SUPPLEMENTO AL VOLUME XIII, SERIE X
DEL NUOVO CIMENTO

N. 2, 1959
3° Trimestre

Questions of Linguistics.

M. HALLE

Massachusetts Institute of Technology - Cambridge, Mass.

Introduction.

We begin our consideration of the linguist's approach to the problem of speech communication by inquiring into the nature of the data that constitute the subject matter of linguistics. We want to know what kind of problems are of special interest to the linguist, for only if we understand this will we be in a position to appreciate the reasons for the ways of the linguist which frequently seem strange to the outsider.

As a first answer it might be proposed that linguistics is concerned with characterizing the class of acoustical signals which men make in speaking. The natural way of going about this would be by investigating in detail the anatomical structures in man that make it possible for him to emit this special set of signals. One would investigate the human vocal tract: the larynx, the pharynx, the nasal cavity, the mouth, the tongue, the lips, etc., and one would attempt to make statements about the motor capabilities of these organs. Once one had learned all there is to know about these physiological aspects of the problem, and, provided one knew a great deal of acoustics, one could give the desired description of the acoustical signals which such a mechanism was capable of emitting. One might further investigate the analogous mechanisms in other animals and might succeed in showing how the latter differ from those of man and how this difference accounts for the differences in the respective acoustical outputs. The results of this inquiry would explain why the acoustical signals emitted by men in speaking differ from those of other animals.

This is a very important area of study, and linguistics is vitally interested in these questions. Yet these questions do not exhaust the problems of concern to the linguist: they are but a small part of the puzzles that the linguist would like to solve. As a matter of fact, if linguistics were limited to a consideration

of these problems, there would hardly be any need for a separate discipline, since all of the above problems are dealt with by physiology and acoustics.

What makes linguistics as a field of enquiry quite different from physiological acoustics is the fact that what is commonly referred to as « linguistic behavior » covers a much broader area than the acoustical properties of speech, though—as I have already said—it specifically includes the latter. Let me now describe a few of these additional problems.

We have all had the experience of hearing people speak with a foreign accent. Thus, for instance, we all know people who are physiologically normal, who yet find it difficult to distinguish sounds that we ourselves have no difficulty whatever in distinguishing. For instances no English speaker would ever confuse the words « bitch » and « beach »—not even under conditions of high noise, as G. A. MILLER has shown. Yet a speaker of Russian or Italian would find it extremely difficult to keep them consistently apart. Clearly the difference in the behaviour of English and foreign speakers is not physiologically determined, because the foreigner can—when his attention is drawn to it—make the required distinction. The difference in behavior is, of course, due to the fact that English, Russian, and Italian are different languages, and that different languages use different sounds.

It may, therefore, be proposed that adult speakers have established a particular behavior pattern of their vocal organs and that this behavior pattern accounts for the observed difficulty. Differences in language may, therefore, be equated with different habitual movements of the tongue and lips and with different co-ordinations of these movements. In other words, one might conceivably explain linguistic differences on a physiological-acoustical basis, provided one allowed for some learning.

This, however, is not really an adequate explanation. Consider, for instance, the manner in which Latin is spoken by priests of different nationalities. An English-speaking priest may read mass with a sound repertory that is 100% English, and a French priest may read the same mass with a sound repertory that is 100% French. Yet there is no sense to the statement that the language of the mass is anything but Latin.

An attempt may still be made to save the view of language as a purely physiological-acoustical phenomenon by saying that, *e.g.*, the English priest uses the sounds of English with the statistics appropriate for Latin. This, however, is hardly a good solution since it raises a host of extremely difficult problems. *E.g.*, it raises the question of how it is possible to identify an utterance as English on the basis of a very short sample, which might be totally atypical. But even if this were possible, there are aspects of linguistic behavior which cannot be explained in terms of physiology and acoustics alone, regardless of the refinements introduced. I shall now give a few examples of this.

A joke quite popular among elementary school children in America is the

following question and answer: *Why can't one starve in the desert? — Because of the sand which is there!* The pun is based on the fact that word boundaries are not always marked acoustically, and *sand which is* is frequently indistinguishable from *sandwiches*. Yet word boundaries are crucial in understanding the message correctly, and given enough context the speaker of English will know how to assign word boundaries even if they are not acoustically marked.

Word boundaries, moreover, are not the only boundaries which have no acoustical signal and which affect the behavior of the speaker. Consider the following ambiguities:

Old | men and women	Old men | and women
He rolled | over the carpet	He rolled over | the carpet

which are due to differences in phrase structure that are not marked acoustically.

I should like also to draw attention to another type of behavior. Every speaker of a language can perform rather elaborate transformations upon sentences. Thus, for instance, given a simple declarative sentence there is a standard way of converting it into a « yes or no » question; or given an active sentence there is a standard way for converting it into a passive. As an illustration of the latter take the sentence: *A committee opposed the change in the bill* which can be readily transformed into *The change in the bill was opposed by a committee.* In order to explain how to perform this operation we would normally use such terms as *noun phrase*, *verb phrase*, *transitive verb*, etc., in the obvious way. It is important to note, however, that here, too, there is no such thing as an acoustical signal for these categories, yet the categories are essential in order to explain the speaker's behavior.

Consider again the sentence, *The change in the bill was opposed by a committee.* The choice of *was* as against *were* is governed by the number (singular or plural) of the head of the first noun phrase; *i.e. change.* But the head of the noun phrase, which itself is a noun phrase, does not have any acoustical marker to distinguish it from other noun phrases.

It must also be noted that the head of the noun phrase governs the choice of *was* as against *were* quite independently of the number of intervening words; *e.g., The change in the bill for the promotion of the study of the mating calls of rhinoceri... etc...* ***was*** *opposed by a committee.*

Engineers and other non-linguists have usually neglected problems of the kind just surveyed, considering them either outside of their ken or relatively unimportant refinements. Linguists, on the other hand, have been keenly interested in such problems. The standard grammars of the different languages always try to do something towards solving such problems. Unfortunately the standard grammars fail to be consistent or to make clear the basis on which they operate. In what follows I shall try to present in outline a descriptive

framework for language which I believe to be free of, at least, the most glaring of these failings. The exposition will begin with a review of some recent work of N. CHOMSKY and will go on to a discussion of the phonic aspects of language, which were not considered by CHOMSKY.

