

CHAPTER 10

Reference and Truth

10.1 Introduction

A crucial part of semantic theory is to explain how reference and truth value are attached to linguistic expressions. Common sense tells us that linguistic expressions say things about the world. When I say *Russell was a great philosopher*, I am referring to Russell, a real person in the real world. When I point at something and say *That is a chicken*, my deictic term *that* is intended to refer to the thing in the world that I'm pointing at. Sentences intuitively have truth values by virtue of how they relate to the world. *Snow is green* is false because snow isn't green. *My name is Russell* is true if uttered by somebody named Russell (e.g. Rosalind Russell, Kevin Russell, Russell Stover), and false if uttered by anyone else. *It's raining outside* is true if at the moment of utterance there is rain coming down outside the enclosure in which the speaker is located and false otherwise.

The predominant traditions in Anglo-American semantics and philosophy of language take for granted this common-sense position. They therefore consider it the task of semantic/pragmatic theory to explain how linguistic expressions say things about the world and have truth values based on their relation to the world.

The next two sections will show that a mentalistic theory of language proves not so easy to reconcile with this common sense position on reference and truth. After exploring a couple of not very promising alternatives, I will conclude that *it is necessary to thoroughly psychologize not just language, but also "the world."* The intent behind this perhaps mysterious statement will become clearer as we go along, and many perils must be avoided in order to make sense of it.

I must make clear from the outset that my goal is not to show that the common-sense view of reference and truth is false and should be rejected out of hand. Consider by analogy the notion of *sunset*. The scientific view of sunsets does not say that the sun doesn't really go down and therefore that there is no such thing as a sunset. Rather, it says that (a) a less intuitive account in terms of

the earth's rotation yields a better explanation in terms of physics as a whole; and (b) human psychology is nevertheless such that, no matter how much physics we know, the sun still compellingly *appears* to be going down. Try to watch a sunset and see it (feel it!) as the earth rotating. And "What a beautiful sunset!" is not to be superseded by a scientific reduction.

Similarly, in approaching reference and truth, my goal here is to show that for scientific purposes a position at odds with intuition yields a deeper understanding of the phenomena. But I also wish to show why the common-sense view is nevertheless so intuitively compelling and why it therefore cannot be eliminated from everyday discourse.¹

10.2 Problems with the common-sense view: "language"

Here are four rather typical statements of the standard position.

... linguistic expressions refer to things out in the world. (Abbott 1997)

In general, to determine whether a sentence is true or false, two things are necessary: (1) you must know what the sentence means and (2) you must face the sentence with some situation in the real world and see whether it corresponds to the meaning of the sentence. (Bach 1989: 8)

What is a semantic theory? Following Tarski I view semantic theories as theories that deal with concepts relating language to the world (in a broad sense): "We shall understand by semantics the totality of considerations concerning those concepts which, roughly speaking, express certain connexions between the expressions of a language and the objects and states of affairs referred to by these expressions." (Tarski 1936: 401)

... Reference is a relation between a term and an object it refers to, and satisfaction is a relation between a formula and an object ... satisfying it ... Truth is definable in terms of reference and satisfaction because it has to do with objects and their relations to language. Truth holds ... of a given sentence *s* iff the objects referred to in *s* possess the properties (stand in the relations) attributed to them by *s*. (Sher 1996: 531)

The first family of theories can be labeled "referential" or "denotational." This kind of theory is outward looking; its main emphasis is on the informational significance of language, its aboutness. Meaningfulness lies, according to this view, in the relations of symbols and configurations thereof to objects of various kinds. The study of meaning is the study of such relations. This tradition is the basis of the semantic techniques that

¹ The approach to reference and truth developed in the present chapter has roots in Jackendoff (1983; 1987; 1992a: ch. 8); here I go somewhat beyond those treatments. Lakoff (1987) develops some parallel critiques of the main philosophical positions on truth and reference (what he calls "objectivism"), but he does not to my way of thinking establish sufficiently detailed foundations for his alternative.

have been developed within mathematical and philosophical logic.² (Chierchia and McConnell-Ginet 1990: 46)

These statements accord nicely with common sense; the latter two sharpen the common-sense position. But I would like to dig a bit deeper and question how one is supposed to construe “language” on the one hand and “objects” on the other.

Let us first think about the construal of “language.” As will be recalled from Chapter 9, Frege wished to eliminate personal associations or “ideas” from semantics, and therefore he settled on a theory of meaning whose central notions were the reference of an expression and a speaker-independent “sense.” Hence Frege and much of the tradition following him take language to be independent of its human users: it relates directly to the world. This is the usual construal of the quotes above, and might be diagrammed as Fig. 10.1.

Later semanticists (e.g. Kripke 1972; Lewis 1972; Stalnaker 1984) replace this simple conception with an approach in which language maps to “possible worlds,” as in Fig. 10.2. Among other things, this move enables semantic notions like analyticity and necessity to be captured in terms of “true in all possible worlds.”

