ZELLIG SABBETTAI HARRIS

Zellig Harris died in his sleep on May 22, 1992. He served as president of the Linguistic Society of America in 1955, and was one of the half dozen or so American linguists whose work has had the greatest influence both in his own country and abroad. This obituary fills a gap that should have been filled much earlier, and it is an honor to have been invited to write it.

Harris was born on October 23, 1909, in what was to become the Soviet Union, but he was brought to Philadelphia when he was four. His entire academic career lay at the University of Pennsylvania, as student and professor, until he formally retired in 1979. For years, however, he commuted to Israel, where his wife Bruria Kaufman, a theoretical physicist and former assistant of Einstein, was a professor at the Weizmann Institute. I wish that I could refer to other memorials which would say something of the man himself. I met him once, and found him courteous and charming. Instant tributes in the *Linguist list*, by Bruce Nevin and Michael Kac, include some relevant anecdotes.[1] But it would be impertinent for an obituarist so little connected with him to attempt to say more.

A list of Harris's academic writings has been published by Konrad Koerner (1993); to complete it we should add, in particular, a single book on politics, published well after his death (Harris 1997). He was in the beginning a Semitist, and, as a later reviewer was to recall perhaps with some nostalgia, a good Semitist (Cantineau 1954:4). His first book (and doctoral thesis) was a grammar and glossary of Phoenician (1936), by all accounts excellent and still cited for points of detail. Before that he had already contributed much to the analysis of the then new material in Ugaritic. His next book (1939) was a study of the early history of the Canaanite branch of West Semitic, to which the Phoenician dialects, with Hebrew, Moabite, etc., belong. This too is an exemplary philological study, of a kind that he would not publish again. Harris was then thirty, and there seems little doubt that, had he continued in the Semitic field, he would have been a leader in it.

But by 1940 his attention had already turned to general linguistics, and in a critical review of Gray's *Foundations of Language* (1940) he set out many of the ideas that from then on were to guide his scholarly life. With other linguists of that period, he stresses the importance of synchronic structure and 'the structural method'. He talks of the need to 'organize data by their place in the structure' (218), and argues that this structure 'can be described only in terms of the formal, not semantic, differences of its units and their relations' (223). 'Particularly undesirable are psychological explanations' (225). 'The work of linguistics is reducible, in the last analysis, to establishing correlations' (228). In a review of Trubetzkoy's *Grundzüge der Phonologie* (1941), naturally more favorable, he objects nevertheless to the notion that the 'language structure' (translating *Sprachgebilde*) can be studied independently of the 'speech act' (*Sprechakt*); the former is 'merely the scientific arrangement' of the latter (345). He pleads again for the selection of distributional, this time in preference to phonetic, criteria.

From then on, the development of his thought may helpfully be divided into three, in part overlapping, phases. The first begins in 1942 and may be seen to end in 1955, with articles on the morpheme. This is the period, above all, of *Methods in Structural*

[1] May 29 and May 31, 1992; also, in print, Nevin 1992. I am very grateful to Alan S. Kaye and Anna Morpurgo Davies for sending me copies; also to her and to Henry M. Hoenigswald for help generally with this obituary.

https://doi.org/10.1017/S0097850799130200 Published online by Cambridge University Press

Linguistics (1951a), which should be read with, in particular, his article 'Distributional structure' (1954a), put first in the most accessible collection of Harris's papers (1981). The central aim throughout this phase was to establish the basic units of a language on the evidence of distributional patterns. Phonemes are identified to account for regularities over sounds that can be distinguished in speech; morphemes to account for patterning over longer stretches; syntactic units to account for patterns over morphemes. The resulting 'description of the language structure' (Harris 1951a:372) inexorably took the form of what became known, later in the 1950s, as a GENERATIVE GRAMMAR. It was a description, still abstracted from consideration of meaning, which identified the units of a language and the distributional relations that could hold between or within them. It could therefore be viewed, as Harris put it in a paper towards the end of this period, as 'a set of instructions which generates the sentences of a language' (1954b: 260). This is, of course, equivalent to the justly famous formulation, again in the 1950s, of his then protégé Chomsky.

