THE CASE FOR CASE REOPENED

CHARLES J. FILLMORE
University of California at Berkeley

Several years ago I wrote a fairly long article (Fillmore 1968) with the punning title, "The Case for Case." In it I presented a notion I referred to as deep structure cases, together with a number of proposals on how the notion could be incorporated into a generative grammar. Since that time, I have become aware—through the many published and unpublished reactions to the piece that have reached my desk—not only of some of my own mistakes, but also of a number of ways in which readers have misunderstood my intentions or have acquired expectations for the theory which went beyond anything I had in mind.

Here is what I am going to try to do in this chapter: I will locate the concept of deep cases within traditions of semantic and grammatical inquiry; I will review a few of the basic assumptions of case theory and will point out what I originally thought of as its advantages; I will discuss some of the most important challenges to the theory, including the serious one of how one can know what the cases are and how many of them there are; and I will fail to offer a satisfying solution to this problem. I will instead propose a new interpretation of the role of cases in a theory of grammar and a new method of investigating the question of their number and identity. I consider this new interpretation as a position in the theory of grammatical relations and as a position in semantic theory with which one could associate this slogan: meanings are relativized to scenes.

To illustrate my points, I will mainly use the same examples that I have used in earlier writings, including, of course, many that I borrowed from other
authors. This is not because I can’t think of any new ones, but because I wish to show with examples that might be familiar some of the ways in which the new interpretation is different.

Within semantics in general, the notion of deep cases is a part of what might be called internal, as opposed to external, semantics; that is, it concerns, not the semantics of truth or entailment or illocutionary force, but rather the semantic nature of the inner structure of a clause. Within internal semantics, the concern is syntagmatic rather than paradigmatic; that is, deep cases are among the types of semantic relations that elements of sentence structures have with each other in context, rather than with the system of contrasts and oppositions that differentiate constituents paradigmatically. The concern is with the inner structure of clauses rather than with the semantics of interclausal connections through the devices of coordination and subordination.

Within grammatical theory the concept of deep cases can be thought of as a contribution to the theory of grammatical levels, to the theory of grammatical relations, to the description of valences and collocations, and to the general theory of the functions of sentence constituents. The proposal is that there is a level of structural organization for a sentence which is distinct from what is usually thought of as a semantic representation and which is distinct likewise from the familiar notions of deep and surface structure syntactic representations. The theory is concerned with the determination of the nuclear grammatical relations in a sentence—subject, object, and indirect object—in that it addresses itself to the question of how particular aspects of the meaning of an utterance determine which constituent will appear as the (deep structure) subject, which as the object. The theory of cases can also be seen as offering at least part of the semantic valence descriptions of verbs and adjectives, comparable to the syntactic valence descriptions given by certain European linguists (Tesnière 1959, Helbig 1971, Helbig and Schenkel 1969, Emons 1974). And, lastly, it may add something to the theory of the functions of sentence constituents. Katz (1972:113) has distinguished three kinds of sentence-constituent functions: the grammatical, illustrated by the notions of ‘subject’ and ‘object’; the rhetorical, illustrated by such oppositions as ‘given’ versus ‘new’, ‘topic’ versus ‘Comment’, etc.; and the semantic, illustrated by such notions as ‘agent’, ‘recipient’, ‘means’, ‘result’, etc. In his discussion of these functions, Katz claims that Chomsky’s failure in the Extended Standard Theory (Chomsky 1970) is one of confusing the grammatical and the rhetorical functions, and that my failure in “The Case for Case” is that of confusing the grammatical and the semantic functions. There is, I believe, a fourth way of looking at the functional structure of the parts of sentences, something for which such words as orientation and perspective suggest what I have in mind. The parts of a
message can be divided into those that are "in perspective" and those that are "out of perspective". My current position is that it is the orientational or perspectival structuring of a message which provides the subject matter for the theory of cases, and that the case notion figures very differently in grammatical description from what I originally had in mind.

One essential feature of the theory of deep cases is the case frame (Fillmore 1968:27), the function of which is to provide a bridge between descriptions of situations and underlying syntactic representations. It accomplishes this by assigning semantico-syntactic roles to particular participants in the (real or imagined) situation represented by the sentence. This assignment determines or constrains the assignment of a perspective on the situation by means of what I have called subject selection principles and the case hierarchy.

Some of the Subject Selection principles seem to be language-universal. Thus, given certain qualifications for the interpretation of ergative systems, one candidate for a universal Subject Selection Principle is this: If there is an agent which is brought into perspective, the nominal which represents it must be the (deep) subject.

Other Subject Selection principles appear to be language-specific. Japanese, according to Kuno (1973:31), and German, according to Rohdenburg (1970), do not allow certain kinds of enabling or occasioning causes to be chosen as subjects of their sentences, whereas English does tolerate such choices, as in sentences like those in (1):

(1)  
    a. Fifty dollars will buy you a second-hand car.  
    b. The smell sickened me.  
    c. The accident killed the woman.

Still other Subject Selection principles appear to be word-specific. This appears to be true, for example, of (at least) one member of the pair regard and strike, as in Chomsky's examples (1965:162) given here in (2):

(2)  
    a. I regard John as pompous.  
    b. John strikes me as pompous.

