

XXX

## THE TWO SYSTEMS OF GRAMMAR: REPORT AND PARAPHRASE

### 1. INTRODUCTION

#### 1.0 Summary

- 1.1. An Extension of Transformational Analysis
- 1.2. An Extension of Free Variation in Structural Linguistics

### 2. UNRESTRICTED FREE VARIANTS FOR RESTRICTED OPERATORS

- 2.1. *Be-ing*
- 2.2. *Very*
- 2.3. Adverbs of Manner
- 2.4. Subjunctive
- 2.5. Time-Conjunctions
- 2.6. Comparatives
- 2.7. Special Comparatives
- 2.8. *Wh-*
- 2.9. Vocabulary

### 3. UNRESTRICTED MORPHOPHONEMIC SOURCE FOR RESTRICTED OPERATORS

- 3.1. Plural
- 3.2. Tense

### 4. RESTRICTION DETERMINED BY ENVIRONMENT

- 4.1. Subjunctive
- 4.2. Adverbs of Manner
- 4.3. Reciprocals
- 4.4. Sentence Nominalization: *Sn*

### 5. PREDICATIONAL SOURCE FOR INCREMENTS

- 5.0. Introduction
- 5.1. Noun Adjuncts
- 5.2. Verb Adjuncts
- 5.3. Compound Tense; Auxiliaries
- 5.4. Verb-Operators
- 5.5. Subordinate Conjunctions
- 5.6. Moods
- 5.7. *And, or*

### 6. RESULT

- 6.1. Summary of Reductions
- 6.2. An Unrestricted Subset of Sentences
- 6.3. A Sublanguage for Objective Information

### 7. THE PREDICATE SYSTEM

- 7.0. The Metalanguage
- 7.1. The Base
- 7.2. The Operators
- 7.3. The Values
- 7.4. Morphemic Shape
- 7.5. String Shape
- 7.6. Relation to Linguistic Analysis

### 8. THE MORPHOPHONEMIC SYSTEM

- 8.0. Introduction
- 8.1. The Morphemic and String Changes
- 8.2. The Interpretation of T
- 8.3. Creation of Subsets
- 8.4. String Change
- 8.5. String Properties under T
- 8.6. Invariants
- 8.7. Structure of {S-I}
- 8.8. Derivation

### 9. BEYOND THE PRESENT ANALYSIS

#### 1. INTRODUCTION

##### 1.0. Summary

From an attempt to isolate the independent elements of sentence construction, we arrive at two different and separately acting grammatical systems, which contribute to this construction: a system of predicates (with *and, or*) and a system which can be considered an extension of morphophonemics (which is change purely of phonemic sequences). The predicate system carries all the objective information in the sentence, and the most natural interpretation of its structure is that of giving a report. The morphophonemic system is interpretable as being paraphrastic, and changes at most the speaker's or hearer's relation to the report. The grammar of the language as a whole is simply the resultant of these two systems.

The method used to obtain this result arises from the characteristic problem of structural linguistics: that not all combinations of elements, but only certain ones, occur as acceptable sentences. This limitation is clearly related to the information-bearing power of language. We see this relation in the fact that elements whose environments are complementary to each other (e.g. *am* and *are* in *I am going, You are going*) cannot be the bearers of informational differences (unless their different neighbors are then zeroed); and elements whose environments differ in particular ways have corresponding differences in meaning.

If element A requires element B in its environment, or if all of the environments of A have certain similarities and differences with respect to all of the environments of B, then we say that A is (syntactically) dependent on B.

The question now arises whether we can isolate all these dependences (i.e. restrictions on element-combination) into particular paraphrastic transforms of the restricted sentences, leaving other transforms of the same sentences without the restrictions. This is achieved by showing that one can obtain, within transformational methods, unrestricted free or automatic source-sentences for all sentences which contain restricted word-combinations (§§ 2, 3), and by showing that forms which are restricted to particular subsets of words (having particular properties) can be derived without resort to this dependence (§ 4).

The set of restrictionless transforms that is obtained contains paraphrases of all sentences of the language. It is a sublanguage, closed under the non-paraphrastic transformations (6.2, 3), and in it the choices and the relative positions of words are not affected by any dependences aside from a few very general formation-rules (7.2, 7.5). Within these basic structures of sentences, and within the limits of the vocabulary, the word-combinations are determined solely by the combinations of information (more precisely, the report) that is being expressed. This sublanguage can be put into an operator form far simpler than is possible for the language as a whole (§ 5). The operators in the construction of a sentence are apparently linearly ordered. They consist of predicates (verb; or *be* plus adjective or noun or preposition) or *and*, *or*; each resultant of an operator is formed by placing the operator word after its first argument, thus forming the word sequence that is the sentence.

The sublanguage carries all the objective information, or report, which is carried in the language, and can be used without the rest of the language. Indeed, every sentence of the language can be decomposed (by an algorithm) into a sentence of the report-sublanguage (containing all that is reported in the original sentence) plus various paraphrastic transformations. But if the sentence contains pronouns or other referentials (including covert ones) whose antecedent can be in some other sentence of the discourse, this antecedent cannot in general be determined by a sentence-decomposing algorithm. A discourse in the language can be reduced to one in the sublanguage by reducing each sentence to its source in this sublanguage. Many sentences, more precisely word-sequences that can be read as a sentence, are grammatically ambiguous, i.e. have more than one source from which they could be transformationally derived. This is due to transformational degeneracies: different transformations which yield the same word-sequence, or zeroings and pronounings of different material. In many cases, a comparison of the alternative sources of a sentence with those of neighboring sentences in the discourse shows which of its sources best fits which source of the neighbors. The report which is carried in the original discourse is preserved in the reduced discourse, aside from ambiguities; but now various processings of

the information in the reduced discourse can be performed in orderly ways that were unavailable for the original discourse.

In the sublanguage the connection between the simple syntactic relations in it and the meaning of words and classes becomes much sharper than in the language as a whole. Investigation of what kind of information can be borne in language, and how language can be modified to carry other kinds of information, is made possible by analysis of the structure of this report-sublanguage, in which every item of structure is relevant to the informational burden that language carries. The few types of predicates (7.2) which impose and preserve partial orderings in their operands,<sup>1</sup> together with *and*, *or*, are found to be sufficient and necessary for the information that language can carry. And the simple method of placing the fixed-argument operator-word (the predicate or *and* or *or*) between the argument words (7.5) is found sufficient to enable a parentheses-less string of words (i.e. a sentence) to indicate the operator-structured information that language carries.

The whole of the rest of the language can be derived out of this sublanguage, as source, by a system of paraphrastic transformations (morphophonemics of syntactic sequences), which themselves have a particular grammatical structure, and a subjective (or purely paraphrastic) semantic character. They bring in no new independent words (or new report) into the sentence, but are especially useful for abbreviation. It is these paraphrastic transformations that bring in grammatical ambiguity and complex (and simultaneously differing) string relations (affixes and positionings) among the words of the sentence.

One can, of course, make a description of the whole language, but it is now possible to see that there is no coherent structure in the grammar of a whole language. The whole grammar is a resultant of two quite different structures, the second operating on the prior one. Each system is coherent, as being derivable from a tightly interrelated set of primitives, in a way that the whole grammar is not; and each has a natural interpretation. Together they account for the forms and the interpretation (meaning) of the whole language.

### 1.1. *An Extension of Transformational Analysis*

In transformational analysis, we define over the set of sentences of a language some of them wholly acceptable and some less so, an equivalence relation 'transformation' (written  $\leftrightarrow$ ) between two subsets of sentences, each being of a particular sentence-form.<sup>2</sup> This relation preserves the acceptability-ordering of the corresponding sentences of the two forms. It has been found that there are two types of transformations: One, the paraphrastic (T), holds between two sentence-forms, A, B, each having the independent variables  $x_1, x_2, \dots, x_n$ , when each ordered set of n-tuples of values of these variables

determines the same acceptability-ordering for the sentences it produces out of A as for those it produces out of B.

The other, the incremental, holds between two sentence-sets,  $B_m^x$  of  $m$  variables, and  $C$  containing these  $m$  variables and at least one additional variable called  $C$ 's increment, when, given an ordered set of  $m$ -tuples of values of the common  $m$  variables, the acceptability-ordering which this set of  $m$ -tuples determines for the sentences of  $B$  is related in certain readily summarizable ways to the acceptability-ordering it determines for the sentences of  $C$ . The chief relation is the following: Let  $A$  be a sentence-set of the variables  $x_1, \dots, x_{m-1}$  with an acceptability-ordering for the  $m-1$ -tuples of values,  $B$  a sentence-set of the same variables with an incremental  $x_m$ , and  $C$  a sentence-set of the same  $m$  variables with an incremental  $x_{m+1}$ . Then for each value of  $x_{m+1}$  in  $C$ , the acceptability-ordering for the  $m-1$ -tuples of values of  $x_1, \dots, x_{m-1}$  in respect to each value of  $x_m$  is the same as it is in  $B$ .<sup>2a</sup> Thus the ordering of the sentences of  $C$  consists, in this case, of an ordering of  $x_m$  for each value of  $x_{m+1}$  and an ordering of the  $m-1$ -tuples of values for each value of  $x_m$ , the latter ordering being carried over from  $B$  and there from  $A$ .

Thus we have a paraphrastic transformation between  $NV$  *Ally* and  $N$ 's *Ving is A*, as in:

The ball rolls slowly, and: The ball's rolling is slow  
(both normal word-choices, even though the second is less comfortable),

The rock rolls slowly, and: The rock's rolling is slow  
(both normal word-choices),

The photon rolls slowly, and: The photon's rolling is slow  
(both peculiar or nonsensical).

In contrast, there is no transformation between  $N_1VN_2$  and  $N_2VN_1$ :

A truck carried a boat, and: A boat carried a truck  
(both normal), but

Man invented ether, and: Ether invented man  
(the first normal, the second nonsensical).

And we have an incremental transformation in the increment *slowly* (or *is slow*), as is seen in the fact that, with *roll* as value of  $x_m$  we have the same acceptability-ordering for *The ball rolls slowly*, *The rock rolls slowly*, *The photon rolls slowly* as for *The ball rolls*, *The rock rolls* (both normal) and *The photon rolls* (peculiar or nonsensical). In contrast, *The man slept* is normal, and *The microbe slept* is peculiar, but both *The man slept slowly*, and *The microbe slept slowly* are peculiar. Thus *slowly* (as  $x_{m+1}$ ) preserves the acceptability of sentences with *roll* (as  $x_m$ ) but reduces it for sentences with *sleep* (as  $x_m$ ). Increments impose an ordering upon their immediate argument

(which is the  $x_m$  in this case), and the acceptability-ordering of their resultant sentences ( $C$ , here) is a product of the ordering of  $x_m$  under the increment, and the ordering of the  $m-1$ -tuples for each  $x_m$ .

The increment in  $C$  may therefore be considered an operator (note 1) operating on the  $x_1, \dots, x_m$  sentence-form and altering in a particular way the acceptability produced there by the  $m$ -tuples of values.

In some cases of  $A \leftrightarrow B$ , one or more of the variables in  $B$  has a domain  $e$  which is only a proper subset of its domain  $d$  in  $A$ : we write  $A(d)$  for  $A$  over domain  $d$ , and  $B(e \subset d)$ . In such cases, the transformation holds only for the smaller domain  $e$  (of  $B$ ); or else we can define a directed transformational operator  $A \rightarrow B$ , more precisely  $A(d) \rightarrow B(e \subset d)$ , which derives  $B(e \subset d)$  out of  $A(d)$  as a transform of  $A$  over the smaller domain  $e$ .

This is not the only situation in which a directed transformation,  $\rightarrow$ , is defined. In some transformational work there is a different, though not conflicting, interest in defining an operator  $\rightarrow$ . Because there are degenerate transformations, so that there can be  $T_i: A \rightarrow B$  and  $T_j: C \rightarrow B$ , we may find that the domain of some variable  $x_i$  in the  $A$ -form is a proper subset of the domain of  $x_i$  in  $B$ . In these degenerate cases, the domain of  $x_i$  in  $B$  is completely covered by the domains of  $x_i$  in  $A$ ,  $C$ , and in any other proposed sources for the form  $B$ . The situation is different for the analysis proposed in this paper. Here if the domain of  $x_i$  in  $B$  is only a proper subset of the domain of  $x_i$  in  $A$ , then there are no transforms, different from  $B$ , which cover the residue of the domain of  $x_i$  in  $A$ . In such a case we must take  $A$  as the source, from which  $B$  is derived. For if we did not, then indeed  $A$  could be derived from  $B$  (instead of  $B$  from  $A$ ) for that part of the domain of  $x_i$  which is common to both  $A$  and  $B$ ; but the  $A$  form for the residue of the domain of  $x_i$  would still have to be formed in  $A$ , since by assumption it could not be derived from any other form over that residual domain.

The methods proposed in the present paper examine the domains of the transformations, in particular the incremental ones, and thereby determine a particular organization of the transformations. The chief method used here will show that, given an incremental  $A(d) \rightarrow B(e \subset d)$ , we can find some sentence-form  $B^*(d)$  (of the same independent variables as  $B(e \subset d)$ , but without the reduction of domain in any variable), such that  $A(d) \leftrightarrow B^*(d)$  is an unrestricted incremental transformation and  $B^*(d) \rightarrow B(e \subset d)$  is a restricted paraphrastic transformation.<sup>3</sup> Finding  $B^*(d)$  makes it possible to separate increments from domain-restriction, i.e. to remove domain-restriction from the increments.

## 1.2. An Extension of Free Variation in Structural Linguistics

From the point of view of linguistic structure the methods proposed here

investigate free variation among morphemic elements. Structural linguistics deals with the distribution of elements, i.e. the environmental restrictions on their combinability in making sentences of the language. The fundamental relations are: two elements A, B, may be complementary variants of each other (if all sentence-making environments of A differ in part from all those of B); or they may be free variants of each other (if B occurs in the same environments as A, and if sentences which differ only in having B instead of A are considered structurally equivalent); or else they are distinct ('contrastive'). In phonology, complementary variants are well known (e.g. English unaspirated *t* after *s* but aspirated *t* after word-initial), as are free variants (e.g. English released and unreleased *t* before word-end). In morphology, only complementary variants are well known. They occur with small phonemic difference (e.g. *knife-* before plural, but *knife* otherwise) and with large (e.g. *be* after *to* and the auxiliaries or before *-ing*, *am* after *I*, *are* after *you* and plural, *is* otherwise). Complementary variation is known not only for A, B at identical positions of their respective environments, but also in more complicated cases; e.g. the discontinuous morphemic entity, which consists of a plural suffix after the noun and a (perhaps different) plural suffix after the adjective is complementary to the non-discontinuous plural suffix after an adjective-less noun: the former occurs only when the environment contains noun plus adjective (*-s ... -s* in *Il y a deux erreurs graves*), and the latter only when the environment contains only a noun and no adjective (*-s* in *Il y a deux erreurs*).

Free variants are hardly considered in morphology. There are a few cases with small phonemic difference; e.g. *ekənamiks* and *iykənamiks* for *economics* pronounced with initial *e* or *i* *y*; or *I'll* and *I will*. And there may be a few with large phonemic difference: e.g. *etcetera* and *and so on*. But there is no recognition of any large system of free variation, or of free variation which involves various positions distributed throughout the sentence, and so on. One reason for this non-recognition was the lack of instances where such free variation was of use in the formulation of morphological structure. Another was the question of what criterion would determine that a given A and B were in free variation to each other, i.e. what would determine that sentences with A were to be considered equivalent to those having B instead of A.

The syntactic relevance of recognizing morphological free variants may be seen for example in the fact that this relation is necessary in order to obtain the analyses proposed in this paper; and the morpheme sequences which are found to be free variants of each other are not similar in phonemic composition, and are in some cases not similar in the sentential position they occupy, or even in their immediate syntactic statuses. As to the criterion for

determining that sentences containing A are equivalent to those containing B (as free variant) instead of A: it is that replacing A by B does not change the acceptability-ordering of the sentences. In other words, sentences with A are paraphrastic transforms of sentences having B instead of A.

In particular, there is one type of paraphrastic transformation which assures the possibility of the type of analysis proposed in this paper. This is due to the fact that not only does every language contain meta-linguistic sentences (including definitions and grammatical statements), but also every sentence  $S_i$  can have conjoined to it or operating upon it all the grammatical and definitional statements about it.  $S_i$  with these added meta-statements can then be considered as the source for  $S_i$ , which is derived from the source by zeroing the known metastatements.<sup>3a</sup> If a restricted or non-predicate apparent increment I has no free variant which is unrestricted and of predicate form, we can always operate on the I-bearing sentences with a metastatement M defining the I. M need not contain the I restriction, and can be in the form of a predicate on the I-bearing sentence. Under the M operator, the I would be a morphophonemic entity due to the M which defines it and states its presence in the operand of M. The I-bearing sentence is then a morphophonemic variant of the sentence having the M operator.

The inclusion of free variation as a relation utilized within morphology completes the utilization of the fundamental relations of structural linguistics. In the present paper this relation is used for the existing vocabulary of language. It can later be used (§ 9) in conditions that go beyond the vocabulary, somewhat as phonemic long components used complementary variants beyond the conditions of sequential phonemes.

## 2. UNRESTRICTED FREE VARIANTS FOR RESTRICTED OPERATORS

The whole work of distributional linguistics has been to replace restricted elements by less restricted ones. In transformational analysis, a restricted transformational operator is one whose domain is only a proper subset of the domain of its operand: e.g. the apparent transformation  $NV \rightarrow N$  *be Ving* does not apply to certain  $V$  ( $\nexists$  *I am knowing*). The major method of the present paper involves a reformulation of each restricted incremental operator, in such a way that the restriction to a subdomain applies only to a paraphrastic transformation which follows after the increment has operated. This is achieved mainly by finding an unrestricted free variant for the restricted increment (§ 2), secondarily by adding to the grammar unrestricted morphophonemic sources for the restricted forms (§ 3), and finally by finding for the grammatically restricted lists (sub-classes) of words certain envioning words whose presence can determine the subclass properties (§ 4).



2.1. *Be-ing*

The *be-ing* operator which yields, e.g. *I am writing poetry* from *I write poetry* has a restricted domain in that it does not operate on certain verbs:  $\nexists$  *I am knowing English*,  $\nexists$  *I am owning a car*. If we seek free variants for this operator on the verb, we find them in a few perhaps stilted operators on the sentence such as *be in process*, *be on*: *My writing poetry is in process*. Although these are in themselves far less comfortable than *be-ing*, the acceptability-ordering (and nuance) of sentences with *be in process* (or its equivalents) and of those with *be-ing* is approximately the same for the same sets of word-values.<sup>4</sup> The passage between these sentence-forms is accomplished by well-established transformations over these domains. First the argument-skipping transformation (8.1 (9)):

My writing poetry is in process  $\rightarrow$  I am in process of writing poetry.

as in:

My driving is slow  $\rightarrow$  I am slow in driving.

And then:

I am in process of writing poetry  $\rightarrow$  I am writing poetry.

by zeroing of *in process of* as a constant of this form, i.e. as the only segment which is zeroable to produce *N is Ving* with unchanged object of *Ving*. The derivational connection between *be in process* and *be-ing* is supported by the fact that *be in process* has the same domain restriction<sup>5</sup>: One would not say *My knowing English is in process*, *My owning a car is in process*. However, what makes this derivation of interest to our present purposes is the fact that whereas the domain restriction of *be-ing* is more or less an unbreakable grammatical rule, in the case of *be in process* it appears rather as a matter of selection (co-occurrence), which can be changed by adding operators that would affect the reasonableness of the selection, or by cultural changes in what it is acceptable to say. Thus we can extend the selectional domain by adding something which really treats the owning of a car as a process, e.g. *My owning a car is finally in process*; this may then extend with some difficulty to *I am finally owning a car*, although in the latter both the acceptability and the meaning are more uncertain. Also if one can some day trace the course of neurophysiological processing of knowing, which so far has not been seen as a process, we may be able to say *My knowing English is in process*. We thus have an increment *be in process*, which makes no restriction on its operand sentences, except for selection; and a paraphrase (zeroing to *be-ing*) which freezes this selection as a grammatical restriction. Extensions of the selection, as in *My owning a car is finally in process*, are not readily zeroed to *be-ing*.

A major respect in which the source proposed here is unrestricted is in the possibilities of further increments on this increment. *Be-ing* cannot be nominalized: From *I write* we have *My writing is regrettable*, but for (1) *I am writing*,  $\nexists$  *My being writing is regrettable*. In other forms, *be* is nominalizable by *-ing*: *My being here is regrettable*. This restriction does not apply to the source form, which can be nominalized under operators: From (2) *My writing is in process* we have (3) *My writing being in process is regrettable*. Here as in many other examples below the source form is uncomfortable, but becomes more comfortable when something is added, as in (4) *My writing being still only in process is regrettable*. The only source from which the expanded form (4) can be directly derived, by certain incremental operators, is (3); hence (3) must exist. The reason for (4) being more comfortable than (3) is that (3) is derived from (2) which is overshadowed by its short transform (1), whereas (4) is derived from (5) *My writing is still only in process*, which is not as fully overshadowed by its short transform (6) *I am still only writing*. And the reason for (5) being less overshadowed than (2) is presumably that the increments *still only* which operate on the 'continuative' operator of (1), (2) as their argument, are more explicitly related to their argument in its source form (*is in process*, in (5), where they appear as adjuncts on it) than in its short transform (*is ... ing*, in (6)).