In formal semantics (Montague 1973; Partee 1975; 1976), “world” comes to be replaced by “model,” a set-theoretic construct that enables the theory to be completely formalized. In principle formal semantics is neutral about its metaphysics. But on the whole “language” is taken to retain its common-sense status as something “out there in the world.” For instance, the four quotes above are from formal semanticists. David Lewis is equally plain:

I distinguish two topics: first, the description of possible languages or grammars as abstract semantic systems whereby symbols are associated with aspects of the world; and second, the description of the psychological and sociological facts whereby one of these abstract semantic systems is the one used by a person or population. Only confusion comes of mixing these two topics. (Lewis 1972: 170)

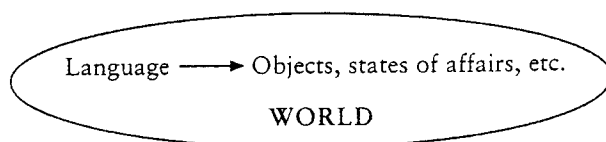


Fig. 10.1. An “objectivist” or “realist” view of language

² Chierchia and McConnell-Ginet’s “second family” of theories are “psychologistic” or “mentalistic,” presumably including the theory to be developed here. The burden of the present chapter is to surmount their criticisms of such theories, to be summarized below. Their “third family” consists of the “social” or “pragmatic” theories discussed in section 9.5.

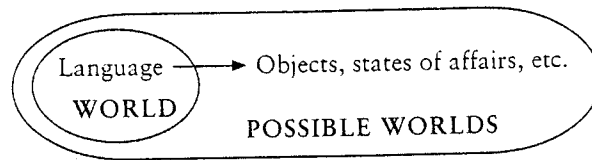


Fig. 10.2. Language in possible worlds semantics

This view of language is of course profoundly at odds with the outlook of generative grammar, which places language in the f-mind of language users. A few formal semanticists have worried about this inconsistency (e.g. Partee 1979; Bach 1986a; Zwarts and Verkuyl 1994), and some formal semantic approaches such as Discourse Representation Theory (Kamp and Reyle 1993) lend themselves better to mentalistic interpretation than others. But by and large the issue has been neglected.³

So let us address it. How is the realist view of language to be reconciled with the mentalist approach? One approach would be to jettison the mentalism of generative linguistics, but retain the formal mechanisms: to take the position that there is an objective “language out there in the world,” and that this is in fact what generative grammar is studying. But this disconnects generative linguistics from all sources of evidence based on processing, acquisition, genetics, and brain damage. Good riddance, many people would say. For example, Katz (1981) retreats (well, he would say “advances”) from his early mentalist position (Katz 1966) to the view that language is an abstract object, independent of the f-mind; he takes the study of language in the mind only to concern issues of performance.

But look what this forces us to give up. As Chapter 4 stressed, the fundamental motivation for positing Universal Grammar and for exploring its character comes directly from the observation that languages come to be instantiated in the f-mind by virtue of being learned by children. Without this mentalistic boundary condition, the subject matter of linguistics becomes limited to the mere description of languages, and the study of Universal Grammar becomes at best an exercise in statistical tendencies and/or formal elegance. Some people don’t mind such an outcome, and often when one is enmeshed in the gritty details of an endangered language it makes little practical difference. However, for reasons detailed in Part I, I think that abandoning the mentalist outlook gives up one of the major conceptual advances of our time.

An alternative position might be that of Frege (1892): language is indeed “out in the world” and it refers to “objects in the world”; but people *use* language by

³ Macnamara and Reyes (1994) is an important (though completely non-mainstream) exception, attempting to build a formal semantics based on psychological principles.

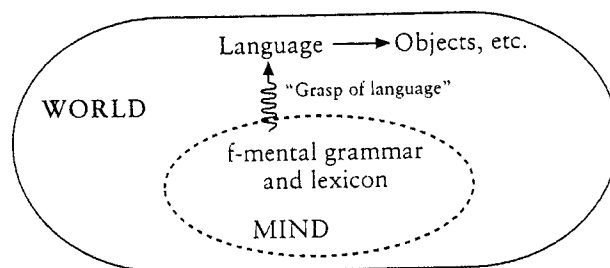


Fig. 10.3. The mind grasping "language in the world"

virtue of their grasp of it, where "grasp" is a transparent metaphor for "the mind holding/understanding/making contact with" something in the world. Fig. 10.3 might schematize such an approach.

Generative linguistics, it might then be said, is the study of what is in the mind when it grasps a language.⁴ This would make it possible to incorporate all the mentalistic methodology into linguistics while preserving a realist semantics.

One might interpret Katz's program this way. He is personally interested only in the part of language that is an abstract object "in the world"; but one could conceivably be concerned with the mental side of language as well. However, this approach faces a curious methodological problem. We can determine properties of "language in the world" only through its manifestations in human linguistic intuition and behavior. Thus we have no independent fix on what parts of human language are due to its mental instantiation and what parts are due to characteristics of the abstract object "Language." Katz, for one, is inclined to attribute all the logical properties of human language to the abstract object and to be noncommittal about the rest. Other decisions might be possible as well. The problem is that they are just that: decisions. There is no empirical way to determine how to divide up the pie.⁵

But there is a still more fundamental problem. What sense are we to make of

⁴ Notice how close this is to the standard phrase in generative grammar, "knowledge of language." This phrase carries with it the implication of an external entity, "language," that is known. Thus Chomsky's disclaimers that there is such an external language are subtly undermined by his choice of terminology.