Lastly, but not surprisingly, Subject Selection principles are capable of changing in time. Jespersen (1924:160) speaks of the changes in the history of English from expressions of the type seen in (3a) and (3b) to expressions of the type seen in (3c) and (3d) as involving changes in the meanings of these verbs:

(3)  
    a. Me dreamed a strange dream.  
    b. Me like oysters.  
    c. I dreamed a strange dream.  
    d. I like oysters.
There is, of course, a way of construing the word meaning so that Jespersen’s description can seem sensible; but one could more revealingly say that what changed were the workings of the Subject Selection principles operating in the language, these changes having to do with the interaction of ordering principles and case-marking principles in a period after which most surface case distinctions on nouns and pronouns were lost.

The advantages that I saw in the notion of deep-structure case were that case-structure descriptions of words and sentences offered a level of linguistic organization at which universal properties of lexical structure and clause organization were to be found, and, moreover, that such descriptions were in some sense intuitively relatable to the ways people thought about the experiences and events that they were able to express in the sentences of their language.

3

It was misleading for me to use the phrase case grammar to describe the proposals I made in “The Case for Case.” My proposals did not cohere into a model of grammar. Instead, they were suggestions about a level of organization of a clause that was relevant to both its meaning and its grammatical structure; that provided a way of describing certain aspects of lexical structure; and that offered convenient classifications of clause types. From the fact that I have frequently received mail asking me questions like, “How does case grammar handle intonation?” I have become aware that my writings somehow gave the impression that case grammar so-called was being presented as a general model of linguistic structure. Nowadays I try to be more careful about the phrase ‘case grammar’.

Criticism of the deep case proposals have come from many places and represent many points of view. Some of them are based on misunderstandings (possibly invited misunderstandings); some of them are valid but require “patching” rather than deep changes in my position; and a number of them are very much worth worrying about.

When I wrote “The Case for Case” the Generative Semantics position had not yet been formulated, or, rather, had not yet been formulated in my hearing except in its preliminary version as Abstract Syntax. In the section of my article which touched on surface case morphology, I described the traditional way of treating cases, which consisted in identifying the cases through their forms and then, one by one, describing their functions within larger constructions. What I suggested was that things should be turned around, so that the organization of the sentence as a whole should be taken as the framework within which the functions of the individual grammatical morphemes could be stated. As a catchword for this position I spoke of the CENTRALITY OF SYNTAX.
Walter Cook, in his article “A Set of Postulates for Case Grammar” (1972), claims to be taking a position in opposition to mine in a postulate which assigns centrality to semantics, not syntax. I suspect that if I knew precisely what it meant, Cook’s is a position I might take, too. But, in any event, it was not an opposition between semantics-at-the-bottom and syntax-at-the-bottom that I had in mind, but rather between analysis that begins with the morpheme and analysis that begins with the sentence.

I have used various forms of arguments in trying to justify specific claims about deep case distinctions. One of these argument forms has two steps. In the first step we recognize an ambiguous sentence whose ambiguity can be accounted for only by assuming that one of its nominals is interpretable as bearing either of two semantic roles in the sentence. As an illustration of this point, we can consider the ambiguous sentence (4a). To see that this sentence is ambiguous, imagine a situation in which you see that I have a letter on my desk, that you see me make a copy of it, and that then you hear me say sentence (4b). If you speak the same version of English that I do, you can’t know for sure, from my instructions, which of the two letters I have in mind.

(4)  
   a. I copied the letter.  
   b. Point to the letter which I copied.  
   c. I copied this from that.

The second step in this argument form consists in exhibiting the same verb in a sentence where it takes two different nominals, each with just one of the two semantic roles sensed in the earlier ambiguous sentence. This property is found in our verb copy, as can be seen in sentence (4c), in which exactly the two previously sensed roles noticed in (4a) are parcelled out among two separate nominals. Having both types of evidence—as provided in the two steps of this argument—makes it believable that we are dealing with two different case roles rather than with simply a matter of vagueness.

In a second kind of argument, also with two steps, I have suggested that we may be dealing with different case relationships whenever we find a single verb collocating with two seemingly disparate classes of nominals in a given grammatical relation, with the nominals from the two classes seeming to exemplify different semantic roles in their sentences. Examples for this point are the possibilities for the subject-nominal in sentences (5a) and (5b). And, again, the argument can be taken as completed if, as step two, we can find a single sentence with the same two roles parcelled out between two separate nominals, as in (5c):

(5)  
   a. My foot hurts.  
   b. This shoe hurts.  
   c. This shoe hurts my foot.
These kinds of arguments have seemed particularly weak to a number of European critics, who have suggested that the method of analysis is based on accidental properties of English words. Other languages, this particular counter-argument goes, might use different words for the different senses of hurt and copy, and so, for purely nonsystematic reasons, the method would yield different results for these other languages.

This objection, too, I take as a misunderstanding. These arguments are not offered as definitions of cases, but rather as steps for pointing out case distinctions in sentence sets where sense differences are only, or are primarily, matched by case role differences. I have been just as interested in similarities in underlying case structure in sentences with different verbs and different relational organization. In a third kind of argument that I have used, different surface verbs are taken from single vocabulary fields—pairs like rob and steal or buy and sell—verbs which have matching (or partly matching) case structures but different assignments of grammatical relations. I feel sure, in other words, that an analyst looking at, say, Swedish, would be able to discover, in expressions of physical pain, roles involving such things as the source of the pain, the location of the pain, the experiencer of the pain, and so on; and this could be done independently of whether the same verb or different verbs were used, independently of whether the same or different choices were made in the selection and orientation of the cases.