A second phase begins in the early 1950s, with the analysis of 'equivalences' in distribution within texts. For example, the sequences *there is often snow* and *it is very cold* would have equivalent distributions in a text in which both precede *in January*. But analyses of individual texts will often reach an impasse. In one, we might have the sequence *the weather is cold at the beginning of the year*. Then, in the same text, we might have a further sequence *the cold weather after Christmas*. On that evidence alone, no further progress can be made. But we can establish more generally that, in the language as a whole, any sequence of the form *the weather is cold* (schematically, *the N is A*) is related, by what Harris called, for the first time, a TRANSFORMATION, to a sequence *the A N*. Given that, we can then speak of a distributional equivalence, within the text itself, between *at the beginning of the year*, which follows the sequence *the weather is cold*, and *after Christmas*, which follows a 'transform' of it. In reading Harris's work on transformations, it is important to appreciate that this is how the notion originated. A transformation was not a rule of grammar that operated on an underlying level of representation. That was to be Chomsky's account, already in the mid 1950s, not his. For Harris it was primarily a relation that could be established between one form of attested sequence, in which specific words, like *weather* and *cold*, are found to 'co-occur' more readily or less readily, and another form of sequence, in which 'co-occurrences' are ranked similarly.

This second phase lasts from Harris's first paper on discourse analysis (1952) until the end of the late 1960s, and is marked by two major papers on transformational analysis (1957, 1965). The first coincided, in print, with the publication of Chomsky's first book (1957); and, at that time, the similarity of their general ideas, especially concerning the relation between form and meaning, with Chomsky's obvious and acknowledged debt to Harris as the older scholar, tended to obscure the ways in which they had already diverged. Both, among other things, identified a set of 'kernel' structures to which all others could be transformationally related: thus, in Harris's terms, a non-kernel *the A N V* (for example, *the cold weather kills*) to kernel *the N is A* combined with kernel *the N V*. But by the time of the second, which coincided with another major book by Chomsky (1965), it was clear that, while Harris was following the program that he had initiated in the 1950s, Chomsky had for some years taken, and was leading an ever larger army of followers down, a very different path. Ten years later, in the flowering of his last phase, I remember having great difficulty in explaining to younger colleagues, who had not followed Harris's work from the 1950s, what he was actually saying. In their demonology he was the man who had tried, and failed, to develop what

https://doi.org/10.1017/S0097850799130200 Published online by Cambridge University Press

Chomsky (not he) called 'discovery procedures'. Plainly it did not fit at all with that. But one had to free one's mind of every current prejudice to understand the thinking with which it did fit.

Harris's second paper on transformations (1965) was followed by a monograph, *Mathematical Structures of Language* (1968), and together they took analysis of this kind as far, with hindsight at least, as it would go. The paper in 1965 was also, sadly, his last contribution to *Language*. I have not thought it my business to inquire into the circumstances; but from then on, for whatever reason, a journal for which he had written so much for a quarter of a century, and so much of such influence and importance, published him no more. It is also remarkable how many of his later writings were to appear in Europe, with the Éditions du Seuil, in the *Journal of Linguistics*, or with the Clarendon Press, and not, as hitherto almost exclusively, in America.

The basic ideas of Harris's last phase became clear in the 1970s, in papers in English (1976a, 1978) and more fully in translation into French (1976b). In the collection published at the beginning of the next decade (1981), papers from our earlier periods are grouped in sections headed 'structural analysis' and 'transformational analysis'; those of this period under 'operator grammar'. But we are concerned in reality with two complementary insights. The first is that of the progressive simplification, or REDUCTION, of sentence structures. Thus, to take an example whose point it is sadly easy to misunderstand, *The paper was written by Mary* stands in a crucial relation not to the former kernel structure *Mary wrote the paper*, but to an unreduced *The paper was in the state of the writing of the paper by Mary*. The second idea is indeed that, in the description of unreduced forms, a straightforward operator-argument structure will suffice.