1. – Chomsky's analysis.

According to CHOMSKY every language has three distinct sets of rules which operate on three different levels. On the highest level the rules are all of the type « $X \to Y$ » where « $X \to Y$ » stands for « replace X by Y », with the restriction that not more than a single symbol can be replaced in a single rule and that $X \neq Y$.

As an illustration of these rules we can take the following (*):

Sentence → Noun Phrase + Verb Phrase + (Adverbial Phrase)	(1)
Noun Phrase → (Article) + Noun + (Prepositional Phrase)	(2)
Verb Phrase → Verb + (Noun Phrase)	(3)
Adverbial Phrase → Adverb	(4*a*)
» » → Prepositional Phrase	(4*b*)
Prepositional Phrase → Preposition + Noun Phrase	(5)
Article → *the*	(6*a*)
» → *a*	(6*b*)
Noun → *committee*	(7*a*)
» → *change*	(7*b*)
» → *dog*	(7*c*)
» → *walk*	(7*d*)
» → *result*	(7*e*)
» → *bill*	(7*f*)
Verb → *opposed*	(8*a*)
» → *took*	(8*b*)
» → *barked*	(8*c*)

(*) In applying a rule the symbols in parentheses may be omitted. The rules are only partially identical with those that would appear in an actual grammar of English.

Preposition → *of* (9*a*)

» → *for* (9*b*)

» → *in* (9*c*)

The application of these rules yields a partially-ordered set of symbol sequences. We shall call each symbol sequence, a *string*, and the set of such strings generated by the rules, a *derivation*. We may illustrate the process of applying the phrase structure rules by the following derivation:

Sentence	by rule
Noun Phrase + Verb Phrase	(1)
Article + Noun + Verb Phrase	(2)
Article + Noun + Verb + Noun Phrase	(3)
Article + Noun + Verb + Article + Noun + Prepositional Phrase	(4)
Article + Noun + Verb + Article + Preposition + Noun Phrase	(5)
Article + Noun + Verb + Article + Noun + Preposition + Article + Noun	(1)
a committee opposed the change in the bill	(6*b*), (7*a*) (8*a*), (6*a*) (7*b*), (9*c*) (6*a*), (7*f*)

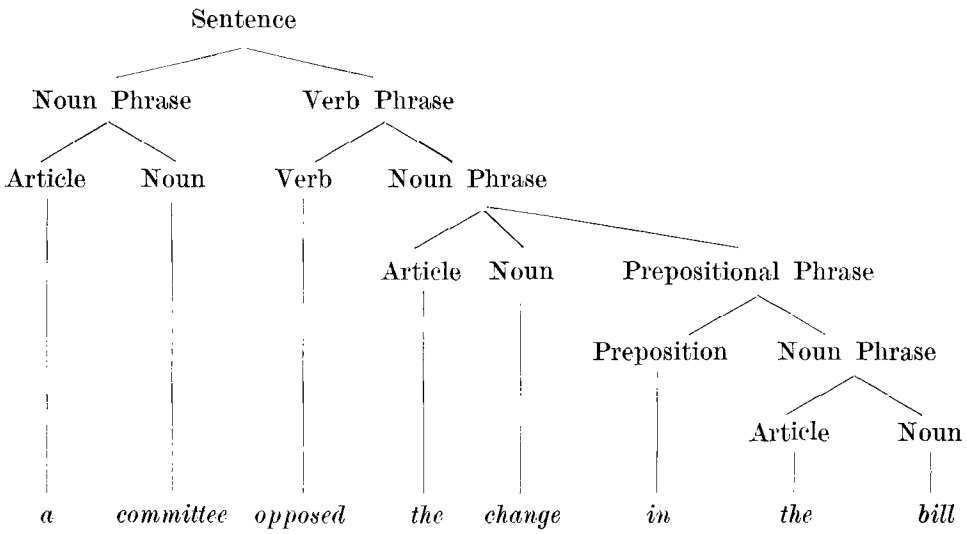
Attention must be drawn to the following facets of the grammar just presented:

1) The order of application of the rules is partly fixed owing to the fact that a given rule can be applied only if the symbol to be replaced—*i.e.*, the one appearing on the left-hand side of the rule—appears in the derivation. There must, therefore, be at least one *initial* symbol which must be supplied to the grammar from the outside and which starts things off. For the present set of rules the symbol « Sentence » will serve this function.

2) In order for the grammar to continue to operate it is necessary that instructions be provided for selecting the next rule to be applied. The instructions must be supplied from the outside. It is by exercising a choice, by selecting one rule from a set of possible alternatives that information is being transmitted. This choice must evidently be made by the user of the grammar, for only he can transmit information.

3) The grammar continues to operate as long as the string contains symbols which themselves appear on the left-hand side of one or more rules. The grammar stops operating when it has produced a string consisting of symbols which occur only on the right-hand side of the rules—e.g., *opposed* in rule (8*a*)—and hence are « irreplaceable. » We shall call these « irreplaceable » symbols, *terminal symbols*; strings consisting of terminal symbols only shall be called *terminal strings.*

It is always possible to convert a derivation into a tree like the one below.



The tree may be familiar to some readers from their school days. It represents what is commonly known as « parsing » or « diagramming » or « immediate constituent analysis » of the sentence. It contains at least a partial answer to the question of whence come the boundaries which in spite of their possible lack of acoustical correlate are nevertheless important factors in the behavior of speakers.