A criticism that has come from some European workers in Valence Theory (Emons 1974 and Panevová 1974) has been that I have obscured the various senses in which it is possible to speak of optional constituents in a sentence. Sometimes some aspect of an event or situation is a part of the speaker’s and the hearer’s understanding of the meaning of a sentence, yet there is nothing in the sentence which expresses it; and sometimes the absence of a constituent in the surface sentence reflects the absence of the associated notion from the conceptualization that is being communicated. It is true that there have been some places where I was wrong in detail in the description of given sentences in just this regard; but my intention was that all of the various senses of optionality could be accounted for by the fact that the system I proposed had CASE FRAMES, indicating the case notions conceptually present in a sentence; CASE FRAME FEATURES, indicating the case notions that could be combined in construction with a given lexical item; and DELETION TRANSFORMATIONS, by which, under various conditions, a given constituent could be, or maybe had to be, absent from the surface structure. I think that in principle my proposals take care of optionality sufficiently well.

Several specific suggestions that I made in “The Case for Case” seem to have been bad ones, and I have been quick to abandon them. I proposed, for example, that all nominals in English sentences are initially provided with
prepositions; this position required that the processes of forming subjects and direct objects had to involve preposition deletion. My motivation was partly that of wanting to make English and Japanese seem more comparable than they might have otherwise. Japanese has POSTPOSITIONS marking all nominal functions in a sentence, INCLUDING THE NUCLEAR GRAMMATICAL RELATIONS (subject and object). I thought that the basic typological differences between the two languages could be captured as follows: In deep structure, English has the verb at the beginning. Japanese, the verb at the end, of each sentence; the language with the verb at the beginning has prepositions on all the nouns; the language with the verb at the end has postpositions on all the nouns. Apart from these differences—which a mirror would show are really similarities—the two languages differ in that the verb-initial language—English—has a Subject-fronting rule, by which the verb ends up in second position, and it has preposition-deletion processes for the subject and direct object. These two language types, then, can be compared with a third type, exemplified by Tagalog, which lacks Subject Fronting and which keeps prepositions in front of all its associated nominals.

As part of an informal general typological sketch, this way of looking at things is not bad, I think; but the specific proposals I had for the assignment of the individual lexical preposition made the system look fairly inefficient. In particular, as I tried to construct a grammar using these proposals, my initial association of the preposition by with the agent case had to yield in favor of a more complicated principle associating by with the highest-ranking case in the sentence, whatever that case might have been. This decision made it possible to account correctly for the by in eaten by George, destroyed by fire, and assumed by everybody (as long as it was constrained to allow known to me, etc.). In the end, however, such provisions did not look any better than an account according to which the preposition by gets introduced by means of a Passive transformation.

In “The Case for Case” I described the agent and the dative cases as necessarily animate. This, by a curious kind of bad logic, led some people, for example me, to assume that obligatory animacy in a verbal complement implied that the associated nominal had to be in one of these two cases. By that kind of reasoning, the subjects of die and melt (to use examples from Huddleston’s 1970 critique) were dative and patient (object), respectively, and the two sentences in (6) would have to be given different case structure analyses:

(6)  
   a. The man died.  
   b. The snow melted.

I am now more careful about keeping relational notions and categorial notions distinct.
Some workers, for example Nilsen (1972), have provided binary opposition analyses of the basic cases, analyzing agent and instrument as animate and inanimate causes, respectively; and experiencer (dative) and patient (object) as animate and inanimate effects, respectively. The initial appeal of this analysis fades when one realizes that it is a confusion of categorial and relational notions, but begins to reappear when one tries to see it as an account of the basic properties of the prototype scenes in terms of which actions and experiences can be structured.

In an attack on case grammar recently published in Germany, Peter Finke (1974) concentrated on what he took to be the categorial implications of cases, and interpreted case theory as a version of many-sorted logic. To say that a verb takes an agent and a patient is to say, Finke would claim, that the two arguments of the verb have to meet different sortal specifications, and that in general cases are to be defined in terms of sortal specifications. Agents are things having one set of properties, instruments are things having some other set of properties, and so on.

Finke’s main point, if I understand him correctly, is that a case grammar makes an ontological commitment about the number of sorts of objects that can exist in a universe, that number equal to the number of cases. If such a strong claim is implied by the theory, according to Finke, then the fact that case grammarians have not been able to come up with a common and stable collection of cases shows that the theory is not to be taken seriously. The reply is, of course, that even if some universe contained only one sort of object—say, human beings—the role-identifying function of the cases could still be maintained. One person could pick up another person, use that person’s body for knocking down a third person, that third person could feel embarrassment, and so on. In a universe with only one sort of object, in short, the case relations of agent, instrument, patient, and experiencer could all be easily imagined. Perhaps it was my misleading statements about animacy that created the impression that case notions were to be understood as categories rather than as types of relations.

A challenge to case grammar that I have recently heard comes from a writer (Raible 1975) who was concerned with how case theory could account for the very complicated set of case functions in a language like Finnish. The author of this challenge had taken my suggestion that when looking at a system of cases we should distinguish case FORMS from case USES, and that we should build a theory of deep cases whose repertory of case concepts was more or less equivalent to the repertory of case uses. Finnish presents special problems, and in two ways. First, because of the system of locative and directional cases which combine, in one and the same surface case category, both locational notions and assumptions about the topology of the reference object (the entity
in terms of which something is being located or with respect to which something is being oriented); and, second, because there is a semantic difference, in what is thought of as the direct object, between choosing the nominative, the accusative, or the partitive case. The difference between choosing the accusative rather than the partitive, for example, appears to involve such notions as definiteness, completion, and totality. My answer has to be, of course, that the morphemes in surface case systems encode more than nominal functions, and that therefore an account of the uses of surface-structure cases requires more than a theory of deep cases.