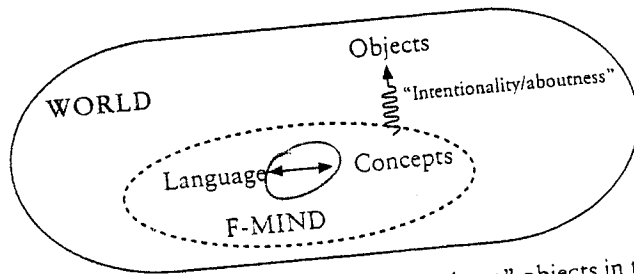
⁵ A similar objection applies to Rey (1996), who wants to "divorce the issue of definitions from the issue of anyone's ability to provide them," and proposes that "the correct definition of a concept is provided by the optimal account of it, which need not be known by the concept's competent users" (293). Here the term "concept" stands for an external abstract object; somehow the "optimal account" of it is supposed to come from science. This may appear apposite for words like *gold*, where science seeks the "true nature" of the substance—but it is surely misguided for such words as *puddle*, *groceries*, and *pie*. Moreover, it directs us away from what language users f-know, which is precisely what a mentalistic account cares about. What is the relation between their f-knowledge and the user-independent definition?

the notion of “grasping” an abstract object? We know in principle how the mind “grasps” concrete objects: by constructing cognitive structures in response to inputs from the senses. This process has a physical instantiation: the sense organs respond to impinging light, vibration, pressure, and so forth by emitting nerve impulses that enter the brain. But an abstract object by definition has no physical manifestations that can impinge on the nervous system. So how does the nervous system “grasp” them? Without a careful exegesis of the term—which no one provides—we are ineluctably led toward a quasi-mystical interpretation of “grasping,” a scientific dead end.

Common sense is also a bit strained when it comes to the provenance of abstract objects. If languages are abstract objects, was Nicaraguan Sign Language lying around in the abstract domain until the 1980s, when it was at last grasped by someone? Come to think of it, has modern English been lying around since the Big Bang, and do the abstract objects also include all the languages ever spoken on other planets, which we humans could not possibly grasp? These questions may seem hopelessly naive, but I don't recall ever seeing them addressed. It seems to me that they have only one sensible answer: the conceptualist view that abstract objects are human creations. But creations out of what? (Dennett's (1991) ironic term “figment” comes to mind.) It becomes clear that we have no way to understand “abstract objects” except through metaphor based on concrete objects (Lakoff and Johnson 1980); the metaphor remains apt only if we don't push it too hard.

Proponents of treating language as an abstract object (Katz included) respond that there *must* be a way for the mind to grasp abstract objects. After all, we do manage to grasp numbers and other mathematical objects, not to mention logical truths. These are surely eternal, abstract, and independent of humans. Two plus two would still equal four, five would still be a prime, and [p and q] would still entail p even if we weren't around to appreciate it. So, the proponents say, whatever mechanism allows us to grasp mathematical and logical truths will do for grasping abstract Language as well. If we can't figure out how to instantiate our grasp of arithmetic and logic in materialist terms, well, so much the worse for science.

This response does raise the stakes. Surely a theory of mind has to explain how we come to understand number and logic. But I am not sure a satisfactory theory is going to come from the assumption that we do so by making mystical contact with abstract objects. Rather, within conceptualist semantics, the problem turns inside out: we should be asking what it is about human beings that (a) permits them to deal with mathematical and logical concepts and (b) leads them to believe that these are true independent of human observers. At the moment I have no proposals to offer; Macnamara (1986) and Lakoff (1987) have interesting



center arrow
between words

Fig. 10.4. Concepts in the f-mind that are "about" objects in the world

discussions on logic, Gelman and Gallistel (1978), Wynn (1995), Makkai (1999), and Wiese (forthcoming) on number.

One way to eliminate the problem of how the mind grasps language is to push language entirely into the mind—as we have done here. We might then arrive at a semantic theory structured like Fig. 10.4.

This is Fodor's position (see section 9.4): for him, language is a mental faculty that accesses concepts (expressions in the Language of Thought). In turn, concepts have a semantics; they are connected to the world by virtue of being "intentional." The trouble, as already complained of in section 9.4, is that one cannot make naturalistic sense of intentionality. It suffers from precisely the same difficulty as "grasping" language in Figure 10.3: there is no physically realizable causal connection between concepts and objects.

In short, there appears to be no way to combine a realist semantics with a mentalist view of language, without invoking some sort of transcendental connection between the mind and the world. The key to a solution, I suggest, lies in examining the realist's notion of "objects in the world," to which we now turn.

10.3 Problems with the common-sense view: "objects"

Consider again the quotes at the beginning of the last section. To be sure, "objects," "states of affairs," and "things out in the world" have an entirely intuitive construal. "Objects" invites us to think of Bertrand Russell, Noam Chomsky, trees, refrigerators, apartment buildings, screwdrivers, and so on. These are things you can see and touch; you can count them, tell two of them apart, in some cases move them around. "States of affairs" invites us to think of snow being white, dogs being animals, Russell's having been a philosopher, the door to my office being open (at the moment of writing), and so on: observable facts about observable objects.

But we refer routinely to all sorts of "objects" that are not so simple to put our hands on. I am deliberately going to overwhelm you with examples, just to give a flavor of the scope of the problem.

- (1) Fictional and mythical characters
 - a. Sherlock Holmes
 - b. the unicorn in my dream last night

Sherlock Holmes exists only as a character in stories, yet one can say true and false things about him, for example it is true that he was English, false that he was Romanian, and—false that he existed! There are no unicorns in “the world,” but one can experience one in a dream “as if it were real”; and within this context, it is true that it did such-and-such and false that it did something else.

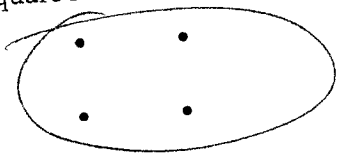
- (2) Geographical objects
 - a. Wyoming
 - b. the Mississippi River
 - c. the distance between New York and Boston

There is nothing tangible about Wyoming, no great geographical features that mark it off, no lines drawn across the landscape (unlike the Land of Oz, where things are all red in one region, yellow in another, and so forth). There may be billboards along the road that say “Welcome to Wyoming”, but these are not what make it Wyoming. It is a purely politically constructed entity, its rectilinear boundaries fixed by a stipulative act.

