

Linguistic Society of America

Review: [untitled]

Author(s): D. Terence Langendoen

Reviewed work(s):

A Theory of Language and Information: A Mathematical Approach by Zellig Harris

Source: *Language*, Vol. 70, No. 3 (Sep., 1994), pp. 585-588

Published by: Linguistic Society of America

Stable URL: <http://www.jstor.org/stable/416495>

Accessed: 12/05/2009 12:12

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=lsa>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit organization founded in 1995 to build trusted digital archives for scholarship. We work with the scholarly community to preserve their work and the materials they rely upon, and to build a common research platform that promotes the discovery and use of these resources. For more information about JSTOR, please contact support@jstor.org.



Linguistic Society of America is collaborating with JSTOR to digitize, preserve and extend access to *Language*.

<http://www.jstor.org>

WILLIAMS, EDWIN. 1981. On the notions 'lexically related' and 'head of a word'. *Linguistic Inquiry* 12.245–74.

Department of English
1215 Patterson Office Tower
University of Kentucky
Lexington, Kentucky 40506-0027
[engl01@ukcc.uky.edu]

[Received 28 February 1994.]

A theory of language and information: A mathematical approach. By ZELIG HARRIS. Clarendon Press, 1991. Pp. xii, 428. Cloth \$69.00.

Reviewed by D. TERENCE LANGENDOEN, *The University of Arizona*

Zellig Harris' linguistic career spanned seven decades, from the 1930s to the present one. For the last four, however, he worked in virtual isolation, interacting almost exclusively with a few of his colleagues and former students at the University of Pennsylvania. In this, as in his other writings of this period, he cites only his own and their work, and that of a few eminent mathematicians and philosophers. On the whole, the favor has been returned. There are almost no citations of H's work in the mainstream linguistics literature, apart from references to his contributions to the founding of generative grammar. Much of the material in this, his last, book will be familiar to those who are acquainted with his work since the mid-1950s, especially Harris 1968, 1982, 1988; but it will be almost totally unfamiliar to the vast majority of linguists who are not. Nothing now can be done to alter H's isolation from the field, except for the field to become more familiar with his work.

From nearly the beginning of his career, H was deeply concerned about linguistic methodology. He justifies this concern in the opening paragraph of the second chapter of the present work (30):

'Linguistics has not in general been one of the sciences in which the relevance and correctness of statements are determined by controlled methods of observation and argumentation. It is therefore desirable to consider what methods are relevant in linguistics, in the hope of establishing criteria for investigation and analysis. Choice of method is not less important than responsibility in data, and the choice should be determined not by personal preference or current custom but by the nature and the problems of the data.'

H states that his 'choice of method arose out of [his] choice of problem' (30), specifically the problem of accounting for linguistic form, and 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' (32). He contends that these departures are not only universal but necessary; they are a necessary consequence of the fact that human languages have no external metalanguage, and hence that 'no external metalanguage is available to the child learning its first language, nor to early man at the time language was coming to be' (32). In order to identify the entities of language, learners can use extralinguistic information, 'such as the occurrence of words in life circumstances that exhibit their meaning' (32), but, as H observes, these resources are not adequate to establish more than the identification of certain words, and are hopelessly inadequate for the identification of many words and of practically

all multiword units. However, given (radical) departures from equiprobability, learners can identify the words and sentences of a language even if they are not told that there are such things as words and sentences, much less what those words and sentences are.

H's theory of linguistic form is a theory of constraints on combination, which together specify 'the various contributions to non-equiprobability that together make up the total non-equiprobability of the system' (33). From this, H draws his major methodological conclusion, namely 'that the constraints ... 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' (33). He uses this redundancy criterion to choose among alternative descriptions, or grammars of a language (33):

'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.'

Such a description H calls a 'least description', or a 'least grammar' (4). He claims that 'no assumption is made that there exists a structure in language, and no appeal is made to any particular principle of structuralism. Such structure as is found comes out in the process of making a least description' (36).

H's application of his least-grammar method to the analysis of English appears as Harris 1982, to which numerous references are made in the present volume. The following illustrations of the method are typical (34–35):

'[I]f a given phoneme- or letter-sequence (say, *attempt*) has been stated to have a particular meaning and particular other words as high-likelihood ('selectional') arguments, one would avoid giving the information twice: once when the word appears in verb position (*They attempt to stop smoking*) and again when in a noun position (*Their attempt to stop smoking failed*); the preferred alternative would be to take the second form as being the same sentence 'nominalized' under a further operator ... [W]hen we find a word—with more or less the same phonemes and clearly related meanings—appearing in different grammatical statuses, we do not simply give it a multiple classification, but rather, as noted above, use one status as derivational source for all the others. Examples are the many words which appear in English as noun and verb (e.g. *to house* from *a house*, and *a day's take* from *take*); or as verb with and without object, or with different object structures (e.g. *read* from *read things*, and *expect John* from *expect John to come*)...

