Another interesting approach, which the present paper remains unexplored, is to calibrate the distributions of the distinctive features in time, as already proposed in Sec. 4-8. Given a long span of text trained phonemically, we write under each column a plot of phonemes, minima, and zeros on tainting its distinctive features in some other sense as in Table A. The horizontal sense of these, minima, and zeros reflects in this way can also be used to ensure "continuity" of the various features. The abilities of such squares may be written, b (m, f), y, 7, b, s, etc. 8 Given that distribution may be provided by the statistical analysis of the acoustic differences between one language another.

The statistical analysis of the phonemes their sequences in connected messages be supplemented by a analysis of dictionary, in order to understand the role of phonemes in the lexical code of a given language. 2 The comparison of the

To B. R. Carpenter's terminology, the occurrences utterance, having been studied in the Russian

It is an exhaustive statistical analysis of the phoneme structure of Russian and morphemes has been prepared by R. M. Abbot. Anomaly within the framework of the morpheme program mentioned in footnote 2.


1. Introduction. Discursive or continuous?

From the time of the transcribable alphabetic writing until the end of the nineteenth century, students of language accepted without the qualification the view that speech consisted of sequences of discrete sounds which are taken as a small number of basic types. In a way, it was held that we speak in a manner similar to the way in which we write, and the idea that there might be any simple correspondence between letters and sounds did not seem as strange in an earlier generation as it does to us. The learned Baron de Remusat 2 believed that the Hebrew letters represented the position of the tongue during the production of the corresponding sounds and to illustrate this, published a set of completely imaginary articulation profiles which have a somewhat gruesome quality about them. 3

The achievements of the natural sciences in the last century permitted as well as forced students of language to make much more detailed observations of the speaking process, with the result that a-deuctive was cast on the validity of a-eo logic all the standard no- tions firmly held by preceding generations. Speech was shown to be not a sequence of clearly separated, discrete events, but rather a continuous flow of sound, an unbroken chain of movements. 4


8. I wish to acknowledge my indebtedness to those of the colleagues and students who have served in some role possible capacity or whose views expressed in this essay were made at various stages of maturation. In particular, I wish to thank Romor Jakobson, whose influence sought to be apparent on every page, and Susan T. Healy, whose illuminating and penetrating discussions of many problems have greatly influenced my views. This work was supported in part by the Signal Corps, the Air Materiel Command, and the Naval Research Laboratory.

Although the view of language as a continuous phenomenon is simple and straightforward, it has certain inherent difficulties which make it undesirable as a basis for description, and investigations of language, phonetics as well as phonicians, have usually preferred to describe language as a sequence of discrete events. 5

Furthermore, it is not necessary that a physical phenomenon be actually discontinuous in order to break it up into a sequence of discrete events. It is possible to divide it into segments if we can show exactly how it is to be done. We shall illustrate one such technique in one segment as they apply to speech (see Chapter 10). At this point in the argument it is only necessary for us to establish the possibility of segmenting into discrete events the continuous acoustical phoneme that is speech. A person taking dictation is performing just such an operation. His hand moves the continuous acoustical wave, his hand writes words (types, if you will) sequences of discrete sym- bols, the letters. The acoustical wave, therefore, must contain units which enable human beings to perform this operation. It may well be that these are, we could presum- ably hold a machine to perform the same operation. In sum, both the continuous and the discrete representation of speech are at least in principle—be achieved by set of

It is necessary for us to know why the discrete picture language is preferable. Our answer is that it enables us to isolate many facts which on the assumption of contin-

...It is in the de...
characteristics of a given number of utterances of a given class, they may be able to make the right predictions, to some extent, even if they have not heard the specific language before. This is because the human mind, even if it is not explicitly trained in a given language, can still recognize certain patterns and make generalizations about the language based on the examples it has encountered. This is known as the cross-linguistic predictability of language, and it is a key factor in the success of natural language processing systems.

The second approach to the problem of identifying phonemes was through a quantitative analysis of the sounds used in speech. This approach was based on the assumption that all languages have a finite number of phonemes, and that these phonemes are used in a consistent and predictable way. This approach involves analyzing large amounts of speech data to identify the most common sounds and their combinations, and then using these patterns to identify the phonemes in a language.

In the end, both approaches have their limitations. The first approach is limited by the fact that it relies on a priori knowledge of the language, which is often not available. The second approach is limited by the fact that it requires large amounts of data, which is often not available either. However, both approaches have contributed to our understanding of the nature of phonemes and how they are used in speech.

The final challenge is to combine these two approaches in a way that allows us to make accurate predictions about the phonemes in a language, even if we have never heard it before. This is a complex problem, but it is one that is worth pursuing, as it has important implications for our understanding of language and the way it is used in communication.
The development of distinctive feature analysis

At our sample. Add a ninth feature to our example—(g) for instance—and three questions, no longer suffice. We now have a choice to complicate our description in one of two ways, either by adopting more answers to some or all of our questions (e.g., instead of asking "continuous? yes or no?"

or we may pose our question in the form "sequential? stop or something between?"

which in turn increases the number of questions (e.g., introducing an additional "yes or no"

"Is it strong?"