The first idea can best be appreciated if we go back to the original notion of transformations. We have seen that, for example, a sequence *the A N V* (such as *The cold weather kills*) can be related by transformation to a pair of sequences *the N V* (*The weather kills*), *the N is A* (*The weather is cold*). The evidence is simply that the cooccurrences of individual words have similar levels of acceptability in all three structures. For example, if *The leafy weather explodes* is less immediately acceptable, so too are *The weather explodes* and *The weather is leafy*. On similar grounds, a passive N_2 *is Ved by* N_1 is a transform of an active N_1 *Ves* N_2, an interrogative *is N Ving*? a transform of a declarative *N is Ving*, and so on. But if the formal relations are similar their semantic correlates are not. There is a sense in which the meanings that we can associate with the kernel structures *the N is A* and *the N V* are preserved, under transformation, in the combination *the A N V*. But the semantic relation between an active and a passive is more problematic, and interrogatives and declaratives contrast. Moreover, there are many formal relations, called in the 1950s 'quasi-transformations', in which patterns of cooccurrence are to varying degrees less regular. For example, the two nouns in *the weather forecast* (*the* N_2 N_1) recur in *The forecast is about the weather* (*the* N_1 *is about the* N_2); but for pairs like *argument* and *children* ratings of acceptability may not be equal (*The argument is about the children*, *the children argument*). Yet many quasi-transformations, like this, were semantically straightforward.

This was clear already in the first of Harris's major papers on transformations (1957), and the problems were not resolved at that stage. But let us compare, for example, *The cold weather kills* with, among its many transforms, *The weather which is cold kills*. The first has a simpler structure, which is accordingly more usual. The structure of the second is more complex and, in consequence, less usual. But, as potential utterances, they are both means by which speakers can say what is, in essentials, the same thing.

https://doi.org/10.1017/S0097850799130200 Published online by Cambridge University Press

Similarly, *The weather forecast is at noon* is a simpler and more usual way of saying what could also be said, in principle, in the more cumbersome form *The forecast which is about the weather is at noon*. Among the mass of transformations, in the sense of Harris's second phase, there are many that have these characteristics. On the one hand, transform *a* is unidirectionally less complex than transform *b*. On the other hand, in saying *a* one is still saying what, in principle, one can also say by *b*.

Hence, directly, the first basic idea of Harris's new phase. If *a* is, at all points where they are different, simpler than *b* we can describe *a* as a simplification or reduction of *b*. Thus, by reduction, $b \rightarrow a$. If what is said by *a* might also be said, in principle, by *b* then, to put it differently, both *a* and *b* give the same information. So, for example, under the reduction of *The forecast which is about the weather is at noon* to *The weather forecast is at noon*, INFORMATION, as Harris uses this term from now onwards, is preserved. It does not matter that there are other transformations, by the criteria of the 1950s and 1960s, which are not reductions, or in which information is lost or altered. It does not matter if, by the same criteria, the relation is not regular. What matters is that simpler forms of words can be described, in general, as reduced alternatives to more complex forms of words.

What then of Harris's new treatment of the passive? It is evident that no speaker would ever say *The paper was in the state of the writing of the paper by Mary*. They would not merely prefer, in practice, to say *The paper was written by Mary*; Harris's unreduced form is not even feasible. But for Harris, if I understand correctly, this was simply an extreme case. In general, as we have said, reduced forms are more usual. Provided information is preserved, and provided no other factor intervenes, speakers will naturally prefer a simpler form. Hence, for example, one would usually say *The cold weather kills*, not *The weather which is cold kills*. It is not clear, indeed, why one should ever say *The forecast which is about the weather is at noon*, instead of *The weather forecast is at noon*. But preferences are matters of degree. At one extreme, an unreduced form might be used quite readily. For example, one might say (reduced) *The paper boy is late*, or alternatively (unreduced) *The boy who brings the paper is late*. At the other extreme, we can envisage unreduced forms, such as *The paper was in the state of the writing of the paper by Mary*, which are not potential utterances at all.

This then was Harris's first central idea: that the forms of words which speakers use are ordinarily reduced from other forms of words that, to varying degrees, they do not use. I stress FORMS OF WORDS; we are again not talking about underlying representations. The second central idea concerns the syntax of, in particular, the unreduced forms.

Let us take, for the sake of clarity, a form that is actually reduced: *Young women helped John*. In Harris's earliest work on syntax (1951a; also 1946), this would have had the structure $N^4\ V^4$: a highest-level constituent of the class *N* (*young women*) followed by a highest-level constituent of the class *V* (*helped John*). The immediate inspiration was in Bloomfield's *Language*; but, by implication, elements were subordinate to a head in phrases of both categories. In his second period, he had continued to address the general problem of subordination, through a method of STRING ANALYSIS (1962). But in the 1970s, just as, by another stroke of timing, Chomsky and his followers were developing an X-BAR system that reflected Harris's account in the 1940s, he himself came up with a far simpler answer. Unfortunately, its very simplicity can make it hard to grasp.

