Molecular Sememics: Toward A Model of an Ordinary Language

Price Caldwell

Some years ago, I attended graduation exercises at a small American university. During the course of the ceremonies, the university Registrar said, "This is undoubtedly the largest class we have ever graduated". I began to wonder what the word "undoubtedly" meant in this instance.

I thought that if he had said, simply, "This is the largest class we have ever graduated" I would have taken it as a statement of documentary fact, coming from the authority charged with assembling such documentary information. Since he didn't say that, I suspected that he didn't know for sure—had forgotten to look at the actual numbers—and was guessing. But since he also did not say, "This is probably (or most likely) the largest class we have ever graduated", I concluded that he wanted to pretend he was not guessing. To put it briefly, his use of the word "undoubtedly" created in me a good deal of doubt as to whether he knew this was indeed the largest class the university had ever graduated.

But while his statement put itself in doubt, it also suggested a rich complex of meaning. There was a range of nuance which went further than his own communicative intent. The meaning that I came to understand, whether correct or not, was antithetical to the dictionary definition of the word—

> un-doubt-ed (un dou'tid), adj. not doubted or disputed;
> accepted as beyond doubt. [ME] --un-doubt'ed-ly, adv.
>
> - Random House

—and was a meaning which involved, in a negative way, the meanings of the other terms which could have been used in its place. Whether consciously or not, the Registrar had considered and rejected those words, and therefore the meaning of the word he did choose meant what it meant by contrast to the other terms in that little momentary "molecule" of possibility. It occurred to me that this tiny and short-lived but dynamic set of choices, within which a specific and highly concrete and subtly nuanced meaning had been born, might itself be the "sememe", the fundamental unit of linguistic meaning, and the ordering principle on which a semantic theory could be based.

The basic notion underlying this "molecule" is, of course, not at all new. It is a revision of the notion of the "contrastive set" or the "paradigmatic series". Early on, many linguists in the Structuralist tradition recognized that the dynamic part/whole logic which operates within such series was fundamental to the functioning of speech. Unfortunately, they were never able to make contrastive sets fit in with any of the structures of syntax, suggest any consistent structures of their own, or even to yield up their common features. In frustration, the effort was abandoned.

But the importance of such sets, or rather of the dynamic of differential meaning to the workings of language cannot be ignored. I have come to believe that the failure to make use of contrastive sets stems from a fundamentally flawed assumption underlying the study of linguistics, the assumption that Language precedes Speech, or (to put it another way) that Competence precedes Performance. The contrastive set, it occurred to me, must be the result of the process of conventionalization, the refereeing and compromising process individual utterances go through on their way to being adopted by the speech community. New meaning is only created within the momentary molecules of original speech.

Contrastive sets, on the other hand, are collocations of repeated molecules whose differential meanings have been compromised and hardened into conventionalities. It is

not surprising, in this view, that they did not generate the familiar categories and structures of language, for they are merely the products of those structures. I propose, therefore, to posit the notion of the molecular sememe, which is a version of the familiar contrastive set revised to the dynamic form it must have taken in Speech. This will require re-examining the assumptions of all structures which are incompatible with it to see whether they can be derived from it. Finally, I must show how a semantic theory can be constructed from it, as well as a possible grammar of an ordinary language such as English. I am submitting here a highly abbreviated outline of the shape the research is taking, and an interim report on some of the implications so far suggested for some of the major problems in semantics.

I – The Redefinition of the Linguistic Sign

"Molecular Sememics" is a term meant to refer collectively to the implications deriving from the revised notion of the contrastive set, the resulting re-definition of the linguistic sign, and the re-evaluations of linguistic methods and structures required by it. The fundamental unit of meaning, in this view, is not a single linguistic item such as a word, a morpheme, or a phoneme, but rather a synthetic structure, the small "molecule" of possible counters from which one is chosen at any decision-point in the creation of an utterance.

I call this molecule the "sememe" because it is the unit of meaning. For the moment let me avoid trying to narrow the definition of "meaning". It will be obvious that I do not mean, simply, "reference", nor do I mean to define it in terms of truth-conditions. The sememe is fundamental to *all* meaning, and so I shall begin with an inclusive sense of it, including reference, syntactic value, logical implication, illocutionary force, and/or emotional import. As it is created in individual speech, it comes into being before conventionalization turns it into either a lexeme or a taxeme.

And I describe the sememe as "molecular" in order to suggest a contrast with the

"atomic" imagery of the popular paradigm, with its feature analysis and other forms of dependence on the analysis of small constituents. For the sememe is a synthetic unit, a minimally complex unit containing parts which participate in meaning but are incomplete in themselves. Such a unit is posited as a logical vehicle for describing the indisputably synthetic and organic operations of the brain which must underlie both perception and speech. Its relation to the familiar term in Chemistry is only analogical; it might just as well be called a "cell", and more obviously evoke the Darwinian paradigm to which it is kin. Perhaps there is something fundamental to intelligence, just as there is in tide-pools according to our Darwinian model, that each organism should seek its own ecological niche by differentiating itself from its fellows. Could it be that neurological phenomena, or meanings, or sememes, do that too?

Everyone has noticed how, on a clean surface, water will bead up in coherent droplets. A few specks of dust or debris, when dropped onto the surface, will instantly move as far away from each other as possible. Whether because of like electrical charges or because of the dynamics of surface tension, the bits of dust will maximize the distances among themselves. So, too, signs and meanings, whenever they are brought into a local relationship, maximize their distances from each other. This phenomenon must be fundamental, underlying the power of contrastive sets and the markedness phenomenon, and giving meaning to the differential that has always been so important to structuralist linguistics since Saussure.

If this is the case, why did Structuralism fail to make the principle of the differential generate all the structures of language? I will try to answer this more fully in the history discussion which follows, but the short answer is that they believed too strongly in the notion of La Langue as in itself the principle of order. Thus, all phonemes were said to have meaning in their difference from all other phonemes; by implication, it was hoped that

morphemes too would prove to have meaning by their difference from all other morphemes. Saussure failed to prove his idea of order because he looked for it on the wrong scale. This of course is to be expected; we all believe that the fundamental principles, whatever they are, must be universal and large. The Molecular Sememe, however, argues for the principle of a local order operating on a small scale.

The Molecular Sememe, then, claims to be a (perhaps neurological) principle of order NOT within the large categories of Language, but within the momentary dynamics of speech. I hope it will be possible later to demonstrate how contrastive sets and "grammatical categories" arise as a result of a process of the conventionalization of successful molecules in repeated use. But prior to that process of conventionalization, we must hypothesize that speech is governed by a molecular order, not a categorical order.

The molecule, then, is to be distinguished from a category, and it is not a categorical phenomenon. Rather, it is a small whole, a universe, an order possessing both a differential and a coherence. Furthermore, it is a complex whole, in which highly subtle nuances of difference are to be recognized, but only among a few (two to four or five) foregrounded choices at a time. Such a structure, I think, is consistent with Connectivist models, recently proposed, of neurological structure, and could also express the Gestaltish qualities which have been observed to pattern human perception. Thus, the molecule can be seen as an expression of the mind's need for an order which is both analytic and synthetic at the same time. If it can also be shown to be a basis on which an ordinary language could be described, it could allow us to make the so-far elusive connections between language and sense-perception, neurology, and psychology. This would bode well for learning theory and the psychology of meaning.

II – Re-evaluating the Structure of Language

Redefining the fundamental unit of meaning as a dynamic structure in Speech, rather than a conventional structure in Language, requires the re-examination of several very durable assumptions with which it is inconsistent. Indeed, it requires a break with both the formalist approaches associated with Generative grammars and some assumptions of the empirical tradition. As a paradigm, however, Molecular Sememics is in some respects a kind of micro-structuralism, faithful to most of the implications of Saussurian structuralism except that the idea of the "structure" must be applied to the molecule rather than to Language itself.

This "except that" is of course a large one. I do not mean to imply that Language is not a structure of some kind. But Molecular Sememics is obviously incompatible with most (if not all) formalist approaches to language, as it sees no necessary logical coherence in Language as a whole. Indeed, the Chomskian principle that Competence precedes Performance assumes a notion of Language as in itself a coherent set of generative rules, while Molecular Sememics sees the principle of order contained in the molecule as prior to any large syntactic or semantic structures of Language. It explains Language, rather, as a compromised and conventionalized collocation of molecules and hierarchies of molecules. Some molecules may be fairly large—large enough to account for the apparent isomorphisms created in the well-formed discourses of competent native speakers. But from the point of view of individual speech, Language itself is really a collection of many languages, certainly not a single consistent, or consistently rule-ordered system.

The claim that the individual physical or acoustic sign is not in itself the sememe implies, as well, a radical criticism of a range of notions in the philosophy of language. The standard notions of reference and meaning, especially as they have come out of the quarrels between realists and nominalists, assume that reference and meaning are properties which belong to individual signs. On the other hand, Molecular Sememics may help

explain more recent events, such as Derrida's efforts to deconstruct language. Molecular Sememics implies that a word is only a conventionalized and ambiguous token, meaningless until it is put into play. Like a candle-flame, a sign may appear simple and durable and nameable. One might take a photograph of it, for instance, and record its existence; but the resulting token should not be taken for the real event, which involves a complex of momentary interactions.

Perhaps the most difficult implication of Molecular Sememics, though, is the one mentioned first. It suggests a radical criticism of the formalist and empiricist methods that have become orthodox in this century. This should become clear in the following capsule history. I regret that, due to the limitations of space, it must appear quite oversimplified and polemical.

III – History

§A – Origins of The Arbitrariness Doctrine

The story of Western linguistics in this century begins in the semiology of Ferdinand de Saussure, whose lectures at the University of Geneva between 1906 and 1911 were reconstituted from student notes by Charles Bally and Albert Sechehaye and published in 1915 as the Course in General Linguistics (English translation by Wade Baskin). Two of his questions were central to this study: What kind of thing is an ordinary language? And how does it create or express meaning?

Of course, his formulations of these questions were dictated by his own place and time. He asked, "What is . . . the integral and concrete object of linguistics?" and "What is the nature of the sign?" These questions were necessary because Saussure needed to distinguish the study of linguistics as a science both from the normative grammars of the Eighteenth Century and the historical studies of philology and phonetics in the Nineteenth

Century. He also needed to distinguish linguistics' special kind of science from the methods of psychology, anthropology, and phonics as practiced at that time.

The problem was that the study of language was muddled by too many approaches. Saussure hoped to isolate a special concept of language, separating the object of his study from both its fluid character as a historical phenomenon and from the unmanageable accidents and varieties of local speech. He wanted to study it as a synchronous phenomenon, a set of relationships existing simultaneously at a single moment.

> Taken as a whole, speech is many-sided and heterogeneous; straddling several areas simultaneously—physical, physiological, and psychological—it belongs both to the individual and to society; we cannot put it into any category of human facts, for we cannot discover its unity.
>
> Language, on the contrary, is a self-contained whole and principle of classification. As soon as we give language first place among the facts of speech, we introduce a natural order into a mess that lends itself to no other classification.
>
> (Course, 9).

Saussure's definition of La Langue, of language as the system which became visible when speech was viewed synchronically as an idealized social object, was indeed brilliant, especially in its contribution to methodology: one could avoid bogging down in the morass of ever-changing meanings, sounds, and grammatical rules. One could instead think cleanly and abstractly about the internal relations among the various components of a

system. Much later, Roland Barthe called the decision "a great novelty" by comparison to the methods of historical linguistics of the last century.

However—and this is a crucial point for Molecular Sememics—Saussure did not credit the idea of La Langue with ontological integrity, or insist that the concept of such an idealized system was axiomatic, the way many of his successors have. I believe he recognized it as a makeshift concept, one with methodological use, a heuristic. He recognized that the "system" was really no more than a set of "habits", seen deliberately as an idealization for the sake of isolating the object of his study from the methods of the other sciences. Individual speech was still the real thing, though, unfortunately, he could see no principle of order in it.

Saussure's tactical redefinition of language, however, was consistent with two extremely powerful concepts that his studies of phonology had made possible: the idea of the arbitrariness of the sign, and the concept of the structure of relations among those signs as capable of creating meaning. Saussure believed that in language there are no positive terms, only differences, and that only the community could make a system out of this collection of arbitrary signs. "The community is necessary" he said, "if values that owe their existence solely to usage and general acceptance are to be set up; by himself the individual is incapable of fixing a single value" (Course, 113).

§B – Mitigations of the Arbitrariness Doctrine

The concept of language as a structure within which arbitrary signs come to have meaning, and on which no individual has any influence, is famous enough not to need elaboration here. It has fruitfully been applied to sociological, literary and political systems by researchers like Claude Levi-Strauss (1962), Roland Barthe (1967, 1970) and Michel Foucault (1966).

In linguistics, however, there has been a constant assault on the arbitrariness doctrine. Roman Jakobson insisted, for instance, that although the purely differential character of the phoneme was the "cardinal element on which everything in the linguistic system hinges", it was not to be found in any other part of the system, not even the morphology (Jakobson 1971, 66). Taking Saussure's example of the contrast between the German singular "Nacht" ("night") and its plural "Nachte", Jakobson granted the differential value—"it is true that the two members of this pair mutually presuppose each other"—but he would not accept Saussure's insistence that "taken in isolation neither "Nacht" nor "Nachte" are anything". For all speakers, Jakobson insisted, "Nachte" is an "independent and direct designation of a concrete plurality" (1971, 64).

Jakobson's argument eventually came to dominate the conversation, and by 1949 he had begun to undermine the arbitrariness of the phoneme itself. He argued that Serbo-Croatian phonemes could be coded as combinations of the presence or absence (+ or -) of six distinctive articulatory features (vocality, nasality, saturation, gravity, continuousness and voicing) (1971, 421). By 1955 he was convincingly arguing that phonological structures universally could be accounted for by some twelve inherent features and a few prosodic features. By subdividing the phoneme, Jakobson reduced the differential character of the phoneme to a bundle of binary oppositions (1971, 498). The phoneme became a mathematically simple set of oppositions, and Saussure's arbitrariness doctrine disappeared in favor of a computational taxonomy.

Thus the concept of the arbitrariness of the sign, no matter how essential it was to the original idea of language as "structure", gave way to a theory of distinctive features as an information-encoding structure which could be merged with purely mathematical notions of communication theory. The resulting notion, that the system of language could be described by logical and automatic rules, encouraged Chomsky's proposition that

Saussure's idealized social object was in fact a set of competency rules unconsciously internalized by the native speaker (Chomsky 1965, 28-29). Later, Jakobson contributed further to the destruction of the arbitrariness doctrine at the level of the morphology through markedness theory, in which pairs of lexical items were seen to stand in a motivated, asymmetrical, binary oppositions to each other. Language began to be seen not as a structure of arbitrary relations, but as a "code" in which binary oppositions at the bottom level generated meanings through a series of generative rules. Since computers are very good at manipulating binary codes, general optimism could easily foresee the day when the generative rules would be exhaustively discovered for semantics as well as syntax, the code could be completely machine-managed, and Artificial Intelligence would help us all get our work done.

§C – The Power of the Formalism

By the early 1970s the optimism generated by Transformational-Generative Grammar had begun seriously to erode. While Chomsky's initial applications of the formalism to the regularities of syntax were highly convincing, subsequent applications to semantic structure were far less successful. However, the concept of language as a code describable by a set of generative rules is still alive among many who still expect a viable Semantic Theory, when one is discovered, to conform to the Generative model.

Just as Jakobson found universal features underlying the phonological system, we must eventually find lexical entries to be describable in terms of "semantic features". Surely, many feel, with the addition of the right combination of "field properties", "contextual constraints" and "felicity conditions", the system can eventually be tinkered with until it works.

This program has met with considerable difficulty, however. More recently, and for

good tactical reasons, Chomsky has seemed to want to neutralize the difficulties of semantic theory by shifting the issue away from any questions of the relationship between language and the world, or of reference and truth-value. Rather, he now speaks of I-language, a purely internal (as opposed to E-language, or external language as he now describes the Saussurian notion) "system of knowledge of language attained and internally represented in the mind/brain" (Chomsky, 1986, 24). In this view, semantics is not a wholly separate set of problems, but in much of its substance, merely an aspect of syntax.

> One can speak of "reference" or "coreference" with some intelligibility if one postulates a domain of mental objects associated with formal entities of language by a relation with many of the properties of reference, but all of this is internal to the theory of mental representation; it is a form of syntax.
>
> (Chomsky, 1986, 45. Emphasis mine)

Some, on the other hand, have argued that the formalist program has proved a failure and should be abandoned. Empiricists such as Talmy Givon (1979) accused T-G grammarians of mistaking the formalism for explanation and of gutting the data base to include only rule-generated sentences. The result, he said, was essentially a tautological argument. Such charges, on top of the generally recognized failure of Generative Semantics, have pushed many linguists to entertain more functional and pragmatic considerations: Speech Act analysis based on the philosophical work of J. L. Austin and Searle; Context Analysis by John Lyons; Pragmatics by Levinson and Leech, and Functional Grammar by M.A.K. Halliday.

The formalism continues to have tremendous authority, though, and many have been

mainly trying to formalize pragmatic and functional principles. Semantics is still studied mainly as an issue of truth-value rather than use. Austin's discovery of "speech acts", performative rather than constative sentences, is a formalization of what has been recognized as a rhetorical distinction since before Aristotle. While recognizing the importance of context in the ascertaining of meaning, Lyons is intent on finding formal ways of dealing with it. Levinson and Leech have likewise felt obliged to formalize pragmatic principles. Such is the power of the formalist program.

On the face of it, formalism derives its great influence from the notion that it is the method of the hard sciences, and that it can express itself in the one language that is more universal than English—that is, formal logic. From this point of view, the problem of language is the problem of describing the set of logical rules which comprises the competence of any human being to speak her mother tongue. This set of rules is also presumed to be a sub-set of the rules of Universal Grammar, the innate rules which underlie the ability of any human child to learn to speak whatever language her parents speak. Thus, the object of study is not language itself, but the "knowledge of language" (Chomsky 1986), an abstract and highly generalized set of logical operations that all human beings are presumed to have inherited because they are human.

But despite the seemingly self-evident truisms formalists have articulated concerning the criteria semantic theories must meet, it has become apparent that formal semantics does not mean to address ordinary-language issues, that is, issues for which it matters whether we are talking about, say, English or Japanese. For Jerrold Katz, for instance, the problem of meaning has to do with explaining things like synonymy and antonymy, ambiguity and redundancy, presupposition and entailment, in abstract logical ways which have nothing to do with which "natural language" these meanings are expressed in. Thoughts are assumed to be independent of language, and the issue of

meaning in speech, that is, meaning which does depend on the immediate utterance being spoken, is ruled out of consideration from the beginning as belonging merely to the realm of pragmatics.

While Molecular Sememics is a theory of speech pragmatics, it rejects the "merely" that usually accompanies that label. In neglecting the vagaries of meaning, the formalism forgets to look for what is essential in language—its ability to construct worlds of discourse rich in subtlety and nuance—in favor of what seems to me a lesser thing, a scheme for describing the relationships of syntax. I think, furthermore, that it is an effort doomed to failure, finally, because it is committed to a consistent description of what is merely conventional, and therefore, by and large, only *mostly* consistent, and then consistent (in overt form at least) in mainly pragmatic and uninteresting ways.

Of course it is not necessary to repudiate the formalism in order to pursue speech pragmatics. Although the molecular sememe suggests an order which is very unlike that of formal logic, I do not intend that my description of it should be illogical. In any case, even among formalists, there are substantial arguments that the mind needs no formal logic in order to reason, that there are heuristic strategies and modelling techniques which can accomplish, within the human mind, what computers require an explicit set of instructions to accomplish. In short, ordinary language itself, as Molecular Sememics describes it, can generate the operators which give rise to formal logic. Thus it is, I believe, a derivative of, not a constituent of, ordinary language.

If Molecular Sememics seems radical, then, it is because it seems to repudiate the foundations upon which science itself seems necessarily to depend. But that is not the case. Molecular Sememics means merely to find the privileged fundamentals, the primes, in the realm of speech rather than in language. This entails, however, a repudiation of Competence as prior to Performance, which, in the terminology of Cognitive Science,

might well seem to be a denial of general truth in favor of the anomalies and accidents of anecdotal evidence. It is not that, though it might seem so. It is, rather, a vote in favor of a local and immediate ordering principle rather than a universal set of rules which mean to govern from on high. It in effect pretends that the person in charge of language is the local schoolteacher or the teenager with the most influential fashion sense, rather than either Universal Grammar or the National Academy. She, not a set of abstract and universal rules, creates the language. Of course she must persuade the local community to speak as she speaks, and her influence may seem limited to small and inconsequential changes in the uses of popular words. But the crucial point is that while she must persuade them to alter the conventions they have known since birth, they will do it willingly and meaningfully: creativity *is* possible and intellectual growth is (usually) welcomed. The conventions are not unbreakable universal contracts, and they do not necessarily follow from some genetic necessity, some natural law.

In arguing that speech precedes language I am not, however, repudiating any notion of a universal grammar. I am only arguing against the notion that a full set of syntactic rules is required for that grammar. The structure of the molecular sememe itself requires an underlying neurological capability which, though I think it is consistent with what is known about neurological structure, remains to be demonstrated in those terms. But it does not dictate any particular set of rules governing the use or application of those molecules. The experience of the speaker in the world and in the company of other speakers is sufficient, given the neurological capability of forming molecular sememes, for any child to learn any language if it is presented to him. For if Molecular Sememics has any validity for English, it is also valid for any other ordinary language. In this sense it is a semiotic theory, and it might well prove to work for bird-song and dog-bark, or for music and art, as well as for the chatter and the poetry of human beings. The Molecular Sememe is a synthetic unit in the

order of Signs, in a generic sense, and not just a function of the lexicon of English.

What will seem most radical about this proposal is its implication that the Universal Grammar may be a (relatively) simple biological or synaptical organization for speech, not a complex set of logical commands in a computable program. Whether the molecular sememe ever in fact proves to be computable or not remains to be seen; I know of no principle that would prevent it, but it would require a logic very different from those in use today. For the molecular sememe must make use of part/whole logic, rather than item/category logic. Whether algorithms exist that can implement that or not remains to be seen. Suffice it to say, in the descriptions that follow, since our object is rhetorical truth, not logic, speech not language, and performance, not competence, we will eschew the language of the formalism.

The most radical implications of Molecular Sememics, then, are owing to its notion that meaning is created in immediate and local speech rather than in the systematics of Language, in a momentary dynamic involving the immediate juxtaposition of a term with its molecular "Others". This does not deny the influence of the speech community in having conventionalized (or 'assigned') the meanings of those terms, but it does argue that meaning does not originate there, and that "meanings" are not, therefore, the "properties" of words. In this perception, then, the failure of current linguistics to provide a believable theory of meaning is seen to be related to the assumptions of formal approaches which have, I think, made us look for the answers in the wrong places. Thus, Molecular Sememics proposes a paradigm which is not based on the "lowest common denominator" procedure of analytic science, and does not assume an atomistic notion of constituency. It rejects not only the terminology of atomism—the "features", "figurae", "phones" and "bits", but also the logical paraphernalia of the formalism, the "properties", "categories", "rules", "constraints", and "criteria" which impose their unrhetorical characteristics on the

object of study.

To many, of course, saying this is tantamount to refusing to abide by the scientific method. I mean no such thing. I simply find it difficult to accept formal logic as the value-neutral medium it is intended to be: for as long as English words provide the content of any proposition, that proposition is as value-laden as any statement within ordinary language. Furthermore, the inclusion of semantic items from ordinary language within propositional logic merely provides the qualitative content which is the sine qua non of any proposition which hopes to say anything real about the world; and so it is in fact welcomed, even though its status is denied.

What formal logic does, on the other hand, is take qualitative content and try to strip it of whatever indeterminacy of implication it might have had. This is considered a necessary and useful thing, without which formal statements would be as ambiguous and useless as ordinary statements. But its real effect is to impose an ideology on the object of study, one which forces it to pretend that its structures are in fact as qualitatively empty as scientific description demands. Thus the category “Direct Object”, say, would indeed describe an abstract commonality shared by all “direct objects”. That is to say, all the underlined items in this list would share some specific syntactical meaning:

Jennie told her parents she was going to Anne’s house.

She said, "I won't be late."

Jennie shared her thoughts with Anne frequently.

She hoped Anne would understand.

But Anne often pled ignorance of Jennie’s meaning.

In fact direct objects have little or nothing in common except that they are usually noun

phrases and that they follow verbs. And often noun phrases which follow verbs are NOT called Direct Objects:

> Anne issued Jennie an ultimatum.
>
> "Go home," she said, "or else."
>
> "You are driving me bananas."

This is not the place to conduct an exhaustive critique of grammatical categories. For the moment I only wish to provoke the reader into entertaining an alternative to the whole analytic procedure, and look at a synthetic unit instead, one which might put the whole paraphernalia of categories and items and rules and propositions into a different light.

In revising the paradigm, Molecular Sememics hopes to provide a description of an ordinary language which will not only be useful to linguistics, but will also suggest how logic, poetry, science, and literary criticism can utilize, to their diverse ends, the structures of language. It wants to ask not the question of formal semantics—"How can we be sure we know what a sentence means?"—but another, more natural one: What kind of a thing must a human language be, if on one level it can provide a ready-made network of conventions rich enough to make any native speaker think he can think, and at the same time provide the flexible precision instrument with which poets and philosophers and scientists continually change our conceptual world?

II – Molecular Sememics and European Structuralism

While the American formalist program has replaced the arbitrariness doctrine with the notion of a code grounded in ontology, Molecular Sememics means to restore a revised sense of Saussure's arbitrariness doctrine. In this sense, Molecular Sememics is

reactionary rather than radical. But in being so, it also hopes to offer a revised idea of the structure of speech.

The idea of the arbitrariness of the linguistic sign has been taken in many senses. One meaning, central to both Saussure and Hjelmslev, was simply that the differential character of the phoneme did not depend on its phonic substance (Saussure 7; Hjelmslev 50). Molecular Sememics argues that this is still true. Jakobson's brilliant reduction of all the world's phonic substances to a handful of acoustic features is taxonomically useful to phonetics, but irrelevant both to phonemics and to morphology. It does not provide a non-arbitrary physical or ontological grounding for linguistic meaning, despite Jakobson's claim of a fundamental similarity between distinctive features and the elementary particles of physics (Jakobson 1971, 425).

Molecular Sememics, then, maintains the core meaning of the arbitrariness doctrine, which is that meaning is derived from the differential character of the linguistic sign, not from anything inherent in the phonic material of the sign or from any logic prior to language. It denies the existence of any a priori inventory of logically possible meanings from which items can be plucked for "lexical entry". But Molecular Sememics does not claim that language springs accidentally from nothing. The molecular sememe itself is an adaptive structure, whose shape reflects and therefore expresses all of the psychological, sociological, and historical factors which affect speech.

This is perhaps the most important claim of Molecular Sememics. The shape of the molecule itself, as defined by the relationships obtaining among the counters belonging to the molecule, is not arbitrary, but motivated. This is not a unique claim, since many linguists have claimed such motivation for "markedness" patterns, and the molecular sememe may well be thought of as a markedness phenomenon. The difference is that markedness patterns are seen as a manifestation of binary logic in semantics. The

molecular sememe is a direct manifestation of, and index to, the world's order as it exists in the unconscious experience of the speaker. However complex and diverse are the psychological, sociological, and historical factors which affect any utterance, they are expressed in the relationships which obtain among the counters of the molecule. Thus the molecule is the "principle of order" for speech that Saussure despaired of finding.

A simple example can be given here. If a baseball announcer narrates a baserunner's stealing second base, he will use the term "safe". A native listener knows that the word "safe" means what it means by contrast to "out" (and not, say, "unsafe"). That is, in this case a simple binary molecule exists—[safe/out]—to express the relationship between these two technical terms. On the other hand, if the narrator describes the batter's efforts, he will appeal to a more complex molecule [strike/ball/foul ball]. Again, each of these terms means what it means by virtue of its membership in the molecule AND by its differential from the other counters. But the semantic relationships among the counters of the molecule directly reflect the rules of baseball, in which three "strikes" are required for an "out", four "balls" are required for a "walk", and "foul balls" are counted as "strikes" the first two times but not thereafter. Thus, the shape of the molecule itself reflects the semantic structure of the world of baseball, or of whatever "world" the molecule belongs to. This example may also be used to illustrate the relationship between speech and language, and to suggest how language change may occur over time. As for the chicken-and-egg question of whether language or speech comes first, Molecular Sememics argues that while the counters which enter into molecular dynamics are lexical items which are already available in language, and that they bring their histories with them to the encounter, it is the use to which they are put, their interaction with the other counters of the molecule, which gives them their local and immediate meaning. It is this meaning, if it is repeatedly elicited in other molecular encounters, which will become conventionalized and seem to "belong",

henceforward, to the word.

Thus, in the evolution of the language of baseball there must have been a point at which "out" (which is normally contrasted with "in") and "safe" (which is normally contrasted with "dangerous" or "in jeopardy") were selected as most appropriate for special use in contrast to each other. This probably occurred spontaneously and only partly consciously in the speech of some early participant in actual play. Finding a niche within the internalized conventions of other native speakers of baseball language, the distinction proved durable. Molecular Sememics contends that all "lexical contents" have come about in exactly this way, and that language change over time is the story of the successive participation of lexical tokens in various molecules, a few of which prove useful enough to bear repetition, and then become conventionalized.

III – The Conventionalizing Process

If, as I believe, the concept of the molecular sememe does provide a principle of order for individual speech, it should no longer be necessary to labor under the assumption that Language is a single consistent structure. We can then interpret its many apparent anomalies as natural, rather than exceptional. The premises of Molecular Sememics suggest that Language is some kind of 'makeshifty' but efficient collocation of individual utterances, rather than a logically consistent ontological structure, the "innate fixed nucleus" of which was given, according to Chomsky, to the human unconscious by a mysterious genetic event.

In short, meaning originates in speech, which is prior to language. But if this is true, then the process through which molecular meanings pass on the way to creating the phonological, morphological, and syntactic levels of language structure must be describable as a process of conventionalization. Even the "phonological system", of all systems

apparently the most systematic, must be shown to be no more than a compromised and conventionalized collocation of a great plurality of local phonemic distinctions.

And this is not easily done. For all arguments about the origins of language, the hard evidence disappeared centuries ago. On the other hand, it's quite likely that the processes which created language in the first place are the same processes which create it now, whenever poets or philosophers or propagandists set out to explore new rhetorical terrain. Let's consider what happens when any new geographical territory is colonized, explored, and conventionalized:

1) The intrepid explorer sets out to find a route from Settlement A to, say, point B on a distant river. His first trip is highly exploratory, and covers a lot of ground very inefficiently. At rivers he tracks up and down looking for an easy crossing; at mountains he goes around rather than over. Withal, there's a fair amount of sightseeing.
2) When a number of trips by a number of people have been made, however, a consensus emerges concerning the best way, and this way is marked by a path. This is the first stage of conventionalization, which is a highly efficient operation. But it means that less off-path ground is explored, and henceforward the territory will be less well known than before, at least until enough settlers arrive to occupy the empty spaces.
3) Meanwhile, other explorers are making paths between points C, D, E, and F. Inns are built at crossings, and sometimes travelers in one direction share the path of travelers in a slightly different direction. Through such cooperation, the number of paths stops growing and begins to shrink.
4) At some point some engineers from C and D build a bridge over the river and engineers from E and F cut a tunnel through the mountain. Many paths are

abandoned as travelers go a short distance out of their way to take advantage of the new facilities.

5) Eventually a highway is built from City A to point B. But it does not follow the original path. It deviates to take advantage of the bridge and the tunnel. Though it is a bit longer than the original path, the time saved with higher speed more than makes up for the loss.
6) Now, if you want to take a trip from any point to any other, you can look at a road map. It represents the highly conventionalized and compromised version of what was once a great many individual tracks.

As this analogy suggests, the process of conventionalization in language begins in the individual locution. If it is an efficient one, repetition will mark it as a path for communication. Later, though, the necessity for compromise with other people using the same means for other kinds of messages creates less efficiency, more overlapping and redundancy. It is at least possible to imagine that the same thing happens at both the phonological level and the semantic levels of language development. For some excellent suggestions about the kinds of inefficiencies which result from the conventionalization process at the rhetorical level, see Brown and Levinson, 'Politeness: Some Universals in Language Usage' (1878, 1987).

The premise of Molecular Sememics is that meaning originates in the individual utterance, and then is conventionalized. An application of this premise to the study of phonology suggests that a radically different attitude might be taken toward the concept of the phoneme. The familiar Jakobsonian concept, that the phonology is a single consistent system within which every phoneme contrasts with every other, lends itself handily not only to the "distinctive feature" logic of the formalism (how else can we reduce everything to a few definable primes?), but also to the notion of the minimal pair, empiricism's favorite

method for isolating “smallest meaningful units”. Within such methodologies, of course, the discovery of synthetic units is unlikely.

But if Speech, not Language, is the phenomenon under scrutiny, then “emic” logic suggests that only the sounds immediately available for use at any decision-point in an utterance participate in the meaning of the utterance. For example, whether one says “I want my mom” or “I want my dad” depends on the local and immediate emic distinction, valid for this utterance only, between /ma/ and /daed/. Whether this distinction is valid for other utterances, and whether it is constituted of one phoneme or more than one, are of no importance to this immediate use. But if in the accumulation of repeated similar distinctions there comes to be a perceivable regularity in the pronunciations of, say, /m/ and /b/, or /n/ and /d/, or /a/ and /ae/, it will be the kind of regularizing (or path-making) which maximizes perceivable differentiation while minimizing the consumption of scarce phonic materials. Habituating an economical distribution of scarce phonic resources is probably the most useful effect of the process of conventionalization.

And indeed it is necessary, because in the molecular logic of speech, minimal contrasts are rare. By and large, speech requires the clarity of maximal contrasts, on both the phonological and semantic levels. But since (within molecular logic) phonemes rarely need to differentiate themselves from forty or more other phonemes—typically, from only from one to three or four others—there is normally room for all kinds of slurring and blurring of pronunciation without the loss of contrastive clarity.

Even when a pronunciation must distinguish itself from many others, the phonology can normally provide an ample number of maximally-contrasting counters. For instance, even in a very large and loose contrastive set—

Denver Tower, Cessna 22 Tango, [__/__] miles east, inbound for landing . . .

—one in which, say, any number from ten to a hundred could occur, clear contrasts are

available in nearly all cases. When they aren't available, trouble results. For instance, the differentiation (mostly stress) between "fifteen" and "fifty" is inadequate in the speech of many foreign pilots. It creates a special problem for air traffic controllers.

This should illustrate the fact that the distribution of scarce phonic resources is not really economical in the way it should be if Language were a single consistent structure, and if phonemes had to maintain contrast with all other phonemes all the time. Luckily, it is not a single structure, and phonemes normally have to maintain contrast with only a few others in any single instance. Thus, the language can normally afford redundancies and overlapping such as those obvious at the "morpho-phonemic" level, where the inflexional suffix [-s] has three different meanings for different morphological environments, and three different pronunciations for different phonological environments. Because these environments are "molecular" and limited to only a few possibilities, there is normally no confusion of meaning.

Thus the distribution of a single marker to different uses or of different markers to the same use is, as we have seen, exactly the kind of redundancy and overlapping that we would expect the conventionalizing process to create, and the molecular structure of speech to permit. Synonymy and homonymity represent the same kinds of redundancy and overlap at the semantic level. Thus, the notion of a conventionalizing process is at least superficially consistent with some of the apparent inefficiencies of language.

IV – Toward a Grammar of an Ordinary Language

Within Transformational-Generative Grammar, syntax is considered to be the fundamental structure, prior even to semantic meaning. Molecular Sememics argues, on the other hand, that this priority reflects only the formalist bias within Generative Grammar. Syntax, to the extent that it is expressed in word order in English and is therefore temporally ordered

(spatially ordered in written texts), is merely the most easily formalized aspect of speech.

In the determination of meaning, in Molecular Sememics' rhetorical view, syntax is neither the most important structure nor the most interesting. It provides only the "chassis" of a sentence, a convenient framework on which to hang the operating mechanisms. For example, it has been often observed that many sentences, though well-formed by syntactic criteria—

*Many of the world's largest banks have their headquarters.

—require the addition of (what traditional rules would call) "optional" parts (in this case an adverbial modifier) to become rhetorically complete:

Many of the world's largest banks have their headquarters in Tokyo.

This suggests that a sentence's structural integrity is not always best expressed in the Base Rules. "Tokyo" here is the molecular term, and clearly the sentence is not complete without it.

Molecular Sememics assumes the primacy of rhetorical intention in the structure of speech. In this view the primary unit is not a syntactic one such as a sentence or clause, but rather a rhetorical unit similar to those found in other speech-based or functional grammars. M.A.K. Halliday, for instance, calls the primary unit an "information unit" and defines it as consisting of an "obligatory New element plus an optional Given" (c.f. Halliday, 274ff).

Halliday's information unit is the closest approximation I know of to the grammatical unit implied by Molecular Sememics. In Molecular Sememics, the

fundamental grammatical unit is a "molecule-selection-and-execution structure". What Halliday calls the "Given element" is that part of the utterance which, in Molecular Sememics' terms, orients, or sets up, or "selects" the molecule. It usually consists of given information or information which appeals uncritically to the speaker's conceptual world. At the same time it indexes the distributional conventions of the language to "select" the terms available for use as counters in the molecule. What Halliday calls the "New element" is probably coterminous with what Molecular Sememics calls the "molecular" term. This term (though it may be a phrase or a larger unit) is molecular in the sense that it is the term which means what it means within the differential dynamics of the molecule. Thus the meaning of the utterance as a whole really depends on the meaning of this molecule as it is marked by the selected term.

Seeing discourse as a hierarchy of "molecule-selection-and-execution structures" (MSES) suggests that a molecular syntax would have three levels of structure: first, structures which sort out the hierarchical relations among the MSES, (coordination, subordination); second, structures which select and identify the molecular term(s) within the MSES (assertion, interrogation, negation); and third, structures within the molecule itself.

At the second and third levels, structures are determined by the various possible dynamics within the molecule itself. For instance, one molecule may operate to synthesize several diverse counters, and another may operate to distinguish several similar ones. One molecule may be selected as a whole by the mention of two unmarked counters, or implied in the assertion of a single marked counter. One utterance may "query" a molecule, by in effect asking for an itemization of its contents; another may "evaluate" a molecule, ranking its contents in a preferred or dis-preferred order. An utterance can select one counter of a molecule and reject the rest, or it can deselect or negate one of its counters and choose the

rest. We may find that in highly efficient texts, as in poetry, we will have to sort out several superimposed molecules operating on several planes of organization at once.

In short, molecular dynamics can be very rich. There is reason to believe that the various dynamics of the molecule can be shown to generate the familiar rhetorical forms, including interrogation, negation, assertion, naming, irony, metaphor, analogy, and perhaps even rhyme. At this point space permits only a few suggestive examples, with which I will conclude this paper. I hope they will hint at both the complexity and the richness to be found in the premises of the Molecular Sememics paradigm.

EXAMPLES

1) When you have read the question, mark the letter corresponding to the answer as given in your exam booklet: A, B, C, or D. (Interrogation and assertion)

A fully explicit, symmetrical molecule such as the molecule [A/B/C/D] interrogated here is indistinguishable from (and probably the origin of) a logical set. Purely logical counters are typically members of explicit, symmetrical molecules. The molecule [question/answer] is also fully explicit and symmetrical. This kind of molecule characterizes the language of technicality.

2) Q: Coffee, tea, or milk?

A: Coffee, please.

The molecule [coffee/tea/milk] is an itemized and, probably, fully explicit molecule. If it is, it suggests the logic that once obtained in the universe of airline food, within which coffee, tea, and milk were the only alternatives. As should be apparent, the primary rhetoric of interrogation is to itemize the contents of the molecule. At the same time, the primary rhetoric of assertion is to name an item as selected in place of others. Hence, assertion is implicitly a denial of the unselected counters.

3) Q: Did you say 423-2345?

A: No, I said 423-2355.

This molecule, which will be indicated by a tone change emphasizing the new "5" in the answer, is [--- --4- / --- --5-]. That is, it is implicit that the molecule is an order of telephone numbers. Thus, a molecule can overlap with or include a set of numbers. Of course, the "dynamics" here are very logical, and do not possess the nuance that an organization of non-abstract terms possesses. It also should be obvious that in this case the other terms—that is, the other numbers—do not interact with the molecular counters.

4) I won't take 'no' for an answer. (Negation)

[I won't take 'no'/(I will take 'yes')]: this molecule should illustrate that the typical molecular strategy of negation is to deselect the explicit term and therefore select the opposing term.

5) No, Mr. Smith is not the one I had in mind.

The molecule here is [Smith/whoever else *is* held in mind]. Since the molecule is binary, the deselection of Smith says a great deal about who is selected. If the hearer can be relied on to know the possibilities of the molecule, then negating one term effectively names (without naming) the other. Gossips and news leakers often use this method of saying without saying.

6) Either fish or cut bait. (Analogy)

This kind of explicit alternative proposition is really a kind of explicit negation, implying that if you reject one alternative you are stuck with the other. At first glance, [fish/cut bait] might seem, on an abstract level, reducible to a simple categorical set (YES/NO). But that analysis, in ignoring its "poetry", would be inaccurate. The molecule really invokes a very rich universe of values (which we might call the language of fishery) within which a far more particular logic operates. Generalized and paraphrased, it means something like

"either engage the main task or support those who are so engaged". But I only know that because of the analogical relations that are possible between molecules of similar shape.

7) Why don't you put a voltmeter on the starter solenoid and see if it's getting any juice? (Metaphor)

If we look at the "molecule-selection" parts of the sentence, we find so much redundancy that the molecular term, we strongly presuppose, will be "electricity". In short, any sort of word that could be possibly taken as a token for the idea of electricity could have been put in this slot. Its value, then, is measured by its differential from the word "electricity". "Juice" is clearly more fun than "voltage" or "current" or "energy" would have been, and, because any term would be interpreted by analogy to electricity, there is no loss of precision. Even counters such as "tingle", "flow", "oxygen" or even "stuff", might be close enough to work.

8) He: Where have you been?

She: I just went out to get a breath of air.

He: You did, like hell.

She: What do you want me to say, Darling?

He: Where have you been?

She: Out to get a breath of air.

He: That's a new name for it. You are a bitch.[1] (Naming)

An implication here is that any token can be made the name of anything if it is put in the position of a strongly expected term (here the reader has been strongly prepared to understand that the female speaker has been out having sex with the 'other man'). As in the previous example, the semantic value of "out to get a breath of air" is given in the

[1] Dialogue from Hemingway's "The Short Happy Life of Francis Macomber".

differential between what those words ordinarily mean and what they mean now, which is, "in bed with Mr. Wilson". This differential adds up to the insouciance with which she (Margot Macomber) equates the one with the other, without any real effort to obscure the true referent. It is what provides the definition of the word "bitch" in the final act of naming.

9) Q: Would you like another drink?

A: Does a goat stink?

A categorical molecule [YES/NO] might seem to be selected. But when it is executed with the answer "Does a goat stink?" then an extremely complex communicative event occurs. The parallelism of the second question to the first means that the answer to the second is also the answer to the first. The rhyme [drink/ stink] supports this parallelism by proposing an isomorphism where none existed before: a molecule is selected in which the *similarity* within a context of difference is proposed, rather than a difference within a context of similarity.

The immediate result of all this is that the answer to the second question (an emphatic "Yes") will be taken in the place of the merely categorical affirmative of the first molecule. A secondary result is that further parallelisms will be sought. If so, they will be found in [you/goat] and [drink/stink]. Thus, the full implication is not simply the categorical /YES/ but the far more (humorously) emphatic and particular "I am a goat, of course I stink, and of course I want another drink". But this is a paraphrase, which doesn't really do justice to the wealth of implication contained in the complex focusing of rich molecules by a lucky question.

V – Conclusion: Further Implications

It should be clear that Molecular Sememics foresees far more avenues of investigation than

it has had time to explore so far. I have already suggested some of its implications for semantic theory, philosophy of meaning, and the psychology of language and perception. It has not escaped my notice that there are also rich implications in the model for topics such as linguistic change (a single word may have successive membership in a variety of molecules) and language learning (molecular structures can subdivide and reform as needed to order a gradually-increasing vocabulary). There is a great deal to be done in all these areas, though, and I apologize that this brief report could do no more than touch on them.

General Bibliography

Barthe, Roland. *S/Z.* Trans. R. Miller. New York: Hill and Wang, 1975.

Brown, Penelope, and Stephen C. Levinson. *Politeness: some Universals in Language Usage.* Cambridge University Press, 1978, 1987.

Chafe, Wallace. *Meaning and the Structure of Language.* Chicago: The University of Chicago Press, 1970.

Chomsky, Noam. Knowledge of Language: Its Nature, Origin and Use. New York: Praeger Publishers, 1986.

_____. Aspects of the Theory of Syntax. Cambridge, Mass: M.I.T Press, 1965.

Eco, Umberto. *Semiotics and the Philosophy of Language.* London: MacMillan, 1984.

Edelman, Gerald M. *Neural Darwinism: The Theory of Neuronal Group Selection.* New York: Basic Books, 1987.

Givon, Talmy. *On Understanding Grammar.* New York, 1979.

Grice, H. P. *Logic and Conversation.* Unpublished MS, from the William James Lectures of 1967, quoted in Brown.

Goffman, Erving. *Interaction Ritual: Essays on face to face behavior.* Garden City, New York, 1967.

_____. "On Face-Work: An Analysis of Ritual Elements in Social Interaction," Psychiatry, 18 (1955), 213-31.

Greimas, A-J. *On Meaning: Selected Writings in Semiotic Theory.* Paul J Perron and Frank H. Collins, trans. Minneapolis: University of Minnesota Press, 1987.

Halliday, M.A.K. *An Introduction to Functional Grammar.* London: Edward Arnold, 1985.

Hjelmslev, Louis. *Prolegomena to a Theory of Language.* Francis J. Whitfield, trans. Madison: University of Wisconsin Press, 1961.

Jakobson, Roman. *Selected Writings*, Vol 1. The Hague: Mouton, 1971.

Philip N. Johnson-Laird. *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness.* Cambridge, Harvard University Press, 1983.

Julia, Pere. *Explanatory Models in Linguistics.* Princeton: Princeton University Press, 1983.

Katz, Jerrold J. *Semantic Theory.* New York: Harper and Row, 1972.

Leech, Geoffrey N. *Principles of Pragmatics.* London: Longman, 1983.

Levinson, Stephen. C. *Pragmatics.* Cambridge: Cambridge University Press, 1983.

Levi-Strauss, Claude. *The Savage Mind.* London: Weidenfeld & Nicholson, 1966.

_____. *Structural Anthropology.* Claire Jacobson and Brooke Grundfest Schoepf, trans. New York: Basic Books, 1963.

Lyons, John. *Semantics.* 2 vols. London: Cambridge University Press, 1977.

Markova, Ivana. *Paradigms, Thought, and Language*. Chichester: John Wiley & Sons, 1982.

Miller, George A. and Philip N. Johnson-Laird. *Language and Perception.* Cambridge: Harvard--Belknap Press, 1976.

Pike, Kenneth L. *Language in Relation to a Unified Theory of the Structure of Human Behavior.* 2nd Ed. The Hague: Mouton, 1967.

Saussure, Ferdinand de. *Course in General Linguistics.* Wade Baskin, trans. New York; McGraw-Hill, 1966.