Zellig S. Harris

Lecture Notes

I Overview

0. From pre-transformational linguistics.

We begin first with the observation that all discourses (and sentences) are sequences of parts (words) which occur in various combinations, an observation common to all grammars. And second, with the remark that not all word combinations are in the set of discourses, a remark which underlies structural grammars because it presents the data of the science as occurrence in the set of discourses (rather than correctness, meaning, etc.). Third, we note that sentences (or discourses) have different likelihoods of occurrence, i.e. that membership in hte set of sentences is graded, an observation that is external to structural grammar but underlies transformational grammar. It will be seen in 3 that this determines for each operator a graded selection among the members of its argument domain. Indeed, for a given argument domain, the distinguishing property of one operator as against another (aside from their different phonemic compositions) is their different selections among members of their argument domain.

1. Argument-requirement.

To the above we add the result that sentences can be said to contain other sentences, i.e. that in an S_i we may be able to identify an S_j together with some additional material X: where X is a change of shape, or some non-sentential word or word-sequence, or some additional sentence, all located within S_i in respect to the S_j in it. This is the basic statement of transformational grammar. It is not an immediate observation but a product of analysis, because it requires evidence that the S_j-portion within S_i is indeed the same sentence as the independent S_j (and not just the same words as S_j but brought together by other factors). The evidence is supplied by a survey of sets of discourses: One way is to take for each X a set of S_i and see if the inequalities of likelihood among them are identical with (or related in an apriori stated way to) the inequalities of likelihood among the corresponding S_j. This is the transformational approach. Another way is to take a set of discourses containing S_i and a set containing S_j, and see if the neighbors of S_i are similar in an a priori stated way to the neighbors of S_j. This is the quasi-empirical string approach.

When these tests are positive we find that there are certain word-sequences A..DE..G such that

$\exists$ A...DE..G

$\exists$ A...D

$\nexists$ E..G

where $\exists$ (or $\nexists$) means occurs (or does not occur) in the set of sentences. In such cases we will say that E..G is an operator (a later entry into the discourse) which requires A...D as its argument, the A...D in A...DE..G being the same as the independent sentence A...D. Each word is assigned to a set on the basis of what it requires as argument. Words whose argument-requirement is zero are elementary arguments. The others are operators, members of various sets.

2. Operator-position: unrestricted source.

Finally, we bring in the observation that given a finite but not small vocabulary, virtually any information can be expressed in a natural language. This is made possible by the use in an arbitrary sentence S_1 of word-sequences in the position of some single words W in it. The word-sequences are mainly built by adding, to a word W_1 that grammatically belongs in the given position of S_1, various modifiers (adjectives, adverbs, etc.) or conjoints (and..., or...); they thus arise by zeroing and permutation from sentences which have been adjoined by bi-sentential operators (and, which, if, etc.) to S_1.

Now as is noted in 3 and elsewhere, the selectional effect of a bi-sentential operator is in general to favor similarities between its two arguments. Thus for given second-arguments which have become modifiers and conjoints of W_1 (and fill out its position in S_1) there is a special selection in the rest of S_1. That is to say, W_1 together with its second-argument attachments has the property of a new word in the W_1 position, with its own selection within the rest of S_1. Any desired selection in S_1 can be obtained by a suitable choice of W_1 and of second-arguments attached to W_1.

In particular, for any word W_i which has a given argument-requirement A_i, and has a given selection a_i among the members of A_i, it is possible to find a word W_j whose argument-requirement is also A_i with such modifiers and conjoints X (or equivalently such adjoined other sentences) such that the selection in A_i which is due to W_j together with X approximates a_i as closely as we wish. More precisely, the inequalities of likelihood among the various members of A_i would be approximately the same for W_j plus X as for W_i. This satisfies the criterion for W_j plus X being a transform (or synonym, or transformational paraphrase, or-in certain situations-definiens) of W_i.

Certain properties of this synonymy-transform, W_j plus X, should be noted. One: If W_j has the same argument-requirement as W_i, then X need merely affect the selection; for grammatically the X does no more than modify the W_j, so that the argument-requirement of W_j plus X remains that of W_j alone, which is that of W_i, i.e. both W_j plus X and W_i will be talking about the same thing. Two: If there are more than one positions or sets of operators requiring A_i, W_j will have to be in one that has a large vocabulary (i.e. in a large set), for with a small vocabulary there may be no adequate starting-point for the selection-adjustments affected by X. Third: W_j must be in the least restricted of all the sets which require A_i as argument. For if W_j is in a less restricted set than W_i, the additional restriction on W_i

can be stated in a metalinguistic sentence conjoined to the W_j: e.g. If in John can swim has the subject John applied to both the operator can and its argument swim, the paraphrase John's swimming is an ability of the subject has is an ability in a less restricted operator-position (which doesn't require that the subject of the argument swim be also the subject of the operator), but the metalinguistic modifier of the subject states that restriction, so that the selection due to is an ability of the subject should be much the same as the selection due to can. In contrast, if W_j is in a more restricted set of operators on A_i than is W_i, no modification can relieve it of that extra relation to its argument, and it will not be able to match the selection due to W_i: e.g. the range of occurrence of is an ability (not specifically of the subject) cannot be equalled by can no matter how modified.

We have thus found that all operators, of no matter what set or position, which have a given argument-requirement, have a synonymy-transform in a least restricted set of these operators. That is to say, for every argument-requirement there is at least one set or position which is grammatically less restricted than any other, and which contains synonymy-transforms of every operator having that argument-requirement. We can call this distinguished set or position the descriptive source of all the other operators, since we can "derive" all others from this source, but not derive all members of this source from the possibly fewer and more restricted members of the other operator sets for the given argument-requirement. A more important way of looking at this result is that we can define a set of processes or form-options sufficient to change the members of the distinguished set into the corresponding (i.e. selectionally-equivalent) members of the other operator sets (while some members exist only in the distinguished set). These processes are the paraphrastic transformations, and are found in 4 to be not arbitrary, but of a consistent and understandable character.