The restriction on the number of symbols that can be rewritten in a single rule guarantees that given a terminal string—*i.e.* a string produced by the application of the phrase-structure rules—it will be possible to discover the associated tree or trees. Since not more than one symbol can be rewritten in a single rule, every line in the derivation must have at least as many symbols as the one preceding it. Since repetitions of lines in the derivation are not admitted ($X \neq Y$), there must be a finite number of lines between the first line and the terminal string. One can, therefore, try out all one-line derivations, two-line derivations, three-line derivations, etc., until one comes upon a derivation having the desired terminal string.

Since there may be more than one derivation yielding the same terminal

string, there may be more than one tree associated with a single terminal string. The fact that some terminal strings have more than one phrase structure representation accounts for the ambiguity of phrases like *old men and women*; *he rolled over the carpet*; etc.

By repeated reapplication of rules (1) and (5) endless sequences of words may be generated. This is not an oversight but rather a reflection of the fact mentioned above that language places no upper bound on the length of sentences or of constituents, although all sentences are finite in length.

We have made much of the fact that terminal strings have phrase structure. It is now necessary to point out that terminal strings are abstract representations of certain features of sentences and that actual sentences are, in fact, not terminal strings. To see this, consider the English verb. Since verbs can be in the present tense as well as in the past we introduce a rule like the following:

$$\text{Verb Phrase} \rightarrow \text{Verb} + (\text{Past}) + (\text{Noun Phrase}) \ (^*) \qquad (3a)$$

We would then also need rules like

$$\textit{oppose} + \text{Past} \rightarrow \textit{opposed} \qquad (10a)$$

$$\textit{write} + \text{Past} \rightarrow \textit{wrote} \qquad (10b)$$

$$\textit{have} + \text{Past} \rightarrow \textit{had} \qquad (10c)$$

$$\textit{think} + \text{Past} \rightarrow \textit{thought} \qquad (10d)$$

$$\textit{be} + \text{Past} \rightarrow \textit{was} \qquad (10e)$$

Rule (10*a*) is within the restrictions imposed on phrase structure rules, for it requires in effect that the symbol « Past » be replaced by *-d*. The other four rules, however, violate the phrase structure constraints. *E.g.*, in (10*b*) the two symbols « *write* » and « Past » are replaced by « *wrote* » in one step, and it is impossible to achieve the same result if only a single symbol were allowed to be replaced in a single rule. Consequently, rules (10*b*) to (10*e*) are beyond the power of the phrase structure level. Since all verbs violating the phrase structure constraints belong to the so-called « strong » or « irregular » verbs of English it may be proposed that these verbs be handled as exceptions; there would then be no need to utilize more powerful devices in the grammar. We shall see, however, that the phrase structure grammar is not powerful enough to handle other, perfectly regular verbal formations in a reasonably economical fashion. The proposal to consider the « strong » verbs as exceptions is, therefore, of little practical importance.

(*) We are disregarding the problems raised by number and person.

Consider now the Verb Phrases:

had opposed	*was opposing*	*had been opposing*
had written	*was writing*	*had been writing*
had had	*was having*	*had been having*
had thought	*was thinking*	*had been thinking*
had been	*was being*	*had been being*

In order to generate the examples in the first column we should need the rule

Verb Phrase → *have* + (Past) + Verb + Perfect Participle + (Noun Phrase) (3*b*)

as well as

oppose + Perfect Participle → *opposed* (11*a*)

write + Perfect Participle → *written* (11*b*)

have + Perfect Participle → *had* (11*c*)

think + Perfect Participle → *thought* (11*d*)

be + Perfect Participle → *been* (11*e*)

In order to generate the examples of the second column we should need the following rules:

Verb Phrase → *be* + (Past) + Verb + Present Participle + (Noun Phrase) (3*c*)

and

Verb + Present Participle → Verb + *-ing* (12)

Finally in order to generate the examples in the third column we need the following additional rule:

Verb Phrase → *have* + (Past) + *be* + Perfect Participle + Verb +
+ Present Participle + (Noun Phrase) (3*d*)

This rule, however, is the sum of rules (3*a-c*). It is, therefore, natural to investigate whether the set of rules cannot be simplified. Examining rules (3*a-d*) we note the following regularities:

a) The symbol « Past » is always associated with the first element of the Verb Phrase.

b) If the Verb Phrase contains the auxiliary verb *have* the symbol « Perfect Participle » appears after the next element of the Verb Phrase.

c) If the Verb Phrase contains the auxiliary verb *be*, the symbol « Present Participle » appears after the next element of the Verb Phrase.

d) If both auxiliary verbs *have* and *be* occur in the same Verb Phrase, *have* precedes *be*.

e) The only element which must appear in the Verb Phrase is (the main) Verb.

f) The auxiliary verbs precede (the main) Verb.

The simplest way of handling these regularities is by positing the following two rules:

$$\text{Verb Phrase} \rightarrow (\text{Past}) + (\textit{have} + \text{Perfect Participle}) + \\ + (\textit{be} + \text{Present Participle}) + \text{Verb} + (\text{Noun Phrase}) \qquad (3')$$

and

$$\text{G} + \text{V} \rightarrow \text{V} + \text{G} \qquad (\text{Z})$$

where V stands for any specific verb (lexical morpheme) like *oppose*, *have*, *be*, *think*, etc., and G stands for a grammatical operator like « Perfect Participle, » « Past, » etc.

Rule (Z) goes clearly beyond phrase structure, for it changes the order of the symbols, and once the order of the symbols in the strings is changed, there is no longer any possibility of associating a tree with a string. We are, therefore, faced with the alternative of either maintaining the phrase structure restriction and thereby greatly complicating our description—*e.g.*, we would be forced to have four separate rules in place of the single rule (3)—or of admitting into the grammar new rules that are more powerful than those of the phrase structure level. There are various reasons why the latter alternative is to be preferred. Accordingly we establish a second grammatical level, which, following CHOMSKY, we call the *transformational level.*

It is not possible here to go into the details of the transformational level. These can be found in CHOMSKY's book *Syntatic Structures.* I should like, however, to draw attention to a few consequences of the decision to introduce the transformational rules.

