Historigraphia Linguistica XX:2/3.355-398 (1993).

A Minimalist Program for Linguistics *

Bruce E. Nevin
Bolt Beranek & Newman/University of Pennsylvania

1. Introduction

Zellig Harris has an established place in what commonly passes for historiography of linguistics. In this now customary view, he was perhaps the most extreme among the so-called Post-Bloomfieldians who sought to devise discovery procedures whereby a grammar could be derived from distributional analysis of a corpus of utterances without reference to meaning. Taxonomic linguistics, as this has been called, would start with a corpus of phonetic records of speech and proceed by alternating steps of segmenting these records, classifying the segments according to their distributions relative to one another, and representing the data in terms of the resulting class-labels for another round of segmentation. It was scruples of empiricism (logical positivism) and behaviorist psychology, it is claimed, that forbade consideration of meaning.[1]

A responsible assessment of Zellig Harris' work cannot fairly begin until it has first been made obvious — not merely contended but made self-evident — that this now familiar picture is fundamentally wrong about Harris: not only that it asserts things that are not true,

* I am grateful for the helpful comments of Professors Michael Gottfried, Henry Hoenigswald, & E. F. Konrad Koerner on earlier drafts of this paper. Responsibility for errors of fact and interpretation rests, of course, with me. [Typesetting errors corrected 3/2015, original pagination not retained.]

1 This characterization is so familiar and well established as scarcely to need citation. It has pervaded even the literature of popularization, for example, Campbell (1972:171):

> Chomsky […] studied at the University of Pennsylvania under the American structural linguist Zellig Harris, who not only deplored the intrusion of meaning into the science of language, but made every effort to shut it out completely by using mechanical methods of description that computers, which were not very intelligent at that time, were able to process.

Campbell gives no source for this view, which evidently appears to him to be part of the unquestioned background of his reading and conversations with linguists (he does cite conversations with Ray Jackendoff).

The pervasiveness of this caricature in the literature may be seen in R. A. Harris (1993), e.g., "a cornerstone of transformational research right back to Harris" (p. 428), as though he were only an historical figure of the 1950s whose work ended there; accounting Harris among the "Post-Bloomfieldians, from the linguistics boneyard" (p. 429); "the asemantic transformational granddaddy, Zellig Harris"(ibid.).

but also that it misses entirely what it was that Harris was after, and why he went after it in the way that he did.

Such an assessment, and even the prolegomenon to it that I have just proposed, is far too large a task to attempt here.[2] Instead, I will select one theme for discussion, namely, Harris' conception of meaning and linguistic information in respect to the methodology that he developed and demonstrated. I will touch on other themes only in passing, partly for the sake of context and partly to indicate the lines that such an investigation might take. (These tangential themes are summarized in Section 11.)

2. Motivation: Absence of an independent metalanguage

We will begin with the motivation for Harris' particular concerns regarding the methodology of linguistics. He observed that linguistics is set apart from the other sciences by a certain ineluctable characteristic of its subject matter. In any science (including linguistics) researchers communicate findings, propose interpretations, and reach agreements in an evolving consensus. To do this, they use language, especially but not at all exclusively the specialized science sublanguage[3] of their field. For any science, then, scientists use their shared language to reach agreements about a shared point of view on the subject matter of the science, prior to and apart from the science itself. But for linguistics the shared point of view (the agreements) attained by means of language cannot be prior to and apart from the subject matter of the science, which is after all language itself.[4] This, then, is the characterizing dilemma at the foundations of linguistics, and

---

2 This task must be taken up by others better situated than I in terms both of knowledge and resources. For an excellent beginning, see Ryckman (1986), on which I have depended heavily for its masterful survey of the literature. Anyone could verify claims made here about Harris by studying what he has written on its own terms, without undue preconception. Ideally, the present brief exploration might provide entrée into what "its own terms" are, on which Harris' oeuvre must be understood, and might indicate some of the hedge of preconceptions surrounding his work that must be cleared away before it canbe rightly understood.

3 For the present, a science sublanguage may be thought of informally as a specialized technical jargon. For more precise definition and discussion of sublanguage grammar and of science sublanguages in particular, see Harris et al. (1989), Harris (1982b, 1988, 1991), Grishman & Kittredge (1986), Kittredge & Lehrberger (1982), and Sager (1986).

4 This is of course, as said above, an evolving consensus, and as we shall see the specialized forms of language used in a science appear to have a constitutive role in the development of the subject matter of the science.
For any science but linguistics, one may use ordinary language as a metalanguage to define the objects and constructions of the science. For example, the talk and writing in which endocrinologists formulate, present, and discuss findings in their field is clearly within the language that they share with non-endocrinologists, however specialized for their work. The talk and writing in which linguists formulate, present, and discuss grammar is likewise necessarily part of the shared "background vernacular" of language. It goes without saying that no science is at liberty to *assume* that its objects and constructions correspond to reality (or, rather, to prior agreements about reality). But precisely such an assumption is entailed covertly if we take the same liberties with ordinary language on behalf of linguistics as scientists in other fields are privileged to do on behalf of their sciences. The unique problem for linguistics is the

Harris' early recognition of it informs all his work as linguist and methodologist.

It is not clear to me how early Harris realized the import of this fact for linguistic theory and methodology. Munz (1972:269) provides some historical perspective on the observation that a language necessarily includes sentences stating anything that can be said about it metalinguistically ("that the language contains virtually unrestricted metalinguistic devices"), and suggests that the importance of this observation did not become clear to Harris until the 1960s. On internal evidence, some of which Munz cites, I am confident that the distinction between linguistic information and other types of meaning was reached quite early. This distinction was already nascent in writings of Bloomfield and in Sapir's trenchant explorations of individual participation in social pattern.[5]

Harris was by no means the first to observe that "one can form in any natural language various sentences that speak about that language (or about any other), about the combinations of partsin occurrences of the language, and the like" (1991:274). However, he took this observation further, recognizing that any grammar (definition, description, etc.) of a language consists of statements about the language that are made in the metalanguage contained in a natural language (or that depend upon statements in the metalanguage):

---

unavailability of a metalanguage for the subject matter of linguistics that is *separate* from and prior to that subject matter. This is of course because the subject matter of linguistics is language itself, and language as a long-observed fact does contain one of its own metalanguages. Harris developed methods for describing language that do not depend upon a metalanguage imagined to be prior to and external to language itself.

5 It may be that Harris' commitment to formal methodology was initially for the sake of the "detached pattern" in language without full apprehension of the eventual fruit of this long line of research. Or it may simply be that he played his cards close to the vest, reluctant to indulge in speculation about theoretical ramifications until results were more firmly in hand. (See the discussion below of explanatory principles external to language, and how easily a plausible "way of talking" can seduce one into prematurely abandoning a line of research leading to a deeper explanation — a peril which he seems to have appreciated at an early stage (e.g. 1941:707). Harris' conception of the proper conduct of scientific inquiry is another theme calling for research and discussion.) But even if the chronology of Harris' motivations should remain indeterminate for the historiographer, nonetheless because of the consistency of those motivations and the integrity with which he enacted them throughout his long career, his work is enormously easier to understand with the benefit of hindsight, knowing now where he got to with it. And with hindsight, the numerous misunderstandings and misconstruals of Harris' work stand in vivid relief. The dilemma posed by the absence of a language-independent, external metalanguage was the motivation, evidently, of Bloomfield's well-known rejection of mentalistic terminology in favor of studying the correlation of linguistic form with the meanings that such terminology was intended to identify, as reflected for example in (1939 loc. cit.): "It is the belief of the present writer that the scientific description of the universe […] requires none of the mentalistic terms, because the gaps which these terms are intended to bridge exist only so long as language is left out of account." See also Wells (1962:708):

> What Bloomfield was unclearly driving at in his well-known attack on mentalism […] was that propositions connecting a way of speaking with a way of thinking are not empirical but a priori propositions, unless there is some logically independent evidence for a given way of thinking other than the way of talking itself.

> There is no way to define or describe the language and its occurrences except in such statements said in that same language or in another natural language. Even if the grammar of a language is stated largely in symbols, those symbols will have to be defined ultimately in a natural language (ibid.).[6]

2.1 Form, information, and meaning

Harris' claim, then, is that linguistics as a science cannot have recourse to a metalanguage external to language and independent of it, and so can only define and describe the informational properties of language by employing those very properties to do so. This is not merely a methodological problem for linguistics, serious though that is; it is an important key to the essential nature of language. As we will see, it means that the information in an utterance is a function of redundancy in language and in the particular utterance. At first blush, this may seem trivially true, almost a tautology. But it has an important consequence as to what is a possible theory of linguistic semantics: the information in utterances cannot be a function of its being encoded into language from some prior representation of meaning, mental or otherwise.

In the conception taken for granted by many writers, language functions as or is used as a code, and that is how linguistic forms come to be correlated with meanings. The correlation itself, which is presupposed without question as something intuitively self evident, entails a distinction between forms and meanings as independent entities requiring correlation. This autonomy of linguistic form with respect to meaning Harris denies.[7]

---

6 A grammar stated in the metalanguage vocabulary and syntax of some other natural language (say, a grammar of Achumawi written in English) could as well be written in the same language, given suitable vocabulary (even borrowed if need be). The vocabulary and syntax required for a least metalanguage are quite limited (Harris 1991, Chap. 10; see also 1991:31-32).
Tabulations and mathematical or other symbolic formulations are routinely "spelled out" in oral presentations of any science. Even mathematics depends upon the "background vernacular"of language, as noted by Borel (1928:160), quoted in Ryckman (1986:289-90). And as Borel observes, the crucial role of language in all this is usually unnoticed.

7 As did Bloomfield, for instance, where he stated (1927:142) "In the study of linguistic forms, therefore, I should not appeal, as Jespersen sometimes does, to meaning as if it were separable from form", or further:

> We assume that each linguistic form has a constant and definite meaning, different from the meaning of any other linguistic form in the same language. […] In the rough, however, our assumption is justified by the mere fact that speakers co-operate in a very refined way by means of language-signals. In describing a language, we are primarily concerned with the working of this co-operation at any one time in any one community, and not with its success and failures or with its changes in the course of history. Accordingly, the descriptive phase of linguistics consists in a somewhat rigid analysis of speech-forms, on the assumption that these speech-forms have constant and definite meanings. (1933:158)
>
> In language, forms cannot be separated from their meanings. It would be uninteresting and perhaps not very profitable to study the mere sound of a language without any consideration of meaning […] In studying a language, we can single out the relevant features of sound only if we know the meaning. This appears plainly when one confronts an unfamiliar language […] It is only the differences of meaning which decide that most of the inevitable variations of sound are irrelevant and only certain ones play a part in communication. In short, the significant sequences of sound (the *phonemes*) of a language are ... those which involve a difference of meaning. (1943:401-402)

It is not that distributional pattern on the one hand and meaning on the other are correlated somehow, as independent factors. What Harris calls *linguistic information* is indistinguishable from linguistic form because, indeed, it is constituted by it. This is necessarily so, for there is no means for representing the information in language that is not (or does not rest upon) language itself: no separate metalanguage.[8]

Linguistic information/form is conventional, socially shared. We as individuals apparently use this linguistic information to help organize our perceptions in associative memory and imagination. The preponderance of our perceptions are "meanings" that by contrast with linguistic information are subjective, idiosyncratic, notoriously not socially available or communicable in any reliable way. Meanings in this broad sense, which we perceive in all our activities and observations, are indeed correlated with linguistic information (identically, with linguistic form), but variably, idiosyncratically, without the reliability of social standardization that is the hallmark of linguistic information/form.

This is not to say that Harris proposes that there are two kinds of meanings, far less that he made such a distinction a priori and decided to investigate one of them, denominated linguistic information. Linguistic information is distinguished as a product of linguistic analysis. Other kinds of analysis, perhaps in conjunction with linguistic methods, might distinguish other kinds of meanings, as suggested for example for ethnographic analysis with respect to style and social or cultural meanings in Harris (1952c).

We presumably do correlate the socially standardized linguistic information in an utterance with the perceptions (current, imagined, and remembered) that constitute "meanings" for usin a more vivid and individualized sense. Certainly, we make just this presumption every time we use language with the intention of evoking such private perceptions in others, or with the intention of divining what perceptions of theirs may be relevant to our purposes. However, and despite our best intentions, we repeatedly find that the world of nonverbal perceptions is by and large much more subject to idiosyncratic variation and difference, between ourselves and others, and even within ourselves, than is the highly patterned domain of social conventions, especially those conventions constituting language. "Oh, is that what you meant", we say.[9]

The notion of the autonomy of linguistic form from meaning is a development in Generativist theory, not traceable to Harris, much less to Bloomfield and Sapir.

