THE ATOMIZATION OF MEANING

DWIGHT BOLINGER
Harvard University

History repeats itself, with variations. Two decades ago American structuralists were trying, with indifferent success, to apply to morphology the same analytical techniques that had proved successful in the analysis of sound. For a number of reasons—including the lack of a suitable theory of meaning—the attempt made no headway at a time when phonology was still scoring advances with help from both acoustics and information theory. Morphemics still remains, in current texts on linguistics, a kind of relic of the 1940's. Now we witness a revived attempt from a different direction, but with essentially the same desire: to try out in a new field the techniques that have been developed in an older one. The new field is meaning, the old one is syntax, and the techniques are those of generative grammar. For the moment, morphemics is only slightly involved, but the signs are clear.

The phonological notion that was carried over to morphology was that of the *eme*. The syntactic notion that is being carried over into semantics is that of the *marker*. Syntactic markers are a refinement of the categories of traditional grammar—Noun Phrase, Adjective, Adverb, Determiner, etc.—and are of proved usefulness, as emes were and still are in phonology.

Back of the emes of phonology lay three millennia of strivings to hammer out a system of writing. No linguist would have been apt to 'discover' the phoneme in the year 1930 if he had lacked an alphabet. Here was a body of practical experience where the main thing needed was a linguist's close and analytical attention. On the side of syntax there was also—though rather less diffused in society than alphabetic writing—a large body of traditional grammar with wide agreement on larger matters, from which a scientific grammar could draw. One of the strengths of generative grammar is expressed in the half-truth that it contains 'nothing really new'; it confirms what we felt all along.

The value of a tradition rests in the centuries of winnowing, sifting, selecting, and rejecting on a vast proving ground where practical value is the test of survival, and many hands have performed many experiments of which the ultimate scientific triumph is only the last. Phonology and syntax inherited a limited set of contrasts, an apparatus that was not overburdened with entities. Does semantics inherit anything? If it does—or if not, and the whole thing has to be contrived—are the items limited, or can they be limited, to a manageable number? To answer these questions we must look at the role played by semantic markers in a particular theory, and at their possible hereditary relation, if any, to semantic work that has gone before.

The obvious choice for such an inquiry is the study by Jerrold Katz and Jerry Fodor, 'The structure of a semantic theory',[1] which has been deservedly influential in promoting the notion of the semantic marker. For K-F, 'Semantic

[1] *Lg.* 39.170–210 (1963). Hereafter diminished, in length but not in credit, to K-F.

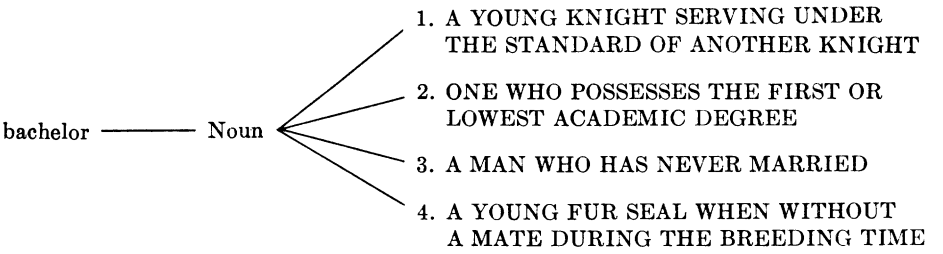

FIG. 1 (185).

markers are the elements in terms of which semantic relations are expressed in a theory' (187), the relational 'atomic concepts' that 'exhibit the semantic structure IN a dictionary entry and the semantic relations BETWEEN dictionary entries' (186). While this puts the case in exclusively decoding terms, and the entire argument is based on how a theory is able to account for the clearing up of ambiguity ('disambiguation' is the term used for this), there are obviously ramifications in other directions, particularly toward the act of encoding. For the moment, what is important is that K-F's statement is concerned exclusively about what a dictionary does or ought to do to enable the rules of the theory—the 'projection rules'—to disambiguate a semantically ambiguous sentence.

Mention of the dictionary suggests that K-F may indeed have in mind a tradition, which is that of lexicography. While they do not claim in so many words that their procedures in setting up definitions are just a refinement of what dictionary-makers have been doing all along, there are some obvious resemblances. We can trace them best by first giving an illustration or two from K-F's treatment.[2]

As an example of one common form of dictionary entry, K-F cite *bachelor*, and diagram it as in Fig. 1. This form, they point out, will not do for a semantic theory, for the characterizations are unsystematic. It must be revised to expose the markers, as shown in Fig. 2. The semantic markers are in parentheses. Brackets enclose the second type of atomic concept, the 'distinguisher', which is the idiosyncratic remainder of a given sense when all the markers have been stripped away. Each sense of the word is exhaustively characterized by tracing the path from grammatical marker (noun) through semantic markers to distinguisher.

[2] This is probably the right place to point out that K-F have tried to protect their flank by disavowing any attempt to BUILD a semantic theory. They undertake no more than to sketch an APPROACH to one. If their examples are matched with convincing counterexamples, they will not be abashed, but will look for better ones or modify some detail of the theory. This of course assumes that the foundations remain intact: it is hard to make alterations in a house whose frame has to be torn out. But the truth of the matter is that without putting together at least a fair sampling of the structure, one can never know whether the theory will account for more than the particular lexical sets from which the examples are drawn. If from a handful of specimens the theory appears as likely to fail as to succeed, there is no defense in promising to patch it up. In this sense, one cannot excuse a taste because it is only a taste. Until more is presented, everything stands or falls by it, and a theory is as subject to criticism in the light of its examples as a principal is liable to damages for the actions of his agent.

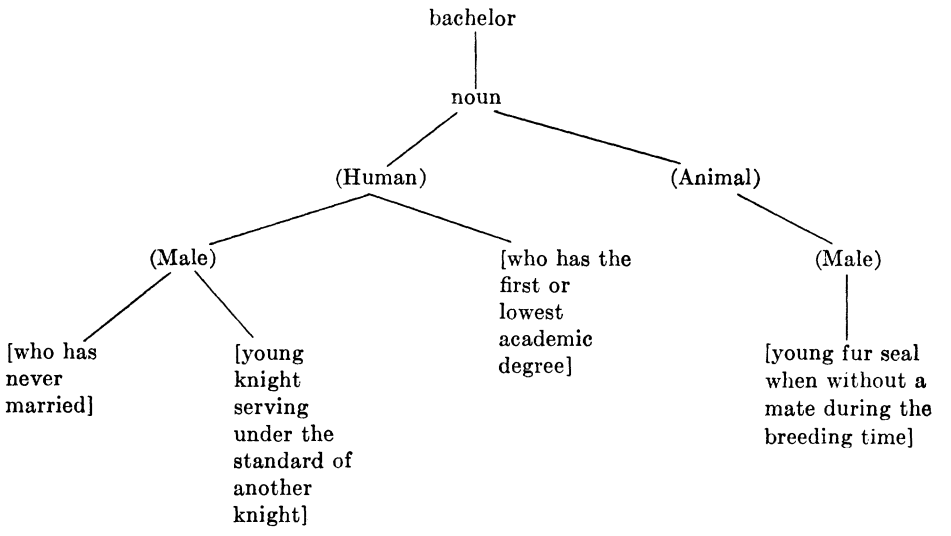

FIG. 2 (186).