Let us start, however, with *young women*. In any account, the presence of an adjective implies or requires that of a noun. But we do not need the traditional parts of speech:

https://doi.org/10.1017/S0097850799130200 Published online by Cambridge University Press

let us therefore say, more generally, that a word of a class of which *young* is a member requires, at some point elsewhere in the sentence, a word of a class of which *women* is a member. As Harris put it, *women* is thus 'prior to' *young*. If x is prior to y, we say, alternatively, that x 'is an argument of' y. So, in *young women*, *women* is an argument of *young*. For the converse relation ('is later than') we can say, alternatively, that y 'is an operator on' x. So, in *young women*, *young* is an operator on *women*. These words can then be assigned to classes as follows. *Young* is an operator, in notation *O*. *Women*, by contrast, is not an operator: it does not in turn require an argument. We can say, alternatively, that it is a 'null' operator: in notation *N*. *Young* therefore is, more precisely, an operator whose argument is an *N*: in notation, O_n.

What holds for *young* and *women* in *young women* also holds for the same words in, for example, *The women were young*. Once again, a word of a class of which *young* is a member requires, at some point in the sentence, a word of a class of which *women* is a member. The rest is, once again, a matter of definition and notation. So, *young* in predicative position, as the construction is traditionally described, is likewise an operator (O) whose single argument may not in turn be an operator (O_n). We can then connect this, terminologically at least, with what everyone else will say about the construction of verbs. In, for example, *The women appeared* the verb *appeared* ('verb', again, in the traditional classification) is, more generally, another word of a class that requires a single argument. This argument may again be an *N*: once more, O_n. In *The women helped John*, the transitive *helped* (again, traditionally 'transitive') is an operator that requires two arguments. Both are again *Ns*: so, in the same notation, O_{nn}. In, for example, *The women knew that John came* the first argument of *knew* is again an *N*; the second, however, is in turn an *O*. So, taking the two in sequence, O_{no}. In, say, *Their helping him surprised John*, the operator *surprised* is similarly an O_{on}; and so on.

This is, precisely, all the system involves. In later work, the relation of operator to argument is described as one of DEPENDENCE. Thus, in *young women*, *young* depends on *women*; and, in *Women helped John*, *helped* depends, as in the system of the Modistae in the thirteenth century, on both *women* and *John*. But other linguists now use that term differently. In Harris's system, all we are concerned with is a partial ordering of elements: *women* is 'prior to' *young*, *women* and *young* are 'prior to' *helped*. Elements are assigned to classes by the numbers and the kinds of element that are prior to them: no element (N), one element (O_n; also O_o), two elements (O_{nn}, O_{no}, O_{on}, O_{oo}) and, to the extent that it may be necessary, so on. If we were to take the set of sentences in general this might not be sufficient. But, for the unreduced set, Harris saw it as entirely sufficient.

These two ideas, of reduction and of priority among elements, were to distinguish Harris's work throughout his last phase. Its centerpiece is a remarkably detailed grammar of English 'on mathematical principles' (1982), described by an obituarist who knew him very well as Harris's 'capolavoro' (Hiz 1994:526). Finally, in his last years and long into retirement, he published two books which together bring his life's work to a real conclusion. The first (1988) is short and based on a series of Bampton lectures at Columbia University. In reviewing it (Matthews 1988) I remarked that, 'as one reads it, everything that matters in his intellectual career falls into place'. The second (1991) is full and comprehensive. I cannot but feel that he departed in peace, with gratitude that he had lived to complete it.