8 On this point, Henry Hoenigswald (p.c.) observes: "Synonymy may not be the best example because of its degeneracy, but do synonyms have near-identical privileges of occurrences because they have near-identical meanings, or are items with near-identical privileges said to have near-identical meanings, and why would it make a difference? The hunt for discourses—words and sentences only insofar as they happen to function as discourses—that accommodate one item but not the other is, in fact, something like a pair test. Harris was creatively generalizing from Bloomfield's notion of `grammatical meaning'—an area where the unity of distributional description (in the grammar) and translating or paraphrasing (in the dictionary; itself a form of distributional statement) is traditional when it comes to things like `particles'."

9 Linguistic information is "in" utterances, socially presented in a way that is simply not available for the memories and imaginings that speakers and hearers may idiosyncratically associate with utterances, despite our assurances to one another that we have reached this agreement or that over our respective perceptions.

> Language is clearly and above all a bearer of meaning. Not, however, of all meaning. Many human activities and states have meaning for us, and only some of this can be expressed in language: feelings and vague sensings can be referred to only indirectly […]; non-public information, such as proprioceptive sensations, can be named only with difficulty; certain kinds of non-language information can be translated directly into language (e.g. graphs and charts); but other kinds (e.g. photographs) can be represented in language only loosely and selectively. Meaning itself is a concept of no clear definition, and the parts of it that can be expressed in language are not an otherwise characterized subset of meaning. Indeed, since language-borne meaning cannot be organized or structured except in respect to words and sentences of language […], we cannot correlate the structure and meanings of language with any independently known catalog or structure of meaning. In each language, we do not know a priori which specific aspects of meaning will be referred to by words, and how much will be included in the meaning of a single word. Even when the meanings are well-defined, it is not always possible for words of a language to mirror in their structural relations the relation among the referents. For example, the relation among the integers is given in Peano's axioms, but language cannot thereupon name them as *one*, *successor of one*, *successor of successor of one*, etc. What languages do in this case is to give arbitrary names to the first ten (or six, or whatever), and then to use these and other names to indicate successive multiples of the initial set: the integers are thus named modulo that initial set, by a language-based decimal or other expansion. Such cases suggest that we have to study the specific words and structures of a language if we wish to see what meanings they cover, and how. (Harris 1991:321)[10]

---

10 Some earlier representative quotations show that this was a view long and consistently held:

> Meaning is not a unique property of language, but a general characteristic of human activity. It is true that language has a special relation to meaning, both in the sense of the classification of aspects of experience, and in the sense of communication. But the relation is not simple. For example, we can compare the structures of languages with the structure of the physical world (e.g. the kind of phenomena that are expressed by differentiation and integration in calculus), or with what we know about the structure of human response (e.g. association, transference). In either case, it would be clear that the structure of one language or another does not conform in many respects to the structure of physical nature or of human response—i.e. to the structure of objective experience from which we presumably draw our meanings. And if we consider the individual aspects of experience, the way a person's store of meanings grows and changes through the years while his language remains fairly constant, or the way a person can have an idea or a feeling which he cannot readily express in the language available to him, we see that the structure of language does not necessarily conform to the structure of subjective experience, of the subjective world of meanings. (Harris 1954:8)
>
> The discrete and socially fixed pre-set words are necessarily arbitrary sound-sequences: they cannot be determined by the continuous and changing world of meanings and by individually differing experience. (Harris 1966:603)

A different perspective, from anthropology and animal ethology, is brought by Gregory Bateson:

> [The] decay of organs and skills under evolutionary replacement [by fitter organs and skills] is a necessary and inevitable systemic phenomenon. If therefore, verbal language were in any sense an evolutionary replacement of communication by means of kinesics and paralanguage, we would

Note carefully that this is not part of an argument for exclusion of meaning from linguistics on the grounds of it being too indeterminate and variable (as the "taxonomic linguistics"straw man would suggest). The point is more fundamental, and has to do with the relation between perception and thought on the one hand and language as a public institution on the other:

> [T]he very simplicity of this system, which surprisingly enough seems to suffice for language, makes it clear that no matter how interdependent language and thought may be, they cannot be identical. It is not reasonable to believe that thought has the structural simplicity and the recursive enumerability which we see in language. So that language structure appears rather as a particular system, satisfying the conditions of Chapter 2 and perhaps also bound by a history, which may evolve not only in time but also by specialization in science languages, and which is undoubtedly necessary for any thoughts other than simple or impressionistic ones, but which may in part be a rather rigid channel for thought. (Harris 1968:216)[11]

It is the constitutive capacity of language, creating a kind of information that cannot exist without language, that Harris is after. But to get at it, he cannot presuppose it, not only because of the characterizing dilemma of linguistics already noted, but also because linguistic information does not exist apart from linguistic form; it is simply not to be found in the world of real, imagined, and remembered nonlinguistic perceptions that constitute "meanings" in the broader sense for each of us idiosyncratically. The illusion that meaning or information can be considered separately and brought into correlation with the forms of a given language is an artifact of translation, either from the investigator's different native language or from some language-like system of mathematics or mathematical logic. In the first case, the investigator's native language constitutes in its structure the linguistic information that the investigator then takes as meanings supposedly given prior to the forms of the language being investigated. In the second case, the investigator can have learned (or created)the formal "language" and the meanings assigned to its elements and structures only by depending upon linguistic information in an antecedent natural

---

> expect the old, preponderantly iconic systems to have undergone conspicuous decay. Clearly they have not […] The logician's dream that men should communicate only by unambiguous digital signals has not come true and is not likely to.
>
> I suggest that this separate burgeoning evolution of kinesics and paralanguage alongside the evolution of verbal language indicates that our iconic communication serves functions totally different from those of language and, indeed, performs functions which verbal language is unsuited to perform. (Bateson 1968:411-412)

See now Harris' discussion of animal communication at (1991:382-3).

11 In the indicated place, Chap. 2 of his (1968), Harris notes "certain apparently universal and essential properties of language, which are observable without any mathematical analysis, and which are such as to make possible a mathematical treatment" (1968:6). These are that language elements are discrete, socially preset in speaker and hearer, and arbitrary, that combinations of these elements are linear and denumerable, that not all combinations constitute a discourse, that operations are contiguous, that the metalanguage is in the language, and that language changes. Each of these properties has metatheoretical consequences for linguistics.

language, and subsequent claims as to the meaningfulness of its formulaic expressions have the same dependency. In both cases we are talking about glosses and not meanings.

The informational properties of language — regularities of linguistic form — are notoriously not all obvious to a superficial inspection of speech. How can we enquire into them and disclose them without begging the question? It was to this question that Harris sought answers.

2.2 Consequences for a science of linguistics

The absence of an external metalanguage has important consequences, then, for linguistics. First, it shows the true motivation for Harris' distributional methodology:

> In the absence of an external metalanguage, the entities of each language can be identified only if the sounds, markers, or words of which they are composed do not occur randomly in utterances of the language. That is, the entities can be recognized only if not all combinations occur, or are equally probable. This condition is indeed satisfied by languages. A necessary step, then, toward understanding language structure is to distinguish the combinations of elements that occur in the utterances of a language from those that do not: that is, to characterize their departures from randomness. (1988:3)

In addition, it imposes a requirement on the results of linguistic analysis:

> This task entails an important demand: it calls for a least description, that is, for a characterization of the actually occurring combinations by means of the fewest and simplest entities and the fewest and simplest rules and conditions of their combination, and with no (or least) repetition. The reason for this demand is that every entity and rule, and every complexity and restriction of domains of a rule, states a departure from randomness in the language being described. Since what we have to describe is the restriction on combinations in the language, the description should not add restrictions of its own. If two descriptions, one more efficient than the other, characterize the same data, then the less efficient description must have overstated the actual restrictions in the language — by overstating and then withdrawing part, or by repeating a restriction, or whatever. (1988:3-4)[12]

12 Harris makes the same point in earlier writings, for example:

> The fact that particular kinds and amounts of redundancy are essential parts of language structure makes it important that a description of language should not add its own redundancy to the picture. A theory of language should not contain elements of wide combinability and then specify which combinations are language. It should contain elements of just such combinability as appears in the language itself. (1968:12fn16)
>
> [T]he grammatical description [must be] as unredundant as possible so that the essential redundancy of language as an information-bearing system […] not be masked by further redundancy in the description itself […] More generally, one must recognize that every new term or category or subclass that is not derivable from the primitives of the system, and also every rule, including every limitation on the carrying-out of a rule, and every ad hoc explanation is a redundancy of the description. (1982a:10-11)

Only certain well-structured aspects of meaning, then, correlate directly with linguistic forms, and the methodological problem[13] is that one cannot presuppose the former (language-correlated structure in the world of perceived meanings) as a way of getting at the latter (the linguistic forms). The reason for this is that it is by way of the latter (language with its public, institutionalized constraints) that people constitute the former in socially standardized ways. Thus, the methodological problem and the aim of Harris' research are two sides of the same coin:

> In a consideration of language structure, the issue of information enters because language is clearly a carrier of meaning and information. In the present theory it enters even more directly, because as we approach a least grammar, with least redundancy in the description of the structure, the connection of that grammar with information becomes much stronger. Indeed, the step-by-step connection of information with structure is found to be so strong as to constitute a test of the relevance of any proposed structural analysis of language. This suggests that the components that go into the making of the structure are the components that go into the making of the information. (1988:57)

The condition that there is no external metalanguage independent of language itself, and the requirement that this condition imposes for a "least grammar", together determine the outline and direction of Harris's research, or at least the last four or five decades of it.

3. Socially constituted meaning

Two presuppositions of the preceding observations should be stated more explicitly. First, Harris shared with Bloomfield[14] and Sapir[15] the recognition that language is a social artifact. The `detached pattern' in language[16] is what makes it learnable as a social given, without which human cooperative action would come virtually to a standstill.

Secondly, I surmise that (like Bloomfield) Harris saw the formal analysis of grammar as an

---

The application of this principle to a critique of alternative theories as a "test of relevance" (1988:57, quoted below) might be a good topic for further research. See also (1991:33-36), quoted in Section 8, below.

13 This methodological problem is usually misconstrued as an exclusion of "mentalist" concepts of meaning on empiricist grounds. Harris was neither a behaviorist nor a logical positivist. His philosophical affinities lay more with the naturalism of Dewey. (On Harris' naturalism, see Ryckman 1986:288-309.)

14 Bloomfield's commitment to social psychology and to the fundamentally social character of language did not end when he exchanged Wundt (in 1914) for the behaviorists (in 1933) as his model of what psychology had to offer the linguist. Then as now, it was clear that linguistics had at least as much, probably more, to offer psychology. See e.g. Bloomfield (1927: 174, 179; 1942:397).

15 Sapir's commitment to the essentially social character of language is perhaps better known than that of Bloomfield. See for example "Language is primarily a cultural or social product and must be understood as such", etc. (Sapir 1929:166).

16 Harris uses the term `detached', with sense similar to Saussurean `arbitrary', in his (1951b) review of Sapir's work. This review deserves careful study not only for its obvious value as an in-depth survey of Sapir's views and methods by one of his students (one whom Sapir regarded, according to at least one of his children, as his intellectual heir (Victor Golla, p.c.; R. Darnell, p.c.)), but especially,in connection with our present interests, for the insight that it gives into Harris' views and methods.

aspect of the study of semantics,[17] and with Sapir he saw that socially instituted form or pattern, especially that in language, constitutes aspects of meaning that could not exist without it. (See Harris 1951b, the long review article on Sapir1949.) Harris saw, as Sapir had seen, that people have historically used language to build up, over generations, types of meaning that did not previously exist and in fact could not exist without language. This linguistic information is conventional, and it is objective to the degree that it is public, transmitted and shared by means of language, as distinct from the private, subjective world of perceptions apart from language. (Individuals in understanding utterances obviously do associate meanings of the latter sort with utterances—that is, with linguistic information—and in a way that is clearly related to it, but not in a socially standardized way.)

3.1 Contrast as social fact

The basis of language structure in social convention is seen first in the fundamental distinction between repetition and imitation. For Bloomfield, it was a "fundamental assumption of linguistics[…] that *in every speech-community some utterances are alike in form and meaning*" (1933:78, emphasis in original). Members of the speech community perceive such utterances as repetitions. Yet they are not repetitions by virtue of imitation. How is repetition different from imitation?

Harris devised a simple operational test, the pair test for what is a repetition, so that the observation that some utterances in a language are "alike in form and meaning", i.e. are repetitions, is an assumption no longer. This test rests on the crucial, social fact that something conforming to a norm or convention may be repeated by producing another token of it, whereas something not participating in an institutionalized type-token relationship may only be imitated. (This differs too from the type-token relation seen in ordinary category perception, where we speak of another X, i.e. another member of category X, rather than a repetition of [category] X.)

Harris equates Bloomfield's `assumption' with the linguist's judgment, on the basis of

17 Bloomfield saw grammar as a topic under the heading of semantics (1933:138):

> When the phonology of a language has been established, there remains the task of telling what meanings are attached to the several phonetic forms. This phase of description is *semantics*. It is ordinarily divided into two parts, *grammar* and *lexicon*.