H contends that systematic application of his method leads to the discovery of four major families of intrasentential constraints (DEPENDENCIES, LIKELIHOODS, REDUCTIONS, and LINEARIZATIONS), and two intersentential ones (WORD REPETITION and STRUCTURE REPETITION). Dependencies assign partial orderings on the words or morphemes in a sentence, by which words or morphemes serve as 'operators' on other words, called their 'arguments' in that sentence. Likelihoods determine the likelihood, for each argument, of other words to appear as operators for it. Reductions reduce words with the highest likelihoods to affixes or zero. Linearizations appropriately linearize the operators and arguments. These four families of constraints 'create, out of the set of all possible permutations of words in a language, precisely the set of possible (grammatical or potentially grammatical) sentences of that language...' (6); they also 'express

and transmit information' (5). In addition, word repetition requires repetition of related words in conjoined sentences, and structure repetition requires repetition of operator-argument relations in related sentences.

Of the four sentence-level constraint families, reductions require the most discussion. H argues that these have significant explanatory power: 'a great many special and seemingly arbitrary or irregular word combinations are seen to result from a constellation of regular processes in a way that explains both their form and meaning' (9). In Ch. 1 H gives two sets of illustrations, drawn from Harris 1982. In the first, he relates the clausal-connective use of *only*, *except*, and *but* to their (apparent) use as prepositions. For example, both *Everyone came except John* and *Everyone came except not John* are derived from the common source *Everyone came, except that John didn't come*, through a series of reductions involving the zeroing of *that*, *not* (optionally), and the repeated verb.¹ In the other, he sketches an account of the relation of passive to active sentences in English. The former he derives by embedding a nominalization of the latter as an argument of an operator: 'Using arrows for derivation we have: *The book was in the state of the taking of the book by John* → *The book was in the state of taking by John* → *The book was taken by John*' (14).

All the forms that appear in the derivations of sentences must be considered sentences of the language, since the reductions, in particular, are merely constraints that legitimate the zeroing of highly redundant elements. Since the zeroing of those elements is not obligatory, the sources of the reduced forms must all be in the language. But, as the immediately foregoing example illustrates, some of these sources are very unlikely. H confronts this problem in the Preface, asserting (v-vi):

'In characterizing ("explaining") the structure of word combinations, a method is used whereby one combination is derived from another (its source). In some cases the derivation for a given form involves reconstructing a source which is not actually said in the language. These sources, however, are never abstract models. They are word combinations made in accordance with the actual grammar of the language, and from which the given form would be derived by the existing derivational relations of the language. That they are not actually said may be due to their complexity or length; in many cases it would be impossible to formulate a regular grammar of the language that would exclude them as ungrammatical. Therefore the unsaid sources can be counted upon as derivational sources for the given word combinations.'

H's reductions, and the sources on which they operate, will strike most contemporary practitioners of generative grammar as unprincipled and ad hoc. However, many of them mirror familiar operations and constructions in generative grammar, and within H's framework they are as principled and motivated as the corresponding operations and constructions are in generative grammar. In the case of H's account of the active-passive relation, the deletion of the object corresponds to NP-movement in the extended-standard-theory version

¹ H also claims that both *There were but two people left* and *There weren't but two people left* derive from the same source by processes which he illustrates on other examples (10-11), but I am unable to determine from his discussion what that source is and how the derivation actually proceeds.

of generative grammar. Generally, what contemporary generative grammar accomplishes by movement, H's approach accomplishes by deletion.

There is much of interest in this book in addition to the development of a particular theory of linguistic form, including the analysis of sublanguages and the relation of natural language to mathematics and to music. However, the concluding paragraph of section 10.7, 'Non-linguistic systems; music', says something so outrageous that I am compelled to quote it in its entirety (318):

'Finally, it seems that the sign language of the deaf does not have an explicit operator-argument partial ordering, nor an internal metalanguage, but rests upon a direct juxtaposition of the relevant referents. This applies to autonomous sign languages, developed by the deaf without instruction from people who know spoken language.'

Lest there be any doubt about the implications of this paragraph, by 'internal metalanguage' Harris means the sentences which constitute the grammar of the language (359).

REFERENCES

- HARRIS, ZELIG. 1968. *Mathematical structures of language*. New York: Wiley-Interscience.
 —. 1982. *A grammar of English on mathematical principles*. New York: Wiley-Interscience.
 —. 1988. *Language and information: The Bampton lectures in America*. New York: Columbia University Press.

Department of Linguistics
 Douglass Building 200 East
 University of Arizona
 Tucson, AZ 85721

[Received 30 March 1994.]

Linguistic individuals. By ALMERINDO OJEDA. (CSLI lecture notes, 31.) Stanford: Center for the Study of Language and Information, 1993. Pp. 205. Cloth \$45.00, paper \$17.95.

Reviewed by KATE KEARNS, *University of Canterbury*

Ojeda argues that the universe of discourse is a mereology structured by the relation of instantiation, this relation subsuming both the familiar relation of instantiation borne to a kind by its instances and the relation borne by quantities of a substance (such as water) to a larger quantity, possibly spatially scattered, of which they are parts—traditionally the 'part of' relation of a mereology of solids.

Ch. 2 ('A theory of linguistic individuals', 17–33) introduces the basic assumptions and definitions of the theory. Chs. 3–5 ('The semantics of countability I', 35–68; 'The semantics of countability II', 69–104; 'The semantics of uncountability', 105–47) explore the denotations of noun forms in detail, defined in terms of atomic, atomistic, and atomless individuals of the mereology.

Atomic individuals are the 'true' individuals, having no instances other than themselves. Atomistic individuals are atomic or polyatomic, and atomless individuals are quantities of substances (or sums thereof) having no principle of individuation; they are therefore conceived of as infinitely divisible, and they