4

There are four more kinds of criticisms of case theory that I wish to mention, two that I do not know how or whether to worry about, and two that seem clearly very serious indeed. The first of the puzzling ones is the collection of criticisms that take the form, "I can do anything you can do"—the argument of MERE NOTATIONAL VARIANCE. One version of the argument goes like this: I have claimed that in the structure underlying expressions containing a verb like English seem there is what I call an EXPERIENCER, the person for whom something seems; and I have claimed that this sometimes is manifested on the surface as a to-phrase, as in (7):

(7)

a. To me, Harry seems intelligent.

b. It seems to me that Harry is intelligent.

This, the argument goes, is pointless. There is no particular advantage to giving theoretical prominence to terms like experiencer, agent, instrument, and the rest. We could just as well state that seem is a MENTAL STATE VERB, meaning that it has a semantic marker 'mental'; and we can then say that selection restrictions for verbs containing this marker require one of their coconstituents to be 'animate' nominal; we can say of the verb seem that it has a selection restriction requiring its associated animate argument to be part of a preposition phrase marked by to; and we could even provide a principle which declared that, for any argument of a mental verb which is obligatorily animate, a semantic marker 'experiencer' can be added to the representation of that nominal in the finished semantic reading of the sentence.

This argument (found in Chomsky 1970, Katz 1972, and Mellema 1974) has something to say for it, of course. It is a version of the view that a change in a scientific paradigm should be tolerated only when the possibilities of the existing paradigm have been exhausted AND when a more satisfactory paradigm can be shown to exist. I sympathize with this view in general; my position, however, is that alternative paradigms—or even alternative notations—should be valued for the kinds of questions they force an analyst to ask. A strong
assumption about deep structure cases forces the analyst to ask certain questions about the number and variety of the semantic functions of the parts of sentences. Whether the answers to these questions can be written down within the terms of some other paradigm does not always seem important to somebody primarily concerned with whether the questions are important ones and whether they can be answered at all.

Next there are the trivializing objections to case grammar leveled by Ray C. Dougherty (1974). Dougherty’s main point is that case grammar so-called offers at best a possibly interesting but theoretically unimportant classification of verbs. Verbs, in this system, are classified on the basis of intuitive judgments about their complements. This—Dougherty goes on—might seem harmless enough, possibly comparable to classifying words by the number of letters, as in the New York Times Crossword Puzzle Dictionary; but it is not harmless, because Fillmore has deceitfully superimposed on this mere taxonomy the trappings of a generative grammar.

If the mere taxonomy charge is valid, then I should probably repeat that the deep case proposal was not intended as a complete model of grammar, but only as a set of arguments in favor of the recognition of a level of case structure organization of sentences. If the mere notational variant evidence is valid, this should not be taken as a reason for dismissing case theory, but rather as a challenge for considering such questions as which of the many possible notational variants of the final correct theory offers the greatest convenience for making generalizations about language typology, about lexical classification, about child language acquisition, and so on. In short, I see the notational and taxonomic possibilities of case theory as standing or falling together. A taxonomy is to be valued if it provides a convenient and revealing conceptual organization of the entities in its realm, in our case something in terms of which grammatical and semantic generalizations can be easily formulated; a notation is to be valued if it allows the formation of such a taxonomy in a simple and straightforward way.

The two sorts of objections I have just been discussing have been that case theory is at best a mere notational variant of some preferred theory, or is at best a mere taxonomy. The critics who have raised these issues have other objections, too, several of them relating to points discussed elsewhere in this paper.

But now let me turn to objections that require action. Stephen Anderson, in a paper called “On the Role of Deep Structure in Semantic Interpretation” (1971), challenged my claim that it would be possible to do without a level of deep structure in the sense of the Standard Theory. The position I had taken was that grammatical theory needs a level of case structure, which I thought of as being, or as being close to, a level of semantic representation; this position recognizes, by means of the grammar’s transformational rules, the levels of
surface structure. But, for reasons like those given by Halliday (1967:39), there was thought to be no need for a separate level of deep structure.

I would say, for example, that in the case of the English verb *break* used transitively, its two arguments could have the case functions agent and patient, the agent being the entity responsible for the breaking, the patient being the entity which broke. In the surface structure, one of these arguments has to become the subject of the sentence. One possibility is for the agent to become the subject, as in (8a); another, which has side effects on the form of the verb, is for the patient to become the subject, as in (8b):

(8)  
a. *John broke the vase.*  
b. *The vase was broken by John.*

I presented these as optional choices provided by a Subject Selection rule. The important point to notice is that, in the history of the formation of a passive sentence, no level of representation was posited according to which the sentence had a subject that was different from the surface subject. It had a case structure and a surface structure, but no intermediate structure within which the notions of subject and object could be defined.

Anderson, defending the Standard Theory, argued on grounds of the simplicity of the semantic component for the existence of deep-structure subjects and objects. His argument took the form of a claim that there are certain semantic generalizations which can be formulated simply only at the standard deep-structure level. There are two situations covered by Anderson’s arguments. The first is that in which certain transformations, like Passive and Dative Movement, are viewable as changing an original assignment of deep-structure subject or object; the second is that in which a given predicate allows more than one way for its arguments to be parcelled out as terms of grammatical relations. In either case, Anderson’s argument is that frequently there is a difference between a holistic and a partitive interpretation given to a noun phrase, depending on whether or not it has one of the primary grammatical relations (subject or object) at the deep-structure level. To illustrate Anderson’s point from familiar examples, we can say that the difference between (9a) and (9b) is that the sentence with *the garden* in subject position conveys the idea of the whole garden containing bees, whereas in the other no such assumption is necessary; and that the difference between (9c) and (9d) is that the sentence with *the truck* in direct-object position conveys the idea of the truck being entirely filled with hay, an assumption not necessary with the other sentence:

(9)  
a. *Bees were swarming in the garden.*  
b. *The garden was swarming with bees.*  
c. *I loaded hay onto the truck.*  
d. *I loaded the truck with hay.*
New examples suggested by Mellema (1974) are the sentences in (10):

(10)  
a. He read from his speech.
b. He read his speech.