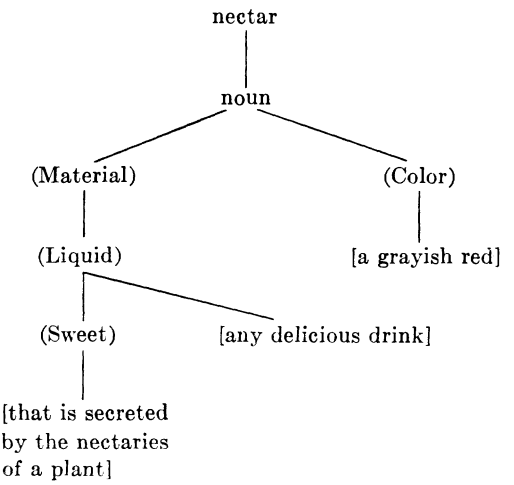

FIG. 3.

The plausibility of this design is of course no accident, for something like it is regularly used in dictionary definitions, even though it did not happen to be used in the definition of *bachelor*. The Merriam-Webster *NID3* defines one sense of *nectar* as 'a sweet liquid that is secreted by the nectaries of a plant'. Presumably this would be traced, along with the other meanings, as in Fig. 3. The markers (Liquid) and (Sweet) are supplied by the dictionary, and serve, respectively, to set off the 'secretion' sense from that of 'a grayish red' and 'any delicious drink'. Probably (Liquid) is too specific (though perhaps needed elsewhere, and so I retain it); I therefore add a superior marker (Material) at the point where the color separates from the other meanings. I am not sure that the markers I have chosen would suit K-F, but I think that the illustration is good

enough to show that the form of the entry is essentially one of markers and distinguishers.

Having settled that there is a practice in lexicography resembling the semantic analysis proposed by K-F, we can now go on to some of the problems raised by attempting to turn that practice into a theory that will cover the field. There will be time as we go along to ask whether the practice in question is the only one, or even the typical one, to be found in dictionaries, which has theoretical interest.

The problem of dualism: I. Markers and distinguishers

The dictionary draws no line between markers and distinguishers. This dualism is created by K-F to make it possible to have 'just enough' markers to do the job. A given theory may have more or less, and hence be more or less sensitive to the disambiguations that a fluent speaker is able to make. It is not enough to define *bachelor*, on one of its paths, as noun—(Human)—(Male), if we expect to account for the speaker's ability to discriminate the sense of *The old bachelor finally died* as 'unmarried man' rather than as 'young knight serving under the standard of another knight'. If we require that degree of sensitivity, then we must elevate (Young) to the status of a marker, for markers are, by definition, the semantic atoms through which disambiguations are effected. How far the theory should go is governed by

> a strategy which seeks to maximize systematic economy: the greatest possible conceptual economy with the greatest possible explanatory and descriptive power. If such decisions are optimally made, there should eventually come a point where increasing the complexity of a semantic theory by adding new markers no longer yields enough advantage in precision or scope to warrant the increase. At that point, the system of semantic markers should reflect exactly the systematic features of the semantic structure of the language (190).

K-F illustrate the process of extracting a marker from the distinguisher by rediagramming to show the new status of (Young), as in Fig. 4.

While K-F admit the possibility of theories of varying degrees of sensitivity, they also, in the passage just quoted, posit an ideal theory whose 'system of semantic markers should reflect exactly the systematic features of the semantic structure of the language'. It would seem that such a theory would need to handle all the disambiguations that the fluent speaker makes on the analogy of K-F's example (Young). The question then arises whether the distinguisher will not keep receding toward the horizon until it vanishes altogether. Let us see if this in fact does happen when we gather more disambiguations for *bachelor*. I shall give the disambiguating sentences, discuss them, and later rediagram.

1. *He became a bachelor.* This rules out the 'man who has never married'—it is impossible to become one who has never done something. We can extract the *-ever* part of *never* from the distinguisher and set up a marker (Nonbecoming).

2. *The seven-year-old bachelor sat on the rock.* The definition 'male who has never married' was deficient. It should have been something like 'adult male who has never married,' and from that expanded distinguisher we now extract the marker (Adult).

3. *Lancelot was the unhappiest of all the bachelors after his wife died.* This seems

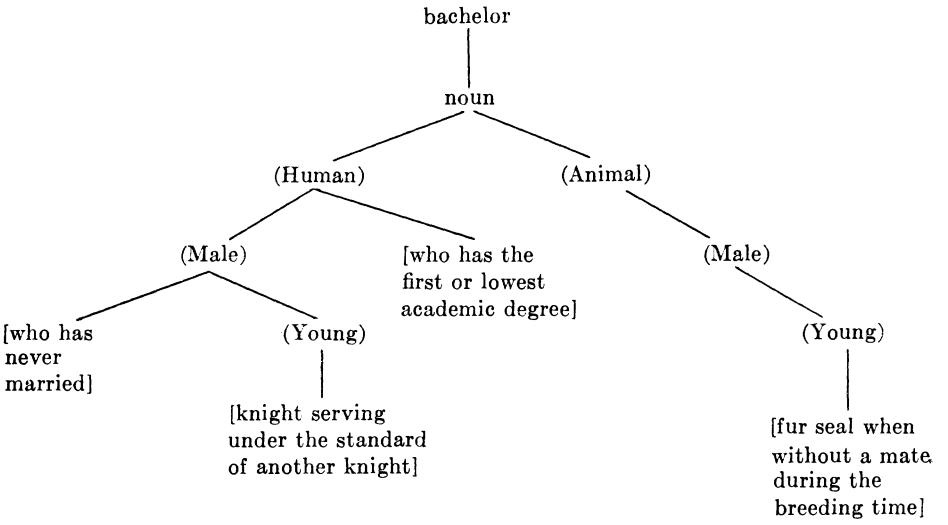

FIG. 4 (190).

to justify raising (Unmarried) to marker status and wipes out the distinguisher on one of the branches: *bachelor*—noun—(Human)—(Male)—(Adult)—(Non-becoming)—(Unmarried).

4. *That peasant is a happy bachelor.* Being a peasant is not compatible with being a knight. There must be a marker of status lying around somewhere. A knight has to be of gentle birth. Let us extract (Noble) from the distinguisher (leaving the degree of nobility for the moment undisturbed as still part of the knight's distinguisher).