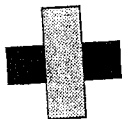
We can touch the Mississippi River, and swim in it. But is the river the water contained in it, the bed of the river, the complex of the two? Exactly where does it end in the Gulf of Mexico, and exactly where does its tributary, the Missouri, end in it? One can draw arbitrary lines on a map, but these are understood as matters of convenience and not some sort of “natural truth about the world.”

The distance between New York and Boston is not tangible. Nor is there an absolute truth about it: how should it be measured? From center to center, from nearest border to nearest border, along some particular highway, on a straight line through the earth’s crust? Much depends on one’s purpose.

- (3) Virtual objects
 - a. the square formed by the four dots below



- b. the horizontal rectangle that goes behind the vertical one below



*still not a square
It must form a square!*

The square and the horizontal rectangle simply aren't there physically. What does it mean to say they are "objects in the world"?

- (4) Social entities
- a. the value of my watch⁶
 - b. the first dollar I ever earned
 - c. Morris Halle's Ph.D. degree
 - d. your reputation
 - e. General Motors
 - f. the score of tomorrow's Red Sox game

These are all intangible entities (one hesitates to call them objects) that are the result of social practices and conventions. As Searle (1995) points out, they are nevertheless treated referentially, as parts of the "real world." I have chosen a wildly heterogeneous group of examples (with the future reference to "tomorrow's game" thrown in for the pleasure of extra complexity); doubtless the reader can come up with many more.

- (5) Auditorily perceived objects
- a. Mahler's Second Symphony
 - b. the words *banana* and *despite*

One cannot identify Mahler's Second as a particular performance, or as a score or a recording. It is in some sense the entity that lies behind all of these, of which these are all realizations. Those of a logical frame of mind might try to treat it as a type rather than a token. But this doesn't really solve the problem: a type of what? There is the type expressed by "performances of Mahler's Second," the type expressed by "scores of Mahler's Second," and the type expressed by "recordings of Mahler's Second," each of which has its own particular tokens; but it is mere sleight of hand to assert that Mahler's Second is, say, the union of these types, or, following Goodman (1968), "the set of performances in conformance with the score." Especially since Mahler's Second existed and was the same symphony before there were any recordings of it. In fact it presumably existed in Mahler's imagination before there were any performances and before the score was completed. And of course many musical traditions are not regulated by written scores at all. The problem only becomes more confusing when we consider variant editions and arrangements of Mahler's Second, where the notes and instrumentation are different. Yet we certainly

⁶ This is a favorite example of Chomsky's, which he uses to illustrate his own skepticism with the notion of "thing in the world," noted in section 9.4. The present section is in a sense an amplification of his point, as well as of a point stressed in Jackendoff (1983).

refer to *Mahler's Second* without any sense of complex circumlocution. It is as much an entity "in the world" as the Sistine Chapel. (Makkai 1999 makes a similar point with respect to poems.)

I leave it as an exercise for the reader to see that words like *banana* and *despite* present the same problems as symphonies. But this is of crucial importance to the theory of reference: exactly what are we *using* to refer to "entities in the world"?

- (6) Other
- a. the set of all possible worlds
 - b. the best of all possible worlds

I am totally at a loss as to what it could possibly mean to claim these are "in the world," particularly since "the world" is included in *them*. But we do refer to them.

Here is the point: The quotes above assert that we refer to "objects in the world" as if this is completely self-evident. It *is* self-evident, if we think only of reference to middle-sized perceivable physical objects like tables and refrigerators. But as soon as we explore the full range of entities to which we actually refer, "the world" suddenly begins to be populated with all sorts of curious beasts whose ontological status is far less clear. For each of the types of entity cited above, one can construct some elaborate story, and some of them have indeed evoked an extensive philosophical literature. But the effect in each case is to distance the notions of reference and "the world" from direct intuition. The cumulative effect of considering all of them together is a "world" in which direct intuition applies only to a very limited class of instances.

10.4 Pushing "the world" into the mind

To sum up so far: The common-sense position on reference, which standard approaches to semantics take as their starting point, suffers from two complementary problems. First, if language is in the minds of language users, it is necessary to invoke some mystical connection from the mind to the world, either at the level of language (Fig. 10.3) or at the level of the concepts the language expresses (Fig. 10.4). Second, the notion of "objects in the world" is itself suspect.

I propose to cut the Gordian knot by abandoning the unexamined notion of "objects in the world," and, for purposes of the theory of reference, pushing "the world" down into the mind of the language user too, right along with language. The change is in how we are supposed to understand statements about reference. (7) and (8) give the two opposing alternatives.

- (7) Common sense realist theory of reference:
Phrase P of language L, uttered in context C, refers to entity E in the world (or possible worlds).
- (8) Conceptualist theory of reference:
A speaker S of language L judges phrase P, uttered in context C, to refer to entity E in [the world as conceptualized by S].⁷

That is, in a conceptualist theory, reference is taken to be at its foundation dependent on a language user—just as relativistic physics takes distances and times to be dependent on an observer's inertial frame. And just as it is often convenient to revert to Newtonian dynamics when relativistic effects can be ignored, we can often take the (7) as a convenient abbreviation for (8), temporarily ignoring the "reference frame"—the presumption of agreement among members of a (relatively) homogeneous speech community.