From Harris's first book to his last is a period of fifty-five years, in which so much has changed. The work of his first phase as a general linguist is now relegated to the

https://doi.org/10.1017/S0097850799130200 Published online by Cambridge University Press

history of our subject, and is sometimes cruelly traduced. That of his second phase was eclipsed, after 1960, by Chomsky's. That of his last phase has, for twenty years, been widely ignored. It is therefore hard to assess his contribution as a whole. But a whole it is. I cited at the beginning his reviews of Gray and Trubetzkoy. Ten years later he published a long appreciation of Sapir's selected writings (Harris 1951b), at once a noble return for Sapir's interest in his own work in the 1930s, and a lucid guide to what he saw as shared aims. Anyone who wants to understand Harris has to read it. He was then in his early forties, and until his late seventies, when he was invited to give the Bampton lectures at Columbia, almost all his writing was severely technical. Those lectures too one must read. They have an elegance that owes nothing to what Harris had, in gently chiding Sapir, called 'the siren of literary effect' (1951b:321, n. 43). In content as in style, they transcend the work in the 1940s and early 1950s for which he is best known. But the underlying problem has not changed.

At heart it is the difficulty of having to deal with language through language. As Harris put it in his last phase, 'natural language has no external metalanguage' (1988: 3). We can write a grammar of English in, for example, French; but not in any system that does not already have 'the same essential structure'. This was a difficulty that all structural linguists faced, but Harris saw it more clear-sightedly perhaps than any other. Hence an insistence on what he had earlier called 'the fact of patterning' (Harris 1951b: 297). If speech were random there would be nothing for us to investigate. But it is not random; its structure lies precisely in its departure from randomness, and it is by searching for patterns of nonrandomness that we can escape from what would otherwise be circularity.

Suppose, for example, that there were no words or morphemes. If so, we would expect the patterning of smaller units to be freer than it is. We would not expect that, in *I went out*, no other consonant could replace the [t] of *went* and still yield an acceptable utterance. We would not expect the [w] of *went* to be replaceable, under the same condition, by so few: by [s] (*sent*) or [l] (*lent*), but not by [k] or [p] and many others. We would not expect, in general, that in scanning such an utterance either forwards or backwards we should find at some points that what comes next is predictable and at other points that it is not. Thus, given *I wen* . . . the probability that *t* will follow is high; but given *I went* . . . that of *ou* is much less. Given . . . *out* the probability that *t* precedes it is quite low; that of *w* preceding . . . *ent out* is again much higher. Such patterning establishes that the units linguists call 'words' do exist. It also suggests a method by which they can be delimited. Let us scan our data, forwards and backwards, as in this example. Within such units we expect what comes next to be relatively predictable. We posit boundaries where it is not.

This method was tested, originally for morphemes, in the paper that marks the end of Harris's first phase (Harris 1955), and from then on he continued to rely on it. But methods are no more than methods; and, in particular, one method does not rule another out. In the 1960s he was exploring both string analysis and transformational analysis. But it was 'necessary to understand that these [were] not competing theories, but rather complement[ed] each other' (Harris 1965:365). In a footnote he deplored the 'absolutist . . . temper' of those who would pit 'one linguistic tool against another'.[2] in later work the first of these 'tools' was to be set aside; the second crucially reshaped. But that does not imply that either was, in an absolute sense, wrong, any more than, say, in writing at a keyboard I imply that pens are wrong.

[2] I am grateful to Terry Moore for drawing my attention to this note.

https://doi.org/10.1017/S0097850799130200 Published online by Cambridge University Press

A tool will also, in the end, determine its own uses. In his first phase Harris sought, for example, to justify pitch phonemes. But he did not claim that every aspect of intonation could be reduced to that model. Some, in applying his method, could be; others could not. In that sense, 'the decision of what to include in the linguistic structure rests with the linguist' (Harris 1951b:303). We develop a method by which we can investigate a specific form of patterning. We see where that method will take us. But methods are once more complementary. Another method may extend the description of linguistic patterns further; another may not take it so far.

Our problem, then, is to develop an account of patterning. But nothing, finally, suggests that the patterns we are studying are anything other than patterns of meaning. In *Methods in Structural Linguistics* Harris laid down that the sole criterion for descriptive linguistics was that of distribution (1951a:5). This is the book of his that has been most read: often, as an added aid to misunderstanding, in an edition with the *Methods in* omitted. So, to put it crudely, Harris excluded meaning from linguistics.