5. *George is one bachelor who is his own boss.* This eliminates the knight, and turns 'serving under' into another status marker that might be called (Dependent).

6. *George is a bachelor in the service of the Queen.* This again eliminates the knight, and yields a marker that shows the direction of the (Dependency) relationship: it is to the person on the next higher rung of the (Nobility) ladder. I suggest (Proximate), dominated by (Dependent).[3]

7. *Knight banneret Gawain is a bachelor.* This eliminates the lower status of the knight bachelor, and admits (Inferior) as a restriction on (Noble).

8. *At some time in his life every man is a bachelor.* This eliminates both the knight and the B.A., because status has no bearing. We can therefore add a generalized status marker to those two other meanings. I call it (Hierarchic).

9. *A bachelor is expected to fight.* This puts the hierarchy in its proper setting, with a superior marker (Military). There is now no distinguisher left on the

[3] The problem of marker dominance is not treated by K-F. Presumably there are set sequences of markers, e.g. (Human) vs. (Animal) must precede (Male) vs. (Female) which in turn must precede (Old) vs. (Young), etc. Otherwise it will be impossible to generalize markers to the point where one like (Inferior), for instance, will serve both for an inferior grade of nobility (when dominated by Noble) and for an inferior academic degree (when dominated by other markers relating to such degrees). A more specific marker will have to be used, increasing the number of markers.

knighthood branch; it reads *bachelor*—noun—(Human)—(Male)—(Military)—(Hierarchic)—(Noble)—(Inferior)—(Dependent)—(Proximate)—(Young).

10. *He's studying hard to be a bachelor.* Again there is suggested a possible hierarchic setting of some kind. I will use (Educand).

11. *Employers prefer married men who are at least bachelors; without the degree you hardly have a chance.* This confirms that the general status marker (Hierarchic) should be repeated on this path.

12. *At the age of twenty-five he ceased to be a bachelor, but he never married.* This has to refer to the knight, and points to a marker akin to the (Nonbecoming) that was added to the 'unmarried man' branch: one cannot become an unmarried man, but can cease to be one; one could both become and cease to be a knight bachelor; one can become but not cease to be a bachelor of arts. The position on the academic ladder is therefore (Permanent), and sets off *bachelor* from *sophomore*, for example. As we already have the marker (Inferior) on the knighthood branch, we can add it here and eliminate the distinguisher. The B.A. comes out *bachelor* — noun — (Human) — (Educand) — (Hierarchic) — (Permanent) — (Inferior).

13. *That pet of mine is always nuzzling me and barking and wiggling his flippers.* I include this to show that the system will need, somewhere, to discriminate (Canine) and (Phocine), whence (Phocine) should be extracted from the distinguisher and set up on this path, beneath (Animal). Without going through another series of disambiguations, we can lay down the path *bachelor*—noun—(Animal)—(Phocine)—(Hirsute)[4]—(Male)—(Adult)—(Young)—(Unmated) as an almost complete characterization of the seal by means of markers alone. The remainder, 'during the mating season,' I shall deal with later.

I have tried to show by this exercise that it is possible to do away with the dualism by converting the distinguisher into a string of markers. It is less awkward theoretically and more satisfying esthetically if this can be done, but for *bachelor* the cost was high—a fivefold increase in the number of markers. It is apparently their unwillingness to pay this price that leads K-F to cling to their dualism, though they recognize that in some definitions no distinguishers are needed; thus (Animal) is not only a marker but also a word, and as a word it will have no distinguisher at the end of the path that corresponds to its meaning as a marker; cf. 187 footnote.[5]

Another fault of distinguishers—at least those in the K-F diagram—is their

[4] These markers are easily established by following K-F's lead with *old*. Their trick for pulling (Young) out of the distinguisher was simply to use its antonym in a sentence: *The* OLD *bachelor finally died.* Similarly, we can take a synonym or an antonym of something in the (Animal) distinguisher, disambiguate with it, and make a new marker: *the phocine bachelor* (establishing Phocine directly), *the hairless bachelor* (establishing Hirsute indirectly), etc.

[5] They also by implication give at least one sense characterization by distinguisher alone. The word *colorful* has as one of its paths (198) 'Adjective—(Color)—[abounding in contrast or variety of bright colors]'. If *colorful* is viewed compositionally in the way that K-F advocate (192), then the only part corresponding to the morpheme *-ful* is 'abounding in', in the distinguisher. The point is not discussed, and there is no way to tell whether it has any significance.

redundancy. When (Young) is taken out of the distinguisher class and turned into a marker, it is eliminated from the distinguisher. But the latter still contains 'knight', which is necessarily (Human) and (Male). With (Human) and (Male) extracted, what is left should be something like 'member of the lowest nobility' etc., which is distinctly more marker-like in appearance, and invites to the extraction of more markers. Even if distinguishers are kept, they need to be cleaned up. That involves simply examining the terms they contain (e.g. *knight*), looking up their definitions in the dictionary and identifying the markers (Human, Male), and pulling those markers out, canceling the ones that are already spread out on the path of the sense being defined.

The chief fault of the marker-distinguisher dualism, however, is that it does not appear to correspond to any clear division in natural language. For a theory that claims to apply to natural language, having as conspicuous a feature as this with no equally conspicuous objective counterpart is a disadvantage.

Or perhaps K-F have wrought better than they wot, and there is a spot for distinguishers, or something like them, though a bit to one side of where they placed it. Two possibilities occur to me, one embracing certain presuppositions of the theory, the other extraneous but still related to common dictionary practices.

For the first, it is necessary to ask whether, if the theory accounts for human behavior, it must not then take stock of that part of human behavior which provides the sense characterizations that we draw upon as raw material for the theory. Speakers make themselves understood. Understanding presupposes disambiguation. Disambiguation presupposes the processes that make it possible. K-F recognize two such processes—the linguistic one, which employs markers, and the nonlinguistic one, which employs the speaker's knowledge of the world. Here, then, is a possible residue that could form the content of lexical distinguishers. If a sentence like *Henry became a bachelor in 1965* is unambiguous because the speaker knows that knighthood died out a long time ago, *bachelor* in the sense of 'knight' could contain something like 'in the Middle Ages' as its distinguisher, or as part of it. Using the distinguisher in this way would enable the theory to sweep up certain usage labels and other lexical material (for example, the *Rom. antiq.* that heads the *Century* definition of *cohort*) now left out of account. It would be knowledge of the world that is general, of course, not something as personal as a disambiguation of *Bessie is a bitch* in favor of 'canine' rather than 'human' because the hearer knows that the speaker does not use bad language.

This solution has the advantage of reducing two dualisms to one, though the one is still formidable: 'knowledge of one's language' vs. 'knowledge of the world'. It also produces the effect, at some points, of looking through the wrong end of the telescope. A label as comprehensive as *Rom. antiq.* represents a universe of discourse, as powerful a classifying tool as a semantic marker, and a bit large for a distinguisher to swallow. No allowance is made for the universe of discourse in the K-F theory, and I am dubious of the distinguisher as a proper place for it. I am also skeptical about the remaining dualism, and will return to these two matters later.

The second solution incorporates a type of definition which is common in dictionaries but which the K-F outline does not touch upon. It can be illustrated by the two words *sun* and *senate*. In the *ACD* definition for the first, sense 1 is 'the star which is the central body of the solar system', and sense 3 is 'a self-luminous heavenly body'. Sense 1 is a particular case of sense 3. In the definition of *senate*, sense 1 is 'an assembly or council of citizens having the highest deliberative functions', etc.; senses 2 and 3 name particular instances. I see no reason why the level of distinguishers could not be assigned to this point where markers can no longer be used, only lists: '*sun* → *Old Sol*, *Betelgeuse*, *Sirius*, *Alpha Centauri* ...' This corresponds to 'NProp → *John*, *James*, *Henry*, *Charles* ...' and covers, of course, those dictionary definitions that semantically parallel proper names.

The problem of the symmetry of markers

The second theoretical question is the need to keep the markers on all the paths as nearly the same as possible. When we accept a given lexical item as 'a word' we commit ourselves to recognizing some degree of internal organization. If the senses are wildly different we ought to suspect that we are dealing with homonyms—in fact, one dividend of a theory of markers could be a measure of homonymy: below a certain number of markers held in common, two senses could be taken to represent two homonyms, not two distinct paths of the same word. The existence of two markers, or of a marker and a distinguisher, that are similar but not the same ought to raise the suspicion that we have overlooked a hidden identity.

I illustrate with the marker (Young). There is an obvious relationship between (Young) and the characterization 'first or lowest' which K-F deposited in their distinguisher for the A.B. (Young) could be regarded as (Early) on a path leading down from (Annuated). The A.B. is (Early) on a path leading down from (Educand) etc. By spinning the thread a bit finer we reduce (Young) to a marker (Early) that can be shared by an additional path. There is no gain in this from my standpoint since I have already had to add (Inferior) to the knight branch and that takes care of the lowly A.B. as well; but with (Young) as the sole extra marker on the knight branch as K-F had it, there would have been an argument for generalizing it to the A.B. The illustration is not the best to be had, but I give it because it does not involve going beyond the markers that K-F set up.

A better one—involving my added markers—is that between (Unmarried) on the (Human) side and (Unmated) on the (Animal) side. If we disregard the particular mating ceremonies of man and seal (as we would need to disregard them, for example, between a high-church wedding and a common-law marriage), it seems an obvious waste to recognize two markers, and I accordingly reduce them to one, (Unmated).

The problem of latent markers

I purposely left one distinguisher on the (Animal) side, 'during the mating season', to illustrate this third theoretical point. Where it was sufficient to designate the human being as (Adult), since human beings do not rut and the

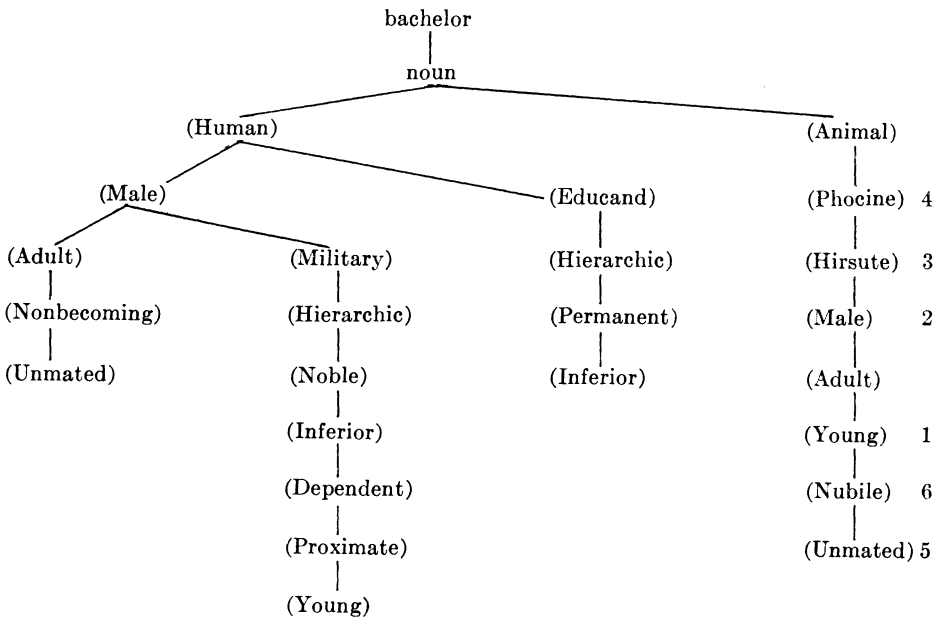

FIG. 5.

The numbers on the right are the order of which the markers appear in the dictionary definition: 1 Young 2 Male 3 Fur 4 Seal 5 Without a Mate 6 During the Mating Season.

whole of their adulthood is characterized by availability for mating, seals do have their season, and (Adult) is not sufficient: one can have a young adult male fur seal without a mate which is still not a bachelor if the time is wrong. 'During the mating season' seems an unlikely candidate for a marker, until we realize that when we call someone—human or animal—a bachelor, we mean that he is without a mate at a time when he is expected to have one: incidentally the whole of adulthood in the human being, incidentally the mating season in the seal. The marker then is (Availability for Mating), or, to use a single appropriate term, (Nubile). If we find that *The migrating bachelors stopped to rest* is unambiguous because we know that the seal's migrating season does not coincide with his mating season, and therefore the bachelors must be human, we are employing our 'knowledge of the world' as we did with *He became a bachelor in 1965*. In other words, (Nubile) can be purified of its temporal associations just as *knight* can, by banishing them to our 'knowledge of the world', and the entry for *bachelor* can now be diagrammed with markers only, as in Fig. 5.

This gives us (Nubile), then, on the (Animal) side. But it is equally appropriate, though unnecessary, on the (Human) side. There it is a latent marker. Shall we add it? This is not a question of CREATING a marker unnecessarily, for it was already required elsewhere. Rather it is a question of how much of the marker system should be made explicit. How cumbersome this would be can be shown by an almost endless number of anomalous sentences: *He walked right through the bachelor* is anomalous because a bachelor is (Solid). *He broke the*

bachelor in two is anomalous because a bachelor is (Pliable). *He welded the bachelor* is anomalous because a bachelor is (Organic). These are markers that dominate (Human) and (Animal). If we are to account for the fluent speaker's ability to recognize an anomaly—as well as an ambiguity—through the markers at his command, then the number is indeed legion. It spreads laterally downward from an apex which I assume to exist on the analogy of the grammatical S = Sentence —perhaps U for (Universe)? Or G for (God)? This is not as absurd as it sounds. The categories have a distinctly Aristotelian flavor.