Bloomfield put no reliance on psychology (either in his later use of behaviorism, or in his earlier use of Wundt) beyond expository convenience (1933: vii-viii), as a kind of place-holder for a science of psychology that did not yet exist. But this did not mean that meaning was to be banished. Pace the now customary allegation that he excluded meaning from the scientific description of language (e.g. Katz 1972:xxii, Fodor 1980: 248), Bloomfield repeatedly made it clear that considerations of meaning are central to both the practice and theory of linguistics, for example (1933:27, 78, 128; 1939:18). See also Koerner (1970), Ryckman (1986:28-33).Furthermore, the notion of meaning involved is not limited those aspects sanctioned by Positivists or Behaviorists:

> The term `meaning', which is used by all linguists, is necessarily inclusive, since it must embrace all aspects of semiosis that may be distinguished by a philosophical or logical analysis: relation, on various levels, of speech-forms to other speech-forms, relation of speech forms to non-verbal situations (objects, events, etc.), and relations, again on various levels, to the persons who are participating in the act of communication. (1939:18)

"situation, meaning, and sounds", when the native speaker's perception that one utterance is a repetition of another cannot be determined or is not trusted (1951a:29fn1):

> When what we obtain is not an admitted repetition, (and, sometimes, even when it is) we have to judge whether utterance *B* is indeed a repetition of utterance *A*, by considering the situation, meaning, and sounds. The validity of our judgment is checked in 4.5 ["Correcting possible errors"] and the Appendix to 4.21 ["On the equivalence of repetitions"]. This is equivalent to Bloomfield's `fundamental assumption of linguistics: we must assume that in every speech community some utterances are alike in form and meaning' …

For Bloomfield, the linguist's determination of *all* repetition (and therefore of the phonemic contrasts of a language) was by assumption, based on the linguist's judgment of "situation, meaning, and sounds"; for Harris, only a residuum of cases was, requiring only passing notice in a footnote. Harris' approach relies upon the social basis for repetition and contrast as a learned norm institutionalized among native speakers, rather than making it a necessary assumption by the linguist. See also Bloomfield (1933: 77, 128).

Harris describes the pair test as follows:[18]

---

18 Harris (1951a:32) gives an earlier formulation. It is not clear to me how early in the 1940s Harris developed this resolution to the problem. The preface of Harris (1951a) was signed in January 1947. Harris (1941) discusses the social basis of phonological contrast without explicit mention of the pair test. [Addendum February 25, 2000]: Harris (1942 sec. 2.0) says: "The resultant phonemic writing must be bi-uniquely correlated with speech, i.e. any utterance can be written phonemically in only one way, and any sequence of phonemes can be pronounced in only one linguistically distinct way." The determination of what is linguistically distinct requires some test of what informants judge to be distinct in their language. Harris (1990) says "that it was possible to describe the entire program from the outset, e.g. in (this) article."]

In this connection, I must emphasize again that (contra Bloch) Bloomfield also held that phonemic contrast is rooted in differences of meaning:

> As long as we pay no attention to meaning, we cannot decide whether two utterances are "the same" or "different" […] To recognize the distinctive features of a language, we must leave the ground of pure phonetics and act as though science had progressed far enough to identify all the situations and responses that make up the meaning of speech-forms. In the case of our own language, we trust to our everyday knowledge to tell us whether speech-forms are "the same" or "different" […] (1933:77)

The last sentence in this quotation posits the native speaker's intuition as criterion. Harris' pair test (embraced by Chomsky 1957, and elsewhere) provides controlled access to the intuitions of informants. The effect of the pair test is to shift the burden from the linguist's perception of differential meaning to the native speaker's perception of repetition. See also Harris (1951a): "In principle, meaning need be involve donly to the extent of determining what is repetition" (7fn4), "Only those tests will be linguistically relevant which will accord with … the speaker's actions. This ultimate correlation is the only one which has so far been found to yield a simple language structure" (31fn7 and supra). By "the speaker's actions" is meant "his use of it, or his acceptance of our use of it" (31).

Harris (1951a) describes a number of other tests for repetition. Although these are not formal instances of conducting the pair test, they rest on the same principle, as indeed do the practical heuristics of language learning: the responses of native speakers tell us what is a repetition and what is not.

> Another essential distributional fact is that some elements are similar to others in terms of certain tests; or similar in the sense that if we group these similar elements into sets (`similarity groupings'), the distribution of all members of a set (in respect to other sets) will be the same as far as we can discover. This reduces ultimately to the similarity of sound segments under repetition, or in the pair test: *x*1 is similar to *x*2 but not to *y*1 if, when one native speaker repeats*x*1*z*, *x*2*z*, *y*1*z*, . . . , a second speaker can guess correctly whether *x*1*z* as against *y*1*z* is being said, but not whether *x*1*z* as against *x*2*z* is being said. We call *x*1 and *x*2 free variants of each other (or members of a similarity grouping). Note that the pair test involves discrimination of sound but not of meaning. (1954:16)

A comment on the last sentence of this passage may be required: although neither the linguist nor the native speakers use the meanings of a pair of forms to judge whether they contrast or are repetitions, the perception of repetition vs. contrast is itself the irreducible, elemental, least datum of linguistic semantics: the perception of those least differences that make a difference in the language.[19]

In any case, it is important to recognize that meaning was seen to be problematic because it was difficult to manage, and not because of Empiricist strictures against mentalist concepts. As Harrisput it (1951a:20, similarly at 7fn4):

> In accepting this criterion of hearer's response, we approach the reliance on `meaning' usually required by linguists. Something of this order seems inescapable, at least in the present stage of linguistics: in addition to the data concerning sounds we require data about the hearer's response. However, data about a hearer accepting an utterance or part of an utterance as a repetition of something previously pronounced can be more easily controlled than data about meaning.

3.2 Deriving new elements from redundancy

The pair test is not itself a distributional test, but rather an operational definition of the least elements, the phonemic contrasts. "The fundamental data of descriptive linguistics are … the distinctions and equivalences among utterances and parts of utterances" (1951a:33). It is because the contrasts are the fundamental (socially determined) elements of language

19 For repetition, allophonic variations are differences that don't make a difference; but for imitation equivalent discernible differences do make the difference between the success or failure of the imitation. Gregory Bateson proposed that a *difference that makes a difference* is the elementary unit of information, and the elementary unit of mind. By this memorable phrase he refers to the transformation of one difference into another around a feedback loop of a cybernetic system. See Bateson (1967, 1979). It is worth noting, though the theme cannot be developed here, that with the later advent of cybernetics and in particular of perceptual control theory (Powers1973, Williams 1992), Bloomfield's identification of science with mechanism (a form of materialism) and of "*mental images, feelings*, and the like" as "merely popular terms for bodily movements" and states, has become greatly more plausible, with the critical difference that steps of linear causation are seen as links comprising circular causation in negative feedback loops. For more on Bloomfield's conception of mechanism vs. mentalism as an issue pertaining to the method rather than the subject matter of science (Bloomfield 1939:13), see Bloomfield (1933: 17, 32-33, 38, 142-4) and Ryckman (1986:22-27). See also Koerner (1970).

that subsequent distributional tests "reduce to" the pair test, as Harris says.

Distributional methods in linguistics identify and analyze patterns of redundancy among the elements of language. The distributional methods themselves rest upon the correlation of redundancy with information shown by Shannon and Weaver in the mathematical theory of communication (Shannon 1949).[20] Distributional restrictions among elements are used to define new elements that are less restricted.

By defining more general (`higher-level') elements in terms of restrictions on prior (`lower-level') elements, those restrictions are removed from the representation of utterances. The restrictions are not removed entirely, but only moved to the definitions of the new, more freely combining elements. The restrictions that define the new elements (phonemes out of contrasts, morphemes out of phonemes, constructions out of morphemes)reflect the contribution of those new elements to the information in utterances. The total redundancy in an utterance is made up of a hierarchy of contributory redundancies. Redundancy among entities on one level constitutes entities of a different type on the next level of generality,[21] which in turn bear their own redundancies. This, in part, is how language creates information.

Beginning in the 1940s, Harris asked: Given a representation of utterances in terms of linguistic elements at a given level of generality, what are the redundancies found among those elements? Obviously,not all possible combinations of phonemes occur.[22] Only certain

---

20 Shannon and Weaver did not, indeed could not say what information is, and avowedly were not able to address questions of meaning in the usual sense of the term. Shannon (1956) provides a useful corrective for overblown enthusiams that persist heedlessly today. See also Ryckman (1986:310-338).

21 It is important to realize that the `higher level' elements are not more abstract than the more restricted `lower level' ones. As Harris rather drily observes (1981:v): "There is an advance in generality as one proceeds through the successive stages of analysis […] This does not mean increasingly abstract constructs; generality is not the same thing as abstraction." Although Harris' specific reference here is to successive stages of syntactic analysis, the distinction that he is making applies rather more obviously to morphology and phonology.

22 This is necessarily so: all combinations are possible only in a code, where different combinations are correlated with elements of a message via a translation table. An example is the combinations of binary bits in the Morse code or in the ASCII code. But as has been pointed out, language is not a code. And even supposing that it were, the structures constituting linguistic information would still have to be investigated in the prior "messages" that were mapped into it via a translation table. By analogy, the linguistic information that is in a Morse-coded message is identical to that in the copy that the sender encoded, and that linguistic information is due to language and not to the code or to the encoding process. In any case, it is an observable fact that only a small percentage of possible phoneme-sequences actually occur in any given language.

We should emphasize that the sound contrasts in spoken language can be represented in various ways, creating records of the language for investigation. The phonemic contrasts could be represented by 8-bit bytes in binary machine code, for example. Harris described approvingly (1951a:148) Jakobson's use of distinctive features, and of course his phonemic long components are even less closely identified with a segmental or strictly alphabetic representation than are distinctive features.

All of these different kinds of representations are fundamentally equivalent, so long as they preserve the phonemic contrasts. The choice of representation is a matter of convenience for the type of investigation intended. For the identification of morphemes and their dependencies, some sort of linear alphabetic or

sequences constitute morphemes and words. But there is greater freedom of combinability between words, and between morphemes, than there is within them.[23] This difference in combinatorial possibilities gives a distributional basis (defined by a stochastic process) for the recognition of morpheme boundaries (Harris 1954:13, 1955, 1967, 1968:24-28).[24]

Given the morpheme boundaries,[25] we observe that not all possible sequences of morphemes occur (or have the same likelihood of occurring). This is for by now familiar information-theoretic reasons: if all combinations of entities in an ensemble are equiprobable, there can be no information in one combination as distinct from another. Determining which combinations can occur is "the central problem of structural linguistics"(Harris 1968:13):[26]

---

syllabary representation appears to be simplest (although this may be in part a reflection of our being accustomed to the conventions of alphabetic writing), without prejudice to multidimensional and multilinear representations for phonological investigations. Whatever the representation, redundancies of combination will correspond to morpheme boundaries, as described here.

23 We can readily see how this comes about historically, as the shapes of words and morphemes (starting from some prior arbitrary shape) are determined by the interaction of phonotactic considerations with socially motivated sound change, whereas phonological factors do not in general determine which pairs of morphemes a speaker brings into contiguity in the course of producing an utterance.

24 The relevance of this finding for the origin and the acquisition of language is not simple. Harris (1991, Chap 12) suggests that it is likely that a certain stock of words developed first, both in language learning by the individual and in the development of language in the species. The institution of contrastive sound category perceptions (phonemes) develops out of this. This institutionalization or learning of phonemic contrasts must be driven by error in communication, since the probability of error is far less in repetition of conventionalized discrete types than it is in imitation of the actual tokens heard in a continuum of non-discrete possibilities. Because of this reduction in the probability of error in communication, a larger and more articulate vocabulary can then develop, and an increase in precision and control of vocabulary items as distinct entities. Only when a system of sound contrasts is established, adequate for distinguishing words of a socially shared vocabulary, is the elaboration of syntax and grammar made possible, and the linguistic information that syntax and grammar in turn constitute in sentences and discourses. Linguistic information is socially transmissible, in that one speaker may repeat information transmitted by another (where at higher levels of generality `repetition' encompasses sentence paraphrase and discourse periphrasis).

It is in interesting question, to what degree learning may recapitulate evolution. There is no a priori need, it seems to me, for two children, even in the same family, to arrive at precisely the same analysis of first-learned vocabulary into a system of contrasts, or to go through precisely the same stages of acquisition (my two daughters did not), so long as in the outcome a shared vocabulary is recognizably distinct for speaker and hearer and repeatable by each in his or her own terms, recognizably to others. This applies to questions of language variety and change as well as to problems of non-uniqueness, as noted earlier. There clearly are acoustic, articulatory, and perhaps other `affordances' (in the sense of Gibson 1982) for terms of contrast, which limit the phonetic diversity that is practically available (see for example work of Stevens, Blumstein, and others, summarized in Lieberman & Blumstein (1988:184-188)). But in the roughly 50 years since Trubetzkoy, research into distinctive features defined on a universal phonetic basis has by no means ruled out the important contribution of social conventionalization to these processes (see Harris 1941).