It may be mentioned that this single distinguished set may not be establishable if, for a given argument-requirement, the selections of one set (in one position), W_i, is sufficiently disjoint from those of another, W_j, as to preclude any adjoined modifiers from being able to adjust the selections of W_j words to being the same as those of W_i words. This situation is rare in natural language, but may well be the case in specialized languages--particularly of the sciences--where it may produce an informational structure different from that of natural language.

This result, that for each argument-requirement there exists a single "source" operator set, is the key step beyond the transformational step stated in 1; and it underlies the entry (operator) -and-reduction grammar presented here. It shows among other things that, surprisingly enough, the complex restrictions that are so characteristic of grammar are not needed for the objective information carried by language.

3. Entry structure.

All discourses are formed by operator words appearing (as later entries) in the distinguished position of 2 in respect to the arguments they require. A subset of discourses are formed by this alone; others undergo the additional reduction-process of 4. In many languages the distinguished position for an operator is the same for all arguments: in English the operator appears after the first of its arguments (which is in general its subject).

Operators on elementary arguments form elementary discourses which contain no discourse as a proper part. Operators on operators are operators on discourses: they thus contain, or extend, these other discourses, and constitute a set of transformations on them.

There is, in the reductionless discourses, little or no further restriction on the operators. But there are inequalities of likelihood of occurrence for each individual operator on the members of its argument domain. For most

In cases where the likelihood property is graded over a set of words, holding strongly for some members and less so for others, the reduction is available for a subdomain which includes the members which have the property strongly, but is cut off rather arbitrarily at some point where the property begins to be held weakly. In this way a graded property is replaced by a sharp division into subsets (having, or not having, the reduction).

The physical content of the reductions is of few types, many of them common to very many languages, and appropriate in a general way to the properties they express. The major types are zero, or short constant phonemic shapes (pro-words), for words having maximum likelihood of occurring at the given point of the given discourse (i.e. for words which are certain to be occurring there), and attachment of special forms of exceptionally common operators to their arguments. There are short forms for common, especially for metalinguistic, operators; and a forward moving of

The reductions constitute a set of partial transformations on the set of discourses. Under them are preserved the argument-requirements of each word in the discourse, the order of operator entry into the discourse, and the inequalities of likelihood of the discourses. The reductions do however bring additional grammatical structure into the language, chiefly the choice of operator sets which receive the special forms, in the interrelating of particular operators by giving them similar forms, and in the creation of special subdomains of arguments to which alone certain special forms are applied.

5. Some properties of the analysis.

This analysis shows that discourses are formed directly from words. Affixes are for the most part reductions of operator words, and sub-sentence constructions are reductions of bi-sentential operators with their second sentences. In the case of languages many of whose words are composed wholly operator sets, each operator has a certain imprecise and changing subset of its arguments on which it has what may be called normal likelihood of occurrence: this subset is called its selection within its argument domain. Certain bi-sentential operators, i.e. those whose argument domain is a pair of operators, favor, for normal likelihood, certain relations (of similarity, etc.) between their two arguments. The inequalities in a set of discourses are preserved (as contributions to the inequalities of the new discourses) when a further operator joins them to form a new set of discourses.

In many languages there is an additional though limited process: Certain operators and their arguments have different (phonemic) shapes when they meet than they do in other combinations. This is called (required) morphophonemics.

4. Optional reductions (variants).

When an operator joins its arguments, either the former or the latter have, in certain situations, a choice among more than one phonemic shape (or, more rarely, position relative to each other). This applies only to the immediate arguments, and not to any arguments of the arguments and so on down. Since in some situations only one shape and postion are available, which we will call the basic form, we can say that in the cases where there is a choice it is a choice between the basic form and a secondary one, i.e. it is the option of changing from the basic form to a secondary one. The discourses in which each entering word remains in its basic form, i.e. in which no choice to another form has been made, are the pure entry ones. In general, the variants consist in giving a reduced physical shape to operators (or arguments) which have a reduced informational contribution in the given discourse position. The situations in question are for the most part special and extreme likelihood properties of the operator or argument either in general or in respect to their particular arguments. Some of the types of special likelihood which occasion the reductions are common to many languages; others are peculiar to one language.

of bound morphemes (i.e. not of free words plus affixes), it is possible to build the syntax as here from the smallest free words, but with an added morphological system which constructs these free words out of bound morphemes in a regular way.

The pure entry discourses from a simply-structured virtually restrictionless sublanguage which is closed under its operations, and which contains all the information given in the language, since the further processings are reductions of shape for reduction of information. The reductions can be considered as deriving the remaining discourses of the language out of this sublanguage. But since there is no order of derivation aside from the order of entry and from the conditions of the reduction, one can look upon these shape differences as merely alternative forms of the sublanguage itself.

The whole theory is in terms of the entries. And the fact that the reductions affect only the immediate joining words (both as bearers and as conditions) simplifies greatly the theory of the language, the formulation of restrictions, and the method of sentence analysis.

Each discourse, including each sentence (which is a transform of a discourse-segment), is built out of ordered entries with possibly shape-variants on the entering words. This gives the effect of two levels or factors in the construction of a sentence. The fact that the reductions of words take place only upon the entry of those words makes possible a directly calculable decomposition of each sentence into these two factors: into the ordered entries (possibly partially ordered for the <u>which</u> operator), and the partially ordered (mostly simultaneous) reductions (if any) at each entry.