2.2. *Very*

While most adverbs of degree (e.g. *considerably*) occur with (modify) all predicate forms, both verb and adjective, the adverb *very* does not occur with verbs. This makes it seem that there is here an intrinsically different modifier for adjectives. We can find, however, free variants of *very*, such as *to a great extent*, which occur freely with all predicate forms: *He favors is to a great extent*, *He is favorable to it to a great extent*, *He is cold to a great extent*. We can then say that there is an unrestricted operator on sentence (in effect, on predicates) *to a great extent*; and that over the adjective subset of predicates (e.g. *cold*, *favorable*) this has a free variant *very*, which provides the same acceptability-ordering as does *to a great extent*. This analysis simplifies the transformational connection in such sentence pairs as

The room darkened to a great extent

and

The room became very dark ( $\leftarrow$ The room became dark to a great extent).

If *very* were not a transform of an adverb which operates also on verbs, we would not be able to derive the sentences with deverbal adjectives (e.g. *He is very fussy*) from their verb sources (e.g. *He fusses to a great extent*).

2.3. *Adverbs of Manner*

About adverbs of manner two restrictions are known: In their predicate form they require the *of*-nominalization of their operand: *His driving of the car is quite slow*,  $\nexists$  *His driving the car is quite slow* (whereas for time-adverbs both nominalizations occur: *His driving of the car is quite frequent*, *His driving the car is quite frequent*). And only a particular subclass of morphemes occurs as manner-adverbs (as others are time-adverbs, etc.), in some cases depending on the verb. Thus *unexpectedly* is an adverb of fact: *He pronounced my name unexpectedly*  $\leftarrow$  *That he pronounced my name was unexpected*. *Continuously* is an adverb of time-distribution: *He rubbed it continuously*  $\leftarrow$  *His rubbing it continued*. *Unfairly* is of fact or of manner, depending on the verb: *He tipped them off unfairly*  $\leftarrow$  *That he tipped them off was unfair*; *He divided the gifts unfairly*  $\leftarrow$  *His dividing of the gifts was unfair (in manner)*. However, we can find for these adverbs certain free variants of them which are not restricted:

He pronounced my name in an unexpected manner.

He rubbed it in a continuous manner.

He tipped them off in an unfair manner.

In all these cases *in A manner*  $\leftrightarrow$  *Aly* (writing *A* for adjective). The restriction of particular *A*, to being manner-adverbs or other adverbs, is therefore not in the increment *in A manner*, which is available for all adjectives within selectional dependence upon the verb, but in the paraphrastic change *in ... manner*  $\rightarrow$  *-ly* which is used only for adjectives that are characteristically manner, or are such for the given verb.

Furthermore, although *of* is required as above when *manner* is absent, we find nominalization without *of* when *manner* is present: *His driving of the car is slow in manner*, *His driving the car is slow in manner*, *The manner of his driving of the car is slow*, *The manner of his driving the car is slow*;  $\nexists$  *His driving the car is slow*. Here we have a free variant of *A* of manner which is not restricted to *of*; the restriction to *of* occurs for the paraphrastic reduction of *in ... manner*  $\rightarrow$  *-ly*.

2.4. *Subjunctive*

In English the subjunctive occurs under certain operators: under operators on sentence, e.g. *I request that he go there* (as contrasted with *I deny that he goes there* (or *will go*)); and in the second operand of conjunctions on sentence-pairs, e.g. *I won't go for fear* (or: *lest*) *he go there* as contrasted with *I won't go because he goes there* (or *will go*). It thus appears as a grammatical property restricted to these and perhaps other situations. However, we can find free

variants in which these operators appear without the subjunctive: *I request his going there*, no different from *I deny his going there*; and *My not going is for fear of his going there*, no different from *My not going is because of his going there*. Hence the increments of the class of the sentence-operator *request* and of the conjunction *for fear* are not restricted to occurring on the subjunctive form of their operand sentence. The domain restriction comes in later, in the paraphrastic free variant of the type

I request his going  $\rightarrow$  I request that he go.

This transformation operates not on all *N's Ving* under sentence operators and conjunctions, but only for the ones under a subdomain of sentence-operators and conjunctions: specifically in English those sentence-operators and conjunctions in which the time-location adverb of the operand sentence is dependent upon the time-location adverb of the sentence-operator or the primary sentence preceding the conjunction.

For *deny*, the time-location adverb of the operand sentence is independent:

I denied yesterday that he went the day before yesterday.

I denied yesterday that he will go tomorrow.

I will deny tomorrow that he went yesterday.

I will deny tomorrow that he will go the day after tomorrow.

But for *request*, the time-location adverb of the operand sentence always indicates a later time than that of *request*; or, less specifically, the operand sentence can be viewed as always containing a (zeroable) time-order adverb *afterwards* or *subsequently*:

I requested the day before yesterday that he go yesterday (or: that he go afterwards).

I requested yesterday that he go tomorrow (or: that he go afterwards).

I will request tomorrow that he go the day after tomorrow (or: that he go afterwards).

$\nexists$  I requested yesterday that he go the day before yesterday.

$\nexists$  I will request tomorrow that he go yesterday.

Similarly, for *because* we have independent time-adverbs:

I went yesterday because he went the day before yesterday.

I went yesterday because he will go tomorrow.

I will go tomorrow because he went yesterday.

I will go tomorrow because he will go the day after tomorrow.

But for *for fear* or *lest* the second is dependent upon the first:

I went the day before yesterday lest he go yesterday (or: afterwards).

I went yesterday lest he go tomorrow (or: afterwards).

I will go tomorrow lest he go the day after tomorrow (or: afterwards).

∫ I went yesterday lest he go the day before yesterday.

∫ I will go tomorrow lest he go yesterday.

We see then that the paraphrastic free variant transformation from *N's Ving* to *that NV* (the subjunctive) occurs in operand sentences under those operators which impose a time-location dependence upon the operand (or which, equivalently, always permit the word *subsequently* (or: *afterwards*) in their operand). There are, of course, various details which are omitted here. There are some verbs, e.g. *request*, whose operand always contains *afterwards* and which never have a tense on their argument. There are some which take *afterwards* usually but perhaps not always, and which may therefore sometimes appear with a tensed argument: *I suggest that he go*, *I suggest that he possibly went by himself* (or: *I suggest that he might have gone by himself*). There are some verbs which have both the *afterwards* and the tensed arguments: *I insist that he go*, *I insist that he went*. In some cases the tensed argument may be of a variant of the untensed: *I prefer that he go*, *I prefer that he went now*. In other languages, the subjunctive is determined by morphemes expressing possibility, not only by morphemes (such as *afterwards*) expressing time-order.

### 2.5. Time-Conjunctions

If we consider the time-ordering conjunctions *before*, *after*, *until*, *since*, *while*, etc., we find complex restrictions as to the tenses of the verbs which they connect. It has been shown<sup>6</sup> that if certain compound tenses and verbs are characterized as being either perfective or else imperfective, these restrictions can be organized into a few readily stateable requirements, e.g. that the verb following *before* or *after* be perfective (while the verb preceding the conjunction can be either perfective or imperfective), and that either both verbs be past or else both verbs be future. (The present tense which is replaceable by a synonymous future tense is considered here as future.) Thus we have:

- (1) He returned (yesterday) before she arrived (yesterday).
- (2) He will return (tomorrow) before she arrives (tomorrow) (or: will arrive).

But

- (3) ∫ He will return (tomorrow) before she arrived (yesterday).
- (4) ∫ He returned (yesterday) before she will arrive (tomorrow).

It should be noted that while (3) could be said to be excluded as having an impossible meaning, (4) is perfectly reasonable but grammatically excluded.

As in the preceding cases, it turns out that the language contains free variants which are not restricted, and which are related to (1)–(4) by known paraphrastic transformations:

His return yesterday was before her arrival yesterday.

His return tomorrow will be before her arrival tomorrow.

(5) His return tomorrow will be before her arrival yesterday.

(6) His return yesterday was (or: will have been) before her arrival tomorrow.

The tenses in (1)–(4) correlate with the time-location adverbs of their verbs. But whereas *before* is restricted to certain tense-pairs, *be before* is not restricted by the time-location pairs whose tenses would have restricted *before*. The time-relation which was reasonable but grammatically excluded in (4) is acceptable for *be before* (6). And the time-relation which was unreasonable and grammatically excluded in (3) is not grammatically excluded though still unreasonable here (5). In (5), as in the case of *be-ing*, we have found for a form which is grammatically restricted a free variant which is only selectionally restricted; and, as for *be-ing*, the selectional restriction is changeable due to new sentence-environments or cultural contexts: (5) can indeed be contained in such a sentence as *Gödel's backward-running time could make one's return tomorrow be before one's arrival yesterday*. Thus *before* plus tense is a free variant of *be before* over part of the domain of *be before*: The subdomain is characterized by both of the time-location adverbs under *be before* being such as determine the past tense, or both determining the future tense, and by the second verb being perfective.

It further turns out that all the other ('time-ordering') conjunctions which have such tense and perfectivity restrictions are free variants (preserving acceptability-ordering) of other subdomains of *be before* plus certain additions. First, *be after* plus reflection of operands in the operator (the operand interchange, 8.1 (8)) is a free variant of *be before* over its whole domain:

Her arrival yesterday was after his return yesterday.

Her arrival tomorrow will be after his return yesterday.

Then, *after* plus tense plus reflecting the two operands in the operator is a free variant of *be before* over the following subdomain: the time-location adverbs are as for *before*, and the first verb (under *be before*) must be perfective. When both verbs are perfective we can form out of

His return was before her arrival

both

He returned before she arrived

and

She arrived after he returned.

When only the second is perfective, then from

His continuing to work was before her arrival

we can form

He continued to work before she arrived

but not

∇ She arrived after he continued to work.

And when only the first is perfective, then from

Her arrival was before his continuing to work

we can form

He continued to work after she arrived

but not

∇ She arrived before he continued to work.

The other time-ordering conjunctions can also be shown to be free variants, over other subdomains, of *be before* plus more complex additions.

The sentences containing *be before* have many other free variants. One of these makes a desirable source. It is the form in which *be before* is simply a verb between two time-nouns:

His return at 3 P.M. was before her arrival at 3:15 P.M.

His return at the hour of 3 P.M. was before her arrival at the hour of 3:15 P.M.

- (7) 3 P.M., which was the hour of his return, is before 3:15 P.M., which was the hour of her arrival.

Each of the first two sentences is derivable from the one below it by known paraphrastic transformations. In (7) we see finally that one of the free variants is different in an important respect: instead of *be before* as a connective verb (transformable into a conjunction) between two nominalized sentences containing time-location adverbs, we now have *be before* as a verb between time-nouns, and two *wh*-conjunctions connecting two sentences with time adverbs to these time-nouns.

The source of (7) is:

- (8) (The hour of) 3 P.M. is before (the hour of) 3:15 P.M.  
 (9) His return was at (the hour of) 3 P.M.

- (10) Her arrival was at (the hour of) 3:15 P.M.

with (9) and (10) being each joined to (8) by *wh*-.

In the case of time-connectives without time-location adverbs, e.g. in

His return was before her arrival

we have to assume that the source contained unspecified time-predicates:

- (8') A time is before a time.  
 (9') His return in the past was at a time.  
 (10') Her arrival in the past was at a time.

If we operate on the ordered (9'), (10') with two *wh*- operators (themselves derived from *and*) on the successive *time* of (8'), we obtain (with automatic *the*):

- (11) The time of his return in the past was before the time of her arrival in the past.

Zeroing the *time of* as a constant in respect to *be before*, we obtain:

- (12) His return in the past was before her arrival in the past,

whence by tense-transplacing and second-operand tensing (which make a conjunction out of the verb *be before* 5.5, 8.1 (10), (12)):

- (13) He returned before she arrived.

If the time and perfectivity conditions for the tense-transplacing are not satisfied, the conjunction form as in (13) is not reached. Furthermore, if the two source sentences under the *wh*- (i.e. (9') and (10') here) do not have past or future (from *is in past*, *is in future* predicates) the derivation cannot reach (13) and we are left with

His return is before her arrival.

In (8)–(10) and (8')–(10') we see how the whole system of inter-sentence time-connectives can turn out to be a free variant of one verb, *be before* (or *precede*) on time-nouns, plus the *wh*-connective.

## 2.6. Comparatives

A somewhat similar situation arises in the case of the comparative. The comparative has certain grammatical restrictions. It involves not only a conjunction between two sentences but also a comparative morpheme placed next to a particular word (the one in respect to which the comparison is being made) in the first sentence<sup>7</sup>:

- (1) More men read books than women (read) magazines.

- (2) Men read more books than women (read) magazines.

In the second operand sentence of each pair, there are unique zeroings (discussed below), which are seen in, e.g.:

- (3) More men read books than read magazines.  
 (4) Men read more books than (they read) magazines.  
 (5) Men read more books than are worthwhile.

but no zeroing in:

- (6) More men read books than there are books which are worthwhile.

After zeroing, the *than*... can be permuted, but not to before the comparison-bearing word in the first sentence:

- (7a) More men read books than women (do).  
 (7b) More men than women read books.  
 (7c) Men read more books than women (read).  
 $\nexists$  Men than women read more books.

Finally, the comparative is restricted in repetition. In the case of other conjunctions (except *wh-*), if a conjunction C has operated on a pair of sentences  $S_1, S_2$ , yielding a new sentence  $S_1CS_2$ , it can then operate on a further pair consisting of this resultant and a third sentence:  $(S_1CS_2)CS_3$  or  $S_3C(S_1CS_2)$ :

(He left because it was late) because he had to be back in time.

In contrast, the comparative repeats only on pairs of pairs, yielding  $(S_1CS_2)C(S_3CS_4)$ :

- (8)  $\nexists$  He is richer than she more than you are richer than me.

but  $\nexists (S_1CS_2)CS_3$ :

- (9)  $\nexists$  He is richer than she more than you are (rich).

There are many free variants to the comparative form. One in particular turns out to be free of the various restrictions of the comparative conjunction.

We consider, for

- (10) He is richer than she (is).

the free variant (via (10<sup>o</sup>) below):

- (10<sup>o</sup>) The amount of his riches is more than the amount of her riches.

or with morphophonemic *exceed* for *be more than*:

The amount of his riches exceeds the amount of her riches.

On the basis of (10<sup>o</sup>) we can form a free variant to (8):

- (11) The excess of (the amount of) his riches over hers exceeds the excess of (the amount of) your riches over mine.

And here we can also form  $(S_1CS_2)CS_3$ , which was impossible in (9):

- (12) The excess of (the amount of) his riches over hers exceeds (the amount of) your riches.

In the new form we have no comparative conjunction: As in (8)–(10) of 2.5, (10<sup>o</sup>) is formed out of a certain elementary sentence form (13) plus two sentences connected to (13) by the *wh-*connective:

- (13)  $N_i$  is more than  $N_j$  (by  $N_k$ ),

where  $N_i, N_j, N_k$  are any numbers; or they may be pronouns and classifiers of the numbers, such as the words *number, amount, degree*. To (13) are joined (by *wh-*):

- (14)  $N_i$  is the amount of his riches ← His riches amount to  $N_i$ .  
 (15)  $N_j$  is the amount of her riches ← Her riches amount to  $N_j$ .

The result of connecting (14) and (15) by *wh-* to (13) is:

$N_i$ , which is the amount of his riches is more than  $N_j$ , which is the amount of her riches.

As in (8<sup>o</sup>)–(10<sup>o</sup>) of 2.5, we can form this set also when the nouns of (13) are unspecified quantities, if we allow these further nouns to be subject and object of *is more than*:

- (13<sup>o</sup>) An amount is more than an amount (by an amount).  
 (14<sup>o</sup>) His riches have an amount.  
 (15<sup>o</sup>) Her riches have an amount.

When (14<sup>o</sup>), (15<sup>o</sup>) are each connected by *wh-* to (13<sup>o</sup>), and addressed respectively to the first two occurrences of *amount*, we obtain (comparably to (11)–(13) of 2.5):

- (10<sup>o</sup>) The amount of his riches is more than the amount of her riches.

and, by zeroing of *the amount of* as a constant of *is more than*:

- (10<sup>o</sup>) His riches are more than her riches (or: than hers).

and by tense-transplacing, second-operand tensing, and zeroing we obtain the comparative conjunction:

- (10) He is richer than she.

The source set of type (13)–(15), (13')–(15') explains the special zeroing of the comparative conjunction. In this source there is under the second *wh-* (i.e. in the (15)-type sentence) an established zeroing of repeated material in positions corresponding to their antecedents in the (14)-type sentence. Thus from

$N_i$ , which is the (number of) men who read books, is more than  
 $N_j$ , which is the (number of) men who read magazines

we obtain by zeroing of the indefinite  $N_i$ ,  $N_j$  and of parallel repetitions:

The (number of) men who read books, is more than read magazines.

Here operand interchange under *wh-* yields the somewhat dubious:

Men read books who are more than read magazines.

With zeroing of *who are* and permuting of *more* to the left of its host (8.4(6)):

(3) More men read books than read magazines.

The permutation of 8.4 (6) can operate not on *more* alone but also on *more than read magazines* (i.e. on *more* together with its adjunct), yielding:

More men than read magazines read books.

The fact that *than...* is never permuted to before *more* (see (7) above) is due to the fact that the permutation operates on *than...* only as right adjunct of *more*. It is *more* that is permutable, and *than...* only with it.

Similarly, from a source

$N_i$ , which is the (number of) books which men read, is more than  
 $N_j$ , which is the (number of) magazines which men read

we obtain (4) via

The books which men read are more than the magazines (which they read).

And from a source

$N_i$ , which is the (number of) books which are read, is more than  
 $N_j$ , which is the (number of) books that are worthwhile

we obtain (5) via

The books which men read are more than are worthwhile.

And from a source

$N_i$ , which is the (number of) men who read books, is more than  
 $N_j$ , which is the (number of) books that are worthwhile

we obtain (6) via

The men who read books are more than the books that are worthwhile.

The *wh-* on two operand sentences with a common N always brings the common N in its second operand to the start of the second operand, where it is pronounced onto the *wh-*. The common N in (13)–(15) is always a quantity (as the common N in (8)–(10) of 2.5 is always a time-noun), and the noun which bears the quantity is automatically brought up after it (by the string effect of *wh-*, 2.8, 8.4 (6)). It is this that puts the compared (quantified) word at the start and that determines what are the corresponding positions for zeroing. When the tense-transplacing and second-operand tensing make the verb *is more than* into a conjunction *than*, the operand sentences undergo a string change and the compared word may no longer be at the start.

Note that for (6) the second *books* is not zeroable under the source *wh-* because it is not in a position corresponding to that of the first *books*, whereas for (5) the second *books* was in a corresponding position and zeroable. The (13)–(15)-type of source shows that the comparative has not merely one compared word, the bearer of *more* in the first compared sentence (i.e. the word quantified by  $N_i$ ), but also a second compared word, in the second sentence (i.e. the word quantified by  $N_j$ ). The zeroing rules under the comparative conjunction, which seem complex and unique, turn out to be almost entirely the normal zeroings under *wh-* in the source form shown here.

Like all paraphrases in this paper, these transformations are not semantically determined equivalences but established transformations, preserving acceptability-ordering. And the intermediate sentences in the derivation, and the source sentences in (13)–(15), exist in the language, even if they are felt to be cumbersome. It should also be noted that the precise form of the elementary sentences, and precisely which paraphrastic transformations are to be used in the derivation, need not concern us in the present problem, though they would concern a detailed grammar of English. All that is important here is that the comparative conjunction, like the time-order conjunctions, can be paraphrastically derived from an elementary sentence form and the *wh-* connective, and that the source forms do not have the restrictions in domain or in further operability which the derived forms have. The special properties of the comparative are thus merely the result of a few established transformations operating on this unrestricted source.

The discussion above shows that the comparative conjunction has a free variant, in the (13)–(15) source form, which does not have the grammatical peculiarities of the comparative: it does not have a comparative marker which has to be placed in the first operand sentence; its zeroings and permu-

tations are mostly the normal ones for the source form; and it is not restricted as to repetition or other further transformability. Also, the source form is semantically more explicit, in that it naturally distinguishes a quantified (and compared) word in the second operand sentence as well as in the first.

As in other cases, we find that many sentences can be said in the source form which cannot be said (or can be said only by some special adjustment) in the more common derived form. These are the sentences which are excluded by the restrictions that are required for the paraphrastic derived form.

Thus, we have seen in (12) that we could form the source for the (SCS)CS comparison, where the short form (9) was excluded: (12) resulted from joining three sentences to the  $N_i$ ,  $N_j$ ,  $N_k$  of (13). If only one sentence, say (14), is adjoined by *wh-* to (13) we obtain (if  $N_j$  is, say, \$5.):

(16) The amount of his riches is more than \$5.

or:

His riches exceed \$5.

where, as in (9) the comparative-conjunction form does not exist:

(17)  $\bar{\exists}$  He is richer than \$5.

We find additional source sentences which lack a comparative-conjunction paraphrase if we note the zeroings involved in (1)–(7). In all these cases, the source proposed here has zeroable material: identical words in parallel positions. Thus the source of (1) would be:

(18) The number of men who read books is more than the number of women who read magazines.

And the source of (2) is:

(19) The number of books which men read is more than the number of magazines which women read.

In these and the other cases the quantified words have the same position in the source sentences, so that the two *wh-* clauses in (18), or in (19), have identical structures.<sup>8</sup> If, however, we take a sentence of this type in which the two *wh-* clauses have different structures, we find that in certain cases the direct comparative-conjunction form does not exist:

The number of men who read books is more than the number of magazines which women read.

$\bar{\exists}$  More men read books than magazines women read.<sup>9</sup>

Similarly, for

The number of books which men read is more than the number of women who read books.

$\bar{\exists}$  Men read more books than women who read books.

To summarize the situation for the comparative: We have found that there is an incremental system, not restricted to any subdomain, which consists of any sentence of the elementary form (13) to which are joined by *wh-* any sentences with quantity-predicates as in (14), (15). Under certain conditions, when there are certain parallelisms (which may be due to zeroings) in the sentences joined to (13), the section preceding *than* is transformed into a sentence, and *than* becomes a conjunction. It is this last transformation, a paraphrastic one, which is restricted; and certain source-sets of the (13)–(15)-type cannot receive this transformation.

The comparative with *less than* is readily obtained from *more than*. And the *as ... as* comparative can also be obtained from *more than*; or *is as (equals)* can be considered another member of the class of *is more than (exceeds)*, except that it involves no *by N<sub>k</sub>*. Hence by the side of (13) there is also

(20)  $N_i$  is as  $N_j$ ;  $N_i$  equals  $N_j$ .

### 2.7. Special Comparatives

Related to the comparative, there are several highly restricted and seemingly arbitrary sentence-forms. These too are found to be free variants of the comparative source (13) or (20) plus (14), (15) of 2.6, with simple additions.

One set is seen in:

(1) He is too ill for you to leave now.  
He is so ill that you should not leave now.

These are derivable by morphophonemic replacements (*too* and *so...that... should not* for *more...than is appropriate*) from

He is more ill than is appropriate for your leaving now (or: for you to leave now)

which is derivable by tense-transplacing and zeroing, as in 2.6, from

(2) The degree of his illness is more than the degree of his illness which is appropriate for your leaving now

which is derived from (13) of 2.6:

(3)  $N_i$  is more than  $N_j$  (by  $N_k$ ); or: A degree is more than a degree (by a degree).



joined by *wh-* with

- (4) He is ill to a degree ( $N_i$ ).  
 (5) His being ill to a degree ( $N_j$ ) is appropriate for your leaving now.

Another set is seen in:

- (1') It's solid enough to break the wall.  
 It is so solid that it can break the wall.

These are morphophonemically derivable (with *enough* and *so...that...should for as...as is appropriate*) from

It is as solid as necessary (or: appropriate) for its breaking the wall (or: for it to break the wall).

from (as above)

The amount of its solidity is as the amount of solidity of it necessary (or: appropriate) for (its) breaking the wall.

which is derived from (20) of 2.6:

$N_i$  is as  $N_j$  (or:  $N_i$  equals  $N_j$ ); or: An amount equals an amount.

joined by *wh-* with

$N_i$  is the amount of its solidity; It is solid to an amount ( $N_i$ ).  
 $N_j$  is the amount of its solidity necessary (or: appropriate) for it to break the wall; Its being solid to an amount ( $N_j$ ) is necessary for its breaking the wall.

Thus the forms with *too* and *enough* are a special case of the (13)–(15) system of 2.6. They are derived from it by regular zeroings and special morphophonemics from the two forms (13) and (20) respectively, when the two *wh-*sentences are:

$N_i$  is the amount of  $S_1n$ ; or:  $S_1n$  is an amount  $N_i$ .  
 $N_j$  is the amount of  $S_1n$  appropriate for  $S_2n^{10}$ ; or:  $S_1n$  being in an amount  $N_j$  is appropriate for  $S_2n$ .

This derivation explains why there are precisely two sets, (1) and (1'), of these sufficiency-comparisons: they arise from the two sources (13) and (20) of 2.6.

Another special comparative is found in the proverb form

- (6) The bigger they are the harder they fall.

This is derivable by somewhat unusual applications of zeroing from three occurrences (a, b, c) of the whole (13)–(15) system of 2.6:

- (a):  $N_i$  exceeds  $N_j$  by  $N_k$ .  
 $N_i$  is how big they are (or:  $N_i$  is the amount of their bigness).  
 $N_j$  is ... (here: some zeroable S containing a quantity).

which transforms into

- (7)  $N_k$  is the excess of how (much) bigger they are (than ...).  
 (b):  $M_i$  exceeds  $M_j$  by  $M_k$ .  
 $M_i$  is how hard they fall.  
 $M_j$  is ... (here: some zeroable S containing a quantity).

whence the transform:

- (8)  $M_k$  is the excess of how (much) harder they fall (than ...).  
 (c): (9)  $N_k$  is as  $M_k$ .

The remaining sentences of the (c) set are (7) and (8), the resultants of (a) and (b); then (c) transforms to

$N_k$  which is the excess of how (much) bigger they are is as  $M_k$  which is the excess of how (much) harder they fall.

Standard zeroings take this into:

The excess of how (much) bigger they are is the excess of how (much) harder they fall.

whence the unusual constant-zeroing of *excess of how (much)* and *is* produces the irregular form (6).

If instead of (9) we had had its paraphrastic transform

As  $N_k$  is so is  $M_k$

we would have obtained

As they are bigger so they fall harder.

This derivation shows why (6) is available only for *more* and *less* (note: *The less he knows the better*) but not for *as* ( $\nexists$  *The as much he knows the as good*): for the derivation uses essentially, in (7)–(9), the *by*  $N_k$ , *by*  $M_k$  which is available only with *is more than*, *is less than*. And it explains the permutation in *the bigger they are* as being due to the *how* in (a), (b).

## 2.8. *Wh-*

The main use of *wh-*, namely that in which the morpheme following the

*wh-* is a pronoun of a preceding (but possibly zeroed) N or PN, has been analyzed as a free variant which morphophonemically replaces *and*, plus a sentence identifying the pronoun with its located antecedent, by *wh-*. Thus:

He found the book which had disappeared.

can be derived by zeroing and the above morphophonemics from such sources as:

He found a book and a book had disappeared and 'book' of preceding sentence indicates the same individual as 'book' of sentence preceding that.

Thus, *wh-* occurs only between two sentences in which the same noun appears with certain restrictions.

The main interest of such derivations for our present purpose is that, differently from *wh-*, *and* is not restricted as to the two sentences which it joins together, and even the added sentence about individual sameness is not grammatically restricted. E.g. the sequence which does not satisfy this restriction: *He found a card and he wrote a letter: and 'house' of preceding sentence is the same individual as 'book' of sentence preceding that* is nonsensical rather than ungrammatical. Thus the increment *and* is unrestricted; and the increment which consists in adding the sentence about sameness is restricted only in meaningfulness of selection. In contrast, *wh-*, which is a free variant of these together, is grammatically restricted: it cannot be said in a grammatical sentence which does not derive from occurrences of a common noun in two operand sentences.

### 2.9. Vocabulary

A more complicated problem is that of the restricted vocabulary. There are many words whose environment is limited in ways that are characteristic of grammatical restriction rather than of meaningful selection. It can be shown that unrestricted free variants are available for these words or their sequences.

A simple example is the case of restricted single words which have unrestricted synonyms. Thus the noun *flock* occurs primarily in *of birds* (not all birds, at that), *of sheep*, *of parishioners*, or in transforms of these. In all its occurrences there are various free variants available, such as *congregation* or *assemblage*, which are not themselves restricted except selectionally. If we take the unrestricted word as the transformational source, we can then say that for a particular subset of the environments of this word there is a free-variant morphophonemic transformation to the restricted form.

More generally, all idioms have unrestricted free variants. The words of an

idiomatic sequence (e.g. *he threw in the towel*) have a special restriction to each other, and sometimes to the environment, in the environments in which they constitute an idiom, as in *In the competition to see who could throw more towels in through the window he threw in the towel by refusing to throw in any towel*. Similarly, in the environments where this is an idiom,  $\bar{A}$  *He threw in two towels*. And in many idioms and metaphors various further transformations do not operate. Such restrictions do not apply to various free variants of the idioms, e.g. in this case *give up*, which preserve acceptability-ordering of their environments, even though these may not be felt subjectively as perfect synonyms of the idiom.

The claim made here for idioms has to do with the availability of specific free variants, and is not based on any general principle that everything can be said in a language. There are indeed things that cannot be precisely said: e.g. it is difficult to state a precise paraphrase of that which is intended in *They came in one after the other* (where the first person came in after no one).

### 3. UNRESTRICTED MORPHOPHONEMIC SOURCE FOR RESTRICTED OPERATORS

In § 2 we saw that many incremental operators with restricted domain had, over their own domain, free variants whose domain in turn was not restricted. There remain certain operators which do not have unrestricted free variants, but which nevertheless permit the formulation of an unrestricted source. This is possible first of all in the case of two or more forms which are complementary in their environments: in that case it has become customary to define a morphophonemic source, which does not exist in at least certain of the environments (where it is marked \*), and to say that the morphophonemic source occurs in all environments, except that where it is marked \* it is automatically (necessarily) changed in shape into the form that is indeed found there. This is done not only in individual cases such as *knife-knives* and in more prominent situations such as irregular verbs (*be*, etc.), but also in many paradigmatic situations such as number and gender agreement, and conjugations.

The morphophonemic source can still be formulated secondly, in a more complicated situation: when two forms contrast in some environments but are complementary in others. This situation, which in phonemics was called neutralization, is uncomfortable if we seek to state what are the independent elements in all sentences. The situation arises, for example, in gender if some nouns are only feminine, e.g. *la rivière*, while other nouns have both masculine and feminine forms, e.g. *le lion, la lionne*. In this case the feminine endings on *rivière* and the words in the scope of its agreement have to be considered

to be simply phonemic portions of *rivière*, while the feminine ending in *lionne* is a morpheme (in effect an adjective) independently adjoinable to *lion*.

### 3.1. Plural

A clear case of this neutralization is the plural affix. In *The books fell* as against *The book fell*, the *-s* is an independent morpheme, in effect an adjective. In *Three books fell* the *-s* is not independent:  $\bar{\exists}$  *Three book fell*, and *books* is a complementary variant of *book* in this environment. We can avoid this situation by noting that for every *N pl.* (noun plus plural suffix), in every environment where it has no quantifier, there is a free variant containing an indefinite plural pronoun, e.g. *two-or-more N pl.*:

I need books.

I need two-or-more books.

We consider a subset of English which has all English sentences except that instead of each sentence containing unquantified *N pl.*, it has the corresponding sentence with *two-or-more N pl.* In this subset, the plural suffix is never independent: it is always determined by such quantifiers as *two*, *three*, *many*, *two-or-more*. We now replace all the cases of *two books*, *many books*, *two-or-more books*, etc., by a newly created morphophonemic source *\*two book*, *\*many book*, *\*two-or-more book*, etc. In English, the plural suffix is a restricted increment, not occurring after *one N*. But in the above subset of English with its morphophonemic source there is no independent plural morpheme: There are only various quantifiers *one*, *two*, *many*, *two-or-more*, and each of these is unrestricted; and there is a pronunciation *two books* for the source *\*two book*, etc. The restricted plural is now no longer due to adding a restricted increment; rather, it is due to restricted morphophonemics, which pronounces the phonemes of *pl.* after certain quantifiers but not after *one*, *half*, etc., and it is due to the zeroing of a particular quantifier once this morphophonemic *pl.* was present, thus producing out of *two-or-more N pl.* a free variant *N pl.*

It is not necessary to suppose that every noun without *pl.* is singular, either in meaning or grammatically. In the grammar, the category 'singular' is no longer relevant (nor is 'plural' except as the domain of a particular morphophonemic change). In meaning, there is no reason to consider unpluralized mass-nouns as either singular or plural: e.g. in *water*, or even *much water*, *more water*; while *the waters* has to be considered as  $\leftarrow$  *\*the two-or-more water*.

### 3.2. Tense

A similar but more complex case is that of tense. The tense of a verb is

dependent on the time-location adverb of that verb: *I went yesterday*.  $\bar{\exists}$  *I went tomorrow*. But tense occurs also as an independent element on verbs lacking a time-location adverb: *I went*. *I will go*. Again we have neutralization. As above, we can find a free variant with general time-location adverb for every tensed verb which lacks a time-location adverb, and we consider the latter to be zeroed from the former:

He went.  $\leftarrow$  He went in the past (or: before now).

He will go.  $\leftarrow$  He will go in the future (or: after now).

He goes.  $\leftarrow$  He goes at present (or: now).

In the longer forms above the tense is always dependent, so that these can be derived from untensed morphophonemic sources, or nominalization:

\*He go in the past; or: His going is in the past.

\*He go in the future; or: His going is in the future.

\*He go at present; or: His going is at present.

In these source sentences the morphophonemically automatic *be* does not have to be *was* or *will be*; and its form *is* is clearly no indication of present time (which would be meaningless here) but is a morphophonemic requirement.

In sentences without time-location adverb (this can occur only when the sentences are operands), and in particular in the sentences as they are about to be nominalized under time-location predicates as immediately above, we cannot have had a tensed source. Here the tenseless morphophonemic form must have been the source: \*He go. Semantically, the time of such forms is indefinite or indeterminate, rather than being some kind of generalized present.

Timeless source sentences actually exist as operands. In *My going may turn out to be a great mistake* there is no indication of the time of *my going*, and there is no reason to say that this sentence is a degenerate grammatical ambiguity from precisely three sources: *My going in the past may...*, *My going in the future may...*, *My going at present may...* The same timeless operand sentence is seen in *I returned because of his departure*. It would go against the conditions necessary for zeroing to derive ... *because of his departure* from ... *because he departed* or from ... *because he will depart*, since such a derivation would lose time-information, whereas morphemes are zeroed only if they can be reconstructed from the remaining environment (although two different zeroings may yield degenerately the same residual sentence). Hence *I returned because he will depart* is a transform of *I returned because of his future departure* and not of *I returned because of his departure*; the latter has no tensed transform for the second operand.

Tenseless operand forms exist for every operator: no operator in English requires a tensed V as argument. Those for which the operand can be tensed, also have a tenseless form of the operand as in *I regret his departure*, *They made him leave*. This explains why every tensed operand has a tenseless apparent transform but not conversely: (1) *That he went is a fact*; (2) *His going is a fact*; *His driving is slow*,  $\nexists$  *That he drives is slow*. For we now see that (1) is a transform not of (2) but of *His going (being) in the past is a fact*. But *is slow* does not operate on time predicates: *His driving in the past was slow* is derived not from *is slow on is in the past*, but from *and* on the two increments *is in the past*, *is slow*, producing *His driving which was in the past was slow*, *His past driving was slow*. Hence *is slow* has only tenseless operands, while *is a fact* has both tensed and tenseless.

The occurrence of tensed forms can be described as a free variant or automatic change (required variant) of time-location adverbs; but doing so requires a rather complicated statement of, first, when this variant occurs and, second, which tense goes with which particular adverbs.

First, when does tensing occur?

Under certain sentence-operators (e.g. *demand*) the argument verb is never tensed, and under certain operators on sentence-pairs (e.g. *in order that*) the second argument is never independently tensed. As for the subjunctive in 2.4, the time-location adverb A of the argument in question is partially though not entirely dependent upon the time-location adverb B of the sentence-operator in the first case or of the primary sentence in the second case: A is never in the past in respect to B. In these cases, an operand sentence which occurs in the nominalized form *N's Ving* can be transformed into the 'subjunctive' form *that NV*, *that N should V*, *for N to V*, but not into the tensed (indicative) form *that N tense V*.<sup>11</sup>

A few other operators, whose arguments are also never past relative to them, have their operands neither nominalized (*N's Ving*) nor subjunctive but only a tenseless *NV*:

They made him go.  
They let him go.

Other sentence-operators never have tensed operands for other reasons. Because predicates of manner do not operate on time-predicates (as in *is slow* above) they have only nominalized operators:

His driving has (or: is in) a hesitant manner.  
The manner of his driving is hesitant.  
His driving is hesitant (in manner).

The operand is tensed only via the tense-transplacing T (8.1 (10)), in which

the predicate becomes an adjunct (adverb; 8.4 (8)):

He drives in a hesitant manner.  
He drives hesitantly.

In all other situations, there are free variants in which the untensed operands occur tensed.

The tense-transplacing transformations (8.1 (10)) move the tensing from the last operator (the one which is not itself an operand) to its first argument. E.g.:

- (3) Her return tomorrow is because of his departure yesterday.  
(4) → She will return tomorrow because of his departure yesterday.  
Her driving will always be slow (in manner).  
→ She will always drive in a slow manner (or: slowly).

A tensed free variant (8.1 (12)) also occurs in the second argument-verb under certain  $V_{ov}$  (7.2, in particular those consisting of *be P*, e.g. *be because of*) and under *be before*:

- (3) → Her return tomorrow is because he departed yesterday.  
(4) → She will return tomorrow because he departed yesterday.  
His return was before her arrival →  
His return was before she arrived.  
He returned before her arrival →  
He returned before she arrived.

Under the comparative, this happens only if the first argument-verb is tensed:

- (5) His riches are more than her riches  $\leftrightarrow$   $\nexists$  His riches are more than she is rich.  
(5) → He is rich more than she is.

Finally, the predicate analysis of §§ 4, 5 makes possible a simple formulation of the occurrence of tensing: In every sentence the predicate operator which is not itself an operand (except of *and*, *or*) is tensed. That is to say that when, in constructing a sentence, we stop the construction after applying a particular predicate operator (with possibly *and*, *or* operators on it), this operator is then tensed: the tensing can be considered an automatic variant (a required morphophonemic form) of placing a period or sentence-intonation upon the word-sequence. It is better to take period as an independent entity, with tense as automatic in respect to it, than to take tense as independent, because tense also occurs in other positions (above) where it is not

independent but a free variant. The tensing should also not be considered to be simply the morphemic realization of the act of asserting, because in its free variant occurrences it is neither more nor less assertion than is its untensed variant: Both forms of the operand are equally asserted in *I announced his having come on time, I announced that he came on time*; the tensed *he came* is not asserted in *I wonder if he came*; and the untensed operand is implicitly asserted in *I made him bring it*.

Second, given that tensing is a free or automatic variant, it remains to state how the particular tense is determined. A verb is tensed with *-ed* if it has on it any past-time adverb or predicate (including the general *in the past* or the like, which is then zeroable). A verb is tensed with (the auxiliary) *will* if it has on it any future-time adverb or predicate (including the general *in the future* or the like which is then zeroable). A verb is tensed with zero (with variant *-s* after third person singular subject) if it has any present-time adverb or predicate (including the zeroable *at present*, etc.) or if it has no time-location operator (in which case it has not been restricted to being in the past or in the future, so that it often means unspecified or distributed as to time). An example of the latter is seen in *His going is in the past*.

There are a number of amendments to the above. If the verb is a sentence-pair operator its tense is partially restricted by the time-location of its two operands:

- His having arrived will spoil their leaving on time soon.  
 His having arrived spoiled their leaving on time soon.  
 ? His having arrived will spoil their having left on time.  
 ? ? His arriving an hour from now spoiled their leaving on time soon.

Under certain forms of certain operators, the tense of the operand is determined not only by the time-location of the operand but also by that of the operator. Thus

- He announced her being ill in the past. → He announced that she was ill, He announced that she had been ill.  
 He announced her being ill in the present. → He announced that she was ill.

The same operands have different tenses in the 'direct discourse' transforms of the above sentences:

- He announced: 'She was ill'.  
 He announced: 'She is ill'.

#### 4. RESTRICTION DETERMINED BY ENVIRONMENT

In § 2 it was seen that certain increments which appeared to be restricted to particular subsets of the domain of a variable could be derived from unrestricted increments. In § 3 it was seen that certain other restricted operators were variants determined by the presence of other operators, i.e. of particular morphemes in the environment (e.g. quantifiers or time adverbs). Here we consider how yet other grammatical forms, which appear to be restricted to words of some particular subset, are derivable as variants of particular morphemes which are present in the environment of words of that subset.

What is eliminated in § 4 is mostly restrictions on certain paraphrastic transformations, rather than on increments. However, the elimination of these restrictions requires certain added complexities in the source form of the increments upon which the transformations operate (e.g. 4.1, 4.3), and it is therefore being considered here.

##### 4.1. Subjunctive

This type of situation has already been met in the subjunctive (2.4), where the non-tensing of the operand was due to a particular dependence between the time-adverb of the operand and that of the operator.

It was seen there that all sentences containing this time dependence had as one of their free variant transforms a sentence which contained the statement of that time-dependence: *subsequently*, or *afterwards*, or the like. In the formulation of grammar there is a great difference between saying that the subjunctive occurs in the operands of verbs that necessarily (or normally) have this time-dependence to them, and saying that the subjunctive occurs in the operands of those verbs that necessarily (or normally) impose the word *afterwards* on them. For in the latter case, we can simply say that the subjunctive is a free variant of the required *afterwards*; and the required *afterwards* can be considered as simply a part of those verbs which always impose it on their arguments. It is true that the requirement of *afterwards* can be considered a classifier of a subset of words, but it can also be considered a morphemic segment of those verbs and thus simply a part of the sentences in which those verbs occur, a part that can be replaced by free variants of it. Then the source form of *request* would be *request-for-afterwards*.

The analysis of *afterwards* will be clearer if we use here the operator notation of §§ 6-7 (see note 1), in which  $X(Y, Z)$  indicates that  $X$  is operating on the pair  $Y, Z$  as its arguments; in elementary sentences the verb is taken as operating on its subject and object (7.1). Then *I request that he go* is in operator notation:

request-for-afterwards (I, go (he)),

and in the actual string of words (7.5):

\*I request-for-afterwards his going

which becomes by length-permutation (8.1 (6)):

I request his going afterwards,

where the presence of the required *afterwards* prevents the morphophonemic operation of tensing and leaves a tenseless transform

I request that he go (afterwards).

#### 4.2. Adverbs of Manner

The appeal to environment has also been met in the predicates of manner (2.3), which were seen to be derived from adjectives (in principle, arbitrary ones) plus the word *manner*: (1) *His driving of trucks was hesitant* ← (2) *His driving (of) trucks was in a hesitant manner* ↔ *The manner of his driving (of) trucks was hesitant*. What makes them adverbs of manner is that the words in ... *manner* had occurred with them, but were zeroable and zeroed. Furthermore, since the nominalization with *of* before the object (1) is required only when *manner* is absent, we see that the requirement is a variant of *manner*: (1) and (2) are free variants of each other. And indeed the verbs which have *manner* as object (not as part of a predicate) also have this *of*-nominalization when *manner* is zeroed: *They imitated the manner of his driving (of) trucks, They imitated his driving of trucks*. Of course, in order to determine the special transformation (in this case, *of*), the determining environment (in this case, *manner*), if it is zeroed, must be present at the moment of application of the special transformation.

#### 4.3. Reciprocals

Free variation to environing morphemes is seen in the case of the reciprocal verbs. For all verbs, a sentence of the form:

A saw B and B saw A

can be transformed into:

A and B saw B and A respectively.

which transforms into:

(1) A and B saw each other.

There is a particular subset of verbs in English, called reciprocal, after which

*each other* can be paraphrastically zeroed:

A met B and B met A.

A and B met B and A respectively.

(2) A and B met each other.

(3) A and B met.

In such a derivation, the zeroing of *each other* in (2) but not in (1) is determined by the presence of *met* in (2) as against *saw* in (1): that is, the determining environment is a particular subset of words, the reciprocal verbs,  $V_{rec}$ , which have to be listed.

However, there is another way of arriving at (3) which has the same grammatical character as in 4.1. We note that the reciprocal verbs, for which (3) occurs in the sense only of *each other*, are precisely the verbs  $V_{rec}$  for which

(4) A  $V_{rec}$  B and A's  $V_{rec}$ ing B implies B's  $V_{rec}$ ing A.

(4') A  $V_{rec}$  B and that A and B  $V_{rec}$  B and A respectively is implicit.

are transforms of

A  $V_{rec}$  B

as in:

A met B and that A and B met B and A respectively is implicit.

Therefore the sentence:

A met B and B met A

→ A and B met B and A respectively

has various transforms of the type (pronounced from (4')):

(5) A and B met B and A respectively, and this is implicit.

Of course, this is not to say that (4) may not occur for verbs not in  $V_{rec}$ , when they have particular A, B or particular conditions (which could be appended to the sentence): e.g. in a given situation we might have (6) *He saw her and his seeing her implied her seeing him*. But (6) would not be a transform of *He saw her*, preserving acceptability-ordering for all choices of subject and object.

We can now say that what is zeroable is the transformational case (as in (4')) but not (6)) of *B and A respectively and this is implicit*, which indeed meets the informationlessness criterion for zeroing. To zero this segment requires no appeal to a determining environment. (3) is then derived not from (2) but from (5).

At this point one might think that all that has been gained is a shifting of the definition of the  $V_{rec}$  subset, from a subset defined by a list to one defined

by having the transforms (4). However, the difference is that now the reduced form (3) is derived by zeroing the zeroable material in (5) and not from any checking of the environment; but the cost is that the source form of *met* is now as in (5). The  $V_{rec}$  can occur without the adjoined implication of (4) and (5), and in that case the zeroing does not take place; in that case the  $V_{rec}$  can only reach the form with *each other*, in which it does not differ from any other verb. Furthermore, as in 2.4, the availability of (4), (5) is not as sharp a grammatical restriction as would be thought from the listed  $V_{rec}$ . There are verbs, e.g. *equal*, which are fully acceptable in forms (4) and (5), and can with hesitation be used in form (3). As before, the source forms have some freedom in selection of co-occurents, while the paraphrastically shortened forms (3) are frozen into grammatical subjects. Here as elsewhere (e.g. 2.1), what is sharply grammatical in the transformational paraphrase system is derived from what is flexibly selectional (co-occurrence) in the source sentences.

#### 4.4. Sentence nominalization. *Sn*

A more important problem replacing subset lists by source environments is the choice of forms (deformations) taken by operand sentences under various sentence-operators.

Just as some operators get the subjunctive (i.e. do not get tense) on their operand, so other operators have other special deformations of their operand. Most of these are variants of the basic *N's Ving* deformation of 7.5. Thus for a certain subset of  $V_{nv}$  operators (7.2) such as *prevent*, the operand  $N_1$ 's *Ving*  $N_2$  is transformable to  $N_1$  from *Ving*  $N_2$  (*They prevented his taking it, They prevented him from taking it*); which is not the case for, say, *regret* (*They regretted his taking it, † They regretted him from taking it*).

A particularly difficult case is the subset of  $V_v$  operators (7.2), such as *undergo*, whose most common sentence form has for the subject of the operator the same individual as the object of the argument: (1) *He underwent (elaborate) testing*. Since the subject of *undergo* is not an independent variable, we cannot say that *undergo* is  $V_{nv}$ , i.e. an operator on a noun (as its subject) and verb (which in turn carries its own subject and object). We have to say that *undergo* is  $V_v$ , i.e. an operator on a verb (which carries its own subject and object), and that by some transformation the object of the operand verb comes to be the subject of the  $V_v$ . The source form would have to be something like, for (1):

$N$ 's (elaborate) testing of him was an undergoing (or: an experience).

Then we would have to say that there is a transformation, for these opera-

tors but not for others, in which the object of the argument becomes the subject of the operator:

He underwent  $N$ 's (elaborate) testing of him

and, with zeroing of indefinite  $N$ 's and of the now repetitive of *him*:

He underwent (elaborate) testing.

In contrast, many (but not all)  $V_v$  operators have the argument-skipping transformation (8.1 (9)), in which the subject (not the object) of the argument becomes the subject of the operator:

Their (elaborate) testing of him was hard work.

→ They worked hard at (an elaborate) testing (of) him.

The number of different operand deformations under different subsets of operators is not large. The detailed derivation of free variants of these deformations leads to uncomfortable and even marginal forms: However, the hope is that one can find free variants which will contain specific morphemes peculiar to the various subsets, such as would make it possible to set up a source in which these morphemes were present as a (classifier) part of the operator and were then replaced by the deformation common to that operator. This is what was done in deriving the subjunctive (on the operand) from *for afterward* affixed to the operator. In this method, the source of *prevent* might be something like *prevent-in-respect-to-the-subject* (of the argument verb), and the source of *undergo* might be something like *an-undergoing-for-the-object* (of the argument verb).

We could obtain the various deformations of the operand sentence quite simply if we were willing to have an operator act on a pair consisting of an argument B and an argument C which must itself be an operator on B. In that case, *prevent*, for example, would be

prevent ( $N, N_1, V(N_1, N)$ )

and *undergo* would be

undergo ( $N_1, V(N, N_1)$ ).

However, this is a less economical description than the one above, since it sacrifices the principle that each variable must be independent, and therefore it does not show the limitation in the range of combinations, or the semantic relations, as directly as does the first description.

## 5. PREDICATIONAL SOURCE FOR INCREMENTS

### 5.0. Introduction

The attempt in § 2 to reformulate restrictions and subsets by finding



free or automatic variants to the various increments has in many cases turned up variants which had the form, in respect to the rest of the sentence, of a predicate operating on sentences. The term 'predicate' is used here to indicate verb or its linguistic equivalent: e.g. in the order of appearance above *be in process*, *be to great extent*, *be of slow manner* (*have slow manner*), *be easy*, *demand*, *be before* (*precede*), *be more than* (*exceed*), *be in past*, *imply*. In all these cases the predicational form could be considered the source, from which the other forms are paraphrastically derived, so that the grammatical restrictions are on the paraphrase, not on the source. This raises the question whether every increment has a variant form that is a predicate, which could then be taken as source even if no restrictional advantages are gained thereby. If there was a single relative position in the sentence, such that all increments could be derived paraphrastically from operators having that position relative to their operand, it would be convenient for the simplicity of a theory of language, and it would show that all other syntactic relations are merely paraphrases of the single relation in the source. This is all the more important when the source is largely restrictionless and yet carries all the objective information carried in language. For then we can see that a single relation of certain argument-specific operators to their arguments suffices to carry the objective information. We therefore consider for each increment the set of free (or automatic) variants which it has,<sup>12</sup> and we survey all the sets to see if there is some one form which is to be found in every set – preferably the form in which the operator is a predicate in the above sense. If in a given set of variants (transforms) we do not find such a form, we will try to see at what morphophonemic or other cost a variant of this form can be added to the set.

In what follows, many of the attempts to derive increments from a predicate source will seem forced. Some derivations from predicate form have grammatical justification, even if they are not entirely obvious: the moods (5.6) and most of the noun-adjuncts (5.1), verb-adjuncts (5.2), and auxiliaries (5.3). The remaining noun-adjuncts (chiefly quantifiers, *the*) and verb-adjuncts are few and special. It can be argued that there is good reason to try to regularize them (to predicate form), and that relevant classifications are brought to light in the process (e.g. about *all*, *the*). However, one might claim that even if this is the case there is no point in forcing the verb-operators (e.g. *begin*, *take a*) and subordinate conjunctions into the framework of predicates. These are major distinct constructions of grammar, and it seems quite unreal to derive them from uncomfortable or non-existent predicates. Even the fact that many verb-operators occur also as predicates, and that many subordinate conjunctions have associated verb forms, does not alter the grammatical specialness of these non-predicate forms.

However, the intention in this chapter is not to claim that the predicates are sources in any sense except the very special one used here, namely that the established transformations plus a heavy use of morphophonemics suffice to produce all the increments out of predicate forms, and that often these forms are clear syntactic sources in that they are less restricted as to domain or further operability. Thus aside from the standard paraphrastic transformations, the only difference between the normal sentences and the predicate system is morphophonemic, i.e. is in phonemic shape and position. The implications for information, and for the relation of information to syntactic structure, are obvious.

When we can show that no more than one type of operator-operand relation is needed for language, we can deduce one of the major properties of language computability. When grammar is presented with several kinds of operators – adverbs, conjunctions, etc. – it is clear that there are restrictions on the operator combinations, and that these restrictions preserve a certain connectedness among the sentence-parts. For example, *I clearly remember its being put in service with great to-do* is transformable to *I remember its being put in service with great to-do with clarity*, but not to *I remember its being put in service with clarity with great to-do*. The lines connecting *remember* to its adjunct *with clarity* and connecting *put in service* with its adjunct *with great to-do* cannot cross. This does not have to be presented as a special condition on sentence-construction. For when all the segments in a sentence are seen to be brought in by a single type of incremental operator, we see that this connectedness is due to the fact that in each sentence-construction the single operator-type can bring in the segments in only one order.

The predicate source is possible because the different morphological and apparently syntactic classes of words – adjectives, subordinate conjunctions, etc. – do not have different combinability in the language than do particular classes of predicates that can be considered to correspond to them. There is no noun-adjunct that cannot be paraphrased by a conjoined sentence containing that noun, nor any subordinate-conjunction-sentence which cannot be paraphrased by a sentence-pair with an inter-sentence verb. So many grammatical constructions turn out to be derivable from predicates, and so many difficult constructions are conveniently explainable in this derivation, that the syntactic relation here is clearly not external to the language. The predicate sources are not morphologically simple. Since the syntax of a language describes the combinings and relative positionings of elements in sentences, we see that the predicate sources are syntactic primitives and not morphological ones, and they represent (by paraphrastic transformation and morphophonemics) all the morphemes of no matter what morphological class which have identical combinability relative to other morphemes. Just as

phonemes are not necessarily phonetically simple, so the syntactic elements determined here are not necessarily morphologically simple, though many of them are.

### 5.1. Noun Adjuncts

We first note briefly the primitive adjuncts of nouns: *the, all, many, only*, etc. The adjuncts, other than these, of a noun  $N_i$  are all derived from predicates of another sentence whose subject is also  $N_i$ , the second sentence having been adjoined by *wh-* (2.8) to the given occurrence of  $N_i$ : *The heavy book fell* ← *The book which is heavy fell* ← *wh-* (*The book fell, The book is heavy*). The numbers and some of the quantifiers can also be derived in this way: *Five books fell* ← *Books which numbered five fell* ← *wh-* (*Books numbered five, Books fell*); although at a later stage the numbers could be derived from repetitions (under *and*) of the singular sentences.

Indefinite and universal quantifiers such as *all* cannot be derived from some number of repetitions under *and*, nor do they appear by themselves as predicates of the nouns to which they are adjoined:  $\exists$  *The men are all*. However, these quantifiers carry explicit or implicit (i.e. zeroed) references to a domain over which they are taken; and indeed these quantifiers can be derived from appropriate predicates connecting their noun to this domain. E.g. for *all*: *All the men in the room left* ← *Men comprising the set in the room left*: *All books must be returned* ← *All books which have been taken must be returned* (or the like) ← *Books covering the set of those which have been taken must be returned*; *All men die* ← *Men exhausting the set of (what is called) men die*. The permutability of *all* (*Men all die; Men die, all of them*) supports such a derivation.

As to *the*, there may be several sources, from one or another of which we can derive each occurrence of *the*, according to its environment.<sup>18</sup> Certain cases of *the* are automatic in the presence of certain adjuncts of the noun. Others are free variant replacements for such adjuncts as *which has just been mentioned*.

### 5.2. Verb Adjuncts

There are a few adverbs of degree which do not have predicate transforms: *I quite forgot; I simply forgot* ( $\neq$  *My forgetting was simple*). As in the case of *very* (2.2), these can be derived from synonymous predicate-transformable adverbs of degree or conjoined clauses which satisfy the conditions for being free variants of them: e.g. *quite* from *completely*, *simply* from *without qualification*.

Almost all adverbs have predicate transforms. That the predicate forms are the source has been shown from 2.3 and from the fact that adverbs

generally operate semantically on all adverbs which are between them and the verb: Each adverb arose as a predicate operating on the previous adverbial predicate: *He has been writing clearly recently*. ← *His writing clearly has been recent*. ← *His writing being clear has been recent*.

A special problem exists in the case of *not*.  $S \rightarrow not\ S$  is an incremental transformation: Not only does it derive one sentence from another, but it preserves subject-matter context (*The values approach infinity* and *The values do not approach infinity* occur in the same subject-matter sublanguage); and like all incremental transformations it preserves acceptability-ordering over a large subdomain (*The flower fell* and *The flower did not fall*, as against *The flower thundered* and *The flower did not thunder*).

First, we note that all negatives are derivable from *not* on the verb (i.e. in the sentence). Thus *We lost no time* ← *We did not lose (even a little) time*; and *non* occurs more acceptably on predicate nouns (e.g. *non-student*, derived from a verb) than on primitive operand nouns (such as *chair*). Secondly, not too much morphophonemics is involved in deriving *not* as adjunct on verb from *\*is not* as predicate on sentence. Such a derivation is indeed needed for *It is not that  $S_1$  (but that  $S_2$ )*, which has to be derived ← *\*That  $S_1$  is not*, by the well-established transformation seen in *It is false that he came* ← *That he came is false*.

If *\*is not* is taken as a predicate which operates on sentence we find that its further transformations are similar to those of other sentence-operators, except for morphophonemic details. E.g. when we operate with an additional increment, say *may help*, on these sentence-operators we obtain:

- \*My driving is not. → My not driving may help.
- My driving is (in the) future. → My future driving may help.
- My driving is slow. → My slow driving may help.

Another peculiarity of *not* appears in the argument-skipping transformation (8.1 (9)) which is seen in

- My driving continued. → I continued driving.
- My driving was slow. → I was slow in driving.

Here *not* is peculiar in that the verb receives no affix (as it receives none under the auxiliaries), and the tense and auxiliaries are placed before the *not*:

- \*My driving was not. → I did not drive.

This leads to certain degeneracies: In *not* operating on *may* operating on *I drive*, we have *not* on *I may drive*, which yields the word-string *I may not drive*, meaning: it is not the case that I may drive, my driving is not possible

or allowable. Here we have *not* placed after *may* by the argument-skipping transformation, though this transformation otherwise puts the last operator before its argument (8.1 (9)). Now, in *may* operating on *not* on *I drive*, we have *may* on *I do not drive* (or: *my not driving*) where the same transformation produces, ambiguously, the same word sequence *I may not drive*, meaning: it is possible that I will (or: do) not drive, my not driving is a possibility.

Aside from this, the different placings of *not* in the sentence, which mean negation of different parts of the sentence, are appropriately derivable from different orderings of the *\*is not* predicate in the ordering of predicates that constructs the sentence. This is the same as we have seen for the different placings of adverbs (also derived from predicates).

A special situation that may be treated here is that of preposition P following verbs. The P may be considered an adverbial operator on the verb in most cases (e.g. *He looked up*), even though the source with the adverb as predicate is asterisked (*\*His looking was up(ward)*): the tense-transplacing transformation is required here. In the forms with an object, two main types are distinguishable: (1) having no permutation of P, e.g. *They looked out the window* (or: *up the street*) ( $\nexists$  *They looked the window out*); (2) having permutation of P, e.g. *They threw out the food*, *They threw the food out*. Here there is in many cases a (not very comfortable) transform (2') *They threw the food so that it was out*. We might distinguish a type (3) which differs from (2) only in not having this transform, e.g. *They looked up the number* (or, ambiguously to the above, *up the street*), *They looked the number up*.

Type (1) can be derived from a prepositional predicate on a pair: (V, N):

out (look (they), window).

producing the uncomfortable source form

Their looking is out the window

and by tense-transplacing

They look out the window.

This is the operator form that would represent case-endings in languages which have such.

Type (2) can be derived from a prepositional predicate on a verb with object:

out (throw (they, food))

producing the source form:

Their throwing of the food is outward

and by tense-transplacing:

They throw the food out

and by length-permutation on short objects (8.1(6))

They throw out the food.

The conjunctive transform (2') results from the fact that the operator *out* is more precisely *out-in-respect-to-the object* in the manner of 4.4 (end).

Type (3) could be derived from a two-morpheme verb:

look up (they, number)

producing

They look up the number

and with length-permutation on short objects:

They look the number up.

### 5.3. Compound Tense; Auxiliaries

There are two other problems of finding a predicate source for verb-operators, which are special to English and related languages: the auxiliaries and the compound tense in *has gone*, *is gone*.

In the matter of *have-en* we note that there are certain environments in which it is not independent. We will consider here the tensed and tenseless positions of *have-en*.

For the tensed position: By itself, *have-en* seems independent: *He has gone*, *He went*. But under certain operators it is not: (1) *He has gone by now*,  $\nexists$  *He went by now*; *At the time of his arrival I had already gone*,  $\nexists$  *At the time of his arrival I already went*. In (1) we can say that *by now* has automatically (morphophonemically) produced the *have-en*. In environments in which *have-en* and the past contrast ((2) *He has gone twice today*, *He went twice today*.) we can then say that some increment like *by now* had been present and had morphophonemically produced the *have-en* form, and had thereafter been zeroed. The source of (1), (2) would then be something like (3) *\*His going is by now*, *\*His going is twice by now today*. The required morphophonemics in (3) would produce *His going has been by now*, whence tense-transplacing yields (1).

For the tenseless position: Not only do we find *have-en* on verbs with *ing*, where tense does not occur, but we find that in some environments the tenseless *have-en* is related to the past-tense rather than to the tensed *have-en*. Thus (4) *I regret his having gone at just that moment* is not the resultant of *regret* operating on (5)  $\nexists$  *He has gone at just that moment*; since (5) does

not exist. If we compare (4) and also *I regret his going right now with He went at just that moment* and *He goes right now*, it is clear that (4) is the result of *regret* operating the source-form of *He went at just that moment*. That is, *have-en* is a complementary variant here of the ordinary past tense.

In (4), the increment which produced *have-en* could not be *by now*, since *just at that moment* would not co-occur with *by now* (hence (5) is unacceptable). Since *his having gone* in (4) can occur with *yesterday* and other past-time morphemes, but not with future-time morphemes, it follows from the method of 3.2 that there had been present here the ordinary time-location *in the past*. Since *in the past* cannot be transformed into the tense *-ed* when, as in (4), it is in a nominalized sentence, it is here transformed into *have-en*. Thus *have-en* is here a variant of *-ed*.

The fact that *have-en* can appear as a morphophonemic variant of two time-operators – *by now* and *in the past*, in two different environments – makes its derivation more delicate. But it shows that *have-en* is not itself an operator (of a physical type which is unusual in the language and hence inconvenient for a compact structuring) but is a morphophonemic change brought in by an operator (or, in different environments by different operators); and we can now choose the source form of these operators to be the predicate form.

The conditions which determine the producing of *have-en* also involve the perfectivity or imperfectivity of the operand verb: e.g. not all environments which require *have-en* on *go* will also require it on *like* (*At the time of his arrival I already liked it*).

As to the auxiliaries. These are words having many of the properties of the operators on verb which, as will be seen (5.4), can be derived with large or small cost from predicates (operators on sentence). The auxiliaries differ (in English) in that their operand verb is affixless (*began going*, *began to go* but *can go*); and they do not occur in deformed operands (e.g. nominalized sentences), hence cannot operate on each other ( $\bar{A}$  *can may go*); also *not* operating on them is placed after and not before them. These properties are also characteristic of tense. But we have seen that the tense can be derived as a transform of time-operators, and that it is only the tense-transform of the time-operators that cannot occur in the deformed-operand forms of sentence: the source form of the time-operators, or other transforms of them such as the *have-en* in (4), do occur there. We will attempt a similar source for the auxiliaries.

For one auxiliary, the case is almost as good as for the past tense, for we have the same dependence upon a set of time-operators. This is *will*, and we can say, as in 3.2, e.g.:

His coming is tomorrow → He will come tomorrow.

and with zeroing:

His coming is in the future → He will come in the future →  
He will come.

For the other auxiliaries we do not have such a convenient classifier relation to a set of predicates which can be said to produce the auxiliary morphophonemically. All we can do is to seek a synonymous predicate whose acceptability-ordering the auxiliary preserves. Then we can propose something like:

(6) For him to go is an ability (or: a capability, possibility) → (7)  
He is able to go. → (8) He can go.

(6)→(7) is the argument-skipping transformation. (7)→(8) is a unique morphophonemic one: Support for it can be found in the use of *is able to* instead of *can* wherever *can* cannot occur or is disappearing from use. This is seen in nominalizations: (9) *He can clear 7 ft., but his ability to do it is less than it was*; or under the past-operator: *He can not clear 7 ft. now but in his college years he could still do so*, which today is more likely to be (10) *He can not clear 7 ft. now but in his college years he was still able to do so*.

Additional support for the predicate source of the auxiliaries is seen in the way many of them have two meanings of the following type: (11) *A boy can jump all day (and not get tired)* ← *A boy's jumping all day (and not getting tired) is a capability (or: possibility)*; *A boy is able to jump all day*. In contrast: (12) *A boy can speak five languages, (but it is rare)* ← *A boy's speaking five languages occurring is a possibility, (but it is rare)*; *The occurrence of a boy's speaking five languages is a possibility*; *A boy's speaking five languages is capable of occurring*. In (11) the action is a capability; in (12) the occurrence of the action is. The source (11) has *is a capability* operating on *jump*; the source of (12) has *is a capability (or: possibility)* operating on *occur* operating on *speak*. The same structure as in (12) appears in (13) *A boy can jump all day (without getting caught)*. The sentence common to (11) and (13) is ambiguous, and the ambiguity is due to the zeroing of *occur* in (13) as in (12).

The analysis of *can* is paralleled for *may*, *must* if we take (with *that* as a morphophonemic constant due to the operator):

He may go. ← That he go is potential.  
He must go. ← That he go is obligatory.

In addition, there are, though somewhat in disuse, *shall* (now synonymous to *will*), and the resultants of past on *will* (*would*), *shall* (*should*), *can* (*could*) and very rarely *may* (*might*). A different kind of source appears in *should* (or *would*, *might*), which is produced as a variant of zero (non-tensing) under the

'subjunctive' sentence-operators and sentence-pair-operators:

- (14) I request that he go; I request that he should go.  
He left early lest he miss her; ... lest he should miss her.

This analysis treats the auxiliaries as merely a subset of predicates (expressing types of readiness) which transform like the verb-operators (5.4) but have morphophonemic properties somewhat like those of the time-predicates. Then when auxiliaries are under an operator requiring a deformation, the operator simply operates on their source and not on the auxiliary form. More precisely, the morphophonemic transformation from the source to the auxiliary form does not occur in the same situations where tensing cannot occur (except for the variant of non-tensing as in (14)). It is the transformation to auxiliary form that is restricted. And, as elsewhere, the source does not have the restrictions that the transform has: it occurs in the predicate position as in (6), and under nominalization as in (9), and under time-operators as in (10).

In view of the secondary status of the auxiliaries, it is not surprising that the subset is not sharply closed. Most of the relevant morphophonemics apply optionally also to *need* (*He need go but once* ← *His going is a need but once*, and the like), *dare* (usually under *not*), and *ought* only under *not* (*He ought not go* ← *That he go is not a duty*, or *an owing*, or the like). Less close are certain specially restricted verb-operators with *to* which are losing the phonemic and morphemic recognizability of their *to*: *used to*, *has to*, *is going to*, *is supposed to*. These too could be derived from synonymous predicates.

In respect to the affixless form which the arguments of the auxiliaries receive, it may be noted that a similar affixless verb-argument appears under the predicates *make*, *let* (*I made him go*, *I let him go*). Compare the *to V* operands under  $V_o$  such as *important* (*For him to go is important*) and  $V_{no}$  such as *prefer* (*I prefer for him to go*); and *that*-operands under  $V_o$  such as *is a fact* (*That he went is a fact*) and  $V_{no}$  such as *know* (*I know that he went*).

#### 5.4. Verb-Operators

A major type of increment in many languages is the verb-operator, which is seen in e.g. *He ceased writing*, and *He delayed writing*, as against *He wrote*. Verb-operators can be derived from sentence-operators (predicates on a sentence) by several paraphrastic transformations, chiefly by argument-skipping (8.1 (9)), as in

- (1) He ceased writing. ← His writing ceased.  
like  
He is hesitant in driving. ← His driving is hesitant;

and, differently, by zeroing of a repeated subject, as in

- (2) He delayed writing. ← He delayed his writing.

If the subject of the operand is different from the subject of the operator, zeroing of course does not occur, and we do not obtain the effect of a verb-operator: *He delayed her writing*.

The great bulk of English verb-operators can be derived from sentence-operators in one or the other of these two ways. For derivations as in (1): Appropriately to their untensed operand (verb plus *to* or *-ing* or nominalization) these verb operators come from sentence-predicates whose argument is not tensed: Thus we have:

He is irresponsible to see her. ← For him to see her is irresponsible.

and also for predicates whose arguments can be tensed or not:

He is likely to see her. ← For him to see her is likely, That he will see her is likely.

He is certain to see her. ← For him to see her is certain, That he will see her is certain.

but not for predicates whose argument must be tensed unless it is nominalized:

It is probable to see her; It is probable that he will see her is probable.

In many cases the verb-deformation after the argument-skipping remains what it was under the predicate increment, e.g.:

He began writing. ← His writing began.

However, there are cases of *to V* after the verb-operator transform, where the operand was *Ving* (and not *to V*) under the predicate source:

He began to write, It is probable for him to write began.

Similarly:

He (just) happened to come on time. ← His coming on time (just) happened.

In some cases there is an added morpheme:

He persisted in talking. ← His talking persisted.

And the predicate may be a noun instead of a verb:

He took the trouble to see her. ← For him to see her took trouble.

The question may be raised whether these are indeed paraphrastic transforms of each other, since there seems to be a noticeable difference in meaning.

However, the acceptability-orderings seem to be preserved throughout. And the difference in meaning is the subjective one of attributing the new predicate to the subject of the operand sentence: but this is precisely what the form change does explicitly, and it is felt, though more weakly, in other cases of the argument-skipping transformation, as in *He is hesitant in driving* ← *He has a hesitant manner in driving* ← *He drives in a hesitant manner*.

The other derivation, of (2), is seen in clear cases such as

He regretted writing it. ← He regretted his writing it.

(cf. *He regretted her writing it*), and also in less clear cases such as

He missed meeting her. ← He missed his meeting her.

where we may be uncertain about *He missed our meeting her*. For several verb-operators it is not clear whether their source predicates are of one kind or the other: If *He stopped their writing us* exists and is derived from *stop* operating on the pair (*he, they write us*), then

He stopped writing us. ← He stopped his writing us.  
= stop (he, write (he, us)).

However, if *He stopped their writing us* is not really acceptable, or if it is derived from *He made them stop writing us* (which is *make* operating on *They stop writing us*), then

He stopped writing us. ← His writing us stopped.  
= stop (write (he, us)).

This is a detailed problem of English transformations; and the language may be changing in this respect so that both solutions may be possible and no solution 'right'. In either case, a predicate source is available.

In some cases a predicate variant can indeed be found (e.g. *is an attempt* as variant of *attempts to*) but with unwontedly and unwantedly complex restrictions. Consider verb-operators like *try*, *attempt*, perhaps restricted to subjects denoting living organisms: *The dog tried to move?* † *The book tried to move*; but without restriction: *The dog tends to move*, *The book tends to move*. If this were derived from an operator on sentence, e.g. *The dog's moving is a try*, we would have an operator restricted to particular subjects of its operand.<sup>14</sup>

try (move (dog))

where *try* restricts the argument (*dog*) of its argument: † *try (move (back))* as above. However, to an uncertain extent *try* may be derived from operators on the pair (N,S) if we accept such sentences as *He tried that the scaffold should*

*remain*: then the restriction to living subject becomes a restriction of *try* in the domain of its own first argument: in

try (he, remain (scaffold))

the restriction of *try* is to its own first argument (*he*), which is the normal situation for operators. Even if such sentences are (or have become) unacceptable in English, their place is taken by a circumlocution: *He tried to have the scaffold remain*. If *have* here is taken as the argument of *try*, then the restriction holds for the subject (argument) of the argument. But if *try to have* is taken as a variant of *try* in all cases where the subject of *try* is not the same as the subject of the argument (*remain* as against *move, go*) of *try*, then we obtain *I try to go* as *try* operating on (*I, go (I)*), producing *I try (for me) to go*, and *I try to have it remain* as *try* operating on (*I, remain (it)*).

When the verb-operators involve a morphologically limited adjectivizing or nominalizing of the verb they may be hard to match with a predicate (on that verb) which might be considered their source. In such cases, a search for a predicate form involves us in what seems to be primarily semantic considerations of synonymy. One can propose various predicate variants for certain increments on verbs, but it is hard to know whether the predicates show the same acceptability-ordering on their verbs as the increments do:

The government is repressive ← The government's repression  
is characteristic (or: is a disposition).

Intermediate transforms are: via argument-skipping to

The government has a character of repression

and, by morphophonemic change between the adjectivizers *of* and *-ive*, to

The government has a repressive character.

By the same transformations plus zeroing, we derive

He is (very) frightening to them ← His frightening them has a  
(great) effect (or: impression).

And:

He is an actor ← His acting is occupational

The most difficult search for a predicate source is in the case of verb operators on nominalized verbs: *have a look, give a look, make a trip*, etc. They almost all have the meaning of some bounded segment of activity, as is also seen in their time-duration adverbs: *He walked all day*, † *He had a walk all day*. We would expect to derive these nominalizations from a predicate (or a choice of predicates) that state the bounding, but words that would preserve

acceptability-ordering are hard to come by in this case:

I had a walk. ← My walking was an event (or: was in an amount).

The problem is one of the availability of words. That an action-bounding operator is involved may be seen not only from the duration-adverbs and the like, but also from the fact that those adjectives of *have a walk* which are not in the adverb-selection for *walk* are explicable as adjectives of the action-bounding operator which had been the source of *have a*:

I had a short walk. ← My walking had an extent and the extent was short.

A special case of verb-operators are those on *be* which permit *be* to be then zeroed: *He appears ill.* ← *He is ill.* These can be derived by argument-skipping and morphophonemics from a predicate: *His illness is apparent.* Many of the predicate forms do not occur except as presumable sources for the argument-skipping transformation: e.g. we have to derive *He seems to be ill.* ← \**His being ill seems*; *He became ill.* ← \**His illness became* (though marginally ∃ *His illness came to be*).

### 5.5. Subordinate Conjunctions

The subordinate conjunctions can be readily put into predicate form. Almost all of them appear in the following transforms:

- (1) He left because she arrived:  $S_1CS_2$ .
- (2) He left because of her arrival:  $S_1P_vS_2n$ .
- (3) His leaving was because she arrived.
- (4) His leaving was because of her arrival:  $S_1nV_{vv}S_2n$ .

In (1) we have a conjunction C; in (2) what would be called a preposition  $P_v$ . In (3) we have a form (due to the transformation of 8.1 (11), cf. (3), (4) of 3.2) which does not really fit any classical grammatical classification since, *His leaving was* does not occur as a sentence in English, so that the following *because* is not a conjunction and indeed has no category; this is an example of how transformational analysis treats structures that cannot be analyzed except ad hoc in non-transformational grammar (8.4 (9)). In (4) we have a verb  $V_{vv}$  on two nominalized sentences, i.e. a predicate whose argument is a verb-pair. We can take (4) as the source and derive the other forms by paraphrastic transformations from it. It will then be found that these transformations operate on only a subset of the verb-pair predicates: some  $V_{vv}$  are not transformed into C, e.g. *spoil, be due to*. Thus:

His leaving spoiled her arrival

His leaving was due to her arrival.

These  $V_{vv}$  can be transformed almost up to (2) by a simulator of tense-transplacing. We add  $S_1C$  before (4), and then zero:

- (5) He left, (with) his leaving spoiling her arrival →  
He left, spoiling her arrival.

When the  $V_{vv}$  contains *be*, tense-transplacing can operate, producing the structure of (2):

His leaving was due to her arrival.

→ He left due to her arrival.

When the second operand is changed to *the fact that S* or the like, instead of nominalized S, we obtain such forms as

He left due to the fact that she arrived,

in which *due to the fact that* may be considered a member of the set of conjunctions, existing in the form (1).

These considerations not only show how members of  $V_{vv}$  can transform toward that subset of  $V_{vv}$  which has a transform called subordinate conjunction; they also help to organize the transformations operating on various subdomains of  $V_{vv}$  (like the zeroing of *his leaving* in (5)), as a step toward showing that some conjunctions are variants of other conjunctions, under suitable transformations.

The deriving of one conjunction from another, or rather from other  $V_{vv}$ , may also be supported by restriction-removal of the kind seen in § 2. Thus *whereas* is restricted to sentence-pairs which differ at two points: *I play violin whereas she plays piano* (the specific contrasting is not essential, as in *I play violin whereas she goes to museums*);  $\bar{A}$  *He went whereas I went*. Somewhat differently, *but* is restricted to sentence-pairs which differ at least in their verb-plus-object segment, and preferably at two points: *I play violin but she plays piano*, but also *I went but (I) missed him* ( $\bar{A}$  *I went whereas I missed him*), *I came late but I left early*, *I smoke a lot but I smoke mild cigarettes*;  $\bar{A}$  *He went but I went*. One might think that this double contrast is a need of the concessive meaning. However, *whereas* and *but* can be shown to be free variants of *despite* (in the case of *but*: with an added sentence about what is the expectation of  $S_2$ , given  $S_1$ ). *Despite* does not have this restriction: *He went despite my going*; and even (with special meaning) *He is educated despite his being educated*. Thus the restriction is made only in the paraphrastic transformations which, for slightly different subdomains yield *whereas, but* out of *despite*. It is not a property of the source increment, the predicate *despite*.

### 5.6. Moods

The above survey of increments has shown that at one cost or another, every



type of increment (except *and*, *or*) can be derived from a predicate on sentences. In addition, certain apparent increments have been shown in transformational analysis to be obtained by paraphrastic transformations from other incremental forms. In particular, this is the case for the various grammatical 'moods'. All question forms, both the yes-no form (*Are you going?*) and the *wh*- form (*When are you going? What will you take?* are paraphrastic transformations of *I ask you whether you are going or not. I ask you when you are going* (or: *whether you are going at time A or ... or you are going at time Z*). *I ask you what you will take* (or: *whether you will take A or ... or Z*). The imperative, e.g. (*Please go!*), is derived from *I command* (or: *request*) *you that you (please) go*. The optative *Would that he returned!* is derivable from *I would that he returned, I wish that he would return*. And so on. That is to say, these forms are obtained not by an increment of an intonation (plus some changes) to an existing sentence, but by paraphrastic transformation from known types of predicate operators.

The sources reached in many of these derivations have the additional advantage of being less restricted, in the manner of § 2. For example, the question form does not appear under sentence-operators except those which produce that form. We have (1) *I ask you: Is she reading it?* ← (2) *I ask you whether she is reading it (or not)*, (2) being indeed the source of (1). But  $\bar{A}$  *He told them: Is she reading it?* and  $\bar{A}$  *He is waiting with the book while: Is she reading it?* However, the source form of the question occurs freely under all sentence-operators:  $\exists$  *He told them (that) I ask you whether she is reading it (or not)*,  $\exists$  *He is waiting with the book while I ask you whether she is reading it (or not)*. The paraphrastic transformation to the question form occurs in certain of the operator-environments in which the source form occurs (e.g. under *and*: *I'm reading it and is she reading it?*) but not in all, as above.

### 5.7. *And, Or*

Two incremental operators have been left without any indication of how they could be derived from predicates: *and*, *or*. One might think of replacing *and* operating on  $S_1, S_2$  by a predicate *conjoins* on  $S_1, S_2$ ; and so for *or*. But this may not fit as simply into the grammar as the previous predicate sources.

In any case, one can show that *and* is an operator only on sentences. It is always possible to derive *and* in predicates from *and* or other operators on sentence. And even *and* on nouns can be so derived. This has already been seen, for example, in the reciprocal verbs (4.3); many other cases of  $N_1$  and  $N_2$  can be derived from  $N_1$  with  $N_2$  or from *the set consisting of  $N_1$  and  $N_2$*  ← *the set contains  $N_1$  and the set contains  $N_2$* . A particular problem here is that of the collective verbs  $V_{coll}$ , e.g. *gather*, which are restricted to subjects or objects consisting of *N and N*, or *N and N and ... and N* (at least 3), or instead

to certain collective nouns  $N_{coll}$  such as *set, crowd, group*. Here one can say that, in the source,  $V_{coll}$  selects only (in effect, is restricted to)  $N_{coll}$  as subject or object. The cases with *N and N* would be derived as follows:

$N$  and  $N$  and  $N$  gathered at the corner.<sup>15</sup>  
 ←  $N_{coll}$  which consisted of  $N$  and  $N$  and  $N$  gathered at the corner.  
 ← *wh*- ( $N_{coll}$  gathered at the corner:  $N_{coll}$  consisted of  $N$  and  $N$  and  $N$ ).

The second operand here can be built up as follows:

$N_1$  is a member of  $N_{coll}$  and  $N_2$  is a member of  $N_{coll}$  and  $N_3$  is a member of  $N_{coll}$ . →  $N_1$  and  $N_2$  and  $N_3$  are members of  $N_{coll}$ .

On this there operates *is complete* or *is exhaustive*, yielding:

$N_1$  and  $N_2$  and  $N_3$ 's membership of  $N_{coll}$  is exhaustive.  
 →  $N_1$  and  $N_2$  and  $N_3$  exhaust (the membership of)  $N_{coll}$ .  
 →  $N_{coll}$  consists of  $N_1$  and  $N_2$  and  $N_3$ .

We now ask whether *and* between sentences is not simply a morphophonemic form of period between successive sentences in a discourse. Such an analysis does not seem adequate, first because it leaves the parallel problem of *or*, and secondly because *and* (and *or*) also operates on untensed (and period-less) sentences under all operators. In particular, *or* is essential under *ask, wonder* and similar operators, and *and* is essential in the derivation of the numbers, reciprocal verb, collectives, etc., and in specifying the operand-domain (the scope of operation) of certain zeroings (and of *wh*- plus pronouns). There are other operations which depend not upon *and* but upon the successive (and addressable) periods among sentences of a discourse: These are, above all, pronouns and the more subtle methods of reference<sup>16</sup>, and also discourse structure.

## 6. RESULT

### 6.1. *Summary of Reductions*

We begin with the set of sentences  $\{S\}$ , as given in transformational linguistics, and with the sets of incremental transformations (or operators) and of paraphrastic transformations  $T$  defined on it. On the set of increments, the following grammatical reductions have been proposed in §§ 2-5:

For each individual incremental operator  $A$ , if it is restricted to a subdomain of values of the variables in its argument or in the operators upon it, we seek a free variant for  $A$  which is not restricted in this way. The unrestricted variant is taken as source; the restrictions are on the  $T$  (§ 2).

For each incremental operator  $A$  which is restricted in operator or argu-

ment domain, but does not have an unrestricted free variant, we seek one or more increments (one of which may be zero) each of which occurs with a subdomain of operator or argument which is complementary to that of the others and to that of A, such that the sum of the subdomains is the unrestricted domain of the variables in question. If we find increments which are only partly complementary, we try to make them completely complementary by assuming a zeroable element to occur in certain environments of one of the increments. An unrestricted morphophonemic source sentence (\*S) is formed for the complementary variants (§ 3).

If a paraphrastic transformation T is restricted to a particular subdomain A' of its operand A, we try to assume a zeroable operator Z (possibly a meta-linguistic one) which A' can accept on selectional (co-occurrence) grounds, such that T can operate on Z and thereby operate on those occurrences of A' which are under Z. E.g., T may be a free variant of Z (§ 4). This is discussed here because it affects the form of the source increment.

We choose a particular sentence-position in respect to arguments, such that, for each incremental operator which does not occur in that position in respect to its argument (when it forms a sentence with its argument) we seek a free variant – if necessary, a metasentence one – that has this position; and if none is available we construct a morphophonemic source (\*S) which does. For English and many other languages, this sentence position is the 'predicate' position, namely that of *V*, *be N(P)*, *be A(P)*, *be P* (§ 5).

## 6.2. An Unrestricted Subset of Sentences

Each of these steps can be carried out over all the transformations of the language. The first three steps produce free or complementary (automatic) variants which are unrestricted, and which can be taken as sources of the restricted ones. The last step gives to all increments a single-source sentence-position. The recognizing or forming of these source-sentences is of special interest because of the following construction:

First, we form the set  $\{S\}^+$  by adding to  $\{S\}$  all the morphophonemic source-sentences \*S introduced by the steps of §§ 2-5 or the conventional methods of structural and transformational linguistics. Then  $\{S\}^+$  contains, for example, the tenseless \**He go*, as well as *He goes*. Second, we form a set  $\{I\}$  consisting of all the source-sentences (asterisked or not) established in the above steps, plus all those other sentences of  $\{S\}^+$  (again, whether asterisked morphophonemic sources or actual sentences) which contain only increments, and are not paraphrastic transforms. These will be unrestricted increments, since the restricted ones will have been treated in §§ 2-5. Then  $\{I\}$  is a restrictionless subset of  $\{S\}^+$ , and is obtained by removing from the sentences of  $\{S\}^+$  all effects (traces) of T, leaving each residual sentence to

consist only of the arguments and incremental operators specified in 7.2.

We now consider the relation of  $\{I\}$  to  $\{S-I\}$ , its complement in  $\{S\}^+$ . The sentences in  $\{S-I\}$  are products of increments and paraphrastic transformations T, while those of  $\{I\}$  contain no T-traces. If for each sequence of increments and primitive arguments which constitutes the non-paraphrastic material of some sentence in  $\{S-I\}$  there exists in  $\{I\}$  a sentence composed of precisely that sequence, then it follows that for each sentence A of  $\{S-I\}$  there is some sentence A' of  $\{I\}$  such that A is a paraphrastic transform of A', and is derivable from A' or from the same increment-sequence as in A', by certain T.

To show this, it is sufficient that, for all pairs of an increment and a T, if there exists a sentence

$$S_i = \text{increment}_i \text{ on } T_j \text{ on } S_k.$$

then there must exist a sentence differing only in the  $T_j$ -trace<sup>17</sup>

$$S_h = \text{increment}_i \text{ on } S_k.$$

That is, there should be no case of a T enabling an increment to operate on a sentence (or more precisely on an increment-sequence) on which it would not operate otherwise. This demand seems to be satisfied in English and in other languages, as has been sketched in §§ 2-5. In contrast, there are many cases in which a T prevents an increment from operating on an increment-sequence on which it otherwise can operate. E.g. if

- (1) He is writing

is derived by zeroing from

- (2) He is in process of writing,

then we have nominalizing sentence-operators on (2), as in (2') *His being in the process of writing is (quite) frequent*, but not on the resultant of zeroing, i.e. not on (1): Hence (1')  $\nexists$  *His being writing is (quite) frequent*. Also: The question does not occur under sentence-operators, as in  $\nexists$  *They suspect is he going* (5.6), but the source of the question does: *They suspect that I ask you whether he is going or not*, or even (though less certainly) *They suspect that I ask you: is he going* or: *They suspect that I ask you: Is he going?* The increment *They suspect* does not operate on the zeroing of *I ask you*, or equivalently the zeroing of *I ask you* does not operate (is not addressable) under the increment *They suspect*.

For every  $S_i, S_h$  pair as above, the question arises whether the  $T_j$  component of  $S_i$  must be ordered in among the incremental components, or whether we can formulate the T components in such a way that they would

operate after all the incremental components. In other words, can we always obtain  $S_i$  by  $T_j$  (possibly reformulated) operating on the corresponding  $\{I\}$ -sentence  $S_n$ ?

Thus, in investigating the relation between  $\{I\}$  and  $\{S-I\}$ , we now consider the placing of the T in respect to the increments. One arrangement would have each T in a sentence operate as soon as its operand has been formed. That is to say, if a T which is defined on arguments X, Y has operated in the construction of a sentence, then it must have operated before any further increments on X, Y. In such an analysis, the operator representation of the sentence requires no special addressing in the argument of the T. Thus, the two (ambiguous) analyses of

I dislike his speaking because she can't.

would be

(3) dislike (I,  $T_z$ (because (speak (he), not (can speak (she))))))

in the sense of *His speaking because she can't is disliked by me* and

(3')  $T_z$ (because (dislike (I, speak (he)), not (can (speak (she))))))

in the sense of *His speaking is disliked by me because she can't speak*. In each case the  $T_z$  zeroes the word which is repeated (in a position corresponding to its antecedent) under the immediate operand of  $T_z$ . Of course, the increment *I dislike* could also have operated without the  $T_z$ . In both analyses we would have

I dislike his speaking because she cannot speak.

which in the sense of (3) is:

(4) dislike (I, because (speak (he), not (can (speak (she))))))

and in the sense of (3') is:

(4') because (dislike (I, speak (he)), not (can (speak (she)))).

In (3) *dislike* has operated on  $T_z$ , whereas in (4) it did not. In some cases, an increment does not operate on a particular T. Such an increment is found only on the T-less operand, as in (1), (2) above:

for (2'): be frequent (be in process (write (he))).

for (1'):  $\bar{\lambda}$  be frequent ( $T_z$  (be in process (write (he)))).

An alternative arrangement to the above is that all T in a sentence operate after all increments have operated. This is a more complex description, for it requires the arguments of each T to be provided with addresses in the sen-

tence, indicating on what segments or operators in the sentence the given T is operating. Thus, in this decomposition, (3) above would appear as:

$T_z$ (because) (dislike (I, because (speak (he), not (can (speak (she)))))).

In both of these types of T-placing some addressing may in some cases be needed to indicate what word in the argument is zeroed. In the second type, however, the restrictions would be due not to an increment refusing to operate on a T (as for (1') above) but rather to the T refusing to operate on (i.e. be addressable to) the operands of certain increments. Thus we would have, for (1'):

$\bar{\lambda}$   $T_z$ (in process) (be frequent (be in process (write (he))))

### 6.3. A Sublanguage for Objective Information

The result indicated above, that  $\{S-I\}$  contains only paraphrases of sentences of  $\{I\}$ , gives a new importance to the traditional linguistic search for removing restrictions. Modern linguistic analysis, whether descriptive, structural or transformational, has always sought to remove restrictions. Of course, the restrictions could not really be removed, but only moved: For example, in phonology, the great number of sounds that are heard in a language are organized into just a few phonemes; but this is at a cost, for whereas there was previously no such distinction as one between an element and its pronunciation (or its gross acoustic shape in the world of physical events), we now have to say that each phoneme may have more than one pronunciation, according to its environment. The total number of sounds remains naturally the same, but the number of statements necessary to describe them all may be greatly reduced, and the writing far simplified.

In the present case, the fact that  $\{I\}$  contains all the increment-combinations, which in interpretation means roughly all the objective information, that can be carried in  $\{S\}^+$  gives a particular importance to shifting all subdomain restrictions out of  $\{I\}$ : it means that the subdomain restrictions are not necessary for expressing the objective information carried by language. That is to say, one can carry all the objective information of language in a system containing no such restrictions. Moving the restrictions is therefore in this case not merely a matter of structural compactness or elegance, but a gain for the interpretation and for the utilization of language-information, because the restrictions have here been moved out of a distinguished part of  $\{S\}^+$ , leaving that distinguished part as a far simpler system which is nevertheless capable of doing all the objective informational work.

This completes the restriction-removal activity of sentence (but not discourse) grammar, in so far as applies to the existing vocabulary of the language. (A step beyond this will be indicated in §9.) This bringing to completion

creates an effect which the various partial restriction-removals could not produce: Metalinguistically, it brings out the semantic interpretation of syntax: one can now see that each syntactic element (variable, operator-type, etc.) enters into combinations in a way that is directly interpretable by the meanings they carry. Linguistically,  $\{I\}$  is a system in which (within the limitations of how information is represented by morphemes) syntactic form and objective information correlate well: every two different pieces of such information are expressed by different sentences; but it is not at all excluded that, even when paraphrastic transformations are eliminated, the complex semantic overlappings in the vocabulary would permit two different sentences to carry the same information.

In studying the relation of  $\{I\}$  to  $\{S\}^+$ , the major question that remains at this point is to what extent the sentences in  $\{I\}$ , which are the source-sentences from which the sentences of  $\{S-I\}$  are derived, are themselves sentences of the original set  $\{S\}$  (and hence free variants of the sentences of  $\{S-I\}$ ), and to what extent, and in what way, they are asterisked morphophonemic source-sentences, formed for  $\{S\}^+$  but not extant in  $\{S\}$ : From the sentences in  $\{I\}$  which are not asterisked, those sentences of  $\{S-I\}$  whose source they are are derived by T. These T also derive some sentences of  $\{S-I\}$  from other ones. If the sentences in  $\{S-I\}$  for which asterisked sources have been formed in  $\{I\}$  differ from these asterisked sources (i.e. are derivable from them) by no more than these same T, it would mean that all sentences of  $\{S-I\}$  (i.e. all the remaining sentences of the language) are derived from all sources in  $\{I\}$  only by T. This would hold whether the source sentences in  $\{I\}$  are real or asterisked. In English, all sentences in  $\{I\}$  are of the form  $*S$  because they lack the required tensing (and in some cases also plural) morphophonemics. Nevertheless, in the case of English and many other languages, such asterisked  $*S$  sources as have to be formed in  $\{I\}$  satisfy the condition of differing from S only by established T.

When this condition is satisfied, then  $\{S\}^+$  is only a slight extension of the set of real sentences  $\{S\}$ , for it involves no extension of the set T which derive all sentences from others, except for adding new morphophonemic changes to the list of morphophonemic T. In fact it is only a regularization of the domain of T, i.e. of the T-relation among sentences. It is then worth calling  $\{S\}^+$  an extended natural language, and  $\{I\}$  (for all its asterisked members) a subset of that language. It may, however, happen in some languages that some of the asterisked sources differ from the sentences in  $\{S-I\}$  (that are to be derived from them) by  $T^*$ , where  $T^*$  are paraphrastic operations which differ from T or which include T as only a proper part. Then  $\{I\}$ , which contains these asterisked sources, can no longer be reasonably viewed as a subset of what we would consider a (slightly extended) natural language. Rather,

$\{I\}$  would have to be considered a projection of  $\{S-I\}$ . This situation might be the case in some languages for the fourth step above (§ 5), when we seek a single positional relation that all increments have to the sentences onto which they are added. If such a relation is not attainable with any morphophonemics that would fit into the given language, it may still be attainable in principle, as a purely syntactic rather than also morphemic relation. In that case, the purely syntactic form could be simply the operator notation, which identifies each increment as an operator on a particular argument.

We now begin with  $\{S\}^+$  as an empirically given set of sentences with the addition of asterisked source-sentences. Transformational theory shows that it contains a base set K of elementary sentences ('sentences of the kernel') from which all other sentences are derived, and two sets of transformations, the incremental and the paraphrastic, by which the sentences of  $\{S\}^+$  are derived from other members of  $\{S\}^+$ . The empirical set  $\{S\}^+$  is closed under these transformations, and it is as such that we call  $\{S\}^+$  a language.  $\{S\}^+$  is also closed under T above, taking  $\{I\}$  as the base set. When  $\{I\}$  is a subset of  $\{S\}^+$ , it is a sublanguage of the language  $\{S\}^+$ . This is so because the operators in  $\{I\}$  are the increments, which derive members of  $\{I\}$  from other members of  $\{I\}$  and ultimately from the base set of elementary sentences (K or  $*K$ , i.e. morphophonemic sources of K), which is contained in  $\{I\}$ .  $\{I\}$  is closed under the incremental operators.

Although both  $\{I\}$  and its complement  $\{S-I\}$  are subsets of  $\{S\}^+$ , the structures (grammars) of each of these are not subgrammars of that of  $\{S\}^+$ . The subsets have been separated on the basis of grammatical properties: presence or absence of T, and hence of all the environmental ('distributional') and string properties (§ 8) that T bring in. Therefore the grammars of  $\{I\}$  and of  $\{S-I\}$  each contain relations which are lacking in the grammar of  $\{S\}^+$  (or in the grammar of the set of real sentences  $\{S\}$ ). As will be seen below, they are each simpler and more transparent systems than is the grammar of the whole language, and have interpretations which do not naturally arise from the grammar of the whole language.

## 7. THE PREDICATE SYSTEM

### 7.0. The Metalanguage

In considering the structure of  $\{I\}$ , we note first that the removal of restrictions and the environment-filling method of § 4 eliminate most of the meta-grammar needed for stating the grammar of  $\{I\}$ . Each increment can be provided with an indication (a subscript) of its argument variables, and every sequence consisting of increments operating on variables that match their arguments is well formed, i.e. is in  $\{I\}$ .

### 7.1. The Base

When the increments are viewed, like the T, as transformations from one subset of sentences (those satisfying the argument requirements) to another (those containing the increment as their latest – i.e. not operated on – operator), then we have to take the elementary ("kernel") sentences or their morphophonemic sources as a base set K or \*K within {I}; from K or \*K, the sentences of {I} are generated by the increments. However, we have seen that all increments except *and*, *or* form their resultant sentences by serving as the verbal segment in respect to their argument. It therefore becomes convenient to look upon the elementary sentences as being themselves the resultants of an operator (their verbal segment) acting upon their primitive arguments (their subject and object nouns). Then the operators of {I} are not only the increments but also the verbal segments of the elementary sentences. The primitive operands are the subjects and objects of the elementary sentences. The operators are now no longer transformations from sentences to sentences, but operations which produce a sentence when carried out on a sentence or non-sentence operand. Indeed, if we can show that the increments affect only the most recently operating (i.e. not already operated on) verbal segment of their operand, we can define the argument of each increment to be certain operators rather than a whole sentence. This analysis presents, of course, a purely syntactic relation; the physical contents – what morphemes in what positions – will be considered in 7.4, 5.

### 7.2. The Operators

For English the primitive arguments are primarily the concrete N, and the main operator sets are:

$V_n$ : e.g. *exist*, *be tall*, *be a mammal*, *be up*.

$V_{nn}$ : e.g. *eat*, *be father of*, *be near (to)*.

$V_{nnn}$ : e.g. *be between ... and*; possibly *give ... to*.

$V_v$ : e.g. *be a fact*, *continue*. Thus *His smoking continues* would be: *continue (smoke (he))*.

$V_{nv}$ : e.g. *think*. In almost all arguments which contain both v and n, the n indicates the subset of nouns which denote living beings: it is only these that can be related to the v by the given operator. Thus *I think he drinks tea* would be: *think (I, drink (he, tea))*.

$V_{nvv}$ : e.g. *tell*.

$V_{vv}$ : e.g. *be because of*, *correlate with*.

$V_{vvv}$ : e.g. *relate ... to*.

$C_{vv}$ : *and*, *or*.

Special subtypes of V with v-arguments carry particular classificatory

morphemes, as in *request-for-afterwards* (4.2), *an-undergoing-for-the-object* (4.4). We may also wish to recognize a V whose argument is a constant, e.g. *rain*, *late* (Whence morphophonemically *It's raining*, *It's late*).

Each operator operates on its arguments to produce a sentence. Those whose arguments include operators (V) are increments. The others form elementary sentences.

The products (sequences) of operators are associative, even though the conjunctive transforms of many  $V_{vv}$  are not. The  $C_{vv}$  and certain of the  $V_{vv}$  are interpretationally commutative in the sense that while  $S_1$  and  $S_2$  is a different sentence than  $S_2$  and  $S_1$  it carries the same objective information. It is not certain that  $C_{vv}$  must be taken as binary rather than n-ary (for  $n > 1$ ) in argument. In constructing a sentence, the operators are apparently linearly ordered. Each can then occur whenever its arguments are present (and not already operated on). The ordering and repeatability possibilities of an operator are determined by its arguments.

To these must be added one operation quite different from these operators. In this operation we pair each occurrence (in a discourse) of an element or sequence of elements A with an address which indicates the position of A in the discourse or in any stated segment of the discourse (e.g. sentence or structurally identified section of a sentence). In the actual form of the sentence and discourse or a sequence of phonemes, morphemes, words, etc., the address of A states that A is the *n*th occurrence, in a given segment of the discourse, of a member of the class of A. In the operator form, if an operator B contains within its argument the address of A, then the address states what position A occupies in the operand of B. This addressing is implicit in the linear ordering of phonemes and words which comprise a discourse, and in the ordering of operators and their arguments, as morphemes, by which we represent the actual sentences and discourses. This addressing is necessary for the meta-discourse operators which are included within the grammar of the language, and for certain paraphrastic T (e.g. zeroing of repeated words in particular relative positions). It is also needed if we try to place all T after all increments (end of 6.2).

Finally: what is produced by the operator is a sentential string, but not quite a sentence because it lacks the sentence-intonation (period) and its attendant morphophonemics (primarily tensing). The placing of the period takes a sentential string out of {I} for then it can no longer be operated on by the operators of {I}, or by the T-operators of {S-I}. Similarly for sentential strings produced in {S-I}.

### 7.3. The Values

The arguments and operators in formulas are class-marks, i.e. variables

over certain domains. Each value  $A_i$  of an operator  $A$  imposes a partial ordering of acceptability (or subject-matter classification) on the values of its argument  $B$ ; but  $A$  apparently leaves unaltered the partial orderings in which  $B$  otherwise participates. That is, the partial ordering of acceptability which  $B$  had imposed upon the arguments of  $B$ , and any partial ordering of acceptability which  $B$  has in respect to any other operator on  $B$  (if  $B$  can be an argument of more than one operator – see immediately below), are invariant under  $A$ 's operating on  $B$ .

In what variable (or variables) of its operand does each value of an operator impose a partial ordering of acceptability on the values of the variable? This variable in the operand would be the effective argument of the operator. The question is important for the fundamental problem of operator-syntax, namely: what is the relevant argument of each operator? And the main, and still partly open, question in whether the increments are linearly or partially ordered is a matter of how increments relate to each other in respect to the arguments on which they impose an acceptability-ordering.

#### 7.4. Morphemic Shape

For the most part, each operator introduces into the sentence one morpheme, and of only the predicate morpheme classes:  $V$ ,  $be A(P)$ ,  $be N(P)$ ,  $be P$ ; aside from *and*, *or*. The *be* is clearly morphophonemic and not an independent morpheme. The operator  $N$  are e.g., classifiers (*mammal*, *fact*) and relations (*father*), and are not the same morphemes as in the concrete  $N$  class. There are, however, many cases in which the demands of simple and consistent syntactic relations make us assume a morphologically complex form as the syntactic source, as when we take, say, *His acting is an occupation*, or the like, as source (via argument-skipping) for *He is an actor*. Where there is a conflict between syntactic regularity and morphological regularity, we see that the operator system is a syntactic one.

The operator system is morphologically convenient in another way too. To a large extent, each morpheme is a value of only one operator: thus the *-ing* of *be-ing* (*He is running*), as also the *-ing* of verb-operators (*He continued running*), are derived from the *-ing* of sentence nominalization (*His running continued*).

#### 7.5. String Shape

Although every sentence and discourse is a sequence of morphemes, the only way in which grammatical descriptions have succeeded in characterizing which sequences are those that constitute sentences has been by finding among the morphemes of the sentence certain relations which are not immediately apparent from their positions. In transformational theory, and espe-

cially in the further reduction of §§ 2–5, this relation is that of operator to argument. A sentence can be fully represented by a linear or partial ordering of operators, each operating on a prior one, with the innermost operating on primitive arguments. The argument of each operator can be written in parentheses. The sentences of a natural language, however, do not use parentheses and do not directly indicate the operator-argument relation. The possibility of representing in a string of words the same thing that the operator notation represents, is achieved in language by means of an apparatus which can best be described in two stages: In  $\{I\}$  it is done by means of a classification of morphemes in respect to their known arguments, and by means of a relative positioning of operator and arguments (see below); these together give the morphemes of  $\{I\}$ -sentences a 'string' relation to each other: what is subject, predicate, object, etc. In  $\{S-I\}$  it is achieved as follows: when  $T$  produces  $B$  out of  $A$ , it makes particular changes in the relative positions of the morphemes of  $A$ , and may give them the appearance of different operator-argument relations. Thus these morphemes appear in changed string relations to each other in  $B$ ; the relative positioning is important because on it are defined certain prominent morphophonemics such as the location of tensing or the agreement of verb with subject in respect to plural. In many of these changes, the participating morphemes receive particular constants (usually affixes) marking the change. We can say that the string relation of  $\{I\}$  are produced by the operators of  $\{I\}$ , and the string changes (to new string relations) in  $\{S-I\}$  are produced by  $T$ .

Given the transformational representation of a sentence, i.e. the operator relations among the morphemes in it, and the  $T$  if any, we can determine the positions and string-relation constants of the morphemes, i.e. we can determine the actual word-sequence that is the sentence. And conversely: Given the word sequence and a grammatical lexicon, i.e. a list of which morphemes belong to which operator (or argument) classification, we can determine (in some cases, ambiguously) the string relations in this sequence and thence the operator relations and the  $T$ .

Here we consider the string shape of the  $\{I\}$  operators in English; the string changes in  $\{S-I\}$  are noted in 8.4. A major advantage in having found (§ 5) a single positional relation in respect to the argument, which the sources of almost all increments would have, is that this gives a single string structure to all source sentences, i.e. to  $\{I\}$ . It shows that a single positional (syntactic) relation suffices for all the operators of  $\{I\}$ . As among these, *and* and *or* differ from all the other operators (the predicates) only in morphophonemic respects, below. Such string complexity as there is in  $\{I\}$  is due to the nested superposing of the syntactic relations of those elements which are in the given sentence both operators and arguments. In  $\{I\}$ ,



the ordered morphemes of an operator A concatenate with its sequence of arguments B, in such a way that the operator morphemes occupy even-numbered positions, and the ordered arguments the odd-numbered positions, in the resultant sequence:

$$A_1 A_2 \dots A_n (B_1, B_2, \dots, B_n B_{n+1}) = B_1 \widehat{A_1} B_2 \widehat{A_2} \dots B_n \widehat{A_n} B_{n+1}.$$

The interpretation, both in the operator notation and in the word-concatenation, is that the  $B_1$  is the subject of  $A_1$  (as predicate) and  $B_2$  the direct object, and the remaining B what may be called indirect objects. Under a predicate operator (but not *and*, *or*), when one of the arguments  $B_p$  is itself a predicate, there is attached to  $B_p$  a constant which depends upon the first morpheme of the operator. This constant indicates that the predicate is also an argument; the main constant is *-ing* or other nominalizing affix (zero under the operator *make*, etc.). If an argument B itself has arguments C, these stay concatenated to B as a single entity in respect to A; and if B receives a nominalizing constant then C receives an adjectivizing constant, primarily: *'s* (mostly for first operand under B), *by* (only for first operand under B), *of* (rarely on the first operand if it occurs on the second operand under B). The choice of constants is free, or depends on C. Thus when B becomes an operand of A, then C receives constants which mark the relation it had to B before B had become an operand. The constants (and positions) indicate the operator-argument relations that had existed in the sentence before A operated on it.

The string structures in {I} are therefore such as the following:

- V(N): exist (gas) = \*gas exist → Gas exists.  
 V(N, N): between and (N, L, M) = \*N between L and  
 M → N is between L and M.  
 V<sub>0</sub>(N, V<sub>1</sub>, V<sub>2</sub>): relate to (I, grow, increase) = I relate... growth  
 to... increase.

Here the arguments of the two operand-verbs  $V_1$ ,  $V_2$ , were not shown, in order to bring out the positional relation of the new operator  $V_0$  to its immediate arguments. If we take the resultant of the above with

- $V_1$ (N): grow (tree) = \*Tree grow,  
 $V_2$ (N): increase (rain) = \*Rain increase,

we obtain

- relate to (I, grow (tree), increase (rain)) =  
 I relate the tree's growth to the rain's increase.

The *-th* and *'s* are parts neither of the operator nor of the argument. They are constants indicating the superposition of syntactic statuses: *-th* nominalization of  $V_1$  under  $V_0$ ; *'s* adjectivization of subjects when their verb becomes

an argument. Similarly, the *-s* and *is* above are neither in the operator nor in the argument, but are morphophonemic effects. See also 8.5 (1), (2).

In the case of all operators but *and*, *or* the words entering the second position in the resultant string are of the predicate morpheme-classes (7.4) and can undergo certain T (argument-skipping, certain zeroings, placing of the morphophonemic tense). In the case of *and*, *or* the morphological class is different and the T which can operate on them are fewer and different (different zeroings and permutations).

The fact that there is one position, relative to all arguments, which is taken by all operators, and the further fact that the operators belong to morphologically recognizable classes of known arguments, makes the word strings which are sentences of {I} readily computable. That is to say, it makes their operator-argument (i.e. syntactic) relations recognizable to the hearer word by word, without waiting for parenthesis-closings. (It will be seen in § 8 that the T bring considerable complexity, and even degeneracies, into the recognizability of the operator-argument relations of the words, but without destroying computability.) To the extent that different morphemes occur as values of the different variables, the operator-argument relations are more immediately recognizable. To the extent that certain morphemes (or merely phoneme-sequences) occur as values of more than one class-variable (e.g. of both  $V_{nn}$  and  $V_{nv}$ ) the operator-argument relations are less transparent in {I}, and in {S-I} ambiguities may result.

### 7.6. Relation to Linguistic Analysis

In the {I} sublanguage more can be said than had been thought sayable in {S}, because regular sources are discovered whose existence had been put into shadow by the more compact but more restricted variants. Such restrictions as exist in {I} are more like the semantically reasonable and changeable restrictions of selection (co-occurrence) than like the grammatical restrictions of unbreakable 'rules'.

The meanings of words and of constructions are seen more sharply than in the language as a whole. For example, when we see the source of the comparative, which is e.g. for *He is richer than she*:

- n is more than m.  
 His riches amount to n.  
 Her riches amount to m,

we see that indeed the comparative does not include *He is rich* as an immediate syntactic source (but only *N is the amount of his riches*), just as it does not include the meaning of that sentence in its direct meaning<sup>18</sup>.

The approach in descriptive linguistics had been to sweep everything



under the carpet, i.e. to obtain regular grammar at the cost of irregular source. Thus the subjunctive could be presented as what occurs after verbs that require the subjunctive; and then the non-subjunctive form *I request his going* is obtained from the subjunctive by a transformation. In the present work, however, we want the source to be as regular as possible, since we want its structure to contain no complexity other than what is required by its informational burden. The only cost that we can admit is in the morphophonemic complexity of the source; for this is unrelated to the objective information, though it complicates the subjective task of coding and decoding. Thus in the case of the subjunctive we would first note that this can be viewed in English as a morphophonemic change from a tenseless (and non-subjunctive) form; then in considering the verbs which require or permit this morphophonemic change, we would try to characterize them not by a list but by the fact that they selectionally always have *subsequently* or the like in their environment (or, indeed, as part of them) in such a way that (as in § 4) the subjunctive form would be a free variant replacement or a free or required accompanying variant of that accompanying word.

As the operators become unrestricted, and as subset listings are replaced by the presence of classifier morphemes in the environment, we begin to have elements whose properties are no longer inherent but due only to their positional relations and whose combinations (i.e. positional relations) are sufficiently regular and simple to permit of something approaching a mathematical characterization. It also becomes possible to find relations of relations, such as elements whose positional relations are the inverse of each other's: e.g. *before* and *after* are inverses of each other when they are taken as  $V_{vv}$  (or as  $V_{nn}$  on two time-nouns) but not when they are taken as conjunctions.

## 8. THE MORPHOPHONEMIC SYSTEM

### 8.0. Introduction

The paraphrastic transformations T, which operate on the sentences of {I} and of {S-I} to produce the sentences of {S-I}, can be considered an extension of morphophonemics. Morphophonemics is the change of the phonemic shape of a morpheme; if automatic, the change appears when the morpheme occurs in the environment of particular other morphemes. The T include such changes (from the source morphemic shape, 7.4); but they also include changes in the positional relations of the morphemes (from those of 7.5). As in morphophonemics, the original syntactic relations of the source sentence are not lost, although secondary syntactic relations are superposed (8.5); and there are various other changes: e.g. at what point of the sentence the hearer is apprised of particular syntactic elements (8.2). All

this applies also to morphophonemics, though in a much simpler way. Hence all T can be considered to be morphophonemic, or perhaps to constitute a corresponding relation of syntaxomorphemics: the change of the morphemic (and ultimately phonemic) shape of a syntactic element or sequence. The change occurs freely in all or in certain syntactic (i.e., now, operator) environments, or requiredly (automatically) in certain syntactic environments.

The difference between the increment system and the paraphrastic T system is roughly that between the directly useable activities of life and the institutional apparatus which channelizes these activities. Like social institutions, the T system structures, facilitates, slants, and petrifies the activities-for-use of the {I} system, and is inflexible, conventional, and in part historically accidental; and in some cases it stands in the way of further development of the use-activities, i.e. of the directly meaningful expressions of {I}. Indeed, the relations between {I} and {S-I} can be studied as a very special case of the relations between activities-for-use and their institutionalization.

### 8.1. The Morphemic and String Changes

In the case of English, the changes which are introduced by each T can be seen from the list of established paraphrastic transformations plus the further cases of T required by the sources proposed above (§§ 2-5).

The (1) zeroing and (2) pronouncing T replace certain morphemes or sequences by zero or pronominal variant of them, usually at the same site in the sentence (though some pronominal variants occur at a stated other site: e.g. the ones after *wh-*). Zeroing can affect the string relations of the remaining words. Thus, certain expected operands are zeroable, e.g. *arrive*, *come* under *expect*, yielding *I expect him to come* → *I expect him*. In the resultant, *him* is the object of *expect*, so that *expect* now occurs with a noun as object as well as with a sentence as object. Note that in *I expect him to come*, *him* is not the object of *expect*, but the subject of the object (*him to come*) of *expect*.

(3) The T which are purely morphophonemic in the ordinary sense of the word replace certain morphemes (not simply phonemes, although this is a matter of definition) by others, and not always at the same site.

(4) An extreme case is the one which produces a single morpheme out of a specially structured sequence: e.g. *the* from *wh-* plus certain second operands (5.1); *wh-* from *and... and... same...* (2.8, 8.2).

The permutational T change the sites of words in the sentence, i.e. their relative positions.

(5) Some permutations do not change the receptivity of the words to other morphophonemics (e.g. tensing and adjunction), and what is subject and object: *So he says* ← *He says so*. Such is the moving of verb-adjuncts and

sentence-adjuncts (8.4) to before the verb or the subject: *He hesitantly drove the truck*, *Hesitantly he drove the truck*.

(6) A major set of these are the length-permutations, which act on post-verb syntactic entities to put short ones before long ones: *They looked up the number*, but *They looked it up*.

(7) There are some restricted permutational changes which affect in unique ways the subject-object relations of their operand sentence: the *It*-form (*That he came is certain* → *It is certain that he came*), and the *There*-form (*There is a man coming*).

(8) Other T change the words' string relations to each other, as in the interchange of two operands (e.g. Passive), or

(9) in the argument-skipping T which shifts the operator to a lower argument:  $A(B(C)) \rightarrow A(C)$ , as in *His driving is hesitant* → *He is hesitant in driving*. Here the *hesitant* which had been the predicate of the sentence \**He drives* appears as the predicate of the noun *He*.

(10) There is a widespread tense-transplacing T which shifts the predicate status (the location of tensing) from the latest operator to its argument, as in *His death was in March* → *He died in March*. There are reasons for taking the former as the source; but in any case two sentences would have to be related by a transformation, whose effect is to move the predicate status from one to the other.

(11) And there is a free variant in which the morphophonemic operation of tensing (or non-tensing in the subjunctive 2.4) together with a constant such as *that* occurs on the V-arguments of many operators: *I deny his going in the past* → *I deny that he went*.

(12) On the second argument-verb under certain operators (see 3.2, (3)–(5)) tensing occurs as free variant (with no constant).

## 8.2. The Interpretation of T

In structural linguistics, morphophonemic changes do not change the meaning of a morpheme, since the meaning of morphemes does not depend on its phonemes, which is all that morphophonemics changes. But there can result a change as to where in the sentence the hearer is apprised of the presence of a particular morpheme. Thus in a question beginning *Will...* the hearer knows nothing of the subject until he hears it, but in a question beginning *Am...* the hearer knows beforehand that the subject is *I*.

The morphemic and string changes of the type of 8.1 produce various (or no) modifications of the information in their operand sentences, but always without altering the objective information carried by their base sentences in {I}. Pronouncing and zeroing neither destroy nor add information; they are carried out in such a way that the dropped morphemes (or their equivalents)

can be reconstructed from the residue. However, they can give an indication, earlier than in the source, of the identity of individual reference of certain words: *I saw a boy and spoke with him* ← *I saw a boy and spoke with a boy and the two boys were the same*. And they can lead to degeneracies (ambiguities): *I met John, and Mary too*. ← *I met John and I met Mary too*, or ← *I met John, and Mary too met John*. In many cases they do little more than abbreviate the sentence: *He explained this and left* ← *He explained this and he left*.

Many zeroings, and certain other T, occur when a particular argument is the expected one for the given operator (or vice versa)—expected in the culture as a whole, or in a particular subject matter, or in a particular conversation, etc. Thus from the source operator *be in A manner*, the *manner* is zeroable if the particular A is one which is more expectably an adjective of manner than some other adjective, e.g. of occurrence (in 2.3). Hence it is zeroable in *He spoke in a hesitant manner* → *He spoke hesitantly*, but not in *He arrived in an unexpected manner* → *He arrived unexpectedly*. This and some other changes, such as the zeroing of *to arrive* under *expect*, and the conditions for *but*, make expectability a property in {S-I}, though not in {I}.

In transformations which are restricted to a subset of the domain, the resultant receives a more restricted meaning, expressing the effect of the increment upon this more limited environment. Thus, whereas the source *be before* simply means 'prior', the time-conjunction *before* has rather the meaning of 'prior within the same subjective time-segmentation', since it only relates two past events to each other, or two future events.

In some T, especially such as bring a word out to the beginning of a sentence, there is an attention-directing effect upon the meaning of a sentence: e.g. *This in particular I like* ← *I like this in particular*, and the passive, etc. In some T, this device is less a matter of directing attention to a particular word in the sentence as of directing to that word the argument status in some further operator: e.g. *I have a book which you want* is most directly derivable from *wh-* on the pair *I have a book, a book you want*, which is derivable from something like  $n-2$ , *I have a book and, n-1, you want a book and, n, book of n-1 is same individual as book of n-2*. (The  $n$  are addresses in the discourse.)

If we consider carefully the meaning effect of all T, we see that they do not alter the objective report that is carried in their operand, and ultimately in the {I} sentence to which they relate; this is hardly surprising since by definition they do not alter the acceptability-ordering of the word  $n$ -tuples. However, while some of the T are indeed simple paraphrases with no appreciable change in nuance, others add discernible nuances which can perhaps be best summarized as subjective, for the hearer or for the speaker. Some of these simply affect when in the course of hearing the sentence the hearer is

apprised of a given content in the sentence. Some affect what the hearer learns from a sentence by bringing in local ambiguity (which is resolved later on in the sentence) or unresolved ambiguity. Some T involve the criterion of what is considered an expectable operator or arguments at a given point in the sentence construction. Others give special syntactic status to the subjective tense, i.e. the relation between when the sentence is said and when what it says occurs. Others call attention to a particular segment of the sentence. There are various kinds of such effects which are due to the T, and the matter requires investigation. It is in any case clear that we can distinguish, on syntactic grounds, three interpretationally-different properties: objective report, subjective discrimination, and pure paraphrase (including paraphrase for convenience of discourse, i.e. of relation to neighboring sentences). Such a distinction in the interpretation of language might not have been thought of without the evidence from transformations. But this fact need not surprise us, since the history of structural linguistics has shown that precise grammatical analysis throws light on the meanings borne in language. And indeed a test of a syntactic method is whether it makes predictable the meanings of sentences, or the modifiability of their meanings — in short, whether it contributes to the elimination of meaning as an independent primitive of linguistics.

### 8.3. Creation of Subsets

The effect of the T upon the sentences of {I} yields certain new features in the grammar of the whole language. One of these is the creation of grammatical subsets of the major word or morpheme classes. In {I} there are virtually no restrictions in the domain of operations, and virtually no subclasses. Such subclasses as there are, are a matter of changeable selection rather than fixed grammatical subclasses: e.g. the living-being or human-like subjects of many increments whose arguments contain both N and V (*know*, etc.). The non-existence of closed subclasses remains unchanged for those T which do not apply to a restricted subdomain. However, those T which are restricted are defined in respect to a particular subdomain. The defining institutionalizes the subdomain and therefore freezes it, creating an unchanging grammatical fact. This holds even for the T-restrictions which are based on {I}-selections. Thus in {I} certain adjectival increments more naturally select *occurrence* than *manner* (2.3); but in T, the zeroing of *in manner* (as against zeroing *in occurrence*) fixes certain *Aly* as being adverbs of manner: *hesitantly*, but not *unexpectedly*.

In some cases, the net effect of several morphophonemic T is to create a subset of operators which are distinct in terms of the overt grammar: e.g. the auxiliaries *can*, *may*, etc.

### 8.4. String Change

It was seen in 8.1 that the T change the morphemic shapes or relative positions of syntactic entities (or both). The result is that new kinds of string relations are created, which were not found in {I}.

(1) Some T (8.1 (5), (6)) produce resultants in which the relative positions of the entities of their operand sentence are changed, without altering the interpretation of the concatenation in 7.5, or the status of the entities as arguments of further T. Thus in *This I say*. ← *I say this*., the *I* remains the subject, and the *say* remains the location for tensing.

(2) Some T produce unique string relations not found in any other sentence-form: the not quite appositional status of *that S* under *It*, and the status of *It*, *There* (8.1 (7)).

(3) Some T only change the form of a syntactic entity without changing the string relations among the entities. Thus in the zeroing *I play piano and he violin* ← *I play piano and he plays violin*, the second argument under *and* has zero form while its arguments in turn remain, producing the new word-concatenation *he violin* whose operator structure is 0 (*he, violin*); 0 here is for zero.

(4) Some T produce word sequences that are new for a particular environment (i.e. under particular operators) but do not constitute new types of word sequence for the language. Thus the zeroing *to arrive* under *expect* produces, e.g., *I expect him*, where *him* is a new type of object for *expect*, but not for other V of the language; hence the change has merely shifted *expect* from being only in  $V_{nv}$  to being also in  $V_{nn}$ .

(5) Many T produce a new string relation, which may be called the host-adjunct relation. When we zero *He left and he returned* → *He left and returned* we obtain, as above, a new form for second arguments under *and*: in this case an operand whose operand in turn is zero. But when we zero and permute *He left and she left* → *He and she left* we obtain a new syntactic entity *and she*, different from the ones found in {I}. For whereas *He left and she left* is the concatenation of words corresponding (by 7.5) to

and (left (he), left (she))

the word sequence *He and she left* has an interrupting operator (*and*) with an incomplete operand. Interruptions of this kind, headed by *and* or *or*, are found to contain a sub-sequence of the second operand of the *and*, a sub-sequence whose internal string relations (or status) are the same as those of the material immediately preceding the interruption. Hence we call the interruption an adjunct of the immediately preceding syntactic entity. Here fits *wh-* on sentence: *He left, which surprised me* ← *wh-* (*He left, His leaving surprised me*) ← *He left and his leaving surprised me*.

(6) Adjuncts, of noun, also result from the T which produces *wh-* out of *and* (8.2) and permutes it, as in *The book which you want is lost*, where *which you want*, transformed from *and you want the book*, has the new string relation of adjunct to *book* (in *The book is lost*).

(7) Many of these *wh-* adjuncts may have their constants zeroed and are then permuted to before their host; where they become a left-adjunct to the host: *The picture which is small is nicer* → *The small picture is nicer*.

(8) Adjuncts, of verb and sentence, also result from other types of T. The tense-transplacing T (8.1 (10)) which shifts the predicate status (both as interpretation of 7.5 and as location for tensing) from the latest operator to its argument leaves the latest operator without any defined string status in terms of 7.5: Hence in *He died in March*, or *He died quietly* (← *His death was quiet*) the *in March* and *quietly* are called adjuncts of the new predicate *died* (which had also been the previous predicate before *was in March* or *was quiet* had operated). Similarly for this T on  $V_{vv}$ : *His leaving was because of her arrival* → *He left because of her arrival*, where *because of her arrival* is now an adjunct. When both this T and the second-operand tensing (8.1 (12)) operate on  $V_{vv}$ , we have an adjunct consisting of a subordinate conjunction plus its second operand: *He left because she arrived*.

(9) When the second-operand tensing operates on  $V_{vv}$  without the tense-transplacing, we obtain a unique string-relation for the subordinate conjunction plus its second operand: *His leaving was because she arrived*.

(10) Finally, the adjuncts (8) above can be permuted to various points of the remaining sentence, their host: *His leaving was because of her return* → *He left because of her return*; *He left because she returned* → *Because of her return, he left*; *Because she returned, he left*. Similarly: *In March he died*. Again, the permuted segment no longer fits into the string relations introduced in 7.5; it has to be called an adjunct on the residual sentence *He left*.

(11) There is one string relation which results from certain T which can with difficulty be fitted into the relations of 7.5. This is found in the resultant of the argument-skipping T that changes  $A(B(C))$  into  $A(C)$ ; the B which is left over can be considered in some cases as an adjunct of A and in other cases the second argument of A: In *He is hesitant in driving* we can take *in driving* as adjunct of *is hesitant* (whereas in *His driving is hesitant*, we had *is hesitant* as operator on *driving*); In *The franc began deteriorating* we can take *deteriorating* as second argument (object) of *began* (whereas in *The franc's deterioration began* we had *began* as operator on *deterioration*).

(12) A change in the status of certain adjuncts occurs as a result of T which zero the host, in particular situations. This appears in *I like the bigger* ← *I like the bigger one*; *His reading is mostly politics* ← *That which constitutes his reading is mostly politics*; *What he found was valuable* ← *That which he*

*found was valuable*. In such cases the string status of subject or object is taken by an adjective or a segment of certain complex operands.

### 8.5. String Properties under T

In all this we see that in certain cases a T creates in its resultant a relation among the concatenated words (or word sub-sequences) of a sentence that differs from anything present in {I}. In addition to the subject, predicate, and objects of 7.5, we have now adjuncts, which can be attached to one of these or to the string as a whole. These adjuncts are formed by various T out of the latest operator in a sentence, or out of a (permuted)  $V_{vv}$  plus its second operand, or out of (permuted) *and*, *or* plus its second operand (after zeroing). The result of this adjunct-making is to leave the first operand sentence (under a one-sentence or two-sentence operator) as a center string with the structure of 7.5, while the operator (together with its second operand, if it has one) becomes an adjunct of the center or of some part of the center. Many of these adjunct forms are marked by certain constants, in addition to being usually permuted next to their host. In their constants, and in their positions next to their host, they bear some resemblance to the operand of an operand in 7.5. There is also some interpretational similarity: In the string form of  $A(B(C))$  in 7.5., the relation of A to B is unaffected by the presence or nature of C (except perhaps for certain problematic cases). And in the case of adjuncts, the string relations among the non-adjunct entities are unaffected (except for certain problematic cases) by whether any of them are hosts to adjuncts, and of what kind.

Many T also bring in certain innovations in the relations among the string analyses of sentences. One such is the ambiguity exemplified in *I met John, and Mary too* (8.2). Different zeroings and permutations, operating on different sentences, may degenerately produce identical strings of words, although in each case the string of words has string relations (and grammatical meanings) appropriate to the particular operator sequence and T which produced it.

Another such innovation is the fact that the resultants of many T exhibit two or more simultaneous syntactic relations in the meaning of the sentence. This is so because the string relations of the sentence on which the T operated have not been effaced in the resultant. The constant of the T are indicators of a change, rather than simply markers of the string relations in the resultant, hence they indicate the string relations in the operand sentence as well as those in the resultant.

Constants which indicate change have already been seen in the operands of an operand (7.5): given  $V_1(N_1, N_2) = N_1 V_1 N_2$ , we find that when  $V_1$  becomes an operand as in  $V_0(V_1)$  we obtain  $V_0(V_1(N_1, N_2)) = N_1$ 's  $V_1$ ing of

$N_2V_0$ . Thus:

- (1) Students composed jazz.
- (2) Students' composing of jazz continued.

where *Students' composing of jazz* is positionally the subject of *continued*, but within it the constants (and positions) show that *students* is the subject of *compose*, and *jazz* its object. In this case, and throughout the sentences of {I}, no two words show two different string relations to each other.

In the resultants of many T, however, the constants and the meaning of the sentence exhibit both the new and the former string relations among the same words (since no new words except for constants are brought in by T). Thus in the passive of (1),

- (3) Jazz is composed by students,

*jazz* is the subject of *is composed* (by position and by the morphophonemics of the plural subject-verb agreement that is associated with the string form of 7.5); but the constants *be-en* on the verb indicate that this string relation is a mutation from the verb-object relation in the sentence on which the passive had operated. And indeed, at the same time that it is subject of *be composed*, *jazz* has the selectional properties and the meaning of an object of *compose*. Similarly the *by* shows that *students*, which is in the adjunct of *be composed*, is the subject of *compose*. Further, in the argument-skipping transform of (2),

- (4) Students continued composing jazz,

*students* has subject relation to *continue*; but the *-ing* on the following verb *compose* (together with the  $V_v$  membership of *continue*) indicates that *students* had the subject relation to *compose* in the sentence from which (3) was derived. If we take the passive of (4) we obtain

- (5) The composing of jazz was continued by students,

where *the composing of jazz* is the subject, and *by students* the adjunct, of *be continued*; but as in (3) the constants *be-en*, *by* show that *composing jazz* (or *the composing of jazz*) is the object, and *students* the subject, of *continue*; and as in (4) the *-ing* shows that *students* is the subject of *compose* in the sentence form prior to the action of the argument-skipping T. And if we take the passive under the increment of (2) we obtain

- (6) Jazz' being composed by students continued,

where *jazz* is marked both as the subject of the operand *be composed* and

as the object of *compose*. And the argument-skipping transform of (6) is

- (7) Jazz continued being composed by students,

where *jazz* is the subject and *be composed* the object of *continue*; but as in (4) *be composed* is also the prior subject of *continue*, in which case *jazz* is the subject and *by students* the adjunct of the operand *be composed* (under *continue*), where the *be-en* shows that *jazz* is the object and *students* the subject of *compose*.

The comments above list only a few of the string-relations among the words in these transforms, but they suffice to show that many different (but not conflicting) string relations appear in the various transforms. The string relations in the {I} sentence from which the transforms are derived is always indicated, as well as all the string relations in each transform along the line of derivation from {I}. If we compare (7) above with (5), we find that the same two T produced them, but in different order of application. The simultaneous systems of string relations in (7) and (5) are however quite different, except that both contain the string relations of their common {I} source. It is reasonable to say that the meanings of all these string relations are present in the sentence (e.g. *jazz* in (7) being the subject of both *continue* and *be composed* and the object of *compose*), but they do not seem to be ordered as meanings, although to arrive at them requires an ordered undoing of the transformations that had produced the final resultant sentence.

If we go from the ordered operator notation (i.e. of transformations) to the natural string of words, the various constants and positionings are the trace, in the word-string, of each ordered transformation upon its operand sentence; they are changes in the sense that each transformation makes a change upon its operand. If we go from the actual word sequence to its transformational representation, then the constants and the positionings have to be taken not as objects in a static description of the sentence, but as the result of ordered changes: certain ones of them have to be reorganized as the latest change (due to the latest transformation); when they are undone we arrive at the operand sentence on which that change had been made, and then have to recognize what is now the latest change, which had produced that (operand) sentence in turn; and so on to the base.

It must be stressed that there are certain properties common to almost all string changes, such that they preserve the general properties of the string structure of {I}, or change them in only a few generally stateable ways. Hence the whole language has an almost coherent string structure which applies to all sentences (though in part vacuously to the sentences of {I}). Nevertheless even here there are some exceptions to coherence, as in the string structure of the moods (5.6).

### 8.6. Invariants

From the foregoing survey of the T, and of how they change the string structure of a sentence, we can see what it is that is preserved under T: It is the operator order of the {I} source, or equivalently the {I}-string relations (7.5) among words of the sentence. This gives the objective grammatical meaning of the syntactic relation among the words, and is reachable from any T resultant by a computation (which may contain ambiguous degeneracies). The T add no new morphemes other than constants marking string-change, and the only added meanings that these string changes bring in are subjective without affecting the objective grammatical meaning carried over from the {I} source.

In addition, the T preserve, by definition, the independent variables and the acceptability-ordering of the sentences (or, equivalently, the selection-ordering for the words in the sentences), and so the meanings of the sentences.

The {I}-operator ordering and objective meaning of a sentence are therefore invariants of sentences under T.

### 8.7. Structure of {S-I}

Just as we can state the sub-language grammar of {I} (7.2), so we can state one for {S-I}. In {S}+ we take {I} as base, and T as the operators. Each T is defined on particular transformations (increments or T) as arguments. Each resultant sentence can be operated on by a partially ordered subset of T, or equivalently by a string relation of its words which includes a particular change from the string relation of the same words in the operand sentence.

The set of T can be organized in more than one convenient way since there are only a few different types of T (partly indicated in 8.1) and there are various similarities among the types. What is important in the kind of analysis proposed here, is that all T be recognized, and distinguished from the increments. How the T are formulated, and whether a given T is to be decomposed into two or more successive elementary T, is not essential to decomposing the grammar of the language into the two systems {I} and {S-I}. As long as there was no evidence concerning what was the core of language, the main goal of structural linguistics was to eliminate restrictions and increase regularity in the relative occurrence (the combinations) of language elements. There was no other criterion for finding what was linguistically essential, out of languages which obviously contained many inessential restrictions and irregularities. For this reason, every reduction of the data to more regular and basic elements had to be considered, even if it seemed to be based on chance similarities in the data. However, given a base whose reality is supported

by absence of restriction, simplicity of structure and reasonableness of interpretation, we are no longer in the position of simply extracting every irregularity possible toward a residue that is defined only by its irreducibility. We now are given, for analysis, both an initial language and a final base to which to reduce it. The bases – the primitive arguments N, or the elementary sentences K, or the paraphraseless sublanguage {I} – are indeed irreducible, but they are so in respect to explicit operations. In organizing the data as a set constructed from a base (or reducible to it) by certain operators, we no longer go merely by what eliminates an irregularity, but rather by what is the best way of arranging operations that will reach from one structure to the other. Differences of opinion as to what are the base operations (e.g. whether the numbers should be derived from a count on occurrences of *and*), or inefficiencies in their action, are no longer crucial; one can leave room for historical chance, for unused productivity, etc.

### 8.8. Derivation

In {S}+ we speak of deriving the sentences of {I} from its elementary sentences K or \*K, because each derivation adds an increment while preserving the properties (operator order, string relations though in modified form, selection and meaning) of the operand sentence. We speak of deriving the sentences of {S-I} from those of {I} because (a) in those T-derivations which apply to only a subdomain of their operand sentences (in {I} or in {S-I}) the sense of the derivation can only be from the form which applies over a larger domain to that which holds for a smaller domain (even if the former is morphemically complex or uncomfortable); and (b) the other T-derivations, where there is no reduction of domain, are of the same kind as (a).

The course of derivation also has good interpretation, in that the interword grammatical relations which are passed through in the various stages of a derivation are preserved in the simultaneous string-relations of the resultant sentence (8.4), and in that the traces of intermediate derivational stages are left in the lexical shift of words. Thus the meaning of something like 'bounded event' which adheres to *talk* in *His talks are inspirational* comes from the predicate in *His talking was in a piece* (or: *was an event*) which is the source of *He had a talk, He gave a talk* (5.4), which is a stage in the derivation. As far as we can see, no relation and no meaning is present in a sentence except those that are due to the primitive arguments, the ordered {I}-operators and the T which appear in its derivation.

The existence of one apparent derivation for a sentence does not always mean that there is no other derivation for the same sentence, which may avoid some difficulties present in the first. Thus the reciprocal verbs could be derived from a zeroing of *each other* determined by an environing  $V_{rec}$ , but



more satisfactorily (from the point of view of metatheory) from a zeroing of *and this is implicit* as in 4.3.

The sources for these derivations are set up here only in a structural sense. However, some of them reflect the course of history by reconstructing a form which had diverged into separate forms. Other sources, quite the contrary, reflect the detailed limitations upon the combinability of words: There are some words which have no linguistic utilization (combinability) beyond certain operator or operand environments, and it is this limitation that makes them useable as particular sources or as derivable from particular sources.

#### 9. BEYOND THE PRESENT ANALYSIS

The distinguishing of the {I} and {S-I} systems of language requires much further work. Aside from questions of individual transformations and derivations, there is the general problem of the detailed structure of each of these systems: Do they depend on (or affect) only their immediate argument or also the argument of their argument, and if so how far down? What operations are linearly ordered and what partially ordered? What types of domain restriction appear in T, and what kind of subjective information do the T carry? The distinction between objective and subjective information, which arises out of the {I}-{S-I}-distinction, would also bear consideration. Within {I}, where more is sayable than might at first have been thought for language, the question arises what remains unsayable due to its structural limitation.

There is also the possibility of utilizing and even modifying the structural characteristics of {I}. The removal of transformational paraphrase leaves {I} as an information-processing system of a particular kind — and of a widespread kind, since this is human language. Sharper versions of an information processing structure may be found, if an {I} system is constructed for the sublanguages of particular sciences, where much dictionary paraphrase (which is due to semantic overlapping in the vocabulary) can be isolated and eliminated, so that the correlation of structure and report becomes sharper, and the particular relations of the science would be brought out sharply.

The structure of {I} applies to the morphemes that exist in language, and to the particular meaning-ranges which they have, i.e. to the way information is represented by morphemes. The fact that part of its structure is achieved by utilizing the limitations upon use of morphemes (8.7) suggests that if we could find further limitations upon morpheme use we might obtain a still more compact {I}. Such further limitations undoubtedly exist in language, but they are limitations not upon morphemes but upon (synonymous) morpheme sets. This happens because in many cases we can find morphemes A, B such that B is a free variant (and synonym) of A over only certain but not

all of the operator or operand environments of A: that is, a word may have synonyms in all its occurrences i.e. in all its environments, but not always the same synonym. Such free variation will have been missed in the methods of §§ 2-5, which, for each morpheme, dealt with the whole set of its occurrences, i.e. with the whole set of its environments in all its occurrences.

In establishing useful primitives of morphology, morphemes were freed from the need to be phonemically simple. In the methods of the present paper, useful primitives of syntax are sought by freeing them from the need to be morphemically simple. That is, we did not take each morpheme and classify it with those having similar combinability. Instead we sought morphemes and sequences of morphemes that had free or complementary variation to each other and called them all a single syntactic element. Now we can consider the further possibility of freeing the syntactic primitive from vocabulary simplicity, i.e. from having to represent morphemes each in the full range of its occurrence. We can consider how a given morpheme may be a member of one syntactic element in some of its occurrences, and of another element in others. Of course, methods have to be developed to decide how to partition the range of occurrences of a morpheme without being arbitrary or trivial. The partition would have to be similar for a whole class of morphemes, and each morpheme would have corresponding different synonyms in each subset of its range. But it is not clear that effective tests can be made for this.

If a useable method could be developed for discovering all of these sectional free variations (synonym sets), we would eliminate all sectional synonymity (and homonymity, which does not hold between synonym sets). Such a further reduction may be practicable for particular small word classes (e.g. the  $V_{vv}$  and the subordinate conjunctions which are transforms of them), or within sublanguages of science.

The method envisioned here may be applicable in a special situation. In general, we cannot expect to be able to go beyond morphemes toward freer syntactic primitives. The reason is that morphemes have too much selectional flexibility to be replaceable by some more fundamental relation that would express the restrictions on their selection. Language is everywhere expandable and changeable; hence there is no value of a variable (say, a particular morpheme in the V class) about which we can say that it cannot have a particular other value of a variable in its syntactic environment (say, a particular N morpheme as its subject). However, there may be characterizable subsets of language, such as the sublanguages for particular sciences, in which a part of the vocabulary is restricted as to its syntactically enviring vocabulary. In a sublanguage in which we can delimit the acceptability of syntactically enviring morphemes of a morpheme A in class X, we can



discover whether there are other morphemes B, C in the same class X such that a subset p of the environments of B and a subset q of the environments of C together equal the environment-set of A. If this is found to be the case, we can redefine A as being two homonymous morphemes,  $A_p$  and  $A_q$  in class X,  $A_p$  being a free variant and synonym of  $B_p$ , and  $A_q$  a free variant and synonym of  $C_q$ . If this can be done to an important part of the sub-language vocabulary, then instead of the relevant values of the morpheme-class variables being the morphemes A, B, C, they would be now certain environmental subsets of morpheme occurrences, namely p, q, etc. Here the p value of the variable X would have A and B as its freely-varying phonemic representation, and the q value of X would have A and C as its freely-varying phonemic representation.

When p and q are taken as the values of X, there are no homonymities: for whereas  $A_p$  and  $A_q$  could have been considered homonymous morphemes, now they are simply variant forms belonging to different values. And synonymities have been eliminated: for whereas some of the occurrences of morpheme B had been synonymous to some of the occurrences of morpheme A, now the phoneme-sequence of  $B_p$  and the phoneme-sequence of  $A_p$  are merely variant phonemic forms of a value p. In this part of the vocabulary, we can then characterize the values of the variable (morpheme-class) X not phonemically but syntactically, i.e. by their occurring with particular values of enviroing variables (as for the subsets p of A and B above). In this situation, the phonemic representation of the values, e.g. the free-variant A and B forms of p, becomes part of morphophonemics, in {S-I}, and is no longer part of the informational {I}. For morphophonemics is the change of the phonemic shape of a morpheme (as value of a class-variable) in the environment of certain other morphemes, and the phoneme sequences of A and of B are the change (from zero) of the phonemic shape of p, which is a (morphemic) value of X, in the environment of certain values of other variables.

Syntax is initially bound to phonemes because it starts out to state the independent restrictions on the combinability of phonemes, the sentences being given as phoneme sequences. But in the special conditions considered here, we would be able to define the elements not as phoneme sequences (morphemes) but as values of variables which are characterized, purely syntactically, as occurring with certain other values of other variables. The co-occurrence of these values would be in {I}; the phonemic shapes of these values would now be in {S-I}.

## NOTES

<sup>1</sup> An operator  $X_{Yz}$  (Y, Z) is defined as acting on certain arguments Y, Z (written as subscripts), i.e. on certain variables on which it imposes a partial ordering of values. A sen-

tence is produced by the words of an operator A being concatenated with the words of its argument B; and the words of argument B being in turn concatenated with the words of its argument C, if B have any C; and so on. In the sentence, the word-string produced by the arguments B of A together with the arguments C of B, and so on, are the operand of A.

<sup>2</sup> A sentence-form is a sequence of variables and constants, the constants being particular morphemes and the variables being symbols for word- (or morpheme-) classes. The words in the class constitute the domain of the variable. Replacing the n independent variables of a sentence-form by an n-tuple of values from the stated domains produces a sentence S of that sentence-form.

<sup>2a</sup> In the initial condition, A is not a sentence-set but a set of N-arguments, and in B the  $x_m$  is not an increment but the predicate of an elementary sentence (7.2).

<sup>3</sup> Here A is the source of  $B^*$ , written  $A \rightarrow B^*$ , on the general grounds of transformational analysis (Z. Harris, *Mathematical Structures of Language*, Interscience Tracts in Mathematics, Vol. 21, Wiley, New York, 1968, p. 62-3).

<sup>3a</sup> Op. cit. note 3, §§ 5.4, 5.6. 3-4.

<sup>4</sup> Also, the sense of an ongoing act in the special use of *be-ing* before adjectives and nouns (*He is being clever*, *He is being a man*) is seen in the other form: *His being clever is in process*, *His being a man is in process* (is something going on).

<sup>5</sup> Also by the fact that this finally makes it possible to derive all occurrences of *-ing* from a single source.

<sup>6</sup> Richard I. Kittredge, *Transformations and Discourse Analysis Papers*, University of Pennsylvania 1969. The classifications 'perfective' and 'imperfective', introduced in this connection by Kittredge, are used here tentatively, pending further investigation.

<sup>7</sup> In the sentence after *than* (or *as*) parentheses indicate optional zeroing; but if a word of that sentence is omitted then this is due to required zeroing.

<sup>8</sup> The comparative-conjunction form is also obtained when the quantified words in the two source sentences have different positions, provided that they are the same word so that the second occurrence has been zeroed: *The number of men who read books is more than (the number of men whom) you can count. More men read books than you can count.*

<sup>9</sup> The nearest we come to this is the makeshift and semantically unclear *More men read books than there are magazines which women read*. A more acceptable situation of this type is seen in (6).

<sup>10</sup> *Sn* indicates nominalized sentences. The subscripts identify the sentences.

<sup>11</sup> That the subjunctive does not bring in independent morphemes but only a form that is automatic in respect to its operator is seen in the fact that a subjunctive occurrence of a sentence can be zeroed as a repetition of a non-subjunctive occurrence of that sentence: *He opposed it more than I expected (that he would oppose it)*. For the parenthesized segment to be zeroed, it must consist only of its antecedent (the segment, aside from tense, in *He opposed it*) plus entities that are determined by (and reconstructed from) the residual *I expected*.

<sup>12</sup> I.e. its battery of transformations in the sense of H. Hiz, 'Congrammaticality, Batteries of Transformations, and Grammatical Categories', in *Proceedings, Symposium in Applied Mathematics 12*, American Mathematical Society, 1961, 43-50. The view presented in this section is close to that reached on other grounds by A. K. Joshi in his *Properties of Formal Grammars with mixed types of rules and their linguistic relevance* (University of Pennsylvania 1969), and in A. K. Joshi, S. R. Kosaraju, H. Yamada, *String Adjunct Grammars, Transformations and Discourse Analysis Papers 75* (University of Pennsylvania 1968). Both Mr. Hiz and Mr. Joshi have also contributed valuable criticisms to the present paper.

<sup>13</sup> Beverly Levin Robbins, *The Definite Article in English Transformations*, Papers on Formal Linguistics, Mouton, The Hague, 1968.

<sup>14</sup> Since it will be seen that *move* is the argument of *try*, and *dog* is the argument of *move* (7.1), this means that *try* would have a domain restriction on the argument of its argument.

<sup>15</sup> This cannot be derived from *N gathered and N gathered and N gathered* because of the restriction of *gather* to  $N_{coll}$ .

<sup>16</sup> H. Hiz, *Referentials*, *Semiotica*, Vol. I, 2 (1969), 136-166.

<sup>17</sup>  $S_n$  should be obtainable from  $S_1$  by recognizing the trace of  $T_1$ , removing increment<sub>1</sub> and  $T_1$ , and reapplying increment<sub>1</sub>. It has to be shown that the increment can always be applied even without the intermediate  $T_1$ , and that its meaning effect is then unchanged.

<sup>18</sup> As was noted by Edward Sapir in his paper 'Grading' in D. G. Mandelbaum (ed.), *Selected Writings of Edward Sapir*, University of California 1958, p. 122ff.

## ABOUT LINGUISTICS