
the object of the sentence to have the characteristic ascribed by the
corresponding “be” sentence, and this characteristic persists over a
period of time.

In other words, the linguistic parallelism among these sets reveals an under-
lying conceptual parallelism. Thus it is not accidental that many of the same
lexical items recur in (14)–(17).

The idea behind the conceptual parallelism is this: The characteristics that
things can be conceived to have fall into broad families or “semantic fields”
such as the headings in (14)–(17). Within a field, “be” sentences express simple
characteristics such as being in a particular location, belonging to a particular
person, being of a particular color, or being scheduled at a particular time. But
in addition, the conceptual system contains complex concepts that can be
applied to any field, among which are (1) a change from one characteristic to
another (the “go/change” sentences) and (2) something making something
else have a particular characteristic over a period of time (the “keep” sen-
tences). Because the same abstract system appears in many (and possibly all)
semantic fields, it is convenient for the language to use the same words as it
switches from one field to another.

In fact, almost any characteristic that varies along a one-dimensional range
of values turns out to be expressed in terms of the linear opposites up and
down or high and low. Numbers (and hence prices, weights, and temperatures)
go up and down, military ranks go up and down, pitches on the musical scale
go up and down, and so does one’s mood. However, as is well known, time
concepts are a partial exception to this generalization: in just about every
language, they are expressed by terms that also apply to space; but instead of
up and down, they use a front-to-back continuum, for example before and after
in English.

Similar grammatical and lexical patterns appear in language after language.
Talmy (1978, 1985a) and Langacker (1987) essentially take them for granted
and build theories of conceptual structure around them. Lakoff and Johnson
(1980) and Lakoff (1987, 1990) argue that they are part of a vast system of
metaphor that is inextricably embedded in the conceptual structure expressed
by language. They further argue that the pervasiveness of metaphor makes it
impossible to base a semantic theory on a simplistic notion of “literal truth,”
and impossible to treat metaphor as nontruth-conditional and therefore
marginal.
Jackendoff (1983; 1992, chapter 3) and Jackendoff and Aaron (1991), while acknowledging the ubiquity of metaphor, argue for a finer-grained distinction. The traditional notion of metaphor applies to creative, novel expressions, often with a patent semantic clash, used to make speech more colorful. Lakoff et al. extend the term metaphor to any extension of terms from one semantic field to another. However, the parallels illustrated in (14)–(17) exhibit no semantic clash, and they are the only ways available in English of expressing the concepts in question. Thus they are not metaphorical in the usual sense. Jackendoff argues therefore that the primitives of conceptual structure include a set of precise abstract underlying patterns that can be applied to many different semantic fields such as those in (14)–(17). These patterns are the basic machinery that permits complex thought to be formulated and basic entailments to be derived in any domain. The apparent primacy of the spatial field is then due (1) to its evolutionary primacy and (2) to its strong linkage to perception. (This does not preclude more high-level processes of metaphor as well.)

However this dispute is resolved, all major schools of thought in I-semantics (other than Fodor) agree that the linguistic parallelisms shown in (14)–(17) reflect substantive parallelisms in the concepts these sentences express, and thereby reveal fundamental organization in human conceptual structure.

10 Counterpart of Theory of Reference in I-Semantics

Recall that, in I-semantics, the entities to which language can refer are not entities in “the world” plain and simple, but rather entities in the world as humans grasp it. This claim can now be better appreciated in the light of the last three sections. Consider for instance (17b), The meeting was changed from Tuesday to Monday. Unlike The cat sat on the mat, this sentence points to nothing perceptible. It describes something that takes place only in people’s minds, changing their future behavior. In fact, the idea of named time periods like Monday is purely a conceptual invention – there is nothing perceptually salient about the boundaries of Monday. Yet humans indubitably experience reality as containing Mondays and schedules, and the term Monday is taken to refer.

Still more compelling is the case of possession. If something belongs to you, it doesn’t belong to me; and our behavior is strongly constrained by this knowledge. Who owns what is an important part of the reality we experience. Yet the predicate of possession has little perceptual basis; it involves strictly conceptual notions such as rights and obligations over potential use of an object.