Having in mind the reading-aloud sense of read, the sentence with his speech in direct-object position gives the impression that the entire speech was read, while the other sentence definitely does not.

Without feeling required to concede that these arguments support the version of deep structure that Anderson had in mind, I believe I must concede that a level of representation including the grammatical relations subject and object—or something else which recognizes the kind of close participation in the nuclear part of a clause that we find with subjects and objects—is probably necessary for grammatical theory. This does not mean, however, that a level of representation which recognizes case functions is spurious. We still have, after all, the problem of determining, from an understanding of what is being said, which of the arguments of a multiplace predicate are to appear as its subject, and which, if any, as its object.

The next truly worrisome criticism of case theory is the observation that nobody working within the various versions of grammars with ‘cases’ has come up with a principled way of defining the cases, or principled procedures for determining how many cases there are, or for determining when you are faced with two cases that happen to have something in common as opposed to one case that has two variants. In his careful paper ‘Some Thoughts on Agentivity’ (1973), D. A. Cruse examined a number of linguists’ statements about agentivity, found not only that different linguists disagree on the definition of agentivity, but also that the notions that enter into these definitions do not appear to be comparable enough to lend themselves to a uniform definition of a coherent concept.

I have thought of the problem of determining the cases as being somewhat analogous to that of determining the phonological units of language. Since pie and buy are different words in English, we have good reason to believe that we have a /p/ phoneme that is distinct from a /b/ phoneme. In spy, however, we have a problem. Should we analyze its stopped segment as /p/ or as /b/, or as something distinct from either of these? If it is not a /p/ or a /b/, should it be seen as something which in some sense includes both the /p/ and the /b/ phonemes, something which is exactly neutral between the /p/ and the /b/, or something conceptually distinct from those two phonemes altogether—something taking part in a different phonological system? The data do not dictate the correct answers to these questions. American phonemicists tend to regard the segment as /p/, but Danish phonologists, facing a similar problem in their language, choose /b/. Firthian linguists would tend to regard it as a systematically different item entirely; phonologists who allow certain kinds of abstrac-
tions might regard it as a category that includes both /p/ and /b/, and, if they make special assumptions about the feature composition of phonemes, they would recognize it as a segment containing features shared by both /p/ and /b/ but lacking any of the features which distinguish those two segments from each other.

An analogous array of positions could be taken on the problem of the identity of cases. Huddleston (1970) examines my arguments that in (11a) John functions as agent and that in (11b) this key functions as instrument. Here we are dealing with what might be called ‘indirect cause’ and ‘direct cause’, respectively. A reason for feeling sure that the two roles are distinct is that the same two nouns, preserving their case roles, can also occur together, with open, in a single sentence, as in (11c). The problem, Huddleston suggests, is what to decide about the causal element in (11d). Deciding on the case role of the subject in (11d) is like deciding on the “emic” status of the stopped consonant in spy.

(11) a. John opened the door.
b. This key opened the door.
c. John opened the door with this key.
d. The wind opened the door.

One possibility is that the wind, like John in the earlier sentence, is using its own energy rather than energy provided by something or someone else, and that therefore it is the agent. A second way of looking at it is that, since the wind is the direct cause of the door’s opening, it should be seen as the instrument. A third view is that the wind in (11d) has a role that is distinct from both agent and instrument—a call it force—since we have neither manipulator nor manipulated, but only a self-sufficient force. Another view, of course, is that there is a case feature cause that is one of the features of both agent and the instrument case; in the subject noun phrase for our sentence about the wind opening the door, the cause feature is present, but the features that distinguish an agent from an instrument are absent.

Since there are all of these possibilities to draw from—each sanctioned by one or another respectable linguistic tradition—there is little wonder that scholars who have tried to make the case theory work have not been able to agree. Furthermore, since linguists differ in the kinds of sentences they concentrate on and in the various positions that can be taken on the degree of closeness between case structure and semantic structure, and since not every linguist will think of the full range of possible alternatives, scholars working with caselike notions have all come up with different lists. The shortest list is John Anderson’s (1971): ‘nominative’, ‘ergative,’ and ‘locative’. The longest, I think, is that of William Martin (1972), who has separate case labels for the nouns found in the phrases at the station, on the table, in the box, from the station, off the table, out of the box, to the station, onto the table, and into the box. The principle seems to be that there are at least as many deep-structure cases as there are surface-structure cases or types of prepositional phrases.
Halliday (1967) recognizes a small number of different kinds of clauses, and assigns separate kinds of caselike structures to each of them. The method of L. Stephen Coles (1972), by contrast, is to divide verbs into a number of types (he finds sixteen), specifying a framework of caselike concepts for each type. With verbs of construction, for example, he distinguishes the resulting constructed object, such as the house that got built; the material out of which the thing got built; the people who performed the constructing act; the implements that were used; the time of the event; the location of the event; the time span from the beginning of the task to the end of the task; and so on. In Coles's case, if I understand his purposes, there is no concern with whether the resulting object case with a verb of construction is or is not the same as some particular case in the conceptual framework posited for some other verb class.