Since rule (Z) must precede rules like (10) and (11), the latter together with (Z) are part of the transformational level. This makes it unnecessary to do anything special about the « strong » verbs (rules (10*b*-*d*)), since on the transformational level the prohibition against replacing more than one symbol in a single rule does not hold.

The terminal strings, the final output of the phrase-structure rules, will contain symbols of two types: lexical morphemes like *oppose*, *committee*, *of*, *the*, etc., and grammatical operators like « Past, » « Perfect Participle, » etc. This is due to the fact that at least some grammatical operators cannot be replaced by phrase structure rules; *e.g.*, « Past » is replaced in rules (10*a-e*), which are, however, transformational and not phrase structure rules.

The terminal string corresponding to our sample sentence is therefore represented, with some simplifications and omissions, as follows:

$$a + \textit{committee} + \text{Past} + \textit{oppose} + \textit{the} + \textit{change} + \textit{in} + \textit{the} + \textit{bill}$$

The transformational rules operate on terminal strings *and* the trees associated with them. The notion « head of noun phrase » which we have had occasion to use in the above discussion has an obvious and simple meaning if reference is made to the tree associated with the particular noun phrase. It is a matter of considerable difficulty to give a clear meaning to this notion if one limits oneself only to the terminal string.

Up to this point we have been concerned exclusively with what might be termed abstract properties of language and we have said nothing of its acoustical features. It is now necessary to examine the relationship between the abstract entities that have been described in the preceding pages and the concrete sound waves that comprise the spoken message.

2. – Sounds of speech.

The problem with which we shall be concerned in this lecture is the manner in which the sounds of speech are to be described. In every science the choice of a descriptive framework is an extremely important matter. It is usually not enough that the description reflect the physical facts to a sufficient degree of precision. We would like to describe these facts in such a way as to open up the possibility of saying other things of interest, too. The following example illustrates this point as it may affect the linguist.

English speakers form the regular plural of nouns by adding a *sound* or *sounds* to the singular stem. They add [ɪz] if the noun ends in [s], [z], [š], [ž], [č], [ǯ], (*e.g.*, *busses*, *causes*, *bushes*, *garages*, *beaches*, *badges*); they add [s] if the noun ends in [p], [f], [t], [θ], [k], (*e.g.*, *caps*, *cuff*, *cats*, *fourths*, *backs*); and they add [z] in all other cases.

In stating this we have, however, made a number of decisions regarding the manner in which we shall describe the facts. We have spoken of individual sounds—let us henceforth call them segments—and we have attached labels

to them; *e.g.*, [s], [z]. We have decided in effect to view utterances as sequences of a number of discrete entities. If we were asked why we made this decision we would surely reply that this seems to us to lead to a simple description of all kinds of facts. The questioner being a linguist in disguise might then point out that our description would be even simpler if we had a label for the segments [s], [z], [š], [ž], [č], [ǯ], and another one for the segments [p], [f], [t], [θ], [k]. But this is indeed the case if we describe the segments with the help of any of the standard phonetic frameworks: the first set consists of the noisiest sounds in the English language, variously called *hushing* and *hissing* or *strident* sounds, and the second set contains only voiceless sounds. In other words, the classification of sounds into strident and not strident (mellow), and voiced and voiceless fits well with the above facts.

We can now simplify the previous formulation in the following rules:

R. 1 If the noun ends in a strident consonant, then Plural → [ɪz].

R. 2 If a noun ends in a consonant which is voiceless, but is not strident, Plural → [s].

R. 3 In all other cases, Plural → [z].

In order to obtain simple rules we have described the utterances of English in a very special way. In particular we have regarded the utterances as consisting of sequences of discrete segments, and we have viewed the segments as simultaneous actualization of sets of attributes like voicing, stridency, consonantality, etc.

It is a well-known fact that viewed as an acoustic phenomenon speech is quasi-continuous; in many instances there is no obvious procedure for segmenting the continuous acoustic signal in a way which would correspond with the segmentation imposed by linguistic considerations. The question may, therefore, arise: in what sense can utterances be said to consist of discrete entities in sequence?

While a rigorous segmentation procedure which would show in all cases a one-to-one correspondence with the linguistic representation, may not be possible, it is possible to construct devices which produce speech by utilizing a set of discrete instructions which coincide closely with the linguistic segmentation. The devices I have in mind are of the type of the Bell Telephone Laboratories' Voder or the Haskins Laboratories' Octopus. The signal emitted by these devices is continuous speech, yet the input instructions are discrete. There is, therefore, a good sense in which utterances can be said to be made up of discrete segments.

In addition to viewing utterances as consisting of discrete segments we have also viewed the segments as simultaneous actualizations of a set of attri-

butes. In the descriptive framework with which we will be concerned below, the number of such attributes is quite small, about 15. These 15 attributes are sufficient to characterize all segments in all languages. Since we cannot have knowledge of all languages—*e.g.*, of languages which will be spoken in the future—the preceding assertion must be understood as a statement about the nature of human language in general. It asserts in effect that human languages are phonetically much alike, that they do *not* « differ from one another without limit and in unpredictable ways. » Like all generalizations this statement can be falsified by valid counter-examples. It can, however, not be proven true with the same conclusiveness. The best that can be done is to show that the available evidence makes it very likely that the statement is true. Most important in this connection is the fact that all investigations in which large numbers of languages have been examined—from E. SIEVER's *Grundzüge der Phonetik* (1876) to TRUBETZKOY's *Grundzüge der Phonologie* (1939) and PIKE's *Phonetics* (1943)—have operated with an extremely restricted set of attributes. If this can be done with about a hundred languages from all parts of the globe, there appears good reason to believe that a not greatly enlarged catalogue of attributes will be capable of handling the remaining languages as well.

The phonetic attributes and the segments are devices in terms of which the linguist represents his data. Like descriptive parameters in other sciences, these do not always stand in a simple one-to-one relationship with the observable facts. We have already had to remark on this indirect relation in the discussion of the segmentation of the utterance. A similar situation prevails with regard to the phonetic attributes. The absence of this simple relationship, however, does not mean that there is no specific connection between the descriptive devices and the data of linguistics. In the third lecture I shall attempt to outline this relationsip.