As initial motivation for exploring the conceptualist position, observe that a language user cannot refer to an entity without having some conceptualization of it. Consider an example such as (9).

- (9) I don't know what *that* was, but here it comes again!

In order to utter (9) (and mean it), the speaker must have conceptualized some relevant entity, though certainly without a full characterization. That is, a referent's being in the world as conceptualized is a necessary condition for a speaker to refer. However, being in the *real* world is *not* a necessary condition: speakers can refer to entities like *Sherlock Holmes* and *the unicorn in my dream last night*. And an entity's being in the world is not sufficient for reference either: one has to conceptualize it in at least some minimal way. In short, an entity's being in the real world is neither a necessary nor a sufficient condition for a speaker's being able to refer to it. Rather, the crucial factor is having conceptualized an entity of the proper sort.⁸

Still, I would not blame the reader who is a bit suspicious of this expression "the world as conceptualized by the language user." It smacks of a certain solipsism or even deconstructionism, as though language users get to make up the

⁷ Ch. 2 has discussed the tropes needed in order to relativize the idealized term "speaker of language L" to "non-ideal speakers," and to explain the relative uniformity within a speech-community. Likewise, "judges" is a stand-in for a collection of tasks involving language use, including not only conscious judgment but also appropriate use in normal conversational conditions, evidence of understanding in normal conversational conditions, behavior in psychological experiments, and so on.

⁸ However, conceptualization is not a sufficient condition for reference, because one can conceptualize an entity without speaking of it at all, or even having a word for it. Moreover, there are concepts that do not support reference, for example the concept expressed by the word *if*.

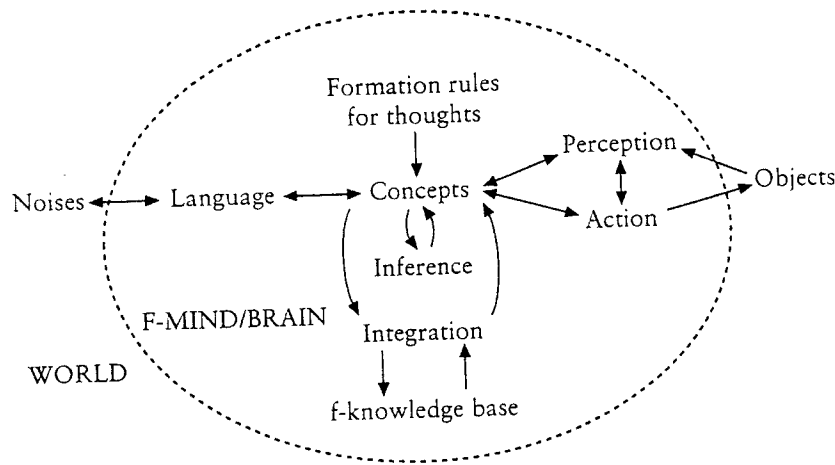


Fig. 10.5. The conceptualist view

world any way they want. Abbott (1997) compares this view to Berkeleyan idealism, which claims that one is referring to one's mental representations rather than to the things represented.

And indeed, there seems little choice. Fig. 10.5, which repeats Fig. 9.1, has no direct connection between the form of concepts and the outside world. On this picture our thoughts seem to be trapped in our own brains.

This outcome, needless to say, has come in for harsh criticism, for example:

Suppose that we adopted an approach . . . that studied meaning by relating symbols to mental representations or mental procedures of some sort, and stopped there. That would amount to limiting the domain of semantics to the relations between a language, which is a form of representation, and another representation . . ., translating our public language into an internal mental code, our "language of thought," say. But how can mapping a representation onto another representation explain what a representation *means* . . .? . . . [E]ven if our interaction with the world is always mediated by representation systems, understanding such systems will eventually involve considering what the systems are about, what they are representations of. (Chierchia and McConnell-Ginet 1990: 47)

. . . words can't have their meanings *just* because their users undertake to pursue some or other linguistic policies; or, indeed, just because of any purely *mental* phenomenon, anything that happens purely 'in your head'. For "John" to be John's name, there must be some sort of *real relation* between the name and its bearer . . . something has to happen *in the world*. (Fodor 1990: 98-9)

Semantic markers [i.e. elements of conceptual structures in the present sense—RJ] are *symbols*: items in the vocabulary of an artificial language we may call *Semantic*

Markerese [here, conceptual structure⁹]. Semantic interpretation by means of them amounts merely to a translation algorithm from the object language to the auxiliary language *Markerese*. But we can know the *Markerese* translation of an English sentence without knowing the first thing about the meaning of the English sentence: namely, the conditions under which it would be true.

Semantics with no treatment of truth conditions is not semantics. Translation into *Markerese* is at best a substitute for real semantics. . . .

The *Markerese* method is attractive in part just because it deals with nothing but symbols. . . . But it is just this pleasing finitude that prevents *Markerese* semantics from dealing with the relations between symbols and the world of non-symbols—that is, with genuinely semantic relations. (Lewis 1972: 169–70)

. . . [T]he brain's causal capacity to produce intentionality [i.e. the "aboutness" of concepts—R] cannot consist in its instantiating a computer program [i.e. in its being nothing but a system of cognitive structures], since for any program you like it is possible for something to instantiate that program and still not have any [intentional] mental states. Whatever it is that the brain does to produce intentionality, it cannot consist in instantiating a program since no program, by itself, is sufficient for intentionality. (Searle 1980: 424)

How is it possible to escape this barrage? Only by going deeper into psychology, and by dealing even more austere with the notion of thought. From the standpoint of neuropsychology, we must recognize that the neural assemblies responsible for storing and processing conceptual structures indeed *are* trapped in our brains. They have no direct access to the outside world. Hence, as stressed in Chapter 9, we must explicitly deny that conceptual structures are symbols or representations of anything in the world, that they *mean* anything. Rather, we want to say that they *are* meaning: they do exactly the things meaning is supposed to do, such as support inference and judgment. Language is meaningful, then, because it connects to conceptual structures. Such a statement is of course anathema to the authors quoted above, not to mention to common sense. Still, let's persist and see how far we can go with it.