There is a difference between a dictionary whose entries are written to account only for disambiguations, and one that must account also for recognitions of anomaly. If the dictionary is to be an auxiliary for a translation machine, where it is fairly safe to assume that no anomalous sentences will be part of the input, then the relatively limited roster of one or two dozen markers per entry may suffice. If it is conceived as a representation of how fluent speakers behave toward anomalies as well as ambiguities, each entry will be interminable. A reasonable solution would be not to use the ordinary dictionary at all, if the model is to serve as a kind of lexical map for a natural language. Instead, it could resemble a thesaurus, where each marker would appear only once, and where each sense of a lexical item would appear as a particular path linking marker to marker.

The problem of abstract characterization

The more specific the marker, the greater the number of markers. This follows if we assume that the concepts in our experience are made up largely of shared elements, and are not overly endowed with unique ones or primes. This assumption I think is inherent in a system of markers. Shall we have two markers, (Liquid) and (Gaseous), or one, (Fluid), crossed by two others, (Dense) and (Diffuse)? There is no gain in the latter unless (Dense) and (Diffuse) have uses other than the one of splitting (Fluid) into two halves. If we find that (Diffuse) is useful in characterizing verbs like *scatter*, *dispel*, *disseminate*, and (Dense) in characterizing nouns like *crowd*, *jam*, *clot*, then there may be an advantage in taking a step back and getting the longer perspective.

The marker (Young) illustrates this point. I have already suggested that it be replaced by (Early), in order to show the symmetrical relationship between the separate paths for *bachelor*. But there is a larger reason. Anything that can be adjusted to a temporal scale can be marked as (Early) or (Late):

	Early	Late
living being (ontogenetic)	young	old
humanity (sociogenetic)	primitive	advanced
perishable product	fresh	stale
other product	new	old
lunar phase	waxing	waning
tide	flood	ebb
fruit	green	ripe
growing thing in general	immature	mature

The interrelations here are shown by various metaphorical transfers, such as *early in life* for *young*, *early civilizations* for *primitive civilizations*, *the fresh of the morning* for *the early part of the day*, *the young moon* for *the crescent phase*. If we are to avoid multiplying markers, then refining them in this way becomes imperative.

Can it be done? It would be rash to assume that it can, until it has been tried on a large scale. The marker theory has to assume that it can, or accept the multiplication of markers.

I am skeptical because I doubt the omnivalence of digits. Certain constructs can be dealt with nicely with a scheme of plusses or minuses like that of distinctive features. A piano, for instance, can be characterized perfectly among musical instruments as + percussion, + keyboard, + string, − wind, + tempered, etc. But a piano is also a piece of furniture, where more gradient things like legs, size, surfaces, inclination (it is not necessary for the instrument to be strictly horizontal or strictly upright), and material become significant.

The problem of gradience: Definition by synonym

Dictionary-makers do not shrink from recognizing gradience. One standard way to deal with it is to define by synonymic overlap. This procedure is so foreign to K-F's outlook that they try to deny its existence (185): 'The explicit inclusion of synonyms in a dictionary entry, which is the common practice of conventional dictionaries, is a redundancy introduced to save the user the effort of discovering the synonyms of a lexical item by comparing its sense characterizations with those of every other item in the dictionary. In short, the practice of listing the synonyms of an item is simply a technique of cross reference.' Nothing here about using synonyms to establish the sense characterization itself. Some examples from *NID3*:

> detract 3: 'divert, draw'
> heavy 5b: 'pregnant, gravid'; 7b: 'dull and confused due to interruption of sleep'; 8f: 'massive, coarse'; 8h: 'steep, acute'; 8i: 'laborious, difficult'

Frequently adjectives are nominalized, as a kind of dodge to avoid the appearance of defining by synonyms. Thus for *petty* 3b: 'reflecting small-mindedness or meanness' (i.e. 'small-minded and mean'); and for *dauntless*: 'marked by courageous resolution' (i.e. 'courageous and resolute'). *NID3* designates its listed synonyms—K-F's 'cross reference'—explicitly, with small capitals; but even these are sometimes pressed into service for a sense characterization. Thus for *encourage* 2: 'to spur on: STIMULATE, INCITE'—three synonyms, two of them explicit. Without the latter two, *encourage* might be taken to apply to a horse; with *stimulate* it is made clear that 'arousal in general' is intended, and with *incite* we learn that the object is normally human, not animal.

When *NID3* uses synonyms to define, there are more often than not two of them: the sense is characterized by an overlap of the semantic ranges of two other terms presumed to be already known, and two are the minimum necessary to have an overlap. Of course it can be argued that this is just a shorthand way of saying 'X has those markers of Y and Z that are not mutually exclusive'—so

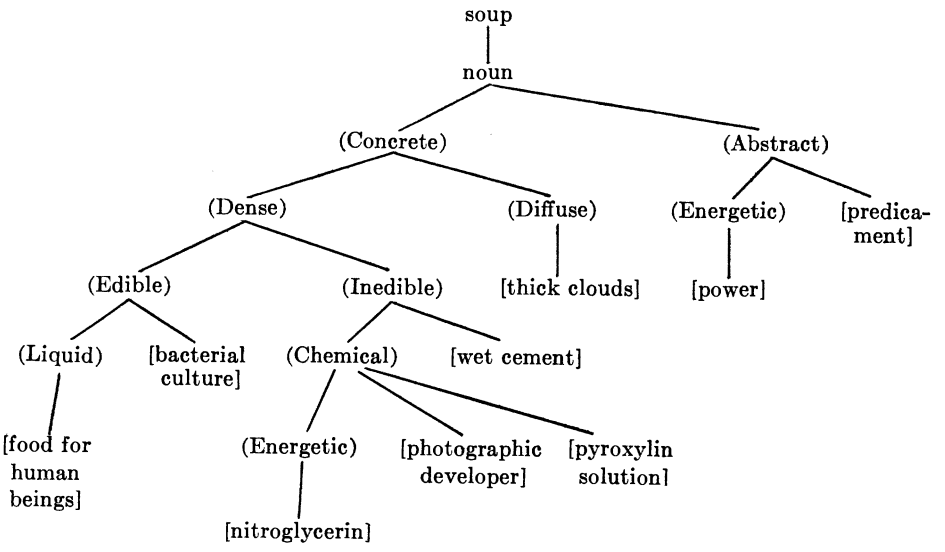

FIG 6

for (Human) under *encourage*. But with *petty* we have a gradience of pejorativeness: *petty* is more pejorative than *small-minded* but less so than *mean*. Definition by synonyms seems not to be a kind of covert definition by markers, but a definition by the semantic range of individual words. If this kind of definition is necessary, then there is a limit to the use of markers, and K-F have not met the challenge of synonymies by calling them 'redundant'.

THE PROBLEM OF METAPHOR

A complete semantic theory must not only map the markers of all senses but show how markers are added and subtracted to alter the senses of words. One corroboration of a marker theory would be its ability to predict semantic shifts, much as a distinctive-feature theory accounts for phonological shifts in terms of change of one feature at a time. It is premature to look for that kind of corroboration or to put the marker theory to so advanced a test. But a more modest requirement is not out of order: to relate the several senses of a word in terms of their probable derivation one from another, so far as this may affect the choice of markers and the sharing of pathways. One would like to see derived senses represented by closely connected paths. A radical departure from this kind of regularity would suggest that semantic derivation is not amenable to rule, and this in turn is a threat to formal treatment of semantics as a whole.

I choose as an example the word *soup*, and diagram its senses according to *NID3*, as best I can in the way that I assume K-F would do it. See Fig. 6.

The metaphorical basis of all but one of the senses is immediately apparent. Clouds are compared to the thickness of soup. Chemicals are compared to its consistency and cook's-broth manner of concoction. Bacterial culture is compared to its consistency and edibility. A predicament is compared to the prover-

bial cartoon of the missionary in the cannibal's cookpot. And power is compared to nitroglycerin which was compared to the food.

From the standpoint of markers, these all represent wide discontinuities. The direct association of predicament and soup kettle spans the widest conceivable gap, that between (Concrete) and (Abstract). So does that of power and nitroglycerin. Looking at these jumps far out of orbit, one can hardly avoid the conclusion of indeterminacy in semantic interrelatedness.

An extreme case is that of thick clouds. Factually, clouds are diffuse. But the intent of calling them soup is to make them appear dense. How, then, does one pick the marker—actually diffuse or metaphorically dense? In shifting its domain from (Dense) to (Diffuse), *soup* 'thick clouds' carries (Dense) along with it, which is just another way of saying that clouds, which cannot be thick, are thick. In a marker theory, which compels us to make all-or-none decisions, this is a dilemma. Markers are atoms. They do not have ranges. A thing is either (Dense) or (Diffuse).

A dictionary is a frozen pantomime. Our problem is only beginning when we consider the pale flowers of that 'nosegay of faded metaphors' that it presses between its pages. A semantic theory must account for the PROCESS of metaphorical invention—all the more so, a theory that stems from generative grammar with its emphasis on creativity. How I make myself understood when I use previously fixed senses that are well known both to me and to my hearer, by a kind of sequence of cancellations, is one thing; how I am understood when I call a chain smoker a fumarole is something else. It is characteristic of natural language that no word is ever limited to its enumerable senses, but carries within it the qualification of 'something like'.

The radical shifts effected by metaphor throw into relief the bold differences between universes of discourse in the totality of language. They raise the question of whether it will ever be possible to attack everything, from slang to technology to science and back again to politics, religion, and the home, with one arsenal of categories.

THE PROBLEM OF SELECTED DATA: CONSTRUCTIVE AND SUBSTANTIVE DEFINITION

Workaday dictionaries do more than one kind of defining. Sometimes they steer it in a more or less marker-like direction, other times by way of overlapping synonyms, others by simile (*cornicle*: a 'horn-shaped' process). Different styles presumably reflect the varying needs of different segments of the lexicon. What will serve for the abstract characterization of *senate* will not serve for the finger-pointing of *Senate*.

This suggests that some kind of survey of the lexicon is needed to determine whether all its entries are really amenable to a single uniform attack. I am inclined to think that K-F's example of *bachelor* is a special kind of word where we 'find' the markers that we have already put in. An *uncle*, similarly, is one who bears a socially defined relationship to *parent* who in turn bears it to an *offspring*—the markers are there because we put them there. It is something

different to find markers in anything that has a life history independent of our naming-operations. A bachelor is a bachelor because he is unmarried, and marriage is an arbitrarily defined social ceremony; we impose the conditions. A bird or a fish is something that we take as we find it, and the markers are adjusted like a suit of clothes, often badly. The fit is crude, metaphorical, subject to revision, and above all subject to change as the entity itself grows or decays through time. This distinction is fundamental. A CONSTRUCTIVE definition applies to a social construct, with markers defined a priori. A SUBSTANTIVE definition applies to the hard objects of the natural world. The contrast becomes vivid when we can express it in minimal terms, when the same object can be viewed either way. A recent review criticizes the definition of *to inspect* as 'to examine carefully' on the grounds that *to inspect carelessly* is normal. But *to inspect* is 'to perform the duties of an inspector', and inspectors, viewed substantively, can be careless. From a constructive standpoint, *inspector* as 'one who examines carefully' is perfectly correct. *Inspector*, *bachelor*, *cousin*, *suffrage*, *baptism*, *corporal*—these are constructive entries and can be dealt with more or less successfully by means of semantic markers. Matters are less rosy with the lexicon of things.

THE PROBLEM OF DUALISM: II. 'KNOWLEDGE OF THE LANGUAGE' AND 'KNOWLEDGE OF THE WORLD'

He became a bachelor in 1965 was disambiguated by the hearer's grasp of history—he knows that chivalry long since passed from the scene. The speaker's 'knowledge of the world' comes in at the point where two possible theories of disambiguation are pitted against each other, one a theory (the common one) that ambiguities are resolved by the context of situation, the other that they are resolved by rules operating on markers that are part of the linguistic apparatus of a sentence. K-F do not deny the role of the nonlinguistic context in resolving many ambiguities; what they deny is that any tightly constructed theory can be built that will reflect its operation, because it involves too much; nothing less, in fact, than everything we know. The example offered to illustrate the point is *Our store sells alligator shoes* vs. *Our store sells horse shoes*. Our knowledge of the world tells us that alligators do not wear shoes, hence the first of these cannot mean 'shoes for alligators', but that horses do, whence the second probably does not mean shoes made of horsehide. (This ignores the fact that it is part of our linguistic equipment that *horsehide* contrasts with *horse* and that *horseshoe* is an existing compound while *alligator shoe* is not, but let that pass for the moment.) In other words, we achieve a disambiguation by way of something that is not a semantic marker.

But why is it not a semantic marker? Where do markers like (Animal), (Physical Object), (Young), and (Female) come from if not from our knowledge of the world? What is strange about (Shoe-wearing) as a semantic marker—not as general, surely, as (Female), but general enough? The discalced branch of Carmelite monks is identified by it, and it crops up every now and then as a mark of status, like horse-riding: 'A Methodist, it was said, is a Baptist who wears shoes ...'[6]

[6] Richard Hofstadter, *Anti-intellectualism in American life* 90 (New York), 1963.

The immigrant quota protecting markers against a contaminating influx from the outer world is no more secure than the social barrier against the upward migration of distinguishers. Ultimately both dualisms are the same.

The problem of the morpheme

The marker is an atom of content. What are the atoms of form? The selection and justification of the latter, and the bonds between the two sets of entities, are a question that K-F barely touch upon.

To reach it I go afield, from a crossroads of linguistic theory where the signs have, I think, been deceptively posted. I refer to a piece of doctrine that one finds repeated in at least every second or third article on grammar in the past few years, and on which K-F are both explicit and emphatic: 'A fluent speaker's mastery of his language exhibits itself in his ability to produce and understand the sentences of his language, INCLUDING INDEFINITELY MANY THAT ARE WHOLLY NOVEL TO HIM ... The striking fact about the use of language is the absence of repetition: almost every sentence uttered is uttered for the first time' (171). K-F then add (171–2),

> In encountering a novel sentence the speaker is not encountering novel elements but only a novel combination of familiar elements. Since the set of sentences is infinite and each sentence is a different concatenation of morphemes, the fact that a speaker can understand any sentence must mean that the way he understands sentences which he has never previously encountered is compositional: on the basis of his knowledge of the grammatical properties and the meanings of the morphemes of the language, the rules which the speaker knows enable him to determine the meaning of a novel sentence in terms of the manner in which the parts of the sentence are composed to form the whole.

If the doctrine of what is 'wholly new' and what is 'familiar' is not to be trivial, there must be some kind of understanding about where the one leaves off and the other begins. In a trivial sense, a stone rolling down a slope and dislodging other stones will form a 'wholly new' pattern with them when they land at the bottom—the novelty is fortuitous. In an almost equally trivial sense, a paragraph of which all has been said before can be juxtaposed to another paragraph of which all has been said before, and the composition will be 'wholly new'. Our attitude toward novelty resembles our curiously puristic notions of 'truth': if anything in a proposition is false, the whole proposition is false; if anything in an utterance is new, the whole thing is new. My example of the two paragraphs may seem extreme, but it must be viewed in the light of K-F's own expansiveness toward the limits of a 'sentence': in making their theory cover all of the linguistic context and exclude all of the situational, they accept the procedure of stringing together unlimited stretches of speech by means of conjunctions: 'replacing discourses or stretches in discourse by single compound sentences' (180). One actually finds such sentences in the speech of the very young, who have not yet learned the best place to put a stop, or in the speech of adults as a way of fending off interruption. But it should be obvious that 'wholly new' in a sense as inclusive as this is trivial. A speaker sitting at the telephone talking with a friend and at the same time looking out the window might be heard to say, *Jack told me he would be home early, and here he comes up the sidewalk.* If *Jack told me he would be home early* and *here he comes up the sidewalk* have been

uttered or heard before, and are hence not new, we cannot claim as new their appearance together when the situation intervenes between them: the man is seen coming up the sidewalk and this serves as a separate stimulus for the speaker—no creation is involved. The only thing potentially 'new' is the *and* between them—a creativity index pretty close to zero. But such an extreme is not what K-F have in mind, as becomes clear from the context in which they use the term 'familiar'. If what is familiar is at the very nethermost level of syntax, the level of the morpheme, then the 'wholly new' does not have to be an inflated sentence but can be any combination of morphemes. And here they err by setting the level too low. There are countless concatenations of morphemes that are familiar, many of them quite large—one intermediate body, with which K-F work and feel no embarrassment in doing so, is that of words. But familiarity reaches much higher. K-F regard the originality of sentences as 'striking'. But what we consider to be striking, what we allow to surprise us, depends on our orientation. I am more inclined to be surprised at the deadly repetitiousness of language. And I am even more surprised at the conformity of linguists to the view that what is 95% old not only in its elements but in much of its internal organization, is to be regarded as 100% new.

To say "X has never been done before" is to claim knowledge beyond human powers. Once we drop our pretensions of omniscience, other ways of viewing disambiguation quickly open up.

I return to the example of *horseshoe*, which I have already quoted from K-F as an instance of possible ambiguity: 'a shoe for a horse' or 'a shoe made out of horse [hide]'. The fact that K-F use it is further proof of their fixation on morphemes as the only level to which one can ascribe 'familiarity'. Yet the morphemes in *horseshoe* are only slightly more separable than those in *pretend*, which could be regarded as ambiguous as between 'to tend beforehand' (no more absurd than 'a shoe made of horse') and 'to feign'. *Horseshoe* as it is used by the fluent speaker of English is virtually as univocal as *spree*.

If that is true, we are now free to expand familiarity upward. Any concatenation that repeats itself is familiar. And since concatenations ARE THE SOURCES of the sense characterizations that we carry in our heads—we learn *bachelor* from contexts like *The old bachelor finally got married*—having sense characterizations among which we must disambiguate presupposes them.

Take an example, a special one, special and atypical to about the same degree as *bachelor* (though in a different way), which I choose because it best illustrates the kind of higher-level familiarity which any theory of natural language must recognize: the noun *spell* in the two senses of 'a period of time' and 'an enchantment'.[7] Familiar concatenations for the first are *a spell of warm weather*, *a cold spell*, *a hot spell*, *a rainy spell*, (*He was there*) *for a spell*, etc. Familiar concatenations for the second are *pronounce a spell*, *cast a spell*, *under a spell*, *release from a spell*, *break the spell* (the last two in the absence of an additional *of hot weather*, for example).

At this point we can argue two ways. We can say, as K-F undoubtedly would

[7] If it is argued that these are 'different entries' because of their etymologies, I would simply get another example. But such an argument would be irrelevant to disambiguation.

say, that *of weather* disambiguates *spell*. Or we can say, as I would say, that there was no ambiguity to begin with: that *spell of weather* is a previously learned unit in the same way that *pretend* is a previously learned unit, and that *of weather* no more disambiguates *spell* than *pre-* disambiguates *tend* by excluding the meaning 'to have a tendency'. The difference of opinion is not one that can be resolved linguistically, but it is one whose philosophical and psychological implications the linguist must face up to. Philosophically, it is the problem of the part and the whole. Psychologically, it is the problem of the point at which integration takes place: do you grasp *A* and *B* in your mind as an integral whole, or do you hold *A* in your mind and *B* in your mind and let them operate on each other? If the real linguistic unit is *a spell of warm weather*, learned as a whole (with *spell* abstracted by adult speakers—when you ask them about words—as a kind of homespun lexicographical exercise which is no more functional in the language than folk etymology), then there was never any ambiguity to disambiguate, and *A* and *B*, for the speaker in that particular act of speech (if not for the linguist), was a false division.

The landscape of frozen forms is a jagged one, here and there rising to great heights of morphemes piled on morphemes, in between sinking to levels only one or two morphemes deep. Disambiguation follows a course that skims the top. At no time does it go morpheme by morpheme. A semantic theory adjusted to natural language must somehow reconcile the way in which human beings operate with wholes and at the same time with forms—morphemes—that they have managed to decontextualize from the wholes. Some day, perhaps, with a little more help from psychology, we shall be ready for such a theory.

Meanwhile, how may one characterize the K-F theory? And what may have inspired it?

It is, first, a picture of semantic units whose outlines are sharp. The marker is an absolute. One either is or is not (Male), or (Young), or (Large), or (Evaluative). If we encounter a sfumato effect it is because we have overlooked a marker or two. Allied to this is a sharpness in the lexical units. They are brittle pieces of crystalline structure that may be picked up but not absorbed: 'The dictionary is something that the speaker learns item by item, more or less by rote' (183 footnote). The mental picture to which this gives rise is that of a transaction in which one comes wholly into possession of a piece of property all at once, like buying a parcel of real estate, rather than the gradual EMERGENCE of meaning, for the speaker, through a long process of decontextualization in which a word is only dimly grasped at first, and slowly, as it gains in contexts, cancels its overextensions one by one. Finally, the lexical units are arbitrary, with no particular relatedness in the system beyond the equally arbitrary sharing of markers, susceptible of being plucked out and replaced with other units without upsetting any internal balances. Fodor, in a recent review, defines the principle of conventionality:[8] 'To say that linguistic symbols are conventional is to say that the

[8] *Lg.* 40.568 (1964).

In principle, probably NO vocable can be replaced by any other without some sort of reverberation in the system. A displacement of just a few rimes would demolish the poetic framework. If the word *too* is replaced, say, by *plethorly*, certain restrictions on the use of

integrity of the symbol depends only upon the consensus of speakers; in principle, any vocable could be employed to refer to any object by appropriately altering the linguistic conventions to which speakers adhere.' This is pure Social Contract, an algebraic Rousseauism applied to semantics, and while K-F did not originate it, neither have they modified it. In a framework so rigid there is room neither for the acquisition of meaning as it actually takes place, nor for extensions of meaning as they really happen.

Allied to the sharpness of outline in the lexical unit is its minimalness. K-F subscribe to disarmament down to morphemes. Their messages are printed on movable type, and after each run the forms are broken up and the type is returned to the font. Morphemes are viewed as monads that associate and dissociate by rule, but whose associations leave no trace. Of course K-F know better than this, and would say so; their theory is not intended to account for idioms nor for the blurrings that one finds in any scientific field where it is useful to make one's theoretical START from atomic entities. But if the carry-over of larger units and residual traces have the importance that I believe they have, a lexical monadism will not do. The assumption of the minimal lexical unit is the weakest point in the K-F theory.

As for the inspiration, it is presented as arising from what dictionaries do, but we have seen that dictionaries do more and also do less. There is less of the sort of marker-like definition that lends itself to the K-F formalization. And there are other ways of defining than by a kind of hierarchic referral system. Dictionaries do not exist to define, but to help people grasp meanings, and for this purpose their main task is to supply a series of hints and associations that will relate the unknown to something known. The orderliness and apparent system in a dictionary are more the result of our instinct to be orderly than of any towering need for system based on the subject matter. The dictionary has done its job when it gives the reader a handhold in his own experience—a pair of synonyms, a diagram, a context, a comparison, tied to any convenient reference post, and to be a good practical lexicographer one needs more to be vastly mindful of the possible associations than to be a powerful theorist. *Instanter* is defined in *NID3* wholly by synonyms; *to entangle* is defined by synonyms plus a 'distinguisher' ('to twist or interweave so as to make separation difficult'); *henna* is 'a shrub or small tree', implying that there is a size which neither *shrub* nor *tree* is adequate to describe; one sense of *insignificant* is 'of little size or importance', familiar terms tied to the unfamiliar ones—all these are scarcely more than homely reminders of things we already know. The success of the dictionary is not achieved in disregard of our knowledge of the world, but through it and because of it.

Of course there are dictionaries and dictionaries, and if we do not make the

---

too, resulting from conflict of homonyms, would disappear (like *excessively, plethorly* would remove the asterisk in *Is he nice?—*Yes, too*; also in **a too large group*). If *tiny* is replaced by *perminute, teeny* will be left high and dry. If all the one-syllable adjectives are replaced with polysyllables so that the *-er* comparative is lost, certain semantic distinctions that have arisen as a result of the option to use either comparison will be lost along with it. As Householder points out, *Journal of linguistics* 1.18 (1965), there are a number of grammars. The grammar in which it makes no difference whether *tiny* is 'tiny' or *perminute* is only one of them.

pretension that our picture of the dictionary is a purer version of the lexical manuals that all students are expected to have at their elbows, then we are free to make any kind of dictionary we want, to suit the use to which it is put. Consciously or unconsciously, K-F have a particular operation in mind, and are describing the kind of dictionary that will serve it.

It seems rancorous at this point to hark back to the accusation that the ideal of generativists is to computerize language. On the side of grammar, the argument has been answered: a formal grammar is one that is self-operative, generating by rule, free of outside interference—such a grammar is LIKE a machine, but not born of one.

But there remains a peculiar sense in which what was untrue, or at least overstated, with a generative grammar, is closer to the truth with a formalized semantics. I refer to the pervasive but never quite avowed conviction that the system can be based on a relatively small number of markers.[9] When K-F write (192) that a formal semantic theory must be able to arrive at semantic interpretations 'without the aid of linguistic intuitions or insights', they merely give another name to all those interpretative leverages that they previously excluded as knowledge of the world, which, if included, would make open-ended and unlimited whatever it is in natural language that corresponds to markers. Whereas generative grammar came into being precisely to give a formal footing to grammatical insights, the K-F theory is at pains to exclude certain kinds of semantic insights. If their exclusion makes the theory that much less adequate to explain the behavior of native speakers, then some other motive must lie back of the selection of a few markers out of a vast sea of intuitions. Native speakers need no drastic reduction. A machine, with its relatively low capacity, does. The job of the programmer is to do the best he can with his machine, and a comparatively small number of markers, chosen for their wide application, is the answer. It is not that these have any better theoretical claim to be called markers, but that they happen to carry the heaviest functional load.

If this is true, the theory of K-F is at best a partial theory of the semantics of a natural language, though it may be a very good theory of how to program for mechanical translation.

[9] A bit of wishful thinking that is well founded, as theories go: 'Economy of description' (small inventory of symbols) is one of I. A. Mel'chuk's 'measures of excellence' according to Householder, ibid. 16.