If it is true that a small set of attributes suffices to describe the phonetic properties of all languages of the world, then it would appear quite likely that these attributes are connected with something fairly basic in man's constitution, something which is quite independent of his cultural background. Psychologists might find it rewarding to investigate the phonetic attributes; for it is not inconceivable that these attributes will prove to be very productive parameters for describing man's responses to auditory stimuli in general. It must, however, be noted that for purposes of linguistics, the lack of psychological work in this area is not fatal. For the linguist it suffices if the attributes selected yield reasonable, elegant and insightful descriptions of all relevant linguistic data.

The attributes in terms of which we shall describe the sounds of speech are due primarily to R. JAKOBSON. Following JAKOBSON, we shall call these attributes *distinctive features*. The distinctive features have been described in

detail elsewhere. We shall, therefore, present here only the articulatory correlates of a few distinctive features (*).

Articulatory correlates of the distinctive features (partial list). (**)

1. Vocalic - nonvocalic. Single vocal cord source and absence of total occlusion in the oral cavity.
2. Consonantal - nonconsonantal. Presence of major constriction in the central path through the oral cavity.
3. Diffuse - nondiffuse. Oral cavity more constricted in front than at velum (backward flanged.)
4. Compact - noncompact. Oral cavity more constricted at velum than in front (forward flanged, horn shaped.)
5. Grave - acute. Major constriction in periphery (lips or velum) of oral cavity.
6. Nasal - nonnasal. Velum lowered.
7. Voiced - unvoiced. Vocal cords vibrating.
8. Flat - natural. Lips rounded.
9. Continuant - interrupted. No stoppage of air flow through mouth.

The first two features produce a quadri-partite division of the sounds of speech into 1) Vowels, which are vocalic and nonconsonantal; 2) Liquids, [r], [l], which are vocalic and consonantal; 3) Consonants, which are nonvocalic and consonantal; and 4) Glides, [h], [w], [j], which are nonvocalic and nonconsonantal.

Like all phonetic frameworks, the distinctive feature system is a catalogue of attributes. The distinctive feature system differs from other phonetic frameworks in that it contains only binary attributes. A segment, *e.g.*, is either voiced or voiceless, and there are no intermediate degrees of voicing ot which cognizance needs to be taken.

The question may well arise whether this is more than an empty trick, since any number of distinctions can always be expressed in terms of binary

---

(*) The fact that in the following list, reference is made only to the articulatory properties of speech and nothing is said about the acoustical properties, is not to be taken as an indication that the latter are somehow less important. The only reason for concentrating here exclusively on the former is that these are more readily observed without instruments. If reference were to be made to the acoustical properties of speech it would be necessary to report on experimental findings of fair complexity which would expand the present lecture beyond its allowed limits.

(**) Each feature is designated by a pair of antonymous adjectives, which, in accordance with the following convention, are used also to designate the segments. If the given description applies to a segment, it is designated by the first adjective; if the description does not apply, the segment is designated by the second adjective.

properties. All phonetic frameworks incorporate a large number of binary attributes: *e.g.*, voicing, nasality, rounding, aspiration, palatalization, etc. It is, of course, possible to replace these attributes by multi-valued properties. No one has ever shown, however, that anything is to be gained by this substitution. The replacement of multi-valued properties by binary features, on the contrary, does result in a gain.

In order to see this we shall examine the so-called *point of articulation*. The « point of articulation » is the place of maximum constriction in the oral cavity, and it has been customary to describe consonants in terms of this point. Thus, for instance, [p] is usually said to have a bilabial point of articulation, [f], a labio-dental point of articulation, [t], a dental or post-dental point of articulation, [k], a velar point of articulation, etc. No limitation is placed on the number of such points. In any given language, however, the number of separate points that need to be recognized is rather small. As a matter of fact, it can be shown that four such points suffice to describe all relevant facts in any known language. Instead of the multi-valued point of articulation dimension, the distinctive feature system contains the two features compact-noncompact and grave-acute, which distinguish the required four classes of segments: [p] is noncompact grave, [t] is noncompact acute, [c] as in *keys* is compact acute and [k] as in *cool* is compact grave.

The distinctive feature system employs less descriptive machinery than do other phonetic systems. Whereas in other systems the number of possible points of articulation is not restricted, in the distinctive feature system there are only as many different classes as are absolutely necessary. The decision to replace the point of articulation by two binary features, however, has other interesting consequences as well; *e.g.*, it makes it possible to explain in a simple manner certain linguistic changes which have puzzled linguists for a long time. One such example we shall examine in some detail.

It has been observed that when sounds change, these changes are gradual. *E.g.*, it is quite common for a voiced consonant to change into its voiceless cognate or vice versa ([v] → [f] or [k] → [g]); it is uncommon, or perhaps even unknown, for a voiceless consonant to change into a vowel ([k] ↛ [u]; [f] ↛ [a]). This observation can be conveniently expressed in terms of distinctive features as follows: a sound change rarely affects more than one feature.

In certain languages it has been found that [k] changes into [p] or vice versa. In terms of the multi-valued point of articulation this change is rather surprising, for [p] and [k] are produced with constrictions at opposite ends of the oral cavity. One might expect a change of [p] to [t] since they have adjacent points of articulation, but it seems rather curious that [p] and [k], which are articulated at such widely separated points should be confused. The distinctive feature system, however, provides a simple explanation for the puzzle. In terms of the distinctive features [p] and [k] differ in only a single feature: [p]

is noncompact and and [k] is compact. Consequently, the change of [p] into [k] is structurally quite similar to the change of a voiced consonant into its voiceless cognate.

The second difference between most standard systems and the distinctive feature system lies in the treatment of the two major classes of segments, the vowels and the consonants. In most standard systems these two classes are described in terms of features which are totally different: consonants are described in terms of the « points of articulation, » whereas vowels are described in terms of the so-called « vowel triangle. » In the distinctive feature system, on the other hand these two classes are handled by the same features: compact-noncompact, (diffuse-nondiffuse) and grave-acute. The distinctive feature system is thus more economical than other phonetic systems (*).

3. – Phonology.

Utterances are represented as sequences of distinctive feature segments. Although in many instances the latter stand in a one: one relationship with the sounds that we speak and hear, there are many instances where this relation is anything but simple. It is the major aim of the present lecture to elucidate this connection. The part of linguistics that is concerned with this problem is called *phonology*.

The phrase structure grammar, which was presented in Sect. 1, contained rules like « Noun → *committee*, *bill*, etc. » - cf., rules (6)–(9). These rules are basically lists of all existing morphemes in the language. Our purpose in preparing a scientific description of a language is, however, not achieved if we give only an inventory of all existing morphemes; we must also describe the structural principles which underlie all existing forms. Just as syntax is not identical with an inventory of all observed sentences of a language; so phonology—*i.e.*, a description of its phonic aspects—is not identical with a list of existing morphemes.

In order to generate a specific sentence it is necessary to supply to the grammar instructions for selecting from the lists of morphemes—*i.e.*, from the morphemes appearing on the right hand side of rules (6)–(9)—the particular morphemes appearing in the sentence. Instead of using an arbitrary numerical code which tells us nothing about the phonetic structure of the morphemes, it is possible—and also more consonant with the aims of a linguistic description—to utilize for this purpose the distinctive feature representation of the mor-

(*) It is curious to note that the Hindu phoneticians had the idea of treating vowels and consonant together over 2000 years ago. Their solution differs from the one proposed here in that it classified vowels as well as consonants in terms of their points of articulation.

phemes directly. In other words, instead of instructing the grammar to select noun (7*f*), we instruct the grammar to select the noun which in its first segment has the features: nonvocalic, consonantal, noncompact, grave, voiced, etc.; in its second segment, the features: vocalic, nonconsonantal, diffuse, acute, etc.; in its third segment, the features: vocalic, consonantal, etc. Instructions of this type need not contain information about all features but only about features or feature combinations which serve to distinguish one morpheme from another. This is a very important fact since in every language only certain features or feature combinations can serve to distinguish morphemes from one another. We call these features and feature combinations *phonemic*, and we can say that in the input instructions only phonemic features or feature combinations must occur.

Languages differ also in the way they handle nonphonemic features or feature combinations. For some of the nonphonemic features there are definite rules; for others the decision is left up to the speaker who can do as he likes. *E.g.*, the feature of aspiration is nonphonemic in English; its occurrence is subject to the following conditions:

a) All segments other than the voiceless stops [k], [p], [t] are unaspirated.

b) The voiceless stops are never aspirated after [s].

c) Except after [s], voiceless stops are always aspirated before an accented vowel.

d) In all other positions, aspiration of voiceless stops is optional.

A complete grammar must obviously contain a statement of such facts, for they are of crucial importance to one who would speak the language correctly.

In addition to features like aspiration in English, which are never phonemic, there are features in every language which are phonemic, only in those segments where they occur in conjunction with certain other features, and are not phonemic in other segments. *E.g.*, in English the feature of voicing is phonemic only in the nonnasal consonants; all other segments except [h] are normally voiced, while [h] is voiceless.

So far we have dealt only with features which are nonphonemic regardless of neighboring segments. There are also cases where features are nonphonemic because they occur in the vicinity of certain other segments.

As an example we might take the segment sequences at the beginning of English words. It will be recalled that the features vocalic-nonvocalic and consonantal-nonconsonantal distinguish four classes of segments: Vowels, symbolized here by V, are vocalic and nonconsonantal; Consonants, symbolized by C, are nonvocalic and consonantal; Liquids [r], [l], symbolized by L, are vocalic and consonantal; the Glide [h], symbolized by H, is nonvocalic

and consonantal (*). We shall be concerned solely with restrictions on these four classes; all further restrictions within the classes are disregarded here.

English morphemes can begin only with V, CV, LV, HV, CCV, CLV, and CCLV: *e.g.*, *odd*, *do*, *rue*, *who*, *stew*, *clew*, *screw*. A number of sequences are not admitted initially; *e.g.*, LCV, HLV. These constraints are reflected in the following three rules which are part of the grammar of English:

Rule MS1: If a morpheme begins with a consonant followed by a nonvocalic segment, the latter is also consonantal.

Rule MS2: If a morpheme begins with a sequence of two consonants, the third segment in the sequence is vocalic.

Rule MS3: If between the beginning of a morpheme and a liquid or a glide no vowel intervenes, the segment following the liquid or the glide is a vowel.

These rules enable us to specify uniquely a number of features in certain segment sequences; *e.g.*,

vocalic	−	−		
consonantal	+		+	

is converted by rules MS1, 2 and 3 into

vocalic	−	−	+	+
consonantal	+	+	+	−

which stands for a sequence CCLV: *e.g.*, *straw*.

The MS rules are partially ordered. If the order is not imposed they will have to be given in a much more complex form. Let us now introduce the convention that whenever a feature is not specified in a segment, a zero shall be written in the appropriate column and row. We shall say, therefore, that a zero stands for an unspecified feature, and a plus or a minus, for a specified feature. In terms of this convention the sequence of columns representing the different morphemes—*i.e.*, the input instructions for phrase structure rules (6)–(9)—will contain many zeros; indeed as many zeros as are compatible with attaining the aims of the grammar.

We define an order-relation between segment-types: We shall say that

(*) We consider the semivowels [j] as in *you* and [w] as in *woo* to be positional variants of the vowels [i] and [u], respectively.

segment-type A is « contained » in segment-type B, if and only if the following two conditions are satisfied: 1) all specified features of A are found with the identical values (the same pluses and minuses) in B; and 2) at least one feature specified in B is unspecified (has a zero) in A. The set of all elements not « contained » in any other element is called the *set of maximal segmenttypes.*

Examples:

		A	B	C
Feature	F1	+	—	+
	F2	0	+	—

A is « contained » in C. The set of maximal segment-types is {B, C}.

		A	B	C
Feature	F1	+	—	0
	F2	0	+	—

all segment-types are maximal.

It has often been observed in linguistics that the primary function of the phonemes of a language is to distinguish one morpheme from another. It is, therefore, natural to require that the set of phonemes of a language be a set of maximal segment-types. In other words, given any two phonemes of a language, it must be the case that for at least one feature, one phoneme has a plus where the other phoneme has a minus, or vice versa.

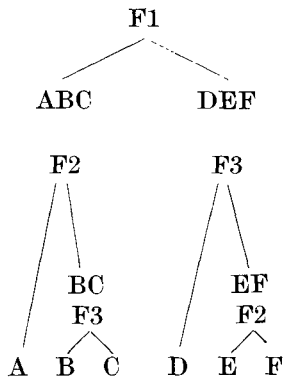
Each specified feature in a segment represents a piece of information that must be provided in the input instructions. If our grammar is a realistic picture of the language, then this information must be supplied by the speaker. Since we speak quite rapidly—at a rate which may be as high as 20 segments per second—it is only reasonable to assume that the number of specified features in the input instructions is consistently kept at a minimum. One way of approaching this desideratum is by minimizing the number of specified features per phoneme. It can be shown that if this condition is imposed on a set of maximal segment-types, it will be possible to map into a branching diagram the matrix representing the set of segment-types, in such a way that if to each node a particular feature is assigned, then each path through the diagram beginning at the initial node and ending at the end points of the branching diagram represents a phoneme.

In order to see what is involved consider the following sets of maximal segment-types.

		A	B	C
Feature	F1	+	−	0
	F2	0	+	−
	F3	−	0	+

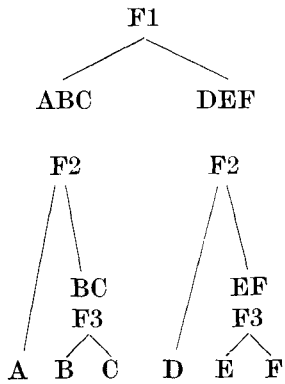
This set of maximal segment-types is not mappable into a branching diagram.

		A	B	C	D	E	F
Feature	F1	−	−	−	+	+	+
	F2	−	+	+	0	−	+
	F3	0	−	+	−	+	+



Note that in the left branch of this branching diagram, F2 precedes F3, while in the right branch the inverse order obtains. Without this reversal in the order of the features, the above set of maximal segment-types is not mappable into a branching diagram.

		A	B	C	D	E	F
Feature	F1	−	−	−	+	+	+
	F2	−	+	+	−	+	+
	F3	0	−	+	0	−	+



This set of maximal segment-types can be mapped into a branching diagram with a unique ordering of the features.

The possibility of mapping a distinctive feature matrix into a branching diagram hinges upon the existence in the matrix of at least one feature for

which there are no zeros. This feature, which must be assigned to the first node, subdivides the segment-types into two classes. The next two nodes must be assigned to features which have no zeros for any of the segments in the two sub-classes. These may be the same or different features. The same procedure must again be possible with regard to the segments in each of the four sub-classes established by the former features; etc. When a sub-class contains a single segment-type, the segment-type is fully specified, and the path through the branching diagram represents exactly its distinctive feature composition. The two conditions establish a hierarchy among the features. This hierarchy, however, need not be complete. For instance, when there are in the matrix two features which contain no zeros, there is no reason to put one feature before the other; any order will be satisfactory. Partial ordering of features for different reasons is illustrated in the second example above.

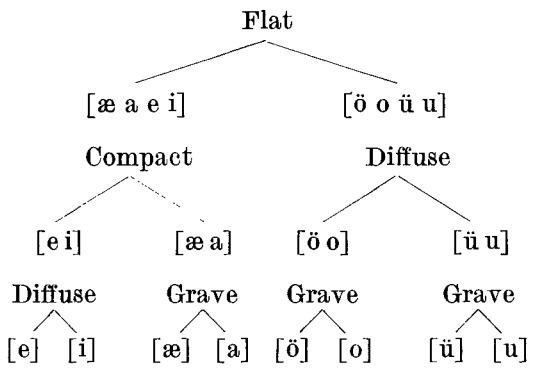
The hierarchy of features established by the two formal conditions imposed on phonemes provides an explanation for a number of observations made by linguists. It accounts, *e.g.*, for the intuition that the distinction between vowels and consonants is somehow more crucial to the phonological system than the distinction between accented and unaccented vowels, or between stops and continuants. Since in all phonological systems it happens to be the case that the features vocalic-nonvocalic and consonantal-nonconsonantal must precede all other features, it is quite natural that the segment classes established by these two features should be felt to be more central than other classifications of segments.

An interesting result of a different sort is obtained in the case of the Finnish vowel system. Finnish has the eight vowel phonemes which can be characterized by means of the following distinctive feature matrix.

	[æ]	[a]	[e]	[ö]	[o]	[i]	[ü]	[u]
flat	−	−	−	+	+	−	+	+
compact	+	+	−	−	−	−	−	−
diffuse	−	−	−	−	−	+	+	+
grave	−	+	−	−	+	−	−	+

Since, however, it is necessary to minimize the number of specified features per segment, we replace certain specified features by zeros as follows:

	[æ]	[a]	[e]	[ö]	[o]	[i]	[ü]	[u]
flat	—	—	—	+	+	—	+	+
compact	+	+	—	0	0	—	0	0
diffuse	0	0	—	—	—	+	+	+
grave-acute	—	+	0	—	+	0	—	+



This replacement of specified features by zero has, however, an interesting parallel. Finnish is one of the languages which possess *vowel harmony*; *i.e.*, there is a restriction on what vowels can occur in a single word. In the case of Finnish, a word can contain as election either from the set [æ, ö, ü, e, i] or from the set [a, o, u, e, i]. The minimal distinctive feature matrix provides us with a very elegant formula for the description of these facts; *i.e.*, a Finnish word cannot contain both grave and acute vowels. The formula holds only for the abstract representation of the phonemes as it is embodied in the matrix, for physically speaking [e] and [i] are both acute. In the construction of the Finnish word, these two phonemes, however, do not behave like other acute vowels. The formal requirements imposed on phonemes force us to treat [e] and [i] as vowels which are neutral with regard to the feature grave-acute, and indeed this is how these phonemes appear to be treated by the language.

The reasons advanced for reducing the number of specified features in the input instructions do not hold only in the case of phonemes. As we have seen in the discussion of the segment-sequences that are admitted at the beginning of an English word, under certain conditions not all features which must normally be specified in a phoneme serve to distinguish one morpheme from another. We have, however, not required that the input instructions consist entirely of phonemes. We can now take advantage of this and leave unspecified in the input instructions all features that are not phonemic. The

rules of the grammar will insure in such cases that unspecified features are specified so as to yield the correct phonetic consequences; *i.e.*, possible English utterances.

The question which we have as yet not discussed is at what point in the grammar must we place the various rules that reflect the constraints on feature combinations. At first sight it may appear desirable to place all of them at the end, after the operation of the transformational rules, since it is only at the end of the transformations that all grammatical operators—*i.e.*, symbols like « Past, » « Plural, » etc.—are converted into features or feature segments. If we were to apply the above rules before the transformations it would be necessary either to apply the same rules again, in order to handle those feature segments that were introduced by the transformational rules, or to specify many more features in the output of the transformational rules. I shall now attempt to present reasons why it is necessary to apply some rules reflecting constraints on feature combinations before the transformations.

Since it is always possible to add new words to the language the lists of morphemes must not be considered closed. The rules which reflect the constraints on feature combinations do not enable us to develop a procedure for discovering the most economical distinctive feature representation for every morpheme; this can be found only by repeated trial and error. Consequently, it is not possible to predict a priori what types of distinctive feature columns will appear in the representations of the different morphemes, for it is conceivable that a new morpheme to be introduced in the future will require for its most economical representation a distinctive feature column that is not otherwise found in the language.

The above fact has important consequences for the construction of the grammar. We have just said in effect that we do not have a way for determining what distinctive feature columns (segment-types) will appear in the terminal strings after the application of the phrase structure rules. In many languages—though perhaps not in all languages—there are certain transformational rules which require that certain features be specified. As an example consider the plural of the English noun « straw » [strˈɔ]. As was shown at the beginning of this lecture the features vocalic-nonvocalic and consonantal-nonconsonantal would be represented in this morpheme as follows:

vocalic	−	−	0	0
consonantal	+	0	+	0

In other words, in the input instruction there would be no statement regarding the nature of the last segment. In order to select the correct plural

ending for this noun, however, it is necessary to know its last segment (*). This information is contained in the rules reflecting the constraints on feature combinations; *i.e.* in rules MS1, MS2, MS3. It is necessary, therefore, to apply these rules before the rule forming the plural of nouns, or more generally before all transformational rules. I believe that the dividing line between the rules that have to be applied before the transformations—let us call them the *morpheme structure rules*—and those that have to be applied after the transformations—let us call these the *phonological rules*—can be drawn by requiring that the application of the morpheme structure rules result in segment-types which are specified to a point where the entire set of segment-types is mappable into a branching diagram in which each segment-type is represented by a distinct path through the diagram, all paths beginning at the initial node, but not necessarily ending in an end point. In other words, at this point the segment-types admitted in the representation are either phonemes or segment-types which are « contained in » phonemes. We shall call the latter segment types *archiphonemes.* Since, however, the entire set must be mappable into a branching diagram a feature specified in a phoneme can remain unspecified in an archiphoneme only if all features below it in the hierarchy established by the branching diagram also remain unspecified.

Since the morpheme structure rules must be applied before the transformations, it is natural to include them in the phrase structure level rather than set up a separate linguistic level containing just these rules. The MS rules must, therefore, be of the same structure as other phrase structure rules; they must, *e.g.*, not violate the restriction against rewriting more than one symbol in a single rule. They can not result, therefore, in the elimination of entire segments from the representation. Such rules, which are necessary in certain instances, will have to be included in another part of the grammar.

All remaining rules dealing with constraints on feature combinations are to be applied after the transformations. Since these rules differ from the transformations in two significant respects—namely, all the rules are obligatory; *i.e.*, require no external instructions to be put into operation; and the rules do not require reference to other, earlier strings in the derivation—it is simplest to set up a special linguistic level containing only these rules. We call this third linguistic level the *phonological level.* The rules of the phonological level complete the specification of the phonetic properties of the utterance in so far as these are governed by the rules of the language. Phonetic properties whose actualization is left to the free will of the speaker are not specified by these or any other rules. They are beyond the purview of the science of linguistics.

(*) The rule governing the selection of the plural endings in English is stated at the beginning of Sect. **2.**

* * *

I want to express my gratitude to my colleagues Y. BAR-HILLEL and M.P. SCHÜTZENBERGER for many hours of discussion and instruction, hours which they might well have spent more enjoyably.

This work was supported in part by the U.S. Army (Signal Corps), the U.S. Air Force (Office of Scientific Research, Air Research and Development Command), and the U.S. Navy (Office of Naval Research), and the National Science Foundation.

BIBLIOGRAPHY

N. CHOMSKY: *Syntactic Structures* (The Hague, 1957).

R. JAKOBSON, C. G. M. FANT and M. HALLE: *Preliminaries to Speech Analysis* (M.I.T. Acoustics Lab. Report 13, 1952).

R. JAKOBSON and M. HALLE: *Fundamentals of Language* (The Hague, 1956).

M. HALLE: *The Sound Pattern of Russian* (The Hague, in press).