10.5 A simple act of deictic reference

Consider about the simplest act of using language to refer to a "middle-sized object": a use of referential deixis such as (10).

(10) Hey, look at that! [*pointing*]

⁹ However, note that Lewis is making the category mistake criticized in section 9.4: he is identifying *Markerese* as a language "like" a natural language, neglecting the fact that *Markerese* is intended as a *component* of natural languages alongside syntax and phonology.

The deictic pronoun *that* has no intrinsic descriptive content; its semantics is purely referential.¹⁰ In order to understand (10), the hearer not only must process the sentence but must also determine what referent the speaker intends by *that*. This requires going out of the language faculty and making use of the visual system.

Within the visual system, the hearer must process the visual field and visually establish an individual in it that can serve as referent of *that*.¹¹ The retinal image alone cannot do the job. The retina is sensitive only to distinctions like "light of such-and-such color and intensity at such-and-such a location on retina" and "dark point in bright surround at such-and-such a location on retina." The retina's "ontology" contains no objects and no external location. Nor is the situation much better in the parts of the brain most directly fed by the retina: here we find things like local line and edge detectors in various orientations (Hubel and Wiesel 1968), all in retinotopic format—but still no objects, no external world. This is all the contact the brain has with the outside world; inboard from here it's all computation.

However this computation works, it eventually has to construct a cognitive structure that might be called a "percept." The principles and neural mechanisms that construct the percept are subjects of intensive research in psychology and neuroscience, and are far from understood. The outcome, however, has to be a cognitive/neural structure that distinguishes individuals in the perceived environment and that permits one to attend to one or another of them. One can stop attending to a perceived individual and then return to it; one can track a perceived individual as it moves through the perceived environment and as it changes properties such as orientation, color, and shape. The cognitive structure that gives rise to perceived individuals is non-linguistic: insofar as human infants and various animals can be shown experimentally to identify and track individuals more or less the way we do, the best hypothesis is that they have percepts more or less like ours.

Of course percepts are trapped inside the brain too. There is no magical direct route between the world and the percept—only the complex and indirect route via the retina and the lower visual areas. So all the arguments directed against

¹⁰ Well, it has a little descriptive content: it denotes something in a relatively distal position, by contrast with *this*.

¹¹ I am bypassing all the fascinating issues of how the hearer follows the pointing gesture to the intended referent. This is quite possibly a feat of which even chimpanzees are incapable (Povinelli et al. 2000), hence another of those cognitive prerequisites to language acquisition that had to evolve in the last 5 million years. Even if the hearer follows the pointing, there is the question of precisely what the speaker has in mind, whether the rabbit or the collection of the rabbit's legs or whatever (Quine 1960). How children manage to accomplish this in word learning is addressed by Macnamara (1982) and Bloom (2000), among many others.

conceptual semantics apply equally to percepts. This may bother some philosophers, but most psychologists and neuroscientists take a more practical approach: they see the visual system as creating a cognitive structure which constitutes part of the organism's understanding of reality, and which helps the organism act successfully in its environment. If there is any sense to the notion of "grasping" the world perceptually, this wildly complex computation is it; it is far from a simple unmediated operation.

And of course a visual percept is what will be linked to the deictic *that* in (9) and (10), through the interfaces between conceptual structure and the "upper end" of the visual system. Thus language has indeed made contact with the outside world—but through the complex mediation of the visual system rather than through some mysterious mind-world relation of intentionality. Everything is scientifically kosher.

A skeptic may still be left grumbling that something is missing: "We don't perceive our percepts in our heads, we perceive objects out in the world." Absolutely correct. However, as generations of research in visual perception have shown, the visual system populates "the world" with all sorts of "objects" that have no physical reality, for instance the square subtended by four dots and the "amodally completed" horizontal rectangle in (3). So we should properly think of "the perceptual world" (or "phenomenal world" in the sense of Koffka 1935) not as absolute reality but as the "reality" constructed by our perceptual systems in response to whatever is "really out there."

Naturally, the perceptual world is not totally out of synch with the "real world." The perceptual systems have evolved in order that organisms may act reliably in the real world. They are not concerned with a "true model of the world" in the logical sense, but with a "world model" good enough to support the planning of actions that in the long run lead to better propagation of the genes. Like other products of evolution, the perceptual systems are full of "cheap tricks," which is why we see virtual objects: these tricks *work* in the organism's normal environment. It is only in the context of the laboratory that their artificiality is detected.

Thus *the perceptual world is reality for us*. Apart from the sensory inputs, percepts are entirely "trapped in the brain"; they are nothing but formal structures instantiated in neurons. But the perceptual systems give us the sense, the feeling, the affect, of objects being "out there." We experience objects in the world, not percepts in our heads. That's the way we're built.¹²

¹² To put it in the terms of Dennett (1991), the perceptual system is a "syntactic engine that mimics a semantic engine," where "syntactic" and "semantic" are intended in Chomsky's and Fodor's sense: syntax = 'formal manipulation internal to the organism' and semantics = 'relating formal objects to the world'.