We shall refer to the first method as an increase in the "accuracy" of measurement; we shall refer to the second method as an increase in the dimensionality of measurement.

In conformity with our requirements that our terms be unambiguously meaningful, the "questions" in the preceding paragraph are of a kind to which phonemes can be provided by physical measurements regarding the presence of certain definite physical properties. Thus if we ask whether a certain phoneme is voiced or we have in mind definite acoustic measurement procedures for determining it.

It is to be noted that the number of questions is considerably smaller than the number of phonemes. The minimum number of "yes or no" questions necessary to identify a phoneme is log n. In the language of computational engineering such a degree would be required to solve for the minimum number of instructions on a natural language we shall not expect minimal redundancy, for redundancy is one of the factors that makes language resistant to noise. On the other hand we would expect natural languages to have excessive redundancy, for a very redundant language is an inefficient language.

The physical properties whose discovery is the purpose of our questions of the type

"Is there fullness of the acoustic measurement procedures? M. Halle and L. G. Jones, The Ragtime Phonemes in preparations (the new M. 4. 10., The Sound Feature of Russian (1936—1942)


R. Jakobson C. G. M. F. and M. Halle, loc. cit.

However, to complicate matters, these new attempts are actually making the number of phonemes necessary for the discovery of distinctive features and to consider the possible relations between them.


P. Lehocky's work in the Proceedings of the American Society for 25 (25) and 26 (1946) have shown that these new attempts provide interesting evidence against this problem.

The corresponding increase in the number of "significant" points of articulation. (As a quirk of fate, all three questions are sufficed for all languages.)

Thus, we now come from the old Hessian grammars but carrying them up to date, Jakobson ordered both vowels and consonants into a series of so-called "points of articulation," the points of articulation and not the entire line of features.

Vivian Jakobson demanded an acute classification of the sounds of speech. Pronunciation is a step in the implementation of this program, for these attempts attempt to state all the dimensions of phonetic description in terms of acoustic as well as articulatory criteria.

Traditional phonetics, with its primary articulatorism, i.e., its stress on major tracts such as the tongue, has been somewhat ironical since, whenever doubt was cast on the validity of these conclusions, the arguments against them were made to rest on the same kind of evidence. Jakobson's fundamental argument was that the monosyllabic language was not to be accounted for by any number of dimensions, but for the deficiency of measurement, i.e., the number of significant dimensions which have in the past been made with regard to each dimension. The dimension was only that it is the simplest possible.

It is true that these dimensions are not the same as those used in the study of speech sounds, and the new attempts have made to rest on the same kind of evidence. Jakobson's fundamental argument was that the monosyllabic language was not to be accounted for by any number of dimensions, but for the deficiency of measurement, i.e., the number of significant dimensions which have in the past been made with regard to each dimension. The dimension was only that it is the simplest possible.

The new attempts have made to rest on the same kind of evidence. Jakobson's fundamental argument was that the monosyllabic language was not to be accounted for by any number of dimensions, but for the deficiency of measurement, i.e., the number of significant dimensions which have in the past been made with regard to each dimension. The dimension was only that it is the simplest possible.

The new attempts have made to rest on the same kind of evidence. Jakobson's fundamental argument was that the monosyllabic language was not to be accounted for by any number of dimensions, but for the deficiency of measurement, i.e., the number of significant dimensions which have in the past been made with regard to each dimension. The dimension was only that it is the simplest possible.
otherwise would require much more compli- cated statements.

Finally comparisons have been made be- tween the economy of a code utilizing the dis- tinctive features and that of a theoretical code having minimal redundancy. The distinctive feature code is, it is expected, somewhat less in economy than the optimal code, but not very far below it.16

IV. PROCEDURE OF ANALYSIS
The procedure of analysis by means of the distinctive features is as follows: The seg- ments which were established or signalized by phonemic differences are subjected to an analysis in terms of the entire list of distinctive features. We obtain answers to questions such as: "Is the segment under consideration voiced? Is it a continuant? Is it for example a fricative? ..." through the entire list of fea- tures. Each segment in each of the words in our catalogue is so characterized. Segments which have the same answers are said to be the same phonemes17 and are symbolized by the same letter.

In the course of this analysis it will turn out that certain questions are not necessary for identification: i.e., in the language under consideration there will be no pair of words which are distinguished by the differences which have been investigated. Thus, for example, we shall find that in the set hi, m, n, s, x, z, all voicestop stops are fricatives (aspirated), while in the set hj, m, n, s, x, z, none of the voiceless stops are aspirated. Since this difference is associated with different contexts, it is not a primary but a redundant difference and can, therefore, be disregarded (at least as far as stops are concerned) for the present purpose if we are to establish the normal conditions for identification.