25 Or most of them: a residue of certain identified kinds of intractible cases are resolved in such a way as to reduce restrictions on distribution of morphemes—that is, on a distributional basis. As with phonemics, a unique solution is neither guaranteed nor required.

26 Harris (1968:18) goes on in a footnote to say:

> It is a problem of finding regularities in those sequences of elements which constitute sentences (or discourses) as contrasted with those which do not. In the denumerably infinite set of word sequences which are discourses […], such regularities must exist, if the elementary grammatical combinations and operations which the discourses exhibit are finite or recursive. (ibid.)

4. Tools of analysis

In the course of work on this "central problem", Harris invented an astonishing range of methods for analysis of language, including an algebraic treatment of constituent structure, with provisions for its inherent weaknesses (1946, 1963),[27] string analysis as a complementary tool compensating for those and other weaknesses of immediate constituent analysis[28] (1959, 1961, 1962), transformational analysis (1952a, 1952b, 1956, 1957, 1965), discourse analysis (1952a, 1952b,1952c, 1989), and sublanguage analysis (1982b, 1985, 1989).[29] It is important to recognize that Harris saw these as tools of linguistic analysis, not as competing theories of language nor even as bearing directly on the question of the logical form of a theory of language. In a well known comparison of string analysis, immediate-constituent analysis (phrase-structure grammar), and transformational analysis, he wrote (1965:238—239):

> Indeed it is the only general problem (about the whole language) which can be formulated in the terms of structural linguistics (phonemes, morphemes, sentences, etc.), since other problems involve external concepts such as the circumstances in which something is said.

That is, it is the only such general problem that can be studied without depending upon the use of linguistic information in a purportedly prior and independent metalanguage.

Harris did not say that problems beyond the scope of structural linguistics were illegitimate or should not be studied, but he did suggest that they would be fruitfully studied with respect to an account of the structure of language in its own terms. Indeed, they can only be studied making use of linguistic information in a metalanguage for the science that studies them, so it is at the least for this reason that work in any such science depends upon results of a science of linguistic information.

27 The relation of Harris' treatment to X-bar notation, in theoretic terms as well as historically, requires more detailed investigation than can be undertaken here.

28 This complementarity was exploited in a series of studies by Aravind Joshi and his students combining rewriting rules with adjunction rules (e.g. Joshi 1969b, Joshi et al.1969, Joshi & Levi 1982), culminating in his Tree-Adjoining Grammar (TAG) formalism (Joshi et al. 1975). Harris' operator grammar may be thought of as exploiting the closely analogous complementarity between dependency grammar and string-adjunction grammar. One advantage of dependency over rewrite rules is that there is no artificial separation of syntax and semantics, that is, no lexical insertion into previously generated structures, no requirement thatsemantic representation and semantic interpretation be separate from syntax. The relation of the autonomy hypothesis to the rewrite rules of phrase structure grammar, possibly as a necessary entailment, is another theme requiring separate treatment.

In this, operator grammar has a kind of philosophical affinity with generative semantics, which has undergone a kind of back-door assimilation into standard generative theory: see Huck & Goldsmith(1992) and e.g. the emphasis on morphology in Chomsky (1992). Operator grammar bears strong affinities with categorial grammar as well.

29 I have not included operator grammar in the enumeration of analytical tools here because (especially in grammars of sublanguages) I see it as a *result* of analysis with those tools. How these tools addressed the "central problem" should be clarified, but that research goes beyond what is undertaken here.

To interrelate these analyses, it is necessary to understand that these are not competing theories, but rather complement each other in the description of sentences.[30] It is not that grammar is one or another of these analyses, but that sentences exhibit simultaneously all of these properties. Indeed, one can devise modifications of languages, say of English, which lack one property while retaining the others; but the result is not structurally the same as the original language.

[…] [E]ach of these major sentential properties can be formulated so as to include the relevant effects of the other properties and so as to describe in its own terms the whole set of sentences.[31] In organizing a language description around one or another of these properties, the main difference lies not in the sentences which can be described, but in the way in which the description relates each sentence to certain others, i.e. in the various subsets of sentences that the description creates […].[32]

---

30 Harris' footnote here gives a rare glimpse of his social philosophy, which was for him a matter of deeply committed practice and not a merely theoretical espousal: "The pitting of one linguistic tool against another has in it something of the absolutist postwar temper of social institutions, but is not required by the character and range of these tools of analysis." It is entirely in keeping with his views, for example, that he did not seek to impose them upon others. The absolutism Harris mentions here is also seen in the recurring demand for `the' `correct' grammar of `the' language. An early instance of this is the controversy over non-uniqueness in phonology, alluded to earlier, and Householder's famous (and oversimplifying) dichotimization of `God's Truth' linguists vs. `hocus-pocus' linguists (see Hymes & Fought 1981:150-151). Harris admitted different, non-unique descriptions by application of alternative distributional procedures, so long as, given the primitive contrasts, "the defining of the elements and the stating of the relations among them be based on distribution, and be unambiguous, consistent and subject to check" (1951a:9). "In any case, there is no harm in all this non-uniqueness, since each system can be mapped onto the others, so long as any special conditions are explicit and measurable" (1954:5; Harris [1951a:32] gives an earlier formulation). The notion that there *is* some sort of harm in non-uniqueness reflects a conception that because the ideal of science is to give a "correct" description eventually, failure to find "the correct" description at every stage constitutes failure to be scientific. Hymes & Fought (1981:148-149) identify this character in Trager's insistance that one start with phonetic data and proceed in rigid, stepwise fashion, never mixing levels, in contrast with Harris' much more flexible (though no less rigorous) approach.

31 Harris' footnote here reads:

> Because of the mass of idiomatic and quasi-idiomatic expressions in language, each type of description has to treat of various special small categories of words, and in some cases even of unique words. But in the case of string and transformational analyses, and less adequately in the case of constituent analysis, the statements for aberrant and idiomatic material can be made in the terms of the given description (constituent, string, or transformation) or in limited extension or weakenings of the rules of that description. In these analyses, the treatment of difficult material does not require us to go completely outside the terms of the given description into the terms of another or into the metalanguage.

In the elided text is a slightly different statement of the footnoted passage: "Each of these properties can be used as the basis for a description of the whole language because the effects of the other properties can be brought in as restrictions on the chosen property". In this formulation, the import of added restrictions is emphasized.

32 Contrast this with the long-running controversies over the generative capacity of one or another formal metalanguage, construed as systems of generative rules for language.

> However, the greatest interest in each of these properties lies not in its utilization as an organizing scheme for grammatical analysis, but in the statements which can be made, uniquely in terms of the given property, about the structure of language.
>
> It should be mentioned that the investigation of the several basic properties of sentences, and the possibility of using each as the central method of sentence analysis, are different from the question of the logical forms of grammar as a theory of language. The properties […] can be studied empirically; and a particular form of grammar can use various of these properties.[33]

The impulse to hypostatize as Language what are really only artifacts or reflections of the tools used for investigating utterances has always been with us. Harris resisted the impulse to rush toclosure because he had larger ends in view.[34] Each of these approaches revealed only some of the characteristic (informational) properties of language; each was incapable of coming to closure over all of these properties, or could be forced to closure only by adding to the redundancy of the description by various ad hoc adjustments and exceptions for difficult cases, which would obscure the essential redundancy that constitutes the information in discourses.

5. Distributionalism

We will turn now to a closer examination of the ways in which Harris extended and refined the methods of linguistic analysis pioneered by Bloomfield and Sapir, which had come to be called distributional analysis.[35] As we have seen, this is a "methodological approach, of defining more freely combining new elements on the basis of occurrence-restrictions of old elements" (Harris 1970:v). Distributional methods disclose, as Harris said in respect to Sapir's way of working, a `detached' pattern in language. In part, this means (as Bloomfield and Sapir also had seen) that factors external to language, such as might be proposed in psychology, cannot legitimately be invoked as *explanatory principles* for the patterning in language, though such factors may have interpretative relevance to the results of formal

---

33 An instance of such a "particular form of grammar" is seen in Joshi's investigation of formal languages generated by mixed types of rules, referenced earlier. Harris's footnote here shows his continuing support and advocacy of Chomsky's divergent views:
This applies, for example, to the formulation of grammar in terms of partially ordered homomorphisms which was sketched in [Harris (1946)] and which has been given an explicit form in Noam Chomsky's rewriting rules; also to the precise theory of generative grammar proposed and formulated by Chomsky in a series of major papers, especially in his [Chomsky (1957)]. Cf. also his interesting [Chomsky (1956)].

34 This touches again on the matter of Harris' conception of the proper conduct of scientific inquiry, an excellent topic certain to be fruitful, but beyond the scope of this paper.

35 The term `distribution' apparently originated with Sapir by way of Swadesh (1934) as a geographical metaphor. In a letter quoted in Diderichsen (1958:158, fn. 4), Swadesh says:
> The source of the usage may have been Sapir, but I do not remember. At the time […] I was not conscious of either adopting or inventing a technical term, but rather used the word simply as a way of describing the spread of occurrences of a sound among the positions within the word. It was an application of the usage represented by `geographic distribution', an expression which was much used by Sapir as by other anthropologists and linguists.

analysis.[36]

Harris knew that such lines of `explanation', if taken up prematurely, could only interfere with the work proper to linguistics:

> [T]alking about function, system, or the like, without defining them in terms of operations and relations, fools even the linguistic worker. For by satisfying him with undefined psychological terms it prevents him from continuing his analysis. (Harris 1941:707)

This view contrasts sharply with the expectation of many linguists that each discovered bit of regularity is an outward sign of some inward and innate explanatory principle, and the perception that researchis worthy of pursuit or support only insofar as it advances or refutes one proposed explanatory principle or another. This way of working fosters fragmentary and mutually inconsistent results. By this I refer not to disagreements among linguists with alternative analyses, but rather to fragments of grammatical description that cannot be reconciled with each other for complete coverage of any single language. See also, for example, Harris (1954:13):

> As Leonard Bloomfield pointed out, it frequently happens that when we do not rest with the explanation that something is due to meaning, we discover that it has a formal regularity or `explanation'. It may still be `due to meaning' in one sense, but it accords with a distributional regularity.
>
> If we investigate in this light the areas where there are no simple distributional regularities, we will often find interesting distributional relations, relations which tell us something about the occurrence of elements and which correlate with some aspect of meaning. In certain important cases it will even prove possible to state certain aspects of meaning as functions of measurable distributional relations.

A footnote in this passage touches (in 1954) on the dilemma due to the lack of an independent metalanguage:

> It should be clear that only after we discover what kinds of distributional regularities there are among successive elements or sections in discourses can we attempt any organized semantic interpretation of the successions discovered. (Harris 1954:15fn19)

5.1 Practical distributionalism

Munz (1972) distinguishes three types or degrees of distributionalism: practical, theoretical, and interpretive. As we noted at the outset of this paper, it is commonly alleged

36 This is not an expression of some sort of disciplinary chauvinism, asserting the priority of linguistics. The unavailability of an independent metalanguage precludes using the findings of a prior science to establish elements of language and their relations. As we have seen, workers in any science presuppose and depend upon linguistic information to communicate findings, propose interpretations, and reach agreements,so results in a proposedly prior science cannot be used for identifying and characterizing linguistic information itself (that is, the elements of language and their formal relations).

that Harris was trying to work out formal analytic techniques that could be applied mechanically (using a computer) to a corpus of linguistic data to produce a grammar.[37] This practical distributionalism is clearly a straw man, at least as regards Harris' work.

The existence of computer programs that demonstrate or test aspects of Harris' work, even the fact that certain aspects of language structure can be (partially) determined by such applications, thus does not gainsay the essential point that practical distributionalism was not his aim.

The situation is somewhat different with sublanguage analysis, especially for sublanguages of science, e.g., Hirschman (1986). Because the metalanguage for sublanguage grammar is external to the sublanguage (being the same metalanguage as for the language as a whole), practical discovery procedures are possible that cannot work over unrestricted language domains. See Harris (1991 Chapter 10) for discussion, and Grishman et al. (1986) for an example of work in progress.

Probably the closest thing to a practical discovery procedure in Harris' oeuvre is his procedure for identifying morpheme boundaries (1955). What resulted from a computational test using very large word lists (from several large on-line dictionaries in standard English orthography) was a good determination of most morpheme boundaries, with a residuum to be resolved by other distributional methods (Harris 1967).

5.2 Theoretical distributionalism

But Harris did not regard the phoneme-to-morpheme prodecure as a practical discovery procedure:

> For methodological purposes and for special problems — though certainly not for practical work — the procedure can … replace the less orderly search for morphemic segments. (1955:32-33)[38]

In Harris' own view, then, this procedure amounts at most to what Munz (1972) terms

37 Probably the best known statement of this view is Chomsky (1957:50-53), which distinguishes three possible requirements for linguistics: a discovery procedure for grammars, a decision procedure for grammars, and an evaluation procedure for grammars. "As I interpret most of the more careful proposals for the development of linguistic theory", he wrote, "they attempt to meet the strongest of these three requirements", namely "a practical and mechanical method for actually constructing the grammar, given a corpus of utterances". In a footnote within this passage, Chomsky claims that "discovery procedures are the explicit goal" of Bloch (1948), Chomsky (1953), Harris (1951a) and (1955), Hockett (1952) and (1947), Wells (1947), "and many other works". I invite the reader now to take an hour in a university library and to read Bloch (1948:5), Harris (1951a:1-2, 1955:33), Wells (1947:193), and Hockett (1947:241), where precisely this is in each case disclaimed. Hockett (1952:27) is especially interesting. Here, in the opening sentence of the article, it is to Harris that Hockett is (mistakenly) attributing the aim of devising a discovery procedure, an aim which he (Hockett) repudiates in favor of an evaluation procedure (p. 29) just such as Chomsky (1957)advocates. This leaves an earlier Chomsky himself as a possible advocate of discovery procedures, but if that was his intention then it is nowhere made explicit in his 1953 paper. See Ryckman (1986:45-53) for a more extended discussion.

38 Another line of inquiry beyond the scope of this paper concerns the range of methodological problems that Harris identified, and his way of working to resolve them.

"theoretical distributionalism". This is the notion that a grammar would result if one used proposed procedures on a sufficiently large corpus for a sufficiently long time, though no one would in fact do so in practice. And this is the ground on which the strongest case might be built for a claim that Harris' aim was the devising of practical discovery procedures.

5.3 Interpretive distributionalism

This leaves for virtually all of Harris' work only `interpretive distributionalism': the view that rules of grammar, however determined, are statements about distribution and are falsifiable with respect to distribution. Munz suggests that this view is merely a minimal requirement for any linguist to be taken seriously. However, because there is no external metalanguage, Munz's interpretive distributionalism must be understood in stronger terms: distinctions among linguistic elements (distinguished for a language in a grammar) must be grounded in the distributional redundancy of previously established elements, resting ultimately on the fundamental social definition of phonological contrasts.

In practice, the linguist working in the field leans on a previously learned language, just as any other scientist does for whatever science domain. The pitfalls of doing so for language description are well known, at least in principle, as is the value of distributional methods as safeguards against projecting inappropriate linguistic categories from one's understanding of one language structure onto that of another (see e.g. 1951a:2-3). Also reasonably familiar is the fact that distributional methods open up for consideration alternatives that might otherwise be neglected (e.g. 1951a:8fn7) and give a principled basis for fundamental decisions (e.g.1951a:32fn8). One form of this is illustrated by Harris' discussion (1954:9-11) of a Cherokee paradigm, where the fact that some elements have no clear English glosses is irrelevant, given distributional support. The converse situation is illustrated by the fact (1951a:193-194, 1954, and elsewhere) that the sl-, gl-, fl- words in English are not analyzable as e.g. an *sl-* prefix plus stem, absent distributional support, despite the correlation of identifiable meanings with word sets like *slide*, *slither,* and *slip*. The sometimes startling originality of Harris' formulations of sources and derivations in operator grammar was enabled, one suspects, by his returning again and again to the distributional facts without prejudice as to their expected interpretation.[39]

5.4 Evaluation criteria

Interpretive distributionalism thus posits a basis or criterion by which to evaluate and

39 Munz writes (1972:273): "The extent of his empirical work is not generally known since in his publications theoretical questions have upstaged the business of writing a grammar. When his students raise concrete troublesome cases he usually replies with a considered answer immediately and is liable to produce the following meeting an envelop[e] filled with dog-eared scraps of paper covered with appropriate examples." Examples used in Harris' writings were usually drawn from newspapers and journals, rather than being made up for the purpose: "It should be noted that the example sentences analyzed here are from current books and periodicals, except for those that are obviously formed as minimal examples of a given construction" (Harris 1982:vi). That this was a practice of long standing is evidenced by e.g. the reference to an "AP Dispatch from Bolivia, July 10, 1944" (Harris 1951a:254fn23).

justify alternative proposed analyses. Absent an external metalanguage, this criterion is as we have seen anecessary one, but given the non-uniqueness of distributionally justified analyses it is not a sufficient one for determining the `best' analysis. Harris repeatedly states clearly that additional criteria are needed of two sorts,which we may here call interior and exterior criteria.

Interior criteria concern relations between one part of the grammar or linguistic description and another. One of the distributionally valid analyses on a given level within the grammar (e.g. phonological)may be more advantageous on the next level (e.g. in determining and specifying morphemes and morphophonemics).

Exterior criteria concern the purposes that one wishes the description to serve. For example, a representation of the phonological contrasts that is optimum for studies of phonological processes may not be the most convenient representation for historical comparison and reconstruction. What is constant under the non-uniqueness of phonological descriptions is the contrasts, which are based in socially established perceptions of what is a repetition.

What is determined here is not `the' grammar of `the' language in some absolute sense,[40] but the optimum description for particular ends, such as fitting into a particular framework in philosophy or psychology to which, for reasons outside of linguistics, one may be committed.

6. Imperfect coverage

Distributional methods avowedly do not provide complete or ideal coverage (1954:12). One reason is the non-uniqueness already discussed. (We should note here that non-uniqueness is of course not at all limited to phonemics. Harris (1951a) took considerable pains to indicate alternative ways of setting up elements at all levels of description.)

An even more fundamental reason is that language is not well defined, a characteristic that has been much discussed in terms of degrees of grammaticality, degrees of acceptability, `squishes', and soon. It may seem that such subtleties are beyond the capacity of distributionalism. It is important, then, to see how Harris addressed these problems without having his theory depend upon a putative independent metalanguage concealed behind elaborated judgments of grammaticality or meaning.

6.1 Graded membership

To accomodate this essential characteristic of language, that the set of sentences is not well defined, Harris refined the definitions of scope and domain used in distributional analysis

40 If one claims that there is such a thing, or if an abstraction over differences among speakers is a useful fiction for some purposes, then the alternatives evaluated here can be seen as alternative representations for "the grammar", equivalent because the differences noted make no difference with respect to it. An evaluation procedure in this sense is not the same as that proposed in Chomsky(1957), which concerns choice among alternative language-like formal systems. In particular, considerations such as generative capacity, strong and weak equivalence, and so on, have no bearing here.

of transformations. In the earlier definition of transformation (1952a, 1956, 1957), if any word-selection that satisfies one sentence form[41] also satisfies a second sentence form, then a transformation exists as a mapping between the two sets of satisfier-sentences (or,at a later stage of analysis, from one set into the other, framing a derivational sequence).

But in the criterion for transformation used first in Harris (1965)—though anticipated as a possible requirement in (1957)—satisfiers of two sentence forms are graded as to acceptability. If the relative acceptability of a pair of sentences satisfying form A is *not reversed* for the corresponding pair satisfying form B (for any pairs of satisfiers), the two sentence forms are transforms.[42] This criterion remains especially important for deciding marginal cases as the gross transformations are analyzedinto elementary increments (ultimately operators) and operations (linearization, transpositions, and morphophonemic reductions).

There are then two points at which Harris' methods depend in principle upon native intuitions. This extension of distributional methodology does indeed call upon intuitive judgments of native speakers,but in a constrained, pairwise manner, and without reference to a prior metalanguage specifying meanings. It is not difficult to draw a parallel between this criterion for transformation and the pair test for repetition, which determines phonemic contrast. (To forestall misconstrual, let it be said once again that even a non-native linguist relies in practice upon intuitive grasp of the language at every point of practical linguistic analysis.)

6.2 From exceptions to extensions

Another reason that distributional methods do not give ideal coverage is the great number of irregularities in any language, exceptions to the regular patterns that they treat so well. Examples include asyntactic utterances, such as "Fire!",[43] and "the mass of idiomatic and quasi-idiomatic expressions in language" (1965:239).

In the work of wrestling the mappings in the set of sentences first into elementary

41 A sentence-form is a sequence of morpheme-class symbols that results from substituting the symbols for the corresponding words in a sentence. Any sentence for which such substitution results in the given sequence of morpheme classes is a satisfier of that sentence form. Two sentence forms are candidates for transformation if the same morpheme classes and constants are involved, perhaps in different order,or with some substring present in one and absent from the other. (In algebraic terms, each morpheme class symbol is a *variable* over the morphemes that satisfy it, and a *constant* is a morpheme that does not vary over the satisfiers of a sentence form, as for example the *it* of it-extraction or *that* marking the argument of a verb like *know*.)

42 Harris discusses various forms that the acceptability gradings may take. For example, sentences may be differentiated as to the subject-matter domains in which they are accorded fullest acceptability. This looks ahead to sublanguage analysis where acceptabilities may become well defined (binary yes-no selection) within the sublanguage domain (see section 10).

43 Harris speculates that these may be of a type of asyntactic speech that antedates the evolution of syntax and that yet persists in language today (1991:367-368). A study of Harris' naturalism might explore the contrast between his view of the conditions for the evolution of language, analogous to those for learning a language, and prevalent innate biologicism.

transformations and then into operator-argument dependencies plus reductions, the residuum of exceptions come to be handled in a natural way as extensions of established relations and reductions. The advantages for this of distributional methods was clear even for morphological analysis (e.g. 1951a:8fn7):

> It may be noted that distributional procedures do more than offer a rigorous alternative to meaning considerations and the like. Distributional procedures, once established, permit, with no extra trouble, the definite treatment of those marginal cases which meaning considerations leave indeterminate or open to conflicting opinion.

It is in the analysis of syntax and information structures, however, that the power of distributional methods to account for exceptions by analogy to regularities is most clearly seen:

> The stock of idioms and similar individual exceptions in language is such that, after all […] generalization is carried out, there always remain many idiomatic exceptions. for these, the only hope is to formulate the general rules in such a way that the individual exceptions can be stated as extensions of the domain of one rule or another, beyond the boundaries allowed in the general rule. (1968:173)

This is the principle of regularization or normalization that is familiar in phonology and in morphophonemics. Harris (1951b) describes this as a characteristic of Sapir's way of working. Here he applies the same method, termed `extended morphophonemics', to syntax. (See, for example, Harris [1965:239fn7], on exceptions, quoted earlier.) The basis for the regularization by analogic extension lies in the mandate for least grammar:

> The same mininalizing considerations that go into identifying the elements go also into identifying the relations among them, these identifications being in any case two sides of the same coin. The analysis of a language form should make maximal use of analyses used for other forms, avoiding as far as possible recourse to new (*ad hoc*) classes, rules, domains. (1991:35)

The analysis of complex transformations into elementary sentence-differences and thence into the sentence-forming steps of operator grammar (word entry, linearization of word dependencies into word strings, and reductions of word shape), is the natural and necessary consequence of acting on that mandate:

> Making a least grammar for an unbounded set of sentences entails that as many words and sentences as possible (and their meanings) be derived from others (by changes and additions), preserving whatever has gone into the prior making of those others. And it entails that the derivation be in small steps repeated in various combinations rather than in large steps each of which accounts for a whole difference between starting-point and end-point. A large step, such as the totality of differences between the active and passive forms of the verb, can be used only at one point of the grammar — in this case, to derive the passive. But if we can obtain

> its effects from a succession of small steps (where the resultants of one step are in the domain of the next), we may find that the same small steps occur elsewhere in the grammar, in different combinations […]. (1991:35)

Some of the derivational sources and some of the intermediate steps in these derivations are sentence forms whose satisfiers are *all* of low or dubious acceptability. More often, sources for individual words or affixes are not attested or naturally accepted in the language today. These, Harris marks with a dagger, and he allows them (exploiting the ill-definition of the set of sentences) only if they are reachable in the derivation by steps that are well attested between fully acceptable sentences in other, different derivations:

> A daggered sentence, postulated as the (entire or partial) source of an attested sentence, differs from it only by the listable changes, made in the same conditions, as those that lead from a source to a reduced sentence when both are attested. Furthermore, the daggered sentences consist of […] words that appear in positions held by those words in the attested […] sentences. No clear line can be drawn between them and the normal sentences[…]. Note also that the set of attested sentences is not well defined but fuzzy. Many `ordinary' sentences are marginal or used by some speakers or writers and not others. The fuzziness of the set is inescapable. There are no objective and precise criteria for deciding which sentences are `acceptable' in the language[…]. Some of the daggered sentences are grating to the ear without[,] however, being excludable from the grammar in any principled way. […] We therefore take these daggered sentences as constituting, together with the normal sentences, an extended set of "grammatically possible sentences." (1982:18)

Harris' image used in a seminar (in discussion of Harris 1969) showed two concentric circles: the regular structure due to operator-argument dependencies, with its close correspondence of form with information, and the extension of this by reductions into the larger set of more idiomatic paraphrases. Superposed over this was a wavy line, approximating the larger circle, showing the fuzzy boundary of questionable sentences and "infrasentences" at the margins of language. Metaphor and other analogic processes result in uncertainty and continual adjustment of these boundaries.

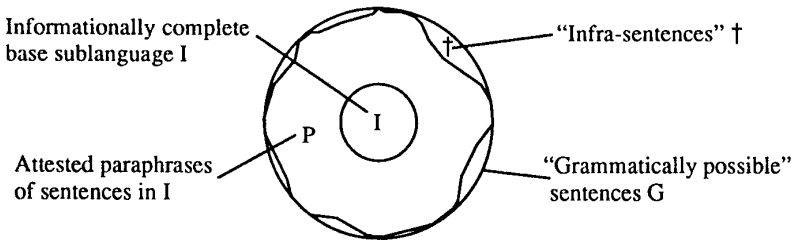

Fig.1

Harris' treatment of exceptions by analogic extension of regularities in the language

provides (and rests upon) insights into processes of language change and variation.

> In spite of the specificity of the processes and conditions, the set of sentences in a language is not well defined: there are many marginal sentences, largely because the normal likelihoods of certain grammatically important words[…], and the domains of certain reductions, are only roughly stateable. There may also be cases of constructions and regularities that do not apply to particular words in their expected domain. Overtly, this situation resembles the `exceptions' so common in the traditional rules of grammar. However, not only do far fewer `exceptions' remain after reduction analysis than in traditional grammar, but those that remain have more assignable causes or systemic properties. Partly, they are due to lack of regularization, by uneven applications of analogy or by retention of old, frequently used forms[…]. And partly they are due to language being constantly in process of change at one point or another, so that when the description is sufficiently detailed one is bound to come upon reductions and analogies that are in the process of extending their domain or of becoming reinterpreted into the operator framework [—] and perhaps differently among different speakers. They may also be due to arrested syntactic change, as in the vagaries of *do* in the course of the history of English[…]. And they may be due in complex ways to the overlapping domains of different reductions. (1991:360-361)

It is important to realize that what is involved here is not merely a capacity of the theory (or of a grammar based in the theory) to accommodate marginal data, but rather a capacity of language which Harris' theory describes. The reality of it in language can be seen in processes of language change and in hearers' ready accommodation to language variation, and the relevance for synchronic linguistic description we have just seen:

> […] [S]ince languages change through time, there may always be some forms which do not fit perfectly into the current grammar. The total system is then simpler to the extent that it minimizes the disturbance due to such on-going changes. This suggests that rules be so stated as to be able to house current derivations in respect to combining-possibilities, phonemic shape, position, etc.—formulations in which the relevant past or the immediate future of forms would be minor changes in condition or domain. (1991:36)

Furthermore, this capacity of language (that is, of its speakers) to accommodate even the irregularities of language as extensions of its regularities, in a manner very like folk etymology, is necessary for language to be learnable and socially transmissible:

> It is a problem of finding regularities in those sequences of elements which constitute sentences (or discourses) as contrasted with those which do not. In the denumerably infinite set of word sequences which are discourses […], such regularities must exist, if the elementary grammatical combinations and operations which the discourses exhibit are finite or recursive. So much so that, as will be seen […], even the grammatical exceptions are only extensions of grammatical

regularities. (Harris 1968:13)

7. `Taxonomic linguistics'

So far in this discussion of Harris' distributionalism, the relation of distributional patterning to meaning has naturally been in the foreground, as indeed it was for Harris. As subsidiary topics,we have touched upon the canard about practical discovery procedures and about Harris' methods being restricted to a fixed corpus. In what has become the customary representation of Harris' work, it is these misapprehensions that have the greater prominence.

To these three forms of distributionalism, then, we may add a familiar straw man, Taxonomic Linguistics: "Its most basic techniques for arriving at […] a grammar are segmentation and classification"(Chomsky 1971:65). "It was believed quite literally that starting with a corpus of utterances, a grammar could be constructed through successive segmentations and classifications" (Newmeyer 1980:6-7).

Setting aside the allegations about discovery procedures, which we have already considered, this description applies superficially to immediate constituent analysis, in terms of which still-prevalent conceptions of syntactic analysis were formed.[44] However, Harris' distributional methods proceed not by segmentation of longer stretches and classification of the resulting segments, but by "defining more freely combining new elements on the basis of occurrence-restrictions of old elements" (1970:v).

> The hierarchical defining of new elements, one level in terms of another, provides a more compact description of the combinations which occur in language. But not only that: it also changes the character of structural linguistics from a science of classificational lists to one of relational types. This development stems from the following: Every linguistic classification is based on a relation in respect to occurrence in combinations. When the classification of the observed data into phonemes, and of these into other entities, and so on, is carried out wherever possible, we find that while the classes are necessarily different in each case, and while the effect of the classificatory relation is in each case the equivalent of certain occurrence-restrictions, the relations which show up in the various cases are only of a few types. And while classes (relations) of physical events with their occurrence-restrictions are the primitives of the grammar, it is the types of relation that are the primitives of the meta-grammar. […] (ibid. vi)

Thus, the pair test for repetition replaces the continua of speech sounds by the relation of contrast.[45] Judgments as to the relative acceptability of homomorphic utterances—the use

44 It is remarkable that even today phrase structure and rewrite rules based on immediate constituency have almost the status of a required convention for the presentation of discussions of syntax, despite their obvious and well known shortcomings. They compel a spurious abstractness of treatment that is congenial to certain kinds of argumentation.

45 Again, that the contrasts may be represented by symbols for segments or classes of segments, and that they indeed apparently must be so represented in order for the linguist to be able to proceed with analysis

of acceptability-gradings as the criterion for transformation given beginning in Harris (1965) was noted earlier—replace an attempt to distinguish categorically what is in the language and what is not.[46] Relations of stochastic dependency among phonemes (that is, among the phonemic contrasts, however represented) define morpheme boundaries, replacing a distinguishing of them in respect to their glossesin some prior language that has been given the role of metalanguage. Relations of dependency between operators and their argument words, under reduction, replace the pseudo-hierarchies of phrase classes of constituent grammar and the string types of adjunction grammar.

8. Summary of Harris' methodology

The chapter on method in Harris' last book (1991:30-49) may be his best summary of his distributional approach. Here, he states unequivocally that the motivation of his methods is the creation of a coherent, consistent, and well founded theory shaped by the character of language itself, rather than being shaped by prior psychological or philosophical commitments. I can quote only part of this summary here:

> In the absence of an external metalanguage, one could seek to identify the entities of language by extra-linguistic means, such as the occurrence of words in life circumstances that exhibit their meaning. Many words may indeed be identified on such grounds. But many other words, and the ordering which makes out of them sentences as against ungrammatical word collections, cannot be be thus identified. In contrast, when only a small percentage of all possible sound-sequences actually occurs in utterances, one can identify the boundaries of words, and their relative likelihoods, from their sentential environment; this, even if one was not told (in words) that there exist such things as words (7.3). And when identified words combine with each other in relatively few regular ways which are used throughout the course of an utterance, one can recognize utterances, long and short, in distinction to non-occurring sequences of words; and one can recognize within long utterances segments (sentences) identical to otherwise-observed short utterances. This holds even if one is not told that there exist such things as sentences (7.7). Given, then, the absence of an external metalanguage, it is an essential property of language that the combinations of words in utterances are not all equiprobable, and in point of fact that many combinations do not appear at all.
>
> It follows that whatever else there is to be said about the form of language, a fundamental task is to state the departures from equiprobability in sound- and word-sequences. Certain problems arise[…]. All these difficulties can be circumvented by seeking out not a list of attested combination but a set of constraints on combination (1.1). A few constraints can together preclude unboundedly many word-sequences while allowing others; and constraints stated

—i.e. that some sort of alphabetic writing is always presupposed in studies of grammar, syntax, and language use—is incidental to the point being made here.

46 The observation that the set of sentences is not well defined must be accommodated by any theory. Munz emphasizes (1972:269) Harris' observation as early as (1957:395) that the set of sentences is ill defined.

on different or partially fuzzy word domains can create marginal forms whose satisfaction of the constraints may then be in question.

If we obtain the occurring combinations, as departures from equiprobability, by stating the constraints that allow them, we are specifying (in the constraints) the various contributions to non-equiprobability that together make up the total non-equiprobability of the system. It is therefore clear that the constraints as formulated should not introduce any additional redundancy, over and above the least needed to account for what is actually present; for the non-equiprobability of language is precisely what we are trying to describe.[…]

If, then, we have for a language different descriptions, adequate to characterize its utterances, but with different amounts of departures from equiprobability ascribed at various points in the course of the descriptions, we opt for the one with least such departures, since that one has clearly added least to the inherent departures in the language being described. The consideration of least non-randomness in description is involved in the discovery of phonemes, which underlies the whole development of a science of language, and it leads to many other methods in the analysis of language. The effect of the least-redundancy test, when applied at each point in the language description, is a grammar with fewest possible and most independent objects (elements and constructions), fewest and least-intersecting classes of objects, fewest and most independent rules (relations, transformations, etc.) on the objects, fewest differences in domain for the rules, and finally, fewest abstract constructs. Such a grammar is least redundant because each object (other than arbitrary ones defined only by their relations), and each rule or relation and classification, and each statement of domain for these, is a constraint on the equiprobability of parts in the utterances which are being described.

At each analysis of a linguistic form one could therefore seek to add as little as possible to the system describing the forms thus far analyzed.[…]

The same mininalizing considerations that go into identifying the elements go also into identifying the relations among them, these identifications being in any case two sides of the same coin. The analysis of a language form should make maximal use of analyses used for other forms, avoiding as far as possible recourse to new (*ad hoc*) classes, rules, domains. To explain a form, then, is to find its maximal similarity to the other forms, with the whole grammar thus becoming a best fit for the data. This means analyzing each form in respect to the whole relevant grammar — and the grammar of a recognizable linguistic community, not of an arbitrary speaker (since language is essentially a public structure). […]

In all this, it should be noted that no assumption is made that there exists a structure in language, and no appeal is made to any particular principles of structuralism. Such structure as is found comes out in the process of making a least description. (1991:32-36)

> The methods […] were determined by the issue of constraints[47] posed in 2.2, and by the nature of language data as seen in respect to that problem[…].
>
> Carrying out these methods does more than provide an analysis of the data. It narrows the range of possibilities toward forming a theory of the constraints[…]. Such selective effect of the data methods upon the ultimate theory happens when the methods are scrutinized for their relevance […], when they are carried out over the whole of an adequate sample […], and as fully as possible — so that the entities reached by the analysis are maximally independent of each other (i.e. are least constrained combinatorially), and when the analyses are made by these methods alone with no appeal to external considerations (such as phonetic similarity, meaning, intention, analogy, or history). This is not to say that such considerations may not have great weight, and may not even be the creators of particular forms and relations; but for purposes of approaching a theory we must first see to what extent a single relevant method can explain the forms, and to what (if any) extent forms due to other influences can be domesticated — reinterpreted — in terms of the central method.
>
> It is especially possible to consider the analysis as laying the groundwork for a theory, when every set of elements and operations or relations that is finally reached is structurally related to a single language-wide property, rather than some of the sets being merely residues created by setting up the other sets. An example of the latter case would be if the set of inter-sentence derivations were not characterizable as reductions but were simply the set of whatever changes one needed to obtain one sentence from another, rather than a motivated, structured set of reductions and transformations.
>
> The reason that the methods help to suggest a theory is that they are not simply empirical or descriptive. They necessarily organize the constraints into a maximally simple system, so that the description of language comes to consist of maximally unconstrained elements and rules (2.3). Such a system may not be identical with the way the speaker produces his sentences or the way the hearer figures them out. But it has a unique least status in respect to the data of the language (2.2). And it has a crucial status in respect to the information carried by language, in that the information is certainly related to the departures from equiprobability in combination; and a most parsimonious grammar reveals most closely just the individual departures from equiprobability in a language (11.4). (Ibid. 48-49)

Harris' footnote (in the last paragraph, just after the reference to section 2.3), succinctly addresses the questions of discovery procedures, evaluation procedures, and some aspects of language universals:[48]

---

47 These are essentially the same constraints as those presented in Chapter 2 of Harris (1968), noted earlier (2.2 and fn. 11).

48 An evaluation metric clearly is available (as described above) if one chooses to work within the metatheoretical framework that Harris advances. One may of course reject this framework, the most

This is not to say that we have here a discovery or decision procedure for grammar. However, given a particular metatheory of language — the need to map the departures from randomness (2.1), which arises from the essential lack of an external metalanguage — we can propose procedures that lead to a least grammar, which constitutes a best fit of the data in the given direction, though not uniquely the best fit. It is nevertheless the case that the departures from randomness in the various languages are so massive and so similar that all grammars exhibit a strong family resemblance, no matter in what style they are stated.

9. Consistency of research into linguistic information

The aim of research into distributional regularities in language, then, was not to expunge meaning from grammar, but to disclose it in grammar. And Harris' interest in identifying the structures in language that carry or constitute information is not a recent change of heart evinced only in his last books, as some might suppose.

For obvious example, this interest is explicitly manifest in the early work on discourse analysis. It is perhaps less obvious in string analysis,[49] but its relevance becomes apparent in consideration of how string analysis is carried out. In the practice of analyzing linguistic data, one repeatedly asks what least subpart one may remove from a given utterance

---

obvious way being to assert that there does exist in fact a universal metalanguage prior to language and external to it. The assumption in Generativist theory that biologically innate neural mechanisms predetermine the form of language (to a range of alternatives for parametric selection) constitutes such an assertion. Harris shows explicitly that, however sufficient this assumption may or may not be, it is not necessary, so that parsimony of theoretical primitives (Occam's razor) urges against it.

49 It has been claimed that string analysis (also termed adjunction grammar or center-and-adjunct grammar), "is, in effect, a variant of phrase-structure description" (Chomsky1975:43, Postal 1964). Joshi et al. (1969) shows that the hierarchy of grammars defined by using the adjunction rules of string analysis is distinct from the Chomsky hierarchy of phrase structure grammars defined by using rewrite rules, with at best only weak equivalence for certain types. It is of course true that "a language cannot be fully described in purely constructional terms without the transform relation" (Harris 1957:338), and that includes string grammar; if this is what which Chomsky (1975:43) means by `in effect', the sharing of this sort of deficiency is a weak form of equivalence indeed. More importantly, it is not inherent in the combinatorial approach,for we see in operator grammar a constructional or combinatorial account that incorporates the results of transformational analysis.
One difficulty with rewrite rules is that they cannot directly show word dependencies, either of adjunction (the head-of relation, properly speaking) or of predication (operator-argument dependencies).Adjunction rules show the head-of relation directly, as a dependency between words that are adjacent at the point in the derivation of a sentence at which the given adjunction is made (though other adjuncts may subsequently intervene).Joshi's tree-adjoining grammar (TAG) formalism exploits this complementarity. TAGs use adjunction rules to adjoin phrase-structure subtrees (rather than the structured strings to which they are equivalent), and use rewrite rules in a very restricted way to generate the center and adjunct strings; see Joshi (1969a, 1972, 1983), Joshi, Levi, & Takahashi (1975, 1982), Kroch & Joshi (1985). In an analogous way, operator grammar could be studied as the combination of dependency grammar rules and adjunction. However, the appearance of an adjunction turns out to be a byproduct or effect of reductions applying to a paratactic secondary sentence interrupting the host sentence just after the head (modified) word. For a succinct description, see Harris (1988:24-25), and for greater detail see Harris (1982:120-127).

without destroying its acceptability.[50] What remains is an elementary sentence called the center or center string; the parts excised are adjuncts, either to the center or to other, previously entering adjuncts. Each substring makes its own identified contribution to the information in the sentence.

String grammar was first developed as a framework for computable syntactic analysis. A string framework for transformational analysis[51] has lent itself well to applications of discourse analysis in which structured databases are generated from free-form sublanguage texts. (This work is partially summarized in Sager 1975; see also Sager(1986).) It is especially relevant to the string approach that members of a `word' class in a sublanguage grammar in many cases comprises several morphemes and words in a general-coverage grammar or in the grammar of a different sublanguage. For example, the phrase *the beating of the heart* is a member of the *SYMPTOM* class in a sublanguage of pharmacology (in, e.g., *Digitalis affects the beating of the heart*), but in a sublanguage of the logically and epistemically prior science physiology this phrase is analyzed further, with *heart* being a member of the *BODY-PART* class and *beating* a member of the *ACTION* class (see Harris 1982b for discussion).

10. Sublanguage analysis and science languages

We will consider last Harris' pioneering of sublanguage analysis in the light of two of his longstanding interests (shared with Bloomfield and Sapir): in auxiliary languages for international cooperation, and in the application of linguistics to the advancement of science. Informally, a sublanguage may be thought of as a form of language specialized for a particular subject matter, especially a technical subjectmatter. Sublanguage grammar is

50 One help in practical linguistic analysis is to choose those excisions that generalize over satisfiers of a sentence form (sentences which are instances of the same sequence of form classes). The effect is that progressive excision does not destroy the central meaning of the sentence, which continues to be conveyed in the remainder after excision and in the center string that remains after all adjuncts have been excised. As always, one may take this semantic effect as the informal criterion, using the formal criterion for verification and for difficult cases. A Kashmiri informant once responded to this analytical procedure,with a chuckle, that one could continue in this way until there was virtually zero information remaining, and as I told him then, this is indeed the point, so that each increment of information is accounted for by an increment to the structure of the sentence. I asked Harris once if he thought string analysis could be done working with an informant on a language that one did not know. He replied unhesitatingly that one would have to control the language oneself to achieve anything like full coverage. Of course, informant work constitutes a process of language learning and not merely acquisition of data, as anyone knows who has done any significant amount of linguistic work in the field. Although one rarely attains native control of a second language, I have found in practice that it is generally pretty obvious what is head and what is adjunct, even in a so-called exotic language.

51 On the relationship between string grammars and transformational grammars see Harris (1961), Joshi (1965, 1969a). In transformational grammar, the adjuncts are derived from conjoined sentences by series of deformations and reductions. Most of the work in developing a transformational grammar from a string grammar lies in defining and refining word subclasses, starting with the gross word classes and superclasses that suffice for string grammar. (See Sager 1975 and references cited there.) The relation to operator grammar is perhaps even more direct, by defining the scope and domain of the reductions, and of course the word dependencies within strings and between adjunct and host.

freed from the proscription against an external metalanguage. This is because sublanguage terms may be defined in ordinary language terms, and the metalanguage for the language as a whole[52] is indeed external to and prior to the sublanguage grammar that it describes. This encapsulation gives rise to capacities and characteristics of sublanguage grammar that are not otherwise available for grammar—for example, the capacity to distinguish nonsense in a well-defined way, and even certain types of practical discovery procedures for sublanguage analysis (e.g. Grishman et al. (1986)).

As it turns out, the information structures in the sublanguage of a field of science closely parallel and (apparently) help to constitute the subject matter of that field: changes in scientists' understanding of the subject matter of a science subfield appear to be closely parallelled by *concurrent* changes in the structure of the sublanguage for the subfield. Indeed, the concurrency suggests that the latter may in part support, facilitate, or enable the former (Harris et al. 1989). The task of translation is far simpler and more direct in science sublanguages than in the general case, and when sublanguage grammar of a given field is compared cross-linguistically (e.g. between English and French or German), the same structures are found, and the same changes of structure concurrent with changes in the field. This indicates that something new has developed here: not a sublanguage of any one language, but a science language.

> [I]f we are freed of the need to state the grammar of the sublanguage in a way that is maximally similar to the grammar of its whole language, we can maximize the regularizations in describing the sentence types of the science information even if this decreases the similarity to the natural languages in which the articles were written [whose analysis yielded the grammar—BN].
>
> We then obtain a representation of the sentences of the science consisting of word classes and subclasses (the latter indicatable by subscripts) which combine into particular formulaic sentence types and formulaic sentences [… which] constitute a canonical form for the information in the science; and the grammar of least constraints which they satisfy constitutes a new linguistic system, differing somewhat from the set of natural languages, and similar in some respects to mathematics. We thus obtain a representation of science sentences by sentential formulas free from the specific words of language. (1991:290-291).

A science language shows certain sentence types whose structures would emerge outside the science language grammar only in discourse analysis. In contrast with the openness of ordinary vocabulary,its vocabulary is effectively closed at any given time. Selection is binary (yes-no), as distinct from the graded selection of ordinary language, and it is this that enables the clear demarcation of nonsense in science. New and more complex argument-dependencies may be emerging in science languages, and other differences and

52 The metalanguage is itself a sublanguage (1991:247-278), as are certain subsets of sentences defined by operator grammar (273-274). The discussion here focuses on subject-matter sublanguages (278ff.), particularly science sublanguages (283-290).

new capacities with respect to ordinary language may be disclosed by further research.

11. Conclusion

In this brief survey of selected aspects of Harris' work I have tried to show that the commonplace view of Harris in the `Neo-Bloomfieldian' stereotype is badly mistaken. According to this standard conception, he sought to devise discovery procedures whereby a grammar could be derived from distributional analysis of a corpus of utterances without reference to meaning. His avoidance of meaning, it is ordinarily claimed, was motivated by empiricist (logical positivist) and behaviorist principles that forbade reference to "mentalist" concepts. Against this view, I have argued that Harris at first avoided reliance on judgments of meaning only because they are too difficult to control (e.g., "struggling with exact meanings (1951a:31fn4)), and that he later showed why meanings are *necessarily* difficult to control in any way that is relevant for linguistics, since (except for sublanguage grammar) there is no means outside of language itself by which to specify them (no external metalanguage). More generally, I have shown how Harris has from an early stage been engaged in the study of linguistic semantics, and that he achieved results that are of very great importance for linguistics and for the advancement of science in general.

I have barely indicated some of the other threads that must be taken up in any responsible assessment of Harris' work, such as his emphasis on social convention as distinct from physical or biological determinants of individual psychology, his perspective on language variation and change, and his suggestions regarding the origins of language, both in language acquisition by individuals and in the evolution of the species. Other themes have suggested themselves just in the course of this writing, including ongoing work on sublanguage analysis, particularly the study of argumentation and proof in languages of science and of the emergence of new capacities not found in ordinary language; Harris' conception of the proper conduct of scientific inquiry, the role of methodology, the range of methodological questions he investigated and his way of working with them; implications that may be drawn from the essential properties of language that Harris identified as making possible a mathematical treatment; the application of Harris' principle of least description to other theories of language as a "test of relevance"(though Harris would never have undertaken or encouraged such comparisons); the use of Harris' theory of linguistic information in connection with ethnographic, sociological, psychological, stylistic, literary, and other studies of language use; the origins of the autonomy hypothesis, especially in relation to the descriptive incapacities and the enforced abstractness of phrase-structure based formalisms; the potential for an international language for scientific cooperation. Countless other lines of research are possible, as with any paradigm of work in science. One might consider in terms of operator grammar issues that have been problematic in Generativist theory at onetime or another, as for example unbounded dependencies (Nevin, MS) or quantifier scope. Harris himself expressed interest in sign language and poetics (p.c. following one of the Bampton lectures at Columbia Univ. in 1986).

I have touched on only a few of the misconstruals of Harris' earlier work by his contemporaries (which have often been taken as definitive) and in more recent writings about American structuralism. Among the direct consequences of this reassessment of Harris' views is the conclusion that he cannot be held responsible for the doctrine of the autonomy of linguistic form. With Bloomfield and Sapir, he saw that linguistic form (pattern,configuration) is identical with that aspect of meaning which is most transmissible precisely because it is socially constituted with language.

Munz (1972), followed by Hymes & Fought (1981), suggests that Harris' views gradually shifted from distributionalism to a concern with meaning during the 1950s and 1960s. One result of the examination undertaken here should be the recognition that the motivation of Harris' distributionalism from an early stage was semantic, namely, the investigation of the meanings that are constituted by linguistic form, and that in this he carried forward like concerns of Bloomfield and Sapir. "It was all there from the beginning; there was no `shifting from distributionalism to meaning', only the constant and unexpected restating and reshaping. After all, the early transformations (with Chomsky and Lukoff) were precisely aimed at the entities whose *distribution* is appropriate to syntax (and semantics). That is why the work was so new and dazzling" (Hoenigswald, p.c.). This continuity has been obscured by the emergence of the `autonomy hypothesis' in generativist theory (really, several distinct hypotheses—see Ryckman 1986:179-196), and the rhetoric of standing on the shoulders of disavowed predecessors. That too, however, is a thread to be unravelled on another occasion.

Author's address:

Bruce E. Nevin
101 Chase Road
Edgartown, MA 02539-7401 U.S.A.
bnevin@alumni.upenn.edu

References

Bateson, Gregory. 1967. "Cybernetic Explanation" *American Behavioral Scientist* 10:8.29-32. (Repr. in Bateson 1972:399-410.)

_____. 1968. "Redundancy and Coding" Chap 22 in *Animal Communication: Techniques of study and results of research* ed. by Thomas A. Sebeok. Bloomington: Indiana Univ. Press. (Repr. in Bateson 1972:411-425.

_____. 1972. *Steps to an Ecology of Mind*. New York: Ballantine Books.

_____. 1979. *Mind and Nature: A necessary unity*. New York: E. P. Dutton.

Bloomfield, Leonard. 1927a. "On Recent Work in General Linguistics" *Modern Philology* 25:211-230. (Repr. in Bloomfield 1970:173-190.)

_____. 1927b. Review of Jespersen, *Philosophy of Grammar* (London: Allen & Unwin, 1924). *Journal of English and Germanic Philology* 26:444-446.

_____. 1933. *Language*. New York: Henry Holt & Co.

_____. 1939. *Linguistic Aspects of Science*. *International Encyclopedia of Unified Science* Chicago: Univ. of Chicago Press.
_____. 1942. "Philosophical Aspects of Language". *Studies in the History of Culture: The disciplines of the humanities*, 173-177. (Repr. in Bloomfield 1970:400-405.)
_____. 1943. "Meaning" *Monatshefte für Deutschen Unterricht* 35.101-106. (Repr. in Bloomfield 1970:400-405.)
_____. 1970. *A Leonard Bloomfield Anthology*. Ed. by Charles F. Hockett. Bloomington: Indiana Univ. Press.
Borel, Emile Felix Edouard Justin. 1928. *Leçons sur la theorie des fonctions.* 3e ed. Paris: Gauthier-Villars & Cie.
Bruner, Jerome Seymour. 1983. *Child's Talk: Learning to use language*. With the assistance of Rita Watson. New York: W. W. Norton.
Campbell, Jeremy. 1982. *Grammatical man: Information, entropy, language and life*. New York: Simon & Schuster.
Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.
_____. 1971. "Onward and Upward With the Arts: John is easy to please" *The New Yorker*, May 8, 1971: 44-87. Interview with Noam Chomsky by Ved Mehta.
_____. 1975a [1955-56]. *The Logical Structure of Linguistic Theory*. New York: Plenum Press.
_____. 1975b. "Questions of form and interpretation", *The scope of American linguistics* ed. by Robert Austerlitz, 159-106. Lisse/The Netherlands: Peter de Ridder Press. (Repr. in *Essays on Form and Interpretation*, 25-59, New York: Elsevier-North Holland, 1977.)
_____. 1992. A minimalist program for linguistic theory. Draft. MIT Linguistics Department.
Fodor, Jerrold A. 1980. "Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology". (Repr. in Jerrold Fodor, *Re Presentations*, 225-253. Cambridge, Mass: MIT Press, 1981.)
Gibson, James J. 1982. *Reasons for Realism: Selected essays of J. J. Gibson*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
Grishman, Ralph, & Richard Kittredge eds. 1986. *Analyzing Language in Restricted Domains: Sublanguage description and processing*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
Grishman, Ralph, Lynette Hirschman, & Ngo Thanh Nhan. 1986. "Discovery Procedures for Sublanguage Selectional Patterns: Initial experiments". *Computational Linguistics* 12:3.205-215.
Harris, Zellig S. 1941. Review of N. S. Trubetzkoy's *Grundzüge der Phonologie* (Prague, 1939). *Language* 17:345-349. (Repr. in Harris 1970:706-711.)
_____. 1946. From Morpheme to Utterance. *Language* 22.3:161-183. (Repr. in Harris 1981:45-70.)

_____. 1951a. *Methods in Structural Linguistics*. Chicago: Univ. of Chicago Press.
_____. 1951b. Review of Sapir (1949). *Language* 27.3:288-333. (Repr. in Harris 1970:712-764.)
_____. 1952a. Discourse Analysis. *Language* 28:1-30. (Repr. in Harris 1970:313-348.)
_____. 1952b. Discourse Analysis: A Sample Text. *Language* 28:474-494. (Repr. in Harris 1970:349-372.)
_____. 1952c. Culture and Style in Extended Discourse. In Sol Tax (ed.) *Indian Tribes of Aboriginal America*, Proceedings of the 29th International Congress of Americanists, Chicago: The Univ. of Chicago Press (210-215).
_____. 1954. Distributional Structure. *Word* 10.2-3:146-162. (Repr. in Harris 1981:3-22.)
_____. 1955. From phoneme to morpheme. *Language* 31.2:190-222. (Repr. in Harris 1970:32-67.)
_____. 1956. Introduction to Transformations. *Transformations and Discourse Analysis Papers* 2. Univ. of Pennsylvania Linguistics Department. (Repr. in Harris 1970:383-389.)
_____. 1957. Co-occurrence and Transformation. *Language* 33:283-340. (Repr. in Harris 1970:390-457.)
_____. 1959. Computable Syntactic Analysis: The 1959 computer sentence-analyzer. *Transformations and Discourse Analysis Papers* 15. Excerpted in Harris (1970:253-277).
_____. 1961. Strings and Transformations in Language Description. Papers on Formal Linguistics 1. Philadelphia: Univ. of Pennsylvania Department of Linguistics.
_____. 1962. *String Analysis of Language Structure.* The Hague: Mouton.
_____. 1963. Immediate-Constituent Formulation of English Syntax. *Transformations and Discourse Analysis Papers* 45. (Repr. in Harris 1970:131-138.)
_____. 1965. Transformational Theory. *Language* 41:363-401. (Repr. in Harris 1970:533-577.)
_____. 1966. Algebraic Operations in Linguistic Structure. Paper read to the International Congress of Mathematicians, Moscow 1966. (Repr. in Harris 1970:603-611.)
_____. 1967. Morpheme Boundaries Within Words: report on a computer test. *Transformations and Discourse Analysis Papers* 73. Univ. of Pennsylvania Linguistics Department. (Repr. in Harris 1970:68-77.)
_____. 1970. *Papers in Structural and Transformational Linguistics*. Dordrecht: D. Reidel.
_____. 1981. *Papers on Syntax*, edited by H. Hiz, Synthese Language Library, D.

Reidel.
_____. 1982a. *A Grammar of English on Mathematical Principles*. New York: Wiley.
_____. 1982b. Discourse and Sublanguage. *In* Kittredge and Lehrberger (1982:231-236).
_____. 1988. *Language and Information*. Number 28, Bampton Lectures in America Delivered at Columbia University. New York: Columbia Univ. Press.
_____. 1991. *A Theory of Language and Information: A mathematical approach*. Oxford: Clarendon Press.
_____. 1990. "La genèse de l'analyse des transformations et de la métalangue". *Langages* No.99 (Sept. 1990), 9-19.
_____, Michael Gottfried, Thomas Ryckman, Paul Mattick, Jr., Anne Daladier, Tzvee N. Harris, and Suzanna Harris. 1989. *The Form of Information in Science: Analysis of an immunology sublanguage*. Preface by Hilary Putnam. *Boston Studies in the Philosophy of Science* 104. Boston: Kluwer Academic Publishers.
Hockett, Charles F. 1947. "Problems of Morphemic Analysis". *Language* 23:321-343. (Repr. in Joos 1957:229-242.)
_____. 1952. "A formal Statement of Morphemic Analysis". *Studies in Linguistics* 10:27-39.
Huck, Geoffrey and John A. Goldsmith. 1993. "Gaps in the Paradigm". (*Contemporary Linguistics*, 1.) Chicago: Department of Linguistics, Univ. of Chicago.
Hymes, Dell, & John Fought. 1981. *American Structuralism*. The Hague: Mouton.
Joos, Martin 1957. *Readings in Linguistics: The development of descriptive linguistics in America since 1925*. Washington, D.C.: American Council of Learned Societies. (4th ed., Chicago, Univ. of Chicago Press, 1966.)
Joshi, Aravind K. 1965. "String Representation of Transformations". (Transformations and Discourse Analysis Papers 58.) Philadelphia: Department of Linguistics, Univ. of Pennsylvania.
_____. 1969a. "String Adjunct Grammars and Transformational Grammars". (Transformations and Discourse Analysis Papers 75II.) Philadelphia: Department of Linguistics, Univ. of Pennsylvania.
_____. 1969b. "Formal Properties of Grammars With Mixed Types of Rules and Their Linguistic Relevance." *Proceedings of the International Conference on Computational Linguistics*, Sänga Säby, Sweden, August 1969.
_____. 1972. "How much Hierarchical Structure is Necessary for Sentence Description?". Plötz 1972:389-398.
_____. 1983. "How Much Context-Sensitivity is Required to Provide Reasonable Structural Descriptions: Tree adjoining grammars." *Natural Language Processing: Psycholinguistic, Computational and Theoretical Perspectives* ed. by David Dowty, Lauri Karttunen, & Arnold Zwicky. New York: Cambridge Univ. Press.
_____, S. R. Kosaraju, & H. Yamada. 1969. "String Adjunct Grammars." (Transformations and Discourse Analysis Papers, 75.) [Dec. 1968; Revised

February 1969.] Philadelphia: Linguistics Department, Univ. of Pennsylvania.
_____, & L. Levy. 1982. "Phrase Structure trees Bear More Fruit Than You Would Have Thought". *American Journal of Computational Linguistics* 6:2.272-284.
_____, L. Levy, & M. Takahashi. 1975. "Tree Adjunct Grammars". *Journal of the Computer and System Sciences* 10:1.136-163.
Katz, Jerrold J. 1972. *Semantic Theory*. New York: Harper & Row.
Kittredge, Richard, & John Lehrberger, eds. 1982. *Sublanguage: Studies of Language in Restricted Semantic Domains*. Berlin: Walter de Gruyter.
Koerner, E. F. Konrad. 1970. "Bloomfieldian Linguistics and the Problem of `Meaning': A chapter in the history of the theory and study of language. *Jahrbuch für Amerikastudien / German Yearbook of American Studies* 15.162-83. (Repr. in Koerner, *Toward a Historiography of Linguistics: Selected Essays*, 157-176 *Amsterdam Studies in the Theory and History of Linguistic Science*, 19, Amsterdam: John Benjamins.)
Kroch, Anthony S. & Aravind K. Joshi. 1985. "The Linguistic Relevance of Tree Adjoining Grammar." MS-CIS-85-16, LINC LAB 03. Philadelphia: Department of Computer and Information Science, Univ. of Pennsylvania.
Lieberman, Philip & Sheila E. Blumstein. 1988. *Speech Physiology, Speech Perception, and Acoustic Phonetics*. Cambridge: Cambridge Univ. Press.
Munz, James. 1972. "Reflections on the Development of Transformational Theories". Plötz (1972.251-274).
Nevin, Bruce E. MS. [1989] Unbounded dependencies in operator grammar. Unpublished paper.
Newmeyer, Frederick J. 1980. *Linguistic Theory in America: The first quarter-century of transformational generative grammar.* New York: Academic Press.
Plötz, Senta, ed. 1972. *Transformationelle Analyse: Die Transformationstheorie von Zellig Harris und ihre Entwicklung/Transformational Analysis: The transformational theory of Zellig harris and its development.* Frankfurt/M.: Athenäum.
Postal, Paul M. 1964. *Constituent Structure*. The Hague: Mouton.
Powers, William T. 1973. *Behavior: The control of perception*. Chicago: Aldine.
Putnam, Hilary. 1975. "The Meaning of `Meaning'". *Language, Mind, and Knowledge*, (*Minnesota Studies in the Philosophy of Science*, vol. 7), Minneapolis: U. of Minnesota Press. Ed. by K. Gunderson.
Ryckman, Thomas A. 1986. *Grammar and Information: An investigation in linguistic metatheory*. Unpublished Ph.D. dissertation, Columbia Univ., New York.
_____. 1991. Zellig Harris' Methodology of Language and Information. Lecture, Boston Colloquium for the Philosophy of Science, 8 October, 1991.
Sager, Naomi. 1975. *Natural Language Information Processing*. Reading, Mass.: Addison-Wesley.
_____. 1986. "Sublanguage: Linguistic phenomenon, computational tool". Grishman & Kittredge 1986.1-18.

Sapir, Edward. 1929. "The Status of Linguistics as a Science." *Language* 5.207-214. (Repr. in Sapir 1949.160166. [Page references are to the reprint.])

_____. 1949. *Selected Writings of Edward Sapir.* Ed. by David G. Mandelbaum. Berkeley & Los Angeles: Univ. of California Press.

Shannon, Claude. 1949. *The Mathematical Theory of Communication*. With an essay by Warren Weaver. Urbana, Ill.: Univ. of Illinois Press.

_____. 1956. "The Bandwagon". (*IRE Transactions on Information Theory* 2:1.3) Cambridge, Mass.: M.I.T.

Wells, Rulon S. 1947. "Immediate Constituents". *Language* 23:81-117. (Repr. in Joos 1957:186-207.)

_____. 1962. "What has Linguistics Done for Philosophy?" *Journal of Philosophy* 59.697-708.

Williams, Gregory. 1992. *A Bibliography of Hierarchical Perceptual Control Theory.* Gravel Switch, Kentucky: CSG Publications.