In speaking about how we experience the world, we are beginning to venture into the sacred realm of consciousness, a topic not for the faint of heart. So let us pause and take stock. We have been attempting to create a conceptualist account of the intuition that linguistic expressions refer to things in the world. The answer so far is that a deictic linguistic expression like *that* in (10) can be linked with a percept. In turn, a percept is an f-mental structure constructed by the perceptual systems in response to stimulation from the outside world. Although a percept does not necessarily correspond exactly to what is "actually out there" (especially with virtual objects), the experience that accompanies having a percept in one's f-mind is that of an object in the world. Hence we experience the deictic expression as referring to this object.

In short, the problem of reference for the intuitively clear cases is not at bottom a problem for linguistic theory, it is a problem for perceptual theory: how do the mechanisms of perception create for us the experience of a world "out there"?

Some readers will no doubt find this stance disquieting. My dear friend John Macnamara, with whom I agreed on so much, used to accuse me of not believing there is a real world. Similarly, Jerry Fodor, in response to similar conceptualist arguments by Chomsky (2000), says:

... in what way is the plausible claim that there are banks (that is, that there "really are" banks) not warranted? There is, for example, the bank that holds my mortgage. That is not just a way of talking; they make me pay up every month, cash on the barrel. How on earth could that be so if there really are not any banks at all? (Fodor 2000b: 4)

One reply would be that mortgages are just as much a mental construct as banks. But there is a deeper reply: we are ultimately concerned with *reality for us*, the world in which we lead our lives. Isn't that enough? (Or at least, isn't that enough for linguistics?) To paraphrase Dan Dennett, who subtitles his book *Elbow Room* (1984) as *The Varieties of Free Will Worth Wanting*, it's useful to ask what varieties of reality are "worth wanting." The reality in which you are reading this book and Jerry Fodor is paying his mortgage is certainly worth wanting; my claim here is that this world is a product of our human modes of perception and conception. If you want to go beyond that and demand a "more ultimate reality," independent of human cognition, well, you are welcome to, but that doesn't exactly render my enterprise here pointless.¹³

I think one could be perfectly justified in stopping at this point, satisfied that reference has been "reduced to a previously unsolved problem." But I am going

¹³ Smith 1996 is an attempt to build the metaphysics of "reality for us" from the ground up, one that dovetails attractively with the approach taken here.

to push on a bit, in the interests of uncovering further aspects of the theory of reference.

10.6 The functional correlates of consciousness

Let us confine ourselves here to simple perceptually mediated consciousness, for example the experience of seeing a refrigerator, tasting an apple, or feeling an itch. For present purposes we fortunately can ignore more loaded issues such as consciousness of self and the sense of free will, which, while of paramount importance in the overall conception of human nature, play a lesser role in the theory of reference.

Francis Crick and Christof Koch (1990; Crick 1994) make an admirable case for searching for the “neural correlates of consciousness,” in particular the correlates of visual awareness. Complementary to their search is an inquiry into what might be called the “functional correlates of consciousness,” the structures and processes in the f-mind that give rise to visual awareness (or *accompany* visual awareness, depending on one’s overall theory of consciousness). Works such as Dennett (1991) and Jackendoff (1987), among many others, outline theories of the functional correlates of consciousness. The hope, of course, is that research at the neural level and at the functional level will converge. Here, as in our treatment of language, we will be thinking functionally, without by any means excluding neural approaches.

Suppose (10) (*Hey, look at that!*) is uttered in response to a particularly large and disgusting bug scuttling across the floor. In the course of processing the sentence, the hearer’s visual system must construct a percept that on one hand results in (or corresponds to) the experience of the bug and on the other hand is bound or linked to the deictic pronoun *that*. What features must be present in this percept?¹⁴

The kinds of features that visual psychologists concentrate on are obvious candidates. The percept has a shape, a size, a color, and can be decomposed into shaped and joined parts (a body, eyes, lots of legs). It also has a location, motion, and a “character of motion” (the way the legs move and the body twists in the course of the bug’s moving). Let us call all these features, however they come to be characterized theoretically, the *descriptive* features of the percept. It is these features that the speaker of (9) (*I don’t know what that was!*) cannot reconstruct or report.

¹⁴ I use the term “features” as a shorthand for “distinctions the system must make,” without prejudging whether these distinctions are in binary, digital, or analogue dimensions, and without commitment as to appropriate formal notation.

More basically, the bug-percept is a visual percept. The hearer is not hearing, tasting, smelling, or (we hope) feeling the bug. Thus percepts must be distinguished by modality. Thinking functionally, this could simply be a consequence of the fact that the descriptive features are those appropriate to the visual modules. Similarly, in neural terms, this could be a consequence of where in the brain the percept develops (this is essentially the position being explored by Crick and Koch).

Another basic characteristic of the percept is that it constitutes a figure distinguished from the background on which it is located and it moves. Suppose that I utter (10), and you stare at the patterned rug and see nothing; then suddenly the bug "pops out" in experience (*Oh god, there it is! Eeuw!*). Nothing has changed in "the world" or the retinal image. All that has changed is the organization of percepts "trapped" in the hearer's brain: the bug-percept has emerged as figural. Let us call this figural characteristic the *indexical* feature of the percept. It gives the f-mind a "something" to which descriptive features can be bound. The speaker of (9) has an indexical feature lacking identifiable descriptive features.