On the other hand, if a given minimally different set is a certain distinction is not repre- sented, it cannot be disregarded if it functions distinctively in another context: e.g., in Welsh there is no stress opposites to hi although the difference between [hi] and [hi] functions in many other contexts, and hence a statement about the quality of gravity distinctive fea- ture characterizing the difference between /hi/ and /hi/ must be made.18

The method of minimally different sets avoids the difficulties connected with all phone- ters and eliminates the need for reference to "phonic identity" as was already pointed out in 1933 by Twaddell.19 Since the members of minimally different sets are by definition identical, we are always comparing things which are otherwise the same, and the question of identifying phonemes by complementary distribution is entirely elimi- nated. The much discussed problem of whether we should consider the English [h] and [g] as one or two phonemes does not arise at all, because the answers from analysis to any minimally different set like [hi], p[ru], s[ib], etc., differ completely from those obtained by analyzing sign, in a set like sign, sing, s[ib], sit, etc.

V. IDENTIFICATION OF PHONEMES
The analysis just described permits us to establish the inviolability of the phonemes of the language. We can write our results in the form of a matrix in which each phoneme is given with the distinctive features which are neces- sary for its identification.

In the table below, the phonemes can thus be thought to proceed as follows:

(1) The speech is segmented into principles outlined above.
(2) Each segment is analyzed in terms of dis- tinctive features. In order to establish the correct answer it may often be necessary to refer to subsections.
(3) Each segment is identified by reference to the matrix (table) with the following indication: If the analyzer output for any given feature is positive, disregard all phonemes which in the matrix are marked negative for that feature. If the analyzer output is negative for any given feature, disregard all phonemes which are marked positive for that feature. Do nothing about phonemes for which the particular feature is not distinctive (i.e., which have marks in the table). At the end of such an analysis there will be only one phoneme which has not been excluded from consideration — it is the phoneme under analysis.

VI. CONCLUSION: FUNDAMENTAL PROPERTIES OF NATURAL LANGUAGES
The models of language which has been presented here has the following properties:

(1) Language consists of discrete units of short duration and has physical requirements (the phonemes).
(2) In their function as signaling devices the phonemes can be viewed as simultaneous implementations of a number of attributes — the distinctive features.
(3) The distinctive features are, with a single exception, the distinctive fatures.
(4) No language utilizes all the distinctive features.
(5) No language has as many phonemes as there are possible combinations of the utilized distinctive features.

---

19. "This constitutes a definition of phonemic identity.

---

"In cases where a feature is neutralized in a certain context, e.g., in English in the above example, the question arises: To which segment [hi] or [hi] is the above context to be assigned? I believe that this is to be decided by physical rea- soning: there is no reason to expect that one segment."

---

"Footnote provides two things: it establishes the status of the morphemes in our matrix tables and makes allowance for what is known as "free varia- tion": the phonemes which are not marked as negative may not be present in various utterances of a given phoneme, e.g., quality in vowels in Midwestern dialects of American English."
TABLE I
Matrix Showing the Phonemes and the Distinctive Features of Standard Literary German

|         | m | p | B | f | v | t | d | z | k | g | f | j | L | Ú | Š | Č | Ž | Č | E | Š | Č | Ž | Č | E | Š | Č | Ž | Č | E |
| Vowels  | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| Consonantal vs. | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| non-consonantal | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Consonantal vs. | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| non-consonantal | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Consonantal vs. | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| non-consonantal | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Consonantal vs. | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| non-consonantal | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Consonantal vs. | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| non-consonantal | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

The symbol indicates an intermediate degree of contrast, for the feature of completeness is tertiary for vowels. The vowels are further differentiated into long vs. short. Among short vowels the distinction between /i/ and /e/ is non-phonemic.

This monograph, the first volume of a new linguistic series, contains a joint essay on phonology summarizing various theoretical and empirical investigations that these authors have reported on during the last few years and an individual contribution of Jakobson's that ranges widely over many patterns of language and pathological disturbances, literature, and general symbolic behavior. Both essays are written in a rather picturesque and inexact style which, to me at least, presents a bar to comprehension. It is difficult to determine which statements are empirical hypotheses and which are true by definition, or just what conditions the authors require a phonemic transcription to meet. Furthermore, the justification that the authors give for their own positions is often vague and unconcerning. I think that this is unfortunate, since it seems to me that much can be said for their approach to phonological problems. At the risk of possible misunderstanding of their position, I will try to state what I think is a concentration in the essay entitled "Phonology and Phonetics," and to see where in fact it diverges from certain other phonological theories.

Suppose that we have a set of utterances belonging to some language. Suppose further that we know which pairs of utterances are phonemically distinct in this language; i.e., a test is available for classifying utterances into sets of repetitions. In addition, we have available a set of physically defined features or phonetic qualities that can be used in describing these utterances. We may now divide the utterances into segments, assigning to each segment as its value the set of features that characterize it. The utterances must be segmented in such a way that the following conditions are met:

(1) If two utterances are phonemically distinct, then the sequences of values assigned to these utterances must differ in at least one place.

Practically all approaches to phonological analysis have at least this much framework in common. Differences appear when we try to investigate the nature of the physical factors and the principles by which segments are assigned to the same phoneme. These are of course independent problems. Jakobson and Halle proceed in the following way: First, they present a certain set of physical features, defined independently of any particular language. That is, these features are part of the general definition of Language—they are part of the conceptual framework, the

CE, "On the logical structures of language in their phonemic aspect," in Linguistic Problems (1955) pp. 201-244


NOAM CHOMSKY