is flawed in one crucial respect. Goodman provides compelling arguments that his measure is a general measure of formal simplicity. Chomsky, on the other hand, proposes including in linguistic theory notational conventions whose effect is to contract rules in grammars. The inclusion of a particular convention is not a formal matter. It is based on empirical experience with successful grammars. The net effect of the correct conventions is to convert best empicially adequate grammars for particular languages into shortest empirically adequate grammars. Thus what Chomsky proposes is not a general, formal measure of complexity but an empirical measure of linguistic goodness. His use of the word "simplicity" is unfortunate and has fathered many senseless misunderstandings. Chomsky has reversed the principle of parsimony – for him a theory is simplest because it is best.

39 Goodman, [AMS], [TS]. The last paper provides a good introduction which avoids many of the technical details.

40 Mehta, *loc. cit.*

41 In the Mehta interview Chomsky suggests that he had difficulty finding a publisher for {SS}. This may have been a serious obstacle for the publication of the much longer {LSLT} for a time. Certainly his work on other books has also competed with the updating of [LSLT] for publication. Still the non-publication of [LSLT] must have been the result of choice rather than necessity.

42 Mentalism was common in linguistics – e.g., Sapir, Whorf, deSaussure – but this tradition is sufficiently different from Chomsky that influence is unlikely. Harris is not aware of the source, if any, of this influence on Chomsky.

43 cf Chomsky, [ATS], Ch 1.

44 Chomsky, [LSLT], p. 174.

45 Chomsky, [NRG], pp. 17-20; Chomsky, [LSLT], pp. 377-406.

46 [TT], p. 537.

47 [CO], p. 449.

48 [CO], Loc. cit.

49 [TIR], p. 461.

50 [MSL], p. 1.

51 [MSL], p. 2.

52 [MSL], p. 211.

53 [RP], p. 613.

54 [RP], p. 680.

55 In so characterizing it, I do not mean to demean it. The requirement for "elegant theories" is shared with mathematicians and common among good empirical scientists.

56 The claim that all transformations are paraphrastic is not as strong as it appears. I'll comment on it briefly at the end of this section.

57 [TT], pp. 539-540.

HARRIS ON THE STRUCTURES OF LANGUAGE

John Corcoran

Zusammenfassung

Diese knappe interpretative Betrachtung über die Ideen von Z. Harris zur globalen Struktur der Sprache beginnt mit einer Behandlung der „Kern-plus-Transformationen-Struktur". Der zweite Teil definiert die „Basis-plus-Kodierer-Struktur" und zeigt einige charakteristische Unterschiede zwischen den beiden Strukturen. Der dritte Teil bringt das Konzept der eingebetteten Prädikation. Im vierten und letzten Teil wird die These aufgestellt, daß Harris in seiner Arbeit in den späten 60er Jahren die „Basis-plus-Kodierer-Struktur" mit dem Konzept der eingebetteten Prädikation verbunden hat, um so zu einer neuen Sicht der Struktur der Sprache zu gelangen. Diese scheint bestimmte theoretische Probleme, die mit der Kern-plus-Transformationen Position verbunden waren, zu lösen. Wesentliche Züge beider Positionen haben ihren Ursprung in Harris' Konzept des "extended discourse". Während die frühere Position semantische Erwägungen vermeidet, ist die spätere Sicht in expliziter Weise auf sie gestützt.

Abstract

This short, interpretative account of Zellig Harris' ideas on global language structure begins with a discussion of the development of the kernel-plus-transformations structure. The second section defines the base-plus-encoders structure and lists a few of the characteristic differences between the two structures. The third section treats the concept of nested predication. In the fourth and final section, it is suggested that in his work of the late 60's Harris combined the base-plus-encoders structure with the concept of nested predication to form a new view of the structure of language. The latter appears to resolve certain theoretical problems inherent in the kernel-plus-transformations view. Major features of both views can be seen to be outgrowths of Harris' notion of extended discourse. The earlier view eschews semantic considerations while the later view is explicitly predicated on them.

> A less rigorous statement may be more
> useful even for scientific purposes.
> ... *Bloomfield*

This work is a short, interpretative account of Zellig Harris' views concerning both the global components of a natural language and also the interdependence of the components in the functioning language system. Although in several places the account takes the form of a chronological narrative tracing the development of Harris' ideas, the aim of the work is more interpretative than historical and, consequently, no attempt has been made to decide academic issues. For example, attribution of an idea to Harris does not imply that it is Harris' invention or discovery.

It has been said that Harris is more a methodologist than a theoretician (Hockett, p. 35). Indeed, the foundation of Harris' work rests on several methods for analyzing bodies of linguistic data (Harris, 1951, 1952b, 1955, 1968) and on his view that a language structure arises only as a result of a particular mode of analysis (Harris, 1951, p. 1; 1954, pp. 36-38). Different methods result in different structures. For Harris, to say that a given language has a given mathematical structure is to say nothing more than that the structure describes and organizes results obtained by applying a particular method of analysis to the data concerning the language (Harris, 1954b, pp. 36-38; 1965). Harris has not permitted himself to speculate concerning whether a language has a particular objective structure independent of any particular method of analysis. Nor, as others have pointed out, has he speculated concerning the origin of language, the nature of language learning, the relation of speech and thought (Harris, 1952b, pp. 356-7; 1954b, p. 38) and so on. Harris' emphasis is descriptive and empiristic.

Harris' methodological relativism has been criticized as "game-playing" and praised as modesty. Neither the blame nor the praise seem apt. There is neither sport nor modesty in Harris' insistence that anything that could be called knowledge of language is the result of rigorous application of more-or-less arbitrary methods of analysis to linguistic data. Although this view shows a keen awareness of the conclusions of two centuries of "critical philosophy", some may feel that Harris has not sufficiently emphasized that his outlook applies no less to "physical reality" than to "linguistic reality". For Harris, as for all great linguists, language is as "real" as natural phenomena, indeed it is a natural phenomenon (Harris, 1954, 1959, 1964).

The emphasis in this paper is theoretical and conceptual. Thus the character of this account is alien to the character of the material with which it deals. Indeed one major reason for this work is to bridge the chasm between Harris' empiricistic operationalism and the rationalism of many younger linguists. Without this or some similar work many of Harris' ideas are likely to be overlooked in future developments.

However, a conceptual approach is not entirely alien to the Harrisian corpus taken as a whole. In his later work (esp. Harris, 1969) one can detect at least the germs of a shift from methodology to theory. The older post-Bloomfieldians were describers, analysts and cataloguers. The younger men have opted for theory, sometimes accompanied by speculation which seemed somewhat metaphysical to their teachers. In Harris' earlier work their was *no room* for speculation (e.g., concerning the relation of form and meaning, concerning how "ideas" are converted into sound patterns, etc.). His later work, however, sometimes even seems designed to elicit such speculation. As will be seen below, methodology is still present in Harris' thought but, despite Harris' own claims to the contrary, structure no longer appears to be based on methodology in any convincing sense. If one could extrapolate from the later works one could predict that Harris will deemphasize methodology and enter a period of highly theoretical, perhaps even speculative, work. (If a personal, subjective comment be allowed, I would say that Harris' 1969 paper is the most clearly written of his works and that he seems even more comfortable in his role as a "latent theoretician" than he did in his role as a methodologist.)

1. A ROAD TO KERNEL-PLUS-TRANSFORMATIONS

Harris presupposes that each language is a real, existent system of utterances naturally classed into types. The type-token distinction even on higher levels is thus fundamental to Harris' thought, although the types are not regarded as independently existent entities but only as logical constructs based on speakers' perceptions of whether two concrete utterances are the same or different. To conceive of a language as a whole (system of utterance types) rather than a disconnected set of some sort is to free oneself somewhat of excessively nominalistic scruples so that one can raise questions about the global structure of a language.

Harris became quite clear, contrary to Bloomfield, that regularities in utterances are not limited to sentences, but also are found in structured sequences of sentences or discourses (Harris, 1952b, pp. 356ff). Thus, Harris was one of the first thinkers (cf. Fries, 1952, p. 290) to countenance the linguistic reality of texts, of conversations, and, generally speaking, of a level of linguistic reality above the sentential level (Harris, 1952a). For Harris, then, a description of the phonemes, morphemes, phrases and sentences of a language was not enough – linguists must account for discourses as well (cf. Harris, 1968, pp. 14ff.). The importance of the recognition of extended discourse should not be underestimated. One should realize that at the time there was no thought of "generative grammar" (anywhere outside of logic) and it is difficult, if not impossible, to find unambiguous statements of a *linguistic* view of a language as a set of sentences, much less a set of discourses.

In the early work on discourses Harris saw that constituents of a discourse were interrelated in a way describable by means of mathematical functions. These functions were always unary (one argument) and they were always one-one. Indeed Harris often spoke of them as equivalence relations. Apparently the term *grammatical transformation* is first used in connection with an operation which "transforms certain sentences of the text into grammatically equivalent sentences" (Harris, 1952b). The term *transformation,* it should be noted, had been used in mathematics to denote unary, one-one functions (only!).

Notice that an infinite set can be generated from a subset by means of unary, one-one functions only if *either* the subset is infinite *or* the set of functions is infinite or the functions are iteratively applied to their own resultants. In the 1952 papers there is no conception of iteration of transformations. Thus, in his original view, transformations did not "produce" one sentence from another, in any but a strictly metaphorical sense, but rather a transformation was merely a functional relationship among sentences as they co-occurred in extended discourses. Hardly had this idea been discovered when Harris began changing it. One of the first changes was to divorce the transformation from discourses and to see it merely as a functional relationship among sentences. Such a concept is a natural complement to the view of language as a system. One could conceive of the transformations as providing the set of sentences with essential interconnetion – without the transformations the set of sentences would *not* be a system but merely a disconnected set.

Given the vision of a language as a set of sentences interrelated by transformations, one must ask what one can say about the global structure of the language. Harris conceived of the gedanken-experiment of choosing a sentence and a transformation leading to that sentence and then considering the sentence or sentences that the transformation operated on to produce the given sentence – in other words, Harris was led to ask what one would arrive at by tracing transformations backward. Could one ultimately come back to the given sentence (find a loop) or are there ultimately simple sentences from which the rest are produced by applying the transformations? On the basis of scientific experience with English and several other languages he became convinced that the latter is the case. Sometimes he gives the impression in writing and in conversation that he had proved that there are no loops in a language, i.e., that each sentence is either ultimately simple (transformationally) or else can be obtained as the value of a sequence of transformations applied, to begin with, to simple sentences (Harris, 1957). But, of course, before any of this could be conceivable it was necessary to get away from the idea of a "transformation" as an equivalence relation among sentences *and* to introduce the idea of a binary transformation. By

now, the analogy between "transformations" in linguistics and "transformations" in mathematics is completely lost – but the term continued to be used. (To my knowledge the term "transformation" has never been used in mathematics for binary functions and I doubt that it was ever used to indicate anything other than a *one-one* function from a set *onto* itself, cf. Rogers, 1967, pp. 50ff.)

Harris called the simple sentences *kernel sentences* and he called the set of simple sentences the *kernel* of the language. Once a language is seen to have what is here called a *kernel-plus-transformations structure* (cf. Harris, 1957) a totally new possibility for descriptive linguistics may *seem* to emerge, viz., the possibility of giving a systematic and comprehensive description of a language by characterizing its kernel and its set of transformations. From the present point of view, however, this is an illusion: the transformations are *defined* on the sentences of the language and, therefore, a description of the transformations presupposes a description of the language. Harris seems to have been aware of the foregoing argument but as far as I have been able to determine he has not advanced it in print. (I remember reading the gist of it in one of his mimeographed reports but I can not now locate it in his published writings).

It is true, of course, that Harris had previously mentioned the idea that a grammar may be thought of as a set of instructions for "generating" the sentences of a language (1954a, p. 260). But, it is quite clear from the context that he is not thinking of the transformations as being the instructions. Not until more than ten years later did Harris even use the term "transformational grammar" (1965, p. 383) and even there it is clear that he is primarily interested in transformational analysis rather than transformational synthesis.

Once the kernel-plus-transformations structure is conceived several interesting problems emerge which occupied Harris in the late 50's (Harris, 1957, 1959). The first problem is to describe the kernel and to discover and enumerate the transformations all relative to a given language, e.g. English. A tremendous amount of work was needed to get a clear idea of the extent of the kernel and to get a set of transformations which conceivably might be sufficient. By the way, it might be appropriate to note that at this point Harris' concern with the exact nature of extended discourse (above the sentential level) took a decidedly secondary place to his more immediate concerns with transformations as providing the structure of the sentential level.

The problem of finding the kernel did not prove to be excessively difficult but the problem of finding all of the transformations became very demanding. There were various conjectures about how many transformations were needed and how they should be classified and whether there was a general method for breaking a transformation down into more elementary parts.

At some point or points in this development twin problems began to plague various workers. First, does it make sense to say of every string of phonemes that it definitely is or definitely is not a grammatical sentence of English, and if so what "scientific" criterion could be used? In order to understand the second problem it is necessary to understand that, given the set of sentences of a language, there is literally an uncountably infinite number of functions which, when applied to sentences, form other sentences. Some, in fact the overwhelming majority, are of no interest or relevance to linguistics (Harris, 1964, p. 420). For example, the active-passive transformation *is* a relevant function but the following function, f, (here defined) is *not* – if one applies f to a sentence, s, having a compound subject of the form 'N_1 and N_2' where N_1 and N_2 are both male, human, first names then f(s) is the result of inserting 'or the new boy' after N_2; if s is not of the above kind then f(s) is defined to be s (i.e., the transformation leaves s alone). Obviously f and some of its even more preposterous fellows have no relevance to the structure of English. But how can one give scientifically acceptable reasons for excluding f from the set of *transformations of English?* This actually became a very practical problem as more and more workers became interested in Harris' idea and began proposing various transformations. On what basis could Harris say that a proposed transformation was wrong – and what precisely does "wrong" mean here? The problem of distinguishing the "real" transformations of English from the functions which are accidental or irrelevant I call the *problem of the reality of transformations.* Problems concerning the concept of grammaticality (above) I call the *problem of grammaticality.* In a certain important sense these two problems are the same kind of problem. To raise one is to challenge the validity of the concept of grammatical sentence – is "grammaticality" a clear, precisely applicable concept and what criteria are there for distinguishing grammatical sentences from "arbitrary" strings of phonemes? To raise the other is to challenge the concept of a transformation of English – is "transformation of English" a clear, precisely applicable concept and what criteria are there for distinguishing "real" transformations from "arbitrary" functions relating sentences to sentences?

The problem of the reality of transformations slowly started to seduce Harris away from an understandable aversion to meaning. Harris (and many other linguists) felt that "meaning" was Pandora's box – that once one allowed oneself to start talking about meaning one could hardly avoid making "unscientific" statements, getting involved in paradoxes and contradictions, or, to say the least, violating the canons of scientific inquiry (cf. Harris, 1940, Fries, 1952, p. 293). Several persons toyed with the idea that *all* transformations are *paraphrastic,* i.e., preserve meaning. At this stage in Harris' thought the idea is obviously wrong (otherwise certain questions would mean

the same as their answers and certain sentences would mean the same as their denials). However, Harris was definitely attracted by the idea and at several points actually suggested in print that the sequence of kernel sources of a text is "roughly equivalent in information" to the text itself (Harris, 1959). At one point he admits that it is not always the case that transformations preserve meaning but then goes on to say that such differences in meaning are "either stylistic or subjective". Unfortunately in this same paper (Harris, 1964) he lists the introduction of 'because', 'while' and 'after' as something accomplished by elementary transformations. It is stretching a point to insist that unary transformations preserve meaning, but to say that binaries (S_1, $S_2 \Rightarrow S_1$ because S_2, e.g.) preserve meanings is incomprehensible as it presupposes that ordered pairs of sentences have sentential meanings.

In any case, although Harris was attracted by the idea that transformations should preserve meaning, he was clearly uncomfortable about the idea and, whether he recognized it or not, he had serious problems in saying in what sense binaries preserved meaning. In dealing with the problem of the reality of transformations, however, he took an entirely different tack which, to some readers, may be reminiscent of the notion of "frame" developed much earlier (Fries, 1952). His solution was as follows. In the first place the domain of a transformation must be a set of sentences all in the same form – the form being represented by a sequence of "word-category" symbols possibly with a few "constants", i.e., symbols for tense morphemes and special prepositions. The range of a transformation must also be a "homogeneous" set of sentences and its form must contain no word-category symbols other than those of the "domain-form". Thus each transformation only rearranges the "category" words, possibly deleting a few and possibly adding and deleting constants. *The main criterion* is that the acceptability ordering of the range must be the same as that of the domain – *the transformation must preserve relative acceptability.* Naturally, this presupposes that each homogeneous set of sentences (i.e., all sentences in a given form) can be unambiguously ordered according to acceptability (to the native speaker, perhaps). In a way this is Harris' answer to both questions as he explicitly makes the point that although grammaticality as a definite property is subjective, the relative acceptability ordering is objective. The problems with this latter view are obvious and, to my knowledge, they have not been scientifically dealt with. However, the problem of binaries which we discussed in relation to the question of whether transformations preserve meaning is still with us in regard to the question of whether they preserve relative acceptability. Surely it is gratuitous, if not meaningless, to postulate an objective relative acceptability among ordered pairs of sentences. Even more gratuitous is the view sometimes put forward that it is possible to assign numbers to acceptability orderings which, summed in some way, give the acceptability ordering

of ordered pairs (Harris, 1965). Any such scheme, it seems to me, must be arbitrary, thus undermining its purpose – viz., to give some objective criterion for distinguishing between transformations of English and "arbitrary" functions. It seems to me that Harris had real difficulties in applying his acceptability criterion to binaries – just as he had in maintaining his "view" that transformations preserve "information". It is obvious however that he is attracted by both of these views. Binary transformations form the fly in Harris' ointment.

But there are other flies as well. As people worked to carry out Harris' program, it became increasingly clear that the set of transformations would be very large and very messy.

2. BASE-PLUS-ENCODERS

In order to get a clear understanding of Harris most recent view of the overall structure of a language system, it is worthwhile to first consider some ideas which have been receiving increasing attention from various linguists. These ideas form what might be called the *base-plus-encoders* proposal for the description of particular languages. Although these ideas are comfortably dealt with merely as a proposal for describing languages, they are clearly very closely related to abstract language theories familiar from the speculations of Sapir (1921, pp. 221-226), the early Wittgenstein, and others.

In a way, the base-plus-encoders proposal can be seen having a loose family-resemblance to the kernel/transformations view. "The base" will be seen to correspond to the kernel and "the encoders" will be seen to resemble transformations – both correspondences are, however, superficial. In short, the proposal is to define, as a descriptive device, a set of symbolic objects called *base structures* and then to develop a set of mathematical functions called *encoders* which produce the actual sentences of English (say) when applied to the base structures. The hope is that by presenting the base and the encoders, one can describe exactly the sentences of English as the results of applying the encoding functions to the base structures. The reason I call the base structures "symbolic objects" is because there are many different kinds of things that have been (or might be) proposed to form the base. Keenan (1969) has suggested that logical formulas be used as base structures. Other linguists sometimes seem to suggest that labeled trees be used as base structures. One could even imagine that graphs resembling chemical ciphers might prove useful. In any case, it is a matter of indifference *here* what kind of symbolic object is used in the base. The issue is very pragmatic: arrange the base in such a way that neat, clean, easily computed functions will suffice to produce from it the sentences of the language – as strings of phonemes suitably segmented and marked for stress, intonation and whatever else is needed.

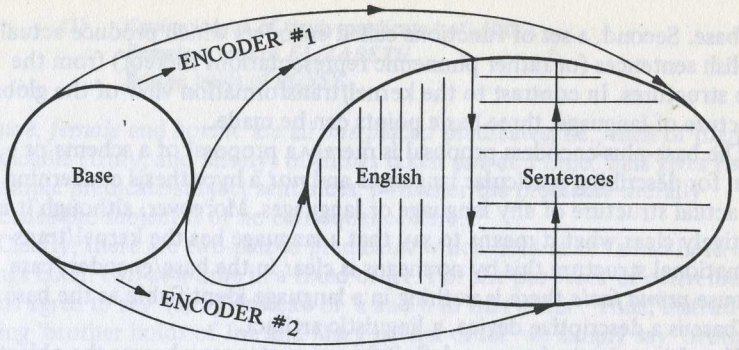

Figure 1

The diagram in figure 1 indicates a base with two encoder functions which together produce the sentences of English. In this particular case the horizontally hatched segment and the vertically hatched segment each represent a total "encoding" of the base. Thus some objects in the base will have two separate encodings ("synonymy"?) and some sentences encode two base objects ("ambiguity"?). [cf. Harris, 1957, § 34].

Naturally, if someone actually presented a base-plus-encoder grammar of English which worked, and then insisted that the base really had no significance beyond being a descriptive device, it would seem appropriate for linguists to receive it with mixed emotions. I can imagine vacillating between "that's wonderful" and "so what?". The point which was hinted at by parenthetical use of the terms 'ambiguity' and 'synonymy' is that one expects the base to have some sort of significance – e.g., that it should represent the absolute language of thought or that it should represent the realm of ideal propositions or that it should represent the field of possible electrosemantic discharges or something. Moreover one also expects the encoders to mirror some sort of process which converts what the base represents into actual sound. The very presentation of the base-plus-encoders proposal is occasion enough for much fascinating speculation concerning the nature of language, the relation of speech and thought, the learning of language by children, the nature of speech defects, and so on. In my opinion this is all to the good. However, let us not forget that all these exciting speculations were only occasioned by the base-plus-encoders view; they are by no means implicit in it. In itself it is a very modest, pragmatic, pedestrian proposal of a scheme for describing the sentences of a language as strings of phonemes without reference to their meanings or uses. The proposal has equal validity as a means of describing the output of an automatic calliope.

As a scheme for describing a language, the base-plus-encoders proposal involves two things. First, a set of non-linguistic symbolic objects called

the base. Second, a set of functions called encoders which produce actual English sentences (or rather phonemic representations thereof) from the base structures. In contrast to the kernel/transformation view of the global structure of language, three basic points can be made.

1 The base-plus-encoders proposal is merely a proposal of a scheme or form for describing particular languages and *not* a hypothesis concerning the actual structure of any language or languages. Moreover, although it is relatively clear what it means to say that a language has the kernel/transformational structure this by no means is clear in the base/encoders case because *prima facie* there is nothing in a language identifiable as the base – the base is a descriptive device, a linguistic artifact.

2 The objects in the base are definitely not sentences whereas the objects in the kernel certainly are. Moreover, the base structures can be as complex as convenience indicates. They can involve nesting, embedding, compound and complex constituents, etc. and, in particular, they can contain other base structures as parts. On the other hand, kernel sentences must be simple and cannot be composed of other kernel sentences.

Finally, because there are only a finite number of words and only a finite number of kernel forms in any language, the kernel must be a finite set of sentences. In contrast it is certainly possible, if not necessary, that an adequate base for a natural language includes an infinite number of base structures.

3 The encoders are functions producing sentences from non-sentences. Transformations, on the other hand, produce sentences from sentences. Since the base structures are not sentences and, consequently, have no relative acceptability ordering, it follows that the condition of preserving acceptability has no application to encoders. Moreover, convenience and other pragmatic criteria are the only guides for distinguishing various encoders, and, in particular, there is no problem of the reality of encoders – they are all artifacts of linguistic description. Incidentally, although there are definitely binary transformations, all encoders are unary – they apply to base structures one at a time.

Finally, because there is no requirement that an encoder operate on a homogeneous set of base structures it is conceivable, at least, that a base-plus-encoders grammar contain only a few encoders whereas it is likely that every language contains hundreds of transformations.

3. NESTED PREDICATION

Some sentences of English seem to express *something being predicated of something else.* In fact sometimes this situation is made explicit as in the following:

(1) *Canine* is true of (truly predicated of) *FIDO.*
Female is true of *ELIZABETH.*
Bovine holds of *BOSSY.*

Canine, female and *bovine* are all *one-placed* predicates, i.e., each of them is predicable (truly, and falsely) of things taken one at a time. This gives rise to simple predicate-subject sentences. [Whether the predicate literally precedes the subject is of no consequence here.]

Clearly, there are two-placed predicates which hold of or fail to hold of things taken two-at-a-time in a fixed order. For the purposes of convenience let us agree to say '*(a, b)*' instead of 'a and b in that order'. Thus, instead of saying 'brother holds of Joe and Mary in that order' we simply say 'brother holds of (Joe, Mary)'. The following are some examples.

(2) *Less-than* holds of *(2, 3).*
Identical holds of *(2, 2).*
Married holds of *(RICH, LIZ).*
Married holds of *(LIZ, RICH).*
Parallel holds of *(LOCUST, WALNUT)**.
Longer-than holds of *(CHESTNUT, LOCUST)**.

This situation generalizes to three, to four and, perhaps, to even higher numbers of places and we generalize our conventional notation accordingly. For example, instead of saying "between holds of Baltimore, Philadelphia and Washington in that order" we simply say "between holds of (Baltimore, Philadelphia, Washington)".

(3) *Between* holds of *(BALTIMORE, PHILADELPHIA, WASHINGTON).*
Mean-proportional holds of *(2, 1, 4).*
Arithmetic-mean holds of *(2, 1, 3).*
Child holds of *(JESUS, MARY, JOSEPH).*
Elected holds of *(THE AMERICAN PEOPLE, NIXON, THE PRESIDENCY).*
Difference holds of *(2,17, 15).*
Married holds of *(REVEREND JONES, ANDY, SUSAN).*

It is easy to supply further 3-placed predicates. Moreover many sentences involving verbs such as *push, move, transfer, return* and *bring* are essentially four-place predication.

(4) *Transfer* holds of *(IBM, JOHN JONES, NEW YORK, ATLANTA)*

* Locust, Walnut and Chestnut are streets in Philadelphia near the University of Pennsylvania.

A significant property of each one-placed predicate is its domain of predication. The domain of predication is not the set of things of which the predicate holds but rather the set of things of which it is literally predicable (truly or falsely). For example, consider the predicate *vertebrate.* It is significant to assert *and* to deny of any animal that it is vertebrate. It is true that dogs are vertebrates; it is false that worms are vertebrates; but to assert or deny that a building, a number, a language, or a machine is vertebrate is to involve oneself in imagery, at best. Restricted domains of predication on the semantic level are reflected in the syntax as selection restrictions. Thus the syntactic fact that *vertebrate* selects for animate reflects the semantic fact that the domain of predication of *vertebrate* is restricted to animals. Let us take note of the facts about *vertebrate* by saying that *vertebrate* is a P (animal) predicate. This notation simultaneously indicates both that *vertebrate* is one-placed and that its domain of predication is restricted to animals. If one assumes that the predicates *true* and *probable* apply only to propositions then one will classify them both as P (proposition).

Naturally the above considerations of domain of predication apply to multi-placed predicates as well as to one-placed ones. For example, *less than, greater than* and *divides* are all P (number, number) predicates and *married, brother, loves* and so on are all P (human, human) predicates. This means that *less than* can be predicated only of pairs of numbers and that *married* can be predicated only of pairs of humans. I should point out that I am not reporting the results of a lot of deep thought in giving these examples – certainly some of them are wrong. The point is that restrictions of these sorts exist, not that any particular example is correct.

Between is obviously a P (location, location, location) predicate. *Mean proportional* is a P (number, number, number) predicate. And *child* is a P (human, human, human) predicate.

So far most of the multiplaced predicates considered have been *uniform.* I.e., all places of a given predicate have one and the same domain of predication. But this is not always the case. For example, *believes* can be thought of as a P (human, proposition) predicate and *infers* as a P (human, proposition, proposition) predicate [*Joe infers the parallel postulate from the Pythagorean theorem*]. *Elected* is a P(group, human, office) predicate.

From these examples one can see that the classification of predicates according to number of places is a gross oversimplification of the fact that predicates are naturally classed according to sequences of domains of predication. It is clearly more informative to talk of Pa predicates, Pab predicates, Pabc predicates, etc. where a, b, c and so on are domains of predication. Even this is an oversimplification because in some cases selection restrictions might be *links* among the domains of predication rather than simply a sequence of such domains. For example, one can at least imagine

a two-placed predicate (pronounced *exceeds,* perhaps) which is predicated of pairs of quantities but only when both are of the same kind. That is, one could literally say that a quart "exceeds" a pint and that a gross "exceeds" a dozen, while one could not literally say that a quart "exceeds" a gross or that a quart does not "exceed" a gross.

Thus far in this section we have discussed single and multiple predication and suggested a scheme for classifying predicates.

Although there is nothing really new above we have prepared the ground for thinking of assertions as expressing nested predications. It is a matter of indifference whether predications of the form $Pa_1\ a_2\ \ldots\ a_n$ are called nested or not. However when we consider the following four sentences, in order to think of them as predications (or expressing predications) they must be thought of as nested in the strict sense, i. e., as having predicates applied to the result of predication.

(5) It is true that 2 is odd.
That 2 is even implies that 5 is odd.
Presently, John stinks.
Yesterday Joe inferred that 5 is odd from the fact that 2 is even.

These sentences might be thought of as expressing the following predications.

(6) True (odd (2))
Implies (even (2), odd (5))
Presently (stink (John))
Yesterday (infer (Joe, odd (5), even (2))

Thus, there are many sentences which are "naturally" thought of as expressing nested predication. Moreover, if temporal adverbs are thought of as predicates which apply to sentences then an even larger set of sentences can be subsumed under this heading.

4. REPORT AND MORPHOPHONEMICS

In this section we present an interpretation of what we take to be the central ideas of Harris' thought through the end of the 1960's. This interpretation is based on the two open lectures given by Harris at the University of Pennsylvania in the Spring of 1969. We have had the opportunity to read Harris' article (Harris, 1969) which was published shortly after the lectures and we will refer to it below. The importance of the lectures is underlined by the fact that they are the only open lectures given by Harris in more than four years. In many ways these new ideas are the most theoretical ever put forth by Harris and they represent a radical departure from his previous attitudes although, as will be seen below, they can be regarded as growing from a combination of older problems and older insights. The two radically new components are the following: first, the view that all statable information

("report") exists in the form of nested predication; second, that within a certain basic sublanguage the syntactical form of a sentence is determined by the semantic form of that which is reported by the sentence (Harris, 1969, p. 2).

Harris' 1969 view is essentially the following. By means of the previously studied linguistic methods it is possible to analyze or decompose each actual sentence in such a way that the ultimate resultant is a nested predication. Moreover, the original sentence will always turn out to be either a strict paraphrase of the ultimate resultant or else an elipsis of the resultant. In no case will any reference or information be found in the original sentence which is not explicitly part of the ultimate resultant. This analysis imposes on a language a base-plus-encoders structure whose base is a sublanguage consisting entirely of nested predications and whose "encoders" are unary functions which do not increase reference or information and which preserve relative acceptability.

The idealization represented in the above picture is very attractive vis-a-vis several problems which have already been mentioned. In the first place validation of transformations by means of acceptability ordering can be saved because all transformations are unary and paraphrastic. In the second place the base of "source sentences" has a "single string structure" (Harris, 1969, p. 60), viz. that of nested predication. In the third place, the problem of accounting for the interpretation (meaning) of the sentences of a language is reduced to the problem of accounting for the meanings of the nested predications and Harris seems to see no serious problems here (1969, pp. 3, 75-78). In the fourth place the base is not some abstract *ad hoc* device but rather it is seen as an objectively important sublanguage of the language under analysis. In fact the base is semantically equivalent to the whole language.

Unfortunately, perhaps, the above ideal falls short of reality. Harris does not conceal the deficit except insofar as he continues to use terminology which is only appropriate if the ideal agrees with reality. He starts with a set S of sentences which he simultaneously (or alternately) thinks of as the language under investigation and as the "set of sentences given in transformational linguistics" (1969, p. 50). By working backward from S certain "starred" expressions occur and these get added to S to form S^+. Now working backward again form S^+ he finds a set I of "sources". This set I is what I have been calling his "base". Harris freely states that I contains starred elements but he still wants to call it a "sublanguage" (not of S but) of S^+ which he calls "an extended natural language" (1969, p. 55). In fact, in one place Harris states that I consists entirely of starred elements.

Thus, it seems that Harris has in fact adopted a base-plus-encoders view wherein the base consists entirely of abstract elements (non-sentences)

which somehow represent (or are) the information expressible in natural language. The fact that the base is not literally a sublanguage has some important implications. In the first place, it constitutes a clear case of postulation of abstract entities for theoretical purposes. In Harris' former views linguistic objects tended to be derived more or less directly from concrete linguistic experience (narrowly conceived). Here we have at least a considerable broadening of what is taken as data.

In the second place, since the transformations are defined on the base either we must abandon acceptability ordering as a criterion of validity of transformations or else we must postulate an acceptability ordering on the abstract entities. If the latter, one might be inclined to order the abstract base elements in accord with the ordering of their images under the transformations – but this would be circular. In the third place, identification of the "information content" of a language with the abstract base brings Harris a step closer to the views advocated by Wittgenstein and Sapir earlier and by Keenan more recently. This step in turn would seem to entail both a more overt consideration of meaning in the framework of linguistics and an examination of the relationship between twentieth century notions of logical form, on one hand, and Harris' "information bearing" base elements on the other. Since the latter seems inevitable, Harris seems to be moving closer to an appreciation of the semantic analysis of language initiated by logicians. Of course, this is not new in linguistics as Lakoff, Thomason, Ross and McCawley have been exploiting logic for some time now. But this use of logic is of an entirely different character than the use of logic in linguistic investigations at the University of Pennsylvania (except for Keenan, of course). The move toward logic may be foreshadowed in Harris' recent abandonment of his older view of quantifier structure (1969, p. 37).

In the fourth place, the relegation of transformations to morphophonemic encoding allows for a complete reexamination of Harris' previous views on transformation structure. It is readily admitted that this structure had become overly intricate and convoluted and that a fresh approach is needed if only to prevent the whole framework from collapsing under sheer weight of technical intricacies. For example, the present framework does not involve binary transformations at all and, therefore, does not require binaries to preserve "meaning" and acceptability ordering. Thus one is not forced to be concerned with pseudo-issues, for example, whether the acceptability of *'Joe died because he ate plastic'* is somehow the same as a kind of joint acceptability of *'Joe died'* and *'Joe ate plastic'*.

In the fifth place, as mentioned above, Harris' "report and morphophonemics" view not only lends itself to speculation but actually seems to suggest it. For example, given this framework one can easily imagine that a speaker formulates a thought as an abstract base element and "feeds" the thought

into a transducer (battery of transformations) which converts it into a sound pattern. One can further imagine that linguistic change takes place in two dimensions – as a result of scientific, cultural and institutional evolution the structure of codable thought (the base) itself undergoes change *and,* on a different level and for different reasons, the sound pattern of a language (morphophonemics) undergoes change. Finally, the new openness to overt consideration of meaning encourages investigation of some of the important issues raised by Harris' early, bold move to treat extended discourses as genuine linguistic elements (in opposition to Bloomfield's "definition" of the sentence as the highest level element). The report and morphophonemics framework requires a theory of sentential meanings to complete the report (or base) component. Once this is partly developed we will be able to investigate more adequately the problem of determining the differences between the kinds of communication possible only in extended discourse and the kinds possible by means of single sentences.

It is perhaps worthwhile to end this paper with an observation already made elsewhere by Munz and others, viz., that in a sense Harris' "report and morphophonemics" returns him to his 1952 situation wherein concern with extended discourse had developed a constellation of ideas and problems concerned solely with the sentential level. Discourse analysis *per se* is hardly mentioned in the 1969 paper which concerns almost entirely the structure of the sentential and sub-sentential levels. In fact, at places the 1969 paper seems almost "inconsistent" with the 1968 book which views a language as a set of extended discourses.

Acknowledgments

This paper originated as a lecture in the Graduate School of Arts and Science of the University of Pennsylvania sponsored by the Linguistics Department, May 1969. It followed Harris' two lectures "Report and Morphophonemics" by a month and it was designed to trace the growth of Harris' ideas on structure from 1951 to the present. After "dividing through" by friendship I was still gratified by the response of Harris' colleagues and students. The present revision is very close to the original lecture except that criticism received after the lecture has largely been incorporated. Despite all of this I have no illusions about the accuracy of the paper, especially on small details, but I do believe that the main ideas are very close to Harris' and that the development is largely historical. Shortcomings, of course, are mine but credit belongs to the following linguists: Jean-Pierre Paillet, James Munz, William Frank, Senta Plötz, Richard Kittredge, Edward Keenan, Aravind Joshi, Henry Hiz and Henry Hoenigswald. Except for the encouragements of James Munz, William Frank and Senta Plötz I would not have thought the lecture worthy of publication.

State University of New York at Buffalo

References

I Harris' Writings Used in This Study:

1940 Review of Gray's *Foundations of Language, Language* 16 (1940), pp. 216-230.

1951 *Methods of Structural Linguistics,* University of Chicago Press, Chicago, 1951. Appeared in 1960 under the title *Structural Linguistics,* University of Chicago Press, Chicago.

1952 a "Culture and Style in Extended Discourse" in *Indian Tribes of Aboriginal America,* S. Tax (ed.), University of Chicago Press, Chicago, 1952, pp. 210-215.

1952 b "Discourse Analysis", *Language* 28 (1952), reprinted in Fodor & Katz (see below), pp. 355-383.

1952 c "Discourse Analysis: a Sample Text", *Language* 28 (1952), pp. 474-494.

1954 a "Transfer Grammar", *International Journal of American Linguistics* 20 (1954), pp. 259-270.

1954 b "Distributional Structure", *Word* 10 (1954), pp. 146-162, reprinted in Fodor & Katz (see below), pp. 33-49.

1955 "From Phoneme to Morpheme", *Language* 31 (1955), pp. 190-222.

1957 "Co-occurrence and Transformation in Linguistic Structure", *Language* 33 (1957), pp. 283-340, reprinted in Fodor & Katz (see below), pp. 155-210.

1959 "Linguistic Transformations for Information Retrieval", *Proceedings of the International Conference on Scientific Information,* 2 (1959), pp. 123-136.

1964 "Transformations in Linguistic Structure", *Proceedings of the American Philosophical Society,* 108 (1964), pp. 418-422.

1965 "Transformational Theory", *Language* 41 (1965), pp. 363-401.

1968 *Mathematical Structures of Language,* Interscience Publishers, New York, 1968.

1969 "The Two Systems of Grammar: Report and Paraphrase", *Transformations and Discourse Analysis Papers,* No. 79, available on request from Linguistics Department, University of Pennsylvania, Philadelphia, Pennsylvania, 19104, U.S.A.

BIBLIOGRAPHIE

II Other Sources Used in This Study:

Fodor, Jerry, and Jerrold Katz, The Structure of Language, Prentice-Hall, Inc. Englewood Cliffs, New Jersey, 1964.

Fries, Charles, The Structure of English, Harcourt, Brace & World, Inc., New York, 1952.

Hockett, Charles, The State of the Art, Mouton, The Hague, 1970.
Keenan, Edward, A Logical Base for a Transformational Grammar of English, Dissertation in Linguistics, University of Pennsylvania, 1969, available on request from Linguistics Department, University of Pennsylvania, Philadelphia, Pennsylvania, 19104, U.S.A.
Rogers, Hartley, Theory of Recursive Functions and Effective Computability, McGraw-Hill, New York, 1967.
Sapir, Edward, Language, Harcourt, Brace & World, New York, 1949, originally published 1921.

STRUCTURAL LINGUISTICS AND THE NOTION OF TRANSFORMATION

Jean-Pierre Paillet

Zusammenfassung

Im Gegensatz zur strukturalistischen Syntax (wie z. B. IC-Analyse), die im wesentlichen auf die Beschreibung der syntagmatischen Struktur von Satzformen beschränkt war, war die transformationelle Syntax von Anfang an so formuliert, daß die Transformationen die Selektionsstruktur von Satzformen unverändert ließen. Der erste Versuch einer formalen Definition von Transformation wurde innerhalb eines rein mengentheoretischen Rahmens gemacht, wodurch aber den Eigenschaften der Selektionsrestriktionen (im besonderen ihrer "Elastizität") nicht Rechnung getragen wurde. Erst durch die Einführung der Akzeptabilitätsordnung wurden Transformationen ins rechte Licht gesetzt. Dadurch wurde in der Folge eine "Operatorensyntax" ermöglicht, wo den Resultaten von verschiedenen Arten von Operatoren verschiedene Akzeptabilität zugeordnet wird. Das Beschreibungssystem, das man so erhält, hat zwei Vorteile: einmal bietet es interessante Analogien mit der Behandlung von "lower-level units" der Sprache (oder: von Spracheinheiten auf niedrigerer Ebene als Satzebene) und zweitens eignet es sich besser als die meisten anderen Modelle für eine Untersuchung der Prozesse, wie sie sich in Sprechakten finden.

Abstract

Whereas structural syntax (e.g. immediate constituent analysis) was essentially limited to the description of the syntagmatic structure of sentence-forms, transformations appear from the start as invariants of selectional structure. The first attempt at a formal definition of transformation was made within a purely set-theoretic framework, which did not do justice to the properties of selectional restrictions (in particular, to their "elastic" nature). The introduction of acceptability orderings placed transformations in their true light. It subsequently allowed the development of an "operator-syntax" where the various aspects of acceptability are assigned to the effects of