The f-mind establishes indexical features in response to perceptual input. But once established, they need not go away in the absence of perceptual input: we intuitively sense that objects continue to exist when we're not seeing them. When an object disappears behind an obstacle and reappears on the other side, is it the same object? To what extent do the descriptive features have to remain constant in order to make this judgment? If an object disappears behind an obstacle and two objects reappear on the other side, which (if any) is the original object? These questions of "tracking" involve the behavior of the indexical feature.¹⁵

Indexical features can undergo "splitting," for instance when we break a lump of clay in half to form two distinct individuals. They can also undergo "merger," as when we mold two lumps of clay into an undifferentiated mass. Or suppose we spot a red circle on one occasion and a blue circle on a second occasion; then on a third occasion we spot what we take to be the red circle again, and there before our very eyes it turns blue. We are perhaps now inclined to think there was only one circle all along: again an indexical merger.

Suppose our bug starts crawling up your leg. A whole new set of descriptive features in a different modality develops in your f-mind. How are you to experience them as a manifestation of the same entity in the world? I suggest we

¹⁵ Baillargeon (1986) and Spelke et al. (1994) pose similar questions to babies. Their results, on the present interpretation, show that babies track indexical features (the identity and number of individuals) but they are far less sensitive than we would expect at tracking descriptive features (exactly what those individuals look like).

think of the descriptive features as being linked to a common indexical feature. Such linking is an act of f-mental construction, not just something to take for granted: think of the (illusory) experience of hearing figures on a movie screen speaking, on the basis of noises emerging from loudspeakers at another location.

Beyond the descriptive features, the modality, and the indexical feature, one other type of feature needs to be considered. My graphic description of hearing (10) in response to a bug may have evoked in you a *visual image* of a bug. And when I spoke of the bug crawling up your leg, maybe you felt one too—you had a *tactile image*. Such experiences arise from cognitive structures in your f-mind; let's call these cognitive structures "functional images" or "f-images." So now I want to ask: How do f-images differ from percepts?

None of the features discussed so far is appropriate to make the difference. Consider indexical features. The experienced bug-image is individuated and stands out as a figure from the (imaged) ground, so the f-image has an indexical feature. What about modality? One can distinguish visual from tactile in images just as well as in perceived objects, so this isn't the difference we want. What about descriptive features? One might be first tempted to appeal to location: the experience corresponding to a percept is "out in the world" and that corresponding to an f-image is "in the head." But I suspect that the tactilely imaged bug crawling up your leg was experienced on your leg, not in your head. So images *may* be in one's head, but need not be.

One might alternatively claim that images are fuzzier and more fleeting than objects, reflecting a difference in their functional structure in the f-mind. But compare for example hearing music way off in the distance, which is indeed fuzzy and fleeting, with hearing in your head some music you know very well, right down to the rasp of John Lennon's voice, for instance. In this case relative clarity goes in the wrong direction: the "real" music is fuzzier than the image. More generally, the range of descriptive distinctions available in any modality of imagery is essentially the same as that in the corresponding modality of perception.

In the old days one might have tried to make the distinction by claiming that there is an "imaging" part of the brain separate from the "perception" part. But we don't believe this anymore (Kosslyn 1980; 1996; Jackendoff 1987): we think that imagery uses most of the same brain areas—as well as the same functional features—as perception. So this way out is not available either.

The distinction I think we need comes from a family of features proposed in Jackendoff (1987: ch. 15, 1997a: ch. 8). These features register not the perceptual qualities of an entity but the associated "feel," so to speak. I called these the "affects" or "valuation" of the percept. One of these features is the distinction

external versus *internal*. Associating *external* with a cognitive structure results in its being experienced as “out in the world,” *internal* results in its being experienced as an image.

Some other features in this family are *familiar* vs. *novel* (which distinguishes not only familiar from novel percepts but also memories from acts of imagination), *self-produced* vs. *non-self-produced* (which distinguishes voluntary images from hallucinations), *meaningful* vs. *non-meaningful* (or *coherent* vs. *incoherent*), and *mattering (positively or negatively)* vs. *non-mattering* (which registers emotional effects such as an object's being delightful or disgusting).¹⁶ These features apply equally to all modalities: visual, auditory, and tactile experiences can all be experienced as real or not, familiar or not, self-produced or not, meaningful or not, and important or not.

registers

Valuations, like size, shape, and color, are subject to illusions. For instance, one can think one saw or heard something real but have just imagined it. Déjà vu is a sense of familiarity associated with a situation known to be novel. Schizophrenics find many more things meaningful than they ought to; perhaps autistics find things less meaningful than they ought to. Dreams feel external and non-self-produced but are of course internal and self-produced. People with phobias attach illusory negative value to all manner of inappropriate entities in the environment. Hence valuations too are constructions of the f-mind.¹⁷

Returning to our bug, it is not simply experienced “out there” automatically: it is “out there” because the percept that gives rise to the experience contains the valuations *external* and *non-self-produced*. Visual processing normally settles on these valuations in response to retinal input: “seeing is believing.” However, it also assigns this valuation to the virtual square in (3) despite there being nothing directly corresponding on the retina. By contrast, the imaged bug has the valuations *internal* and *self-produced*.

To sum up, the character of a consciously experienced entity is functionally determined by a cognitive structure that contains the following feature types:

- An indexical feature to which descriptive features can be attached.
- One or more modalities in which descriptive features are present.
- The actual descriptive features in the available modalities.
- A valuation that registers the status of the entity in a number of modality-independent dimensions.

¹⁶ Damasio's (1994) notion of “somatic marker” might correspond to this feature.

¹⁷ Section 12.4 will show how valuation features are appropriate for the characterization of “intensional” contexts such as the contents of someone else's beliefs.