Returning to categorization, there is nothing in the world that demands a sharp boundary between red and orange, and there is nothing in the world that sharply distinguishes climbing from other kinds of locomotion. It is the human
need to sort the particulars of the world into categories that creates these divisions. In short, many aspects of the concepts expressed by language are purely mental constructs. Nevertheless, they are not fictional or senseless – they are part of the way the world is for human beings.

Physical objects and actions tend to have strong perceptual linkages to the "real" world, social constructs less so. However, it is not a matter for semantics per se how these perceptual linkages are established; it is part of the theory of perception. What is part of semantics is the internal principles of inference among related concepts.

A useful example is the concept of a strike in baseball. What makes a particular event count as a strike is its role in a baseball game; a strike is what there are three of when a batter strikes out. The perceptually verifiable conditions are such things as whether the batter swung at a pitched ball and missed, or whether a pitched ball at which the batter did not swing passed through the "strike zone." It is a matter for perceptual theory to determine how the umpire judges where the ball went and what the batter did. The significance of the physical event – whether it takes place as an appropriate subevent of a baseball game, and how it counts in the progress of the game – are all entirely conceptual. Yet the term strike can clearly be used referentially. Strikes can be individuated, counted, and quantified; not only that, millions of people care about them.

Such combinations of perceptual and conceptual conditions appear to be pervasive in human concepts (see Carey 1985, Keil 1989 for discussion). In particular, it appears that many theoretical and scientific concepts share the general characteristics of strikes, having an elaborate logico-conceptual organization that is grounded in a scattered array of perceptual observations.

A case sometimes offered against this approach to reference is that of logical and mathematical facts, which, though abstract, are often held to be true independent of human observers. (This is, for example, a major argument of Katz (1981) in his rejection of I-semantics in favor of a Platonistic approach.) Within I-semantics, though, the problem turns inside out: what is it about human beings that (1) permits them to grasp logical and mathematical facts and (2) leads them to believe that these are true independent of human observers? At the moment the jury is out; Macnamara (1986) and Lakoff (1987) have interesting discussions.

There remains an issue that some take to be the pre-eminent issue for any theory of semantics, particularly one that situates meanings in the head. This is the issue of intentionality: how a set of formal distinctions internal to the brain can be meaningful, how they can be about "the world." However, in an I-semantic account this issue plays itself out rather differently than usual. I-semantics claims that, although "the world" is experienced as external to the observer, it is full of entities that strictly speaking exist only because of the observer's construal of the world – such as the examples just cited. Consequently, the meaningfulness of mental representations are not to be ascribed to any simple relation between symbols in the head and a preexisting,
precategorized "real world." Quite different solutions are offered by Fodor (1990), Jackendoff (1987a), and Lakoff (1987), even brief sketches of which would take the present essay too far afield.

However this issue comes to be resolved, it is important to notice that it must be confronted by any theory that aspires to account for the human grasp of language. To maintain an E-semantic approach (or a "semi-E-semantic" approach) because it faces no such difficulty simply puts off the issue without solving it.

At the same time, the lack of a firm doctrine on intentionality has not prevented the various schools of I-semantics from producing an array of deep results about the nature of the concepts that language expresses, and the connection of these concepts to human psychology and evolution. It is these results that should command the attention of semanticists in other traditions.

NOTES

1 The list compiled below is by necessity a personal choice. I apologize in advance for any failure on my part to cite work that the reader feels essential.

2 The only major exception is Fodor's (1975) Language of Thought Hypothesis. For reasons too elaborate to go into here, Fodor claims that the meanings of monomorphemic lexical items are undecomposable innate monads. See Jackendoff (1983, Section 7.5; 1992, Section 2.8, chapter 8) for discussion.

3 Typicality judgments occur, for instance, in color concepts, where a focal red is judged more typical than a red tinged with orange. Armstrong, Gleitman, and Gleitman 1983 show that typicality judgments can be obtained even for noncluster concepts; for instance 38 is judged a less typical even number than 4, and a nun is judged a less typical woman than a housewife. Thus typicality in and of itself is only a symptom for a number of underlying phenomena in categorization. But the existence of multiple sources for graded judgments does not undermine the existence of cluster concepts, as Armstrong et al. claim it does. See Jackendoff 1983, chapter 7, note 6 and Lakoff 1987, chapter 9 for discussion.