Another very promising approach, which the present must remain unexplored, is to ·ulate the distributions of the distinctive ures in time, as already proposed in Sