# Chapter 1
# The Epistemologies of Linguistic Science: Reassessing Structuralism, Redefining the Sememe


Since the nineteenth century, students of the science of linguistics (as opposed to grammarians, philologists, natural philosophers, and cultural historians) have wanted to put the discipline onto a solid scientific footing. To be 'scientific'[1] has meant, since the Enlightenment (with its stories of Galileo vs. the Church), to base one's arguments on close observation of nature rather than the appeal to authority, dogma, or the classics of Greek or Roman literature. Yet there can be no such thing as pure observation. It is always conditioned by what constitutes data, what its instruments of observation are, and the problems of assembling particular observations into general ones, so that useful conclusions or deductions – the meaning – can be gleaned from it.

In the study of the behavior of signs, as opposed to natural events, empirical methods are especially problematical. Partly in answer to this complaint, a Structuralist methodology was developed specifically for use in sign systems. It proved an exciting and productive alternative to empiricism for a short time, and then failed. My task here is to revisit the history of Structuralism in order to try to understand the cause of the failure, and to consider whether any part of it can be resurrected. I think it can, and I want to show how.

[1] [Editors' note: In the majority of cases, Caldwell used double quotes to mention a word or to cite a passage. And, as here, he used single quotes for so-called 'scare quotes'. For the sake of uniformity, we have imposed these quotational conventions on the book as a whole – including occasionally altering Caldwell's source texts, to make them conform as well.]



T. P. Caldwell, *Discourse, Structure and Linguistic Choice*, Studies in Linguistics and Philosophy 101, https://doi.org/10.1007/978-3-319-75441-3_1

## 1.1 Linguistic History

### *1.1.1 From Empiricism to Structuralism*

As scientific observation became conventionalized in Western civilization, the practice of observation came to mean a specific procedure: first identifying objective entities and then assembling 'facts' about them. 'Facts' are, I take it, true statements about the behaviors of these entities. Assemblages of facts with common 'attributes' are first used to distribute the 'facts' into 'categories' by a process called 'induction'. But what are these entities? For physicists, it was (for a while) the atom. For chemists, the molecule. For philologists, the word. While human speech is a natural phenomenon, in the sense that it is a manifestation of the physical voice box and articulatory apparatus, human language itself is not. But even in phonetics, a little thought makes it clear that identifying the fundamental entity of any science is not straightforward, but rather something that is complicit with methodological theory. Now, with modern equipment, we can visualize any piece of human speech as a squiggle on an oscilloscope, and even digitize the squiggle. The trouble is – and this is the problem with raw empiricism – every squiggle is different. Before we can use these digitized results, we must find a way of generalizing them into common principles. And even then we find no necessary correlation between the pure sound and the meaning indicated within a human language.

In short, empiricism immediately requires methodological theory. So phonologists were grateful when Baudouin de Courtenay, in the 1880s, proposed the idea of the 'phoneme', a minimal sound element that made a difference in meaning. But this unit was not a sound that was classifiable by its physical attributes, for its physical attributes were irrelevant. This notion turned out to spawn a whole new epistemology for the investigation of sign systems.

The man who most importantly exploited the importance of this discovery was Ferdinand de Saussure, whose book (made from lecture notes by his students), *The Course in General Linguistics*, created the idea of a *semiology*, a science of signs. It had two main features. One was the "arbitrariness doctrine" – the idea that the meaning had no relation to the physical features of the sound. Instead, meaning derived from the differential between one sign and all the others in the language. The other was the idea of the system itself – a synchronic, static slice of language use at any one moment in the ongoing flux of time. To see it as a complex set of abstract relationships rather than a random collection of speech acts, full of slang and dialect, helped greatly to bring order out of chaos.

Saussure's work allowed linguists to look at the problem in a new way – an essentially unempirical way – and became the most important contribution of what came to be known as Structuralist method. The idea of the differential meaning – the "emic" (as in "phonemic") as opposed to the "etic" (as in "phonetic"), in which meanings were systemic yet arbitrary, became the most successful idea in Structuralism.

Structuralism taught us that the structural relations of elements, not their physical characteristics and features, were the important thing. Saussure's great contribution was to see language not as a set of words and their historical meanings but as a system of semantic oppositions and contrasts. Louis Hjelmslev, Saussure's disciple, immediately set out to categorize distinctions between the content plane and the expression plane, between internal and exterior relations among signs, between language as system and as a process, between the syntagmatic and paradigmatic relationships among words. His vision of language as a *system* – by default a *single* system – a structure of interlocking syntagmatic and paradigmatic categories in a two-axis coordinate system – emerged as a compelling vindication of the methodologies then current, and their best illustration. By 1943, Hjelmslev predicted the universal acceptance of a seemingly unarguable set of assumptions – without which, it seemed, science itself could not exist:

> *A priori* it would seem to be a generally valid thesis that for every *process* there is a corresponding *system*, by which the process can be analyzed and described by means of a limited number of *premises*. It must be assumed that any process can be analyzed into a limited number of elements recurring in various combinations. Then, on the basis of this analysis, it should be possible to order these elements into *classes* according to their possibilities of combination. And it should be further possible to set up a *general and exhaustive calculus* of the possible combinations. A history so established should rise above the level of mere primitive description to that of a *systematic, exact, and generalizing science*, in the theory of which all events (possible combinations of elements) are foreseen and the conditions for their realization established. (Hjelmslev 1961, p. 9)

This marked a new concept of science – not exactly empiricism, but the next best thing: systematics. An explanation could claim objectivity if it could boast of systemic consistency. No longer were we to concentrate on words and their histories; we were to see the language at any historical moment as an idealized set of relationships. Every word had a value based on its relation to every other word in the system. Systems of differentials could, linguists believed, provide the basis for semantics, syntax, and even discourse-sized bodies of thought. It came to be realized that the two principles – the differential and the system – were co-dependent. That is to say, the differential was meaningful only in a context, and the concept of the Language as a System provided that context. Or so it seemed.

But results were mixed. For large discourse-level bodies of thought, the "emic" method was successful. Increasingly, the Structuralist method was applied to psychology, history, philosophy, sociology and politics, with exciting results. For Roland Barthes, for example, the ephemera of local fashion and popular art were always to be seen as versions of durable myths: *figures* finding their meaning against the *ground* of universal pattern. But it didn't work for language itself, either for the syntax or the semantics. Even with the best efforts of European linguists, especially those associated with the Prague School, where Saussure's influence was greatest, it couldn't generate the grammatical categories, and it couldn't predict the semantic meanings of words as found in dictionaries.

What could have been the problem? If this key structuralist idea had proved as successful in morphology, syntax, and semantics, as it was in the phonology, likely

there would have been no second-guessing it as a method of inquiry. But it wasn't. While the Europeans continued to work within the Structuralist method, the first instinct among American linguists was to reconsider the method. One flank, led by Leonard Bloomfield, returned to a raw empiricism in which there was no place for theoretical principle, philosophical generalization, or even explanation. The other, led by Jakobson, tried hard to protect the principle of arbitrariness, but had to admit that words and morphemes, unlike sounds, had objective and necessary meanings of their own without regard to their relations to other words. His example was the plural ending in German – a predictable sound that always had an objective meaning: the meaning of plurality. He concluded that what was true of phonemes was not true of any other part of the system. In 1943, he wrote the obituary for a pure Structuralism as the key to the semantics:

> So the phoneme, this cardinal element on which everything in the linguistic system hinges, stands in contrast to all the other integral parts of this system, and has a completely exceptional and distinctive character, a character which is not to be found in any entity analogous to the phoneme in the other sign systems. There is no entity similar in this respect either in the language of gesture, nor in that of scientific formulae, nor in the symbolism of heraldry, the fine arts, or ritual… Only the phoneme is a purely differential and contentless sign. The phoneme's sole linguistic content, or more generally its sole semiotic content, is its dissimilarity from all the other phonemes of the given system.
>
> Therefore language, in the narrow sense of the word, is distinguished from other sign systems by the very basis of its constitution. Language is the only system which is composed of elements which are signifiers and yet at the same time signify nothing. Thus the phoneme is the element which is specific to language… Language (in this sense of being constituted of phonemes) is the most important of the sign systems, it is for us language par excellence, language properly so-called, language *tout court*, and one might ask whether this special status of phoneme language is not due precisely to the specific character of its components, to the paradoxical character of elements which simultaneously signify and yet are devoid of all meaning. (Jakobson 1978, pp. 66–67)

I don't want to oversimplify Jakobson's careful and influential work on the phoneme in this brief summary. It is clear that he wanted badly to protect the Structuralist principle of the differential, but he also wanted to give his study a firm empirical basis. Thus, a 1949 article argued that Serbocroatian phonemes could be coded as combinations of the presence or absence (+ or −) of six distinctive articulatory features, including vocality, nasality, saturation, gravity, continuousness, and voicing (Jakobson 1949, p. 421). This meant that now linguistics could claim to be empirical, based once again on a fundamental entity, the "articulatory feature". The advantage of a computational taxonomy was apparent to all, and it reassured everyone that now linguistics was in the mainstream of modern science. As Jakobson said at the end of his article, "Linguistic analysis, with its concept of ultimate phonemic entities, signally converges with modern physics, which has revealed the granular structure of matter as composed of elementary particles" (Jakobson 1949, p. 425). In short, linguistics could be said to be like physics, in that both of them were based on atoms. That is, language too had its elementary particles: if not atoms, at least phonic elements – culturally distinct articulatory features.

But in 1955, he abandoned cultural distinctiveness by rejecting Saussure's and Hjelmslev's argument that the differential character of the phoneme did not depend on its phonic substance. By then he had found "inherent" features in universal physical articulation. With confidence that the whole world's phonemic production could be accounted for by 12 'inherent' features and a few 'prosodic' features, Jakobson was ready to claim universality for his theory of distinctive features as an information-encoding structure. He called it Markedness Theory. By subdividing the phoneme into distinctive features, Jakobson had redefined the differential as *binary opposition* (Jakobson and Halle 1971, p. 497*ff*). This achievement was an empiricist's dream. He found a way to *encode* the complexity of meaning as a simple mathematical formula.

### *1.1.2 Structuralism As System: Philosophy, History, Anthropology, Sociology*

With Jakobson's apparent success in identifying fundamental entities and returning linguistics to the realm of empirical science, linguists immediately set out to apply his method of feature analysis to the semantic system. Thus, Umberto Eco in his 1976 book *A Theory of Semiotics* concluded that just as a phoneme is a "bundle of more analytical distinctive features … [so] the same internal network of mutually opposed features should also rule the differences between two sememes" (Eco 1976, p. 84). This effort, like Hjelmslev's effort to codify the morphology in the same way, indicates a cardinal tenet of the empirical method – that whatever the structure of language is, it must be *systematic*. That is, it must follow *one* set of rules, because the truth is one, and because the system must be unified.

This emphasis on *systematicity* has become part and parcel of what we understand to be the scientific method. It is not part of the original meaning of empiricism, which simply meant "based on experience without regard to theory". While the idea of Language as a single system of relationships was understood by Saussure as a convenient heuristic which made analysis possible, he understood that it was an idealization, not a fact. What we have to account for is that Structuralism as a method for analyzing systems has failed in describing either the syntax or the semantics of ordinary language, despite Eco, even though it was extremely powerful when used for large philosophical and cultural discourses.

Thus, systematicity and its attributes have been at the focus of much thought since then: witness Structuralist philosophy by Roland Barthes, anthropology by Claude Lévi-Strauss, history by Michel Foucault, and narrative systems by A.J. Greimas. At first it appeared, with the influence of Karl Jung's archetypes, that systematicity implied universality, as if human culture had fundamental grounding in genetics or a kind of 'human nature'. That is to say, structures with well defined centers and settled boundaries. But when their centeredness was questioned by Post-structuralist philosophers such as Derrida, suddenly a great many stable ideas became destabilized, with political and ideological repercussions in every field, even science.

Derrida's argument was one from phenomenology: that large systems are not "centered" because of their own objectivity or universality, but merely because they had been "constructed" within a single political or cultural point of view. By critiquing the oppositions underpinning such systems (male/female, for example, or selfish/altruistic or socialist/capitalist) he showed that they aren't so systematic when you remove their exterior political and cultural supports: rather, they become free-floating, even self-destroying relativities which reveal their counter-arguments as well as their arguments. Post-structuralism became enormously influential in the 1970s as philosophers learned from him how to de-center and deconstruct history, sociology, and political science. As the constructedness of these systems became apparent, various ideologies rushed to re-construct them, especially post-colonialism, Marxism, and feminism. With their emphases on race, class and gender, all of these fields seemed by the 1980s to have become branches of a radical kind of cultural anthropology preaching a kind of "transformational" and "liberational" ideology that is still current in some places.

### *1.1.3 The Move to Computationalism*

In the scientific enterprises, especially in linguistics, systematicity took a different direction. Chomsky took Saussure's idea of *la langue* almost literally, proposing a frankly Cartesian kind of mentalism that almost denied empiricism altogether. For him, Saussure's idea of language as an idealized system was to be taken not as methodological heuristic but as axiom (Chomsky 1965, pp. 28*ff*). It enabled him to convert language competency to a set of logical rules, all syntactical relations being necessary rather than contingent, and thus to see language as a set of blind computational formulae.[2] Insofar as they pretended to be empirical, his formulae claimed to account for language 'facts' in the form of grammatically possible individual sentences. For Chomsky and his descendants, all relations in language are syntactic ones, and sooner or later a Universal Grammar, a UG, will be found to lie at the bottom of all languages human beings are capable of learning. In more recent formulations such as Optimality Theory, the UG is a purely computational set of rules, from which the grammar of any language can be derived by a particular set of constraints on the UG. Now the dominant school seems to be Cognitive Science, whose goal is to describe all mental functioning as based in a purely computational 'mentalese' (*cf*. Pinker 1994) – universal, genetic, hard-wired, and computable.

[2] For criticisms of the formalism from a formal point of view, see Quine (1972).

## 1.2 Reassessing the Problem

One recent reaction to the politicization of post-structuralism and the failure of Structuralism to account for the semantic and syntactic systems of language is to retreat into a kind of pure empiricism. Paul Hopper and Sandra Thompson, for example, seem to have backed away from theory in a radical way, questioning not only the idea of a semantic system, but even the concept of grammar itself, seeing it as a fuzzy, derivative, weak tendency toward patterning perhaps, but not as a coercive set of rules (Hopper and Thompson 1984).[3]

I think they are on the right track but for the wrong reasons. Their instinct is to move away from system toward data. They distrust Structuralism as a method because they identify it with the kind of systemic philosophy that led to Cognitivism. But I think any move back toward a pure empiricism in the study of syntax or semantics would be a mistake.[4] For that means to stay within the *etic* view of things. In order to clarify why it is a mistake I must return to Jakobson.

When Jakobson found that all phonemes could be accounted for as bundles of universal phonetic features, he took himself to be accomplishing two important tasks: to relate physical features to linguistic meaning, and to provide a computational basis for the science. But the 'universal' phonetic features told us only about the commonalities of human mouths over all the world. All people have the same sound-producing structures, so the sounds they are capable of producing are, more or less, the same the world over (not that there isn't a huge variety in their choices of the meaningful ones). The second objection is that Jakobson related these sounds to *information*, but not to meaning. In retreating to the *etic,* Jakobson abandoned the *emic,* and the *emic* is the source of meaning.

In short, the desire for scientific respectability led scientists in the twentieth century mostly in two directions – one the empirical, the other the systemic, both reflecting the received wisdom of classical reasoning as either concrete or universal, and both with the received faith that they are related logically to each other. Indeed, the very idea of induction requires such an assumption. The data, we think, must be a manifestation of the system, and the system must be implied in the data.

But Structuralism offered us an alternative view, and we should not have abandoned it so hastily. It should have taught us that language is a human invention, a manifestation not of universal logic, but of our synthetic minds. The prejudice in favor of systematicity led early structuralism into a dead end. Jakobson couldn't make the differential work as a source of meaning within the assumption that Language was a single system, so he gave it up. I think he should have entertained

[3] Paul Hopper once told me in a personal e-mail that this article had been described by one of his colleagues as the first "postmodern" article in linguistics. This characterization suggests that grammar can be deconstructed by the same methods as any other ideology or discourse, and perhaps should be. My take is that it shows that grammar is a set of conventionalizations rather than a set of rules.

[4] Of course, in many branches of linguistics – dialectology, discourse analysis, and other kinds of applied linguistics – empiricism is still the method of choice, as it should be.

the idea that language expresses the structure of mind, not the structure of logic. Human consciousness is synthetic, *gestaltish,* contingent, and sequential; not analytic, categorical, necessary, or syllogistic.

## 1.3 A Counterproposal: Redefining the Sememe

Jakobson should not have given up the *emic* principle of structuralism in favor of the *etic* principle of empiricism. I want to suggest a way to resurrect the differential as a tool of language study. It requires two simple adjustments to the classic Structuralism program. One is to realize that Language is not a system, but rather simply a collection of tools for creating discourses. It is within particular speech acts, not language as a whole, that we can find semantic consistency. The 'systems' within which differentials define meaning are, I believe, discourse-sized systems. Of course there are words whose meanings carry consistency from one discourse to the next, but that is the result of an unsystematic process of conventionalization.

Louis Hjelmslev identified the two dimensions of discourse as the syntagmatic and the paradigmatic. I propose a small adjustment in our thinking about each of these dimensions. First, we must recognize the contribution of *salience order* to syntagmatic structure. Every act of speech begins by locating itself within something familiar in the ongoing public discourse, and then moving by controlled degrees of increasing specificity to a clear argument or predication. This is more easily seen in Japanese than in English. No Japanese nouns are marked for number, gender, case, or determinacy. They are truly abstract entities which are easily identified in a dictionary, but require additional markings before they are used in discourse.

We should think of English in the same way. When we use nouns in discourse, we mark them with determiners, number indicators, prototypicality,[5] and word order to indicate their rank in a salience order.[6] Salience order is, in essence, the relevance order of the elements in the discourse. The most-salient element – the most-marked or the most-recently-marked element – is the point of focus at any moment in the generation of discourse, and I suggest that it is at that point that Hjelmslev's second dimension, the paradigmatic dimension, becomes operative.

[5] In the sense of Eleanor Rosch's theory of prototypes, as applied to grammatical categories. That is, some nouns are more "nouny" than others. For example, in the phrase "six-foot fence", "foot" is a noun in a subordinate position, and cannot be made plural like nouns in more salient positions can. It can be said to be less prototypically a noun in this position, and this is why I suspect prototypicality indicates rank in discourse salience. See Rosch (1975) and (1978), Lakoff (1987), and Hopper and Thompson (1984).

[6] This may be the answer to Jakobson's argument about German plurals [See p. 18 and p. 68 below – Eds.]. That is, it may be possible to argue that the morpheme indicating plurality does not have an "objective meaning" but is a salience-order indicator in a discourse-level structure. The plural is usually less salient than the singular. The covert structures indicating salience order are more fully described in Caldwell (2002).

The point of the second adjustment is to see that the semantic principle of the differential operates within small universes, not large ones. That is, a word means what it means by its differential not from all the other words in the language, but from *all the other words that could have been chosen in its place at any salient moment in any actual discourse*. Its meaning is formed in its contrast with the other possibilities in a tiny synthetic unit, which I call the Molecular Sememe.[7] Since the discourse itself selects those other possibilities, this differential can express the *local* meaning of the *local* discourse, and contains all the nuance and fineness of distinction that language is capable of expressing at its most concrete levels. Thus, I propose a kind of micro-structuralism that accounts fully for the local meaning of the utterance, even when it involves irony, humor, wit, or any of the subtleties we know discourse is capable of.

The principle of local differentiation also underlies the syntax. The syntax of English, as Hopper and Thompson argue, is not a single coherent system. Its systematicity is weak, better described as a kind of conventionalization rather than something ordered by logical necessity. But there is system to be found at the level of individual discourse, where the contrast is *emic* rather than *etic*. Since all the possibilities within a single set are necessarily of the same grammatical value, whatever that is, we can say that the grammatical value as well as the semantic meaning belong to the sememe itself, *as marked by* the word chosen to represent it.

This structure emerges as soon as we recognize that the ordering principle is not a small, atomistic unit, but rather the salience structure of discourse itself. Ordering at the discourse level is top-down rather than bottom-up. Discourse salience – and its coercive power over semantic possibility – is an expression of the ability of our sensory faculties to organize data as *gestalt* perceptions rather than as analytic categories. Its emphasis on the differential – that is, the *emic* rather than the *etic* – makes use of the mind's ability to spot similarities and differences among complex entities, to register fine distinctions among similar things, and to discover similarities among dissimilar things. It registers human cognitive needs like aim, purpose, emphasis and intention, not just grammatical relation.

In short, both semantic and syntactic structures are to be found, exactly as the value of a phoneme was found, in its contrast to all the other entities with which it is compared. When we consider the whole language as a single system, it cannot be done. But in the case of a single discourse, the differential has far more clarity, and its power to organize both semantic meaning and grammatical value is clear.

Discourse is a highly coercive organization of meaning, governing both semantic and syntactic structures within it. Semantic structure is not consistent among varieties of discourses, but it is coherent within a single discourse. This fact, I argue, allows the construction of a purely synthetic, non-computational, emic model of ordinary language. Though the model I propose is a discourse-level structure rather than a Language-level structure, it should fulfill the expectations of Ferdinand de Saussure for a single account of the syntactic and semantic planes of ordinary language.

[7] The dynamics of the Molecular Sememe are detailed elsewhere. See Caldwell (1989), (2000) and (2004). [Reprinted here in their final form as Chapters 2, 6 and 3 respectively – Eds.]