5

I believe now that there might be a solution. It involves what I said earlier about meanings being relativized to scenes. One way for me to lead into this question is to point out that, as I have conceived them, the repertory of cases is not identical to the full set of notions that would be needed to make an analysis of any state or event. One of the cases in the system I proposed was the agent, identifying the role of an active participant in some event; yet events are not restricted in the number of active participants they can have. For example, in what I shall be calling the commercial event, two different individuals are agentively involved, and the actions of each of the two are part of our understanding of any of the lexical items that can be used for describing this kind of event or any of its aspects. The point I want to make here is that a case frame need not comprise a description of all of the relevant aspects of a situation, but only a particular piece or section of a situation.

One of the individuals in the commercial event—the buyer—hands over some money and takes the goods; the other—the seller—surrenders the goods and takes the money. A complete description of the commercial event would identify the buyer, the seller, the money, and the goods. A prototypic commercial event involves all of these things, but any single clause that we construct in talking about such an event requires us to choose one particular perspective on the event. In the usage that I suggest, any verb identifying any particular aspect of the commercial event will constrain us to bring one or more of the entities in the event into perspective, the manifestation of this choice for English being the selection of grammatical functions corresponding to the notions of underlying subject and direct object. For example, if I wish to take the perspec-

1 A really careful analysis, in fact, would require that reference to the money be made from two points of view, the cash and the value. This would make it possible to distinguish:

i. I bought it with a two-dollar bill.
ii. I bought it for two dollars.
tive of the seller and the goods, I will use the verb sell. Should I wish to take the perspective of the buyer and the money, I will use the verb spend. If I wish to bring into perspective either the buyer and the money or the buyer and the seller, I will use the verb pay. Should I wish to take the perspective of the goods and the money, I will use the verb cost. And so on.

In each of these cases, the speaker is required to construct a sentence in which one of the two or three entities that have been put into perspective becomes the subject; and, maybe, one becomes the direct object. The new question for the theory of cases is this: What do we need to know about the various participant roles in a situation in order to know which of these roles or which combinations of them can be put into perspective, and then, for those which have been put into perspective, which is to become the subject and which is to become the direct object?

The connection with the notion of 'scenes' can be stated this way. The study of semantics is the study of the cognitive scenes that are created or activated by utterances. Whenever a speaker uses any of the verbs related to the commercial event, for example, the entire scene of the commercial event is brought into play—is "activated"—but the particular word chosen imposes on this scene a particular perspective. Thus, anyone who hears and understands either of the sentences in (12) has in mind a scene involving all of the necessary aspects of the commercial event, but in which only certain parts of the event have been identified and included in perspective. The buyer and the goods are mentioned in (12a), the buyer and the money in (12b). In each case, information about the other elements of the scene could have been included—via nonnuclear elements of the sentence, as in (12c) and (12d):

(12)  
\begin{itemize}
  \item a. I bought a dozen roses.
  \item b. I paid Harry five dollars.
  \item c. I bought a dozen roses from Harry for five dollars.
  \item d. I paid Harry five dollars for a dozen roses.
\end{itemize}

One typical way of dealing with conceptually obligatory but superficially optional elements in a sentence is to claim that these elements are present in the deep structure but deleted or given zero representation on the surface structure. [Examples (12a) and (12b) represent different conditions for such deletion operations.] Within the view that meanings are relativized to scenes, however, it may not be necessary to believe that everything that is included in our understanding of a sentence is necessarily a part of the underlying grammatical structure of that sentence; it seems preferable to say that a word like buy or pay activates the scene of the commercial event; that everybody who understands the word knows what are the various components and aspects of such an event; and that a speaker's linguistic knowledge of the verb includes that knowledge of the grammatical ways in which the various parts of the event can be realized in the form of the utterance.
Any particular verb or other predicking word assumes, in each use, a given perspective. The grammatical functions of the nominals that represent the entities that are put into perspective are determined in part by something like a deep case hierarchy. Other parts of the associated scene can be introduced with prepositional phrases, with adverbials of various kinds, and with subordinate clauses. The "circumstantial" constituents of a sentence need not be aspects of scenes that are specifically required by a particular type of situation. Since any event takes place in time, any event sentence can contain a time adverbial; since many kinds of events take place in specific locations, sentences representing such events can contain locative adverbials; and so on.

Now when I say that meanings are relativized to scenes, what I mean is that we choose and understand expressions by having or activating in our minds scenes or images or memories of experiences within which the word or expression has a naming or describing or classifying function. For example, if we have occasion to say that a person has spent a certain amount of time on land, we know that this expression is chosen from a contrast set for which the opposing term is at sea, and that the terms from this set require a background scene involving in some way a sea voyage. Similarly, we know that if we describe somebody as being spry we have in mind a background scene of an age set for which the degree of activity and vigor we have in mind is relatively rare. The point is that, whenever we pick up a word or phrase, we automatically drag along with it the larger context or framework in terms of which the word or phrase we have chosen has an interpretation. It is as if descriptions of the meanings of elements must identify simultaneously 'figure' and 'ground'.

To say it again, whenever we understand a linguistic expression of whatever sort, we have simultaneously a background scene and a perspective on that scene. Thus, in our examples about buying and selling, the choice of any particular expression from the repertory of expressions that activate the commercial event scene brings to mind the whole scene—the whole commercial event situation—but presents in the foreground—in perspective—only a particular aspect or section of that scene.

Languages, and lexical items, differ in interesting ways in the options they present in taking particular perspectives on complex scenes. Consider, for illustrating this point, a scene of a person taking something and causing that thing to come into abrupt contact with something else.