But such a crude view is misleading, as his assessment of Sapir (1951b:30f.), written slightly later, makes clear. What we cannot do is base our account of language on a priori categories. We cannot take what we are seeking to discover as itself a criterion for its discovery. But we do expect that differences in meaning will be explicated by our analysis. Take, for example, *She made him a good husband* and *She made him a good wife* (Harris 1951a:271f.). We can determine, by the technique of substitution on which Harris's account of syntax was initially founded, that the words are related differently. Thus, in the second sentence, *him a good wife* can be replaced by *a good wife for him*, still retaining acceptability. But, in the first, *him a good husband* cannot similarly be replaced by *a good husband for him*. Such results are found to correlate with a semantic difference, whose validity is thereby confirmed. When Harris discussed this example he had devoted himself to structural linguistics for less than a decade. His methods, in syntax especially, were still primitive. In many other respects, a distributional analysis identified both morphemes, such as *of* or *to*, and relations, such as that of a V^4 to an N^4, for which clear semantic correlates were lacking. Hence, yet again, it was vital not to lapse into circularity. But the ultimate challenge, leading finally to his grammar of English in the early 1980s, was to develop forms of description by which meaning would be illuminated.

By the time his project was completed many linguists, not least in America, were quite out of tune with him. One cannot but admire the steadfastness with which he carried it through. [Peter Matthews, *St. John's College*.]

REFERENCES

CANTINEAU, JEAN. 1954. Review of methods in structural linguistics, by Zellig Harris [1951a]. Bulletin de la Société de Linguistique de Paris 50. 2.4–9.

CHOMSKY, NOAM. 1957. Syntactic structures. The Hague: Mouton.

——. 1965. Aspects of the theory of syntax. Cambridge, MA: MIT Press.

HARRIS, ZELLIG S. 1936. A grammar of the Phoenician language. New Haven: American Oriental Society.

——. 1939. Development of the Canaanite dialects: An investigation in linguistic history. New Haven: American Oriental Society.

——. 1940. Review of Gray, Foundations of language. Language 16.216–31.

——. 1941. Review of Trubetzkoy, Grundzüge der Phonologie. Language 17.345–49.

——. 1946. From morpheme to utterance. Language 22.161–83.

——. 1951a. Methods in structural linguistics. Chicago: University of Chicago Press.

——. 1951b. Review of Mandelbaum, ed., Selected writings of Edward Sapir in language, culture, and personality. Language 27.288–333.

https://doi.org/10.1017/S0097850799130200 Published online by Cambridge University Press

——. 1952. Discourse analysis. Language 28.1–30.
——. 1954a. Distributional structure. Word 10.146–62.
——. 1954b. Transfer grammar. International Journal of American Linguistics 20.259–70.
——. 1955. From phoneme to morpheme. Language 31.190–222.
——. 1957. Co-occurrence and transformation in linguistic structure. Language 33.283–340.
——. 1962. String analysis of sentence structure. The Hague: Mouton.
——. 1965. Transformational theory. Language 41.363–401.
——. 1968. Mathematical structures of language. New York: Interscience Publishers.
——. 1976a. A theory of language structure. American Philosophical Quarterly 13.237–55.
——. 1976b. Notes du cours de syntaxe. Trans. by Maurice Gross. Paris: Éditions du Seuil.
——. 1978. Grammar on mathematical principles. Journal of Linguistics 14.1–20.
——. 1981. Papers on syntax, ed. by Henry Hiz. Dordrecht: Reidel.
——. 1982. A grammar of English on mathematical principles. New York: John Wiley.
——. 1988. Language and information. New York: Columbia University Press.
——. 1991. A theory of language and information. Oxford: Clarendon Press.
——. 1997. The transformation of capitalist society. Lanham, MD: Rowman and Littlefield.
Hiz, Henry. 1994. Zellig S. Harris. Proceedings of the American Philosophical Society 138.519–27.
Koerner, Konrad. 1993. Zellig Sabbettai Harris: A comprehensive bibliography of his writings 1932–1991. Historiographia Linguistica 20.509–22.
Matthews, Peter. 1988. Review of Language and information, by Zellig Harris [1988]. Times Literary Supplement. December 23–29. 1430.
Nevin, Bruce. 1992. Zellig S. Harris: An appreciation. California Linguistic Notes 23: 2.60–64.

https://doi.org/10.1017/S0097850799130200 Published online by Cambridge University Press