The English verb hit allows either of two perspectives on such a scene. One is that of the actor and the manipulated object; the other is that of the actor and the affected object. Illustrating the first of these we have a sentence like
(13a); illustrating the second we have (13b):

(13)  
   a. I hit the stick against the fence.  
   b. I hit the fence with the stick.  
   c. I hit the stick.  
   d. I hit the fence.

The event referred to in (13a) and (13b) may be the same, but the perspectives are different.

I will refer to the elements that are brought into perspective—the elements that appear as subjects and direct objects—as **nuclear elements** in the sentence. The first thing to notice about nuclear elements, so defined, is that they are not identical with the obligatory elements of the sentence. In (13a), for example, the prepositional phrase is obligatory, yet it is not a part of the nucleus. Put differently, (13d) can be taken as an in-context abbreviation of (13b), but (13c) cannot be interpreted as an abbreviation of (13a).

A question that must be asked is under what conditions something can be brought into perspective. In the case of a hitting act in which the two objects that come into contact are both inanimate and neither is described as being interestingly affected by the event, it is difficult to imagine any particular salience difference between the two, and hence it is not easy to see the difference between the two perspectival choices. However, just in case one of the two entities brought together in an act of hitting is in some sense inherently more worthy of inclusion in perspective, inherently more salient, the force of the perspective choice begins to become noticeable. One possibility for enhancing the saliency of the affected entity is to make it be, not a fence, but a person. We will find, I believe, that it is in some sense more natural to say (14a) than (14b):

(14)  
   a. I hit Harry with the stick.  
   b. I hit the stick against Harry.

This must be because it is more natural to include human beings within perspective than to leave them out in favor of inanimate objects. It is not that there is not use for sentence (14b), but rather that in that sentence the decision to leave Harry out of perspective has the effect of treating him as a physical object rather than as a sentient being affected by the hitting act.

For scenes involving mechanical actions of the kind just considered, it is fairly easy to set up a conceptual framework which has parts that match caselike notions in a fairly straightforward way. Thus, the thing which gets manipulated is the patient, the thing on which the manipulated thing acts is the goal, and the manipulator is the agent. What we can say about the verb *hit* and a number of other verbs within the same semantic field is that certain perspectives are
available: agent and patient, as in (14a), with agent as subject and patient as direct object; agent and goal, as in (14b), with agent as subject and goal as direct object; and, with the manipulator left out of perspective, patient and goal, with patient as subject and goal as direct object, as in (15):

(15)  
The stick hit the fence.

In sentence (15), the decision to leave the manipulator out of perspective has the effect of taking a smaller perspective on a larger event; one situation justifying this choice would be that in which the stick was thrown into space, so that its coming into contact with the fence could be seen as a separately individuated event.

The verb beat, by contrast, is associated only with a scene in which the agent is holding onto the manipulated object throughout the action. Thus we can say (16a–d), given the special interpretation required for (16c); but we cannot use sentence (16e) for the same scene:

(16)  
a. I beat the stick against the wall.
b. I beat the wall with the stick.
c. I beat the stick against Harry.
d. I beat Harry with the stick.
e. The stick beat Harry.

Since beat necessarily involves an event in the history of the agent, there is no way of leaving the agent out of the perspective.

The verb knock offers a few special problems. Conceptually an act of knocking requires more than one participant, but only the agent need be put into perspective. Thus when we say (17) we have the agent’s knocking action primarily in mind:

(17)  
He knocked on the door with his fist.

Thus we have seen how it is that the notion of perspective can be called on to explain (more honestly, to express) the subtle semantic differences, both in the grammatical organization of clauses having the same verb, and among sentences containing different verbs (hit, beat, and knock) within a single semantic domain.

One of the saliency conditions favoring inclusion in perspective, I have suggested, is humanness. Another is change of state or change of location. When an agent moves a patient against a goal, and as a result the goal participant moves or changes, the element in the goal case has acquired the saliency sufficient for it to be included in the perspective. We saw in (17) that with knock on the agent needs to be included in the perspective; but if in an act of
knocking against a door, the door falls flat, we express this with the two-word verb *knock down* and include the door as direct object, as in (18):

(18) \[ He \text{ knocked the door down.} \]

Consider now the verb *push*, and suppose that I say (19a). Here we have a two-participant scene with a one-place perspective, meaning, if I am right, that it is the agent's action itself that is central. But now suppose that, as a result of my pushing against the table, the table moves. I can now say (19b), with the table in perspective, i.e. with *the table* in direct-object position.

(19) \begin{align*}
\text{a. } & I \text{ pushed against the table.} \\
\text{b. } & I \text{ pushed the table.}
\end{align*}

The fact that some change occurred with the table is what made it natural to include it within the perspective.\(^2\)

Suppose, now, that I manipulate one thing, bring it into contact with another thing, and, as a result, that second thing moves or undergoes a change. The thing which changes, as we have seen, gets included in the perspective, leaving the manipulated object as outside the nuclear system of grammatical relations. It has to be marked, if it is mentioned, with the preposition *with*. The notion that I referred to in my earlier writings as the instrument can thus be seen as a derived notion, involving the relationship between an entity in one event and the event which is caused by that first event. (This interpretation of one part of the notion instrument has been made particularly forcefully in Talmy 1972.) Thus, if I hit a hammer against a vase and the vase breaks, I express that as (20a); if, however, I strike a hammer against a vase and the *hammer* breaks, I express that as (20b):

(20) \begin{align*}
\text{a. } & I \text{ broke the vase with the hammer.} \\
\text{b. } & I \text{ broke the hammer on the vase.}
\end{align*}

\(^2\) One possible interpretation of these phenomena, brought to my attention by students of János Zsiltka when I recently spoke in Budapest about direct object functions, is that instead of speaking here of “choosing” to put some nominal in direct object function by virtue of noticing an enhanced salience, we should speak of a lexicalization process of transitive-ization or some other sort of valence change. Hungarian has many instances of sentence relationships of the sort I have been considering, in which the lexical difference is in the presence or absence of a perfectivizing prefix *meg*- (Zsiltka 1967). I think that the observations about perspective that I have been making in this paper are correctly seen as involving the functions of nuclear grammatical relations, but I agree that in a linguistic description these notions belong more appropriately to an account of aspect and lexicalization patterns than to a description of the reasons speakers have for assigning a grammatical organization to their sentence. I have taken this more “processual” view more as an easy way of talking about the phenomena than as a claim about their place in the operation of a grammar.
In each of the two-event scenes, the first part involves one thing moving toward another; the second part differs from one scene to another. The scene in which the hammer breaks requires the hammer to be realized as the direct object, the vase appearing in a construction with a goal preposition; the scene in which the vase breaks requires the vase to be the direct object the hammer appearing in a construction with the preposition with.

According to this interpretation, the relationship between a change-of-state verb and the entity which undergoes this change of state is reflected, not in the underlying case structure (at least not in the underlying case structure of the first part of the two-event scene), but in the grammatical relation DIRECT OBJECT. Having this in mind, we can now shed some light on the semantic difference between the sentences of (21):

(21)  
a. I cut my foot on a rock.  
b. I cut my foot with a rock.

In the sentence with with the foot has the goal relation to the action, and the rock is treated as the thing which acted against the foot; in the sentence with on the foot has the patient relation to the action, and the rock is seen as the thing against which the foot moved. The thing which underwent the change of state—in each case, the foot—is expressed uniformly as the direct object, independently of its case role in the underlying action scene.

My new way of treating these observations—using both caselike notions and grammatical relations—requires, for verbs of the kind we have been examining, no separate independent instrument case, but instead a process of setting aside by means of the preposition with (in some cases, of) any patient noun which, by virtue of there being another noun phrase with higher salience, does not become part of the sentence nucleus. A result of this change is that there is no need for embarrassment about the fact that the same preposition has both instrumental and noninstrumental uses. With the sentences of (21), and with sentences like (22):

(22)  
I filled the glass with water.

the generalization for the selection of with is uniform: The entity marked by with is the entity which moved with respect to something else but was left out of perspective by virtue of the higher salience of its competitor, the goal entity.

This last point enables us to add to our list of saliency criteria. I have been proposing that the perspective taken in a clause is determined by some sort of hierarchy of importance, which might be called the SALIENCY HIERARCHY. Humanness and change were the two conditions on this hierarchy that we have considered so far. I believe that definiteness and totality can be added to the list. With this decision we may be able to resolve some of the old problems with
smearing mud and loading hay. In sentences (23a) and (23b) we have the idea that we are dealing with an event that more or less completely affects the truck or the wall—in particular, that the truck was filled or that the wall was covered as a result of the action. In the sentences that are typically paired with them, (23c) and (23d), no such assumption is necessary:

(23)  
  a. I loaded the truck with hay.  
  b. I smeared the wall with mud.  
  c. I loaded hay onto the truck.  
  d. I smeared mud on the wall.

The condition seems to be that if, as a result of bringing something into contact with something else, that something else is affected in some complete way, that new status of the goal entity is sufficient for its inclusion in the clause’s perspective. Some verbs allow a choice of perspective—for example, most of those we have looked at so far. Others, as we are reminded in Mellema (1974), have fixed perspectives. Cover, for example, requires that the goal be taken as direct object; and put, by contrast, requires that the patient be taken as direct object. These facts are illustrated in the sentences of (24):

(24)  
  a. I covered the table with a quilt.  
  b. *I covered a quilt over the table.  
  c. I put a quilt on the table.  
  d. *I put the table with a quilt.

Lexically specific requirements of this sort must be included in the technical description of the words; just as with beat, as noted above, the continued relation between the one who does the beating with the implement he uses for the beating must be specified.

What I have tried to do in the last part of this paper is to suggest that some of the problems of case analysis discussed in the earlier parts—in particular the vexing questions of the number and variety of cases—may be drawn a little closer to resolution if we separate from each other two things: one, the role analysis of the participants in a situation, of the kind for which proposals about deep cases have been made; and, two, the conditions under which a speaker can choose to draw certain situation participants into perspective. Since this perspectivizing corresponds, in English, to determining the structuring of a clause in terms of the nuclear grammatical relations, a consequence of this proposal is that it becomes necessary to recognize a level of grammatical structure which makes use of underlying grammatical relations, something I once
thought it important to reject. I believe, however, that a level of case or role analysis is also needed, as a part of a general analysis of the scenes that get communicated with speech; and I believe that what unites these two sorts of structure is the notion of perspective that I have been trying to develop here. We recognize scenes or situations and the functions of various participants in these scenes and situations. We foreground or bring into perspective some possibly quite small portion of such a scene. Of the elements which are foregrounded, one of them gets assigned the subject role—in underlying or logical structure—and one of them—if we are foregrounding two things—gets assigned the direct object role in the clause. Something like a Saliency Hierarchy determines what gets foregrounded, and something like a Case Hierarchy determines how the foregrounded nominals are assigned grammatical functions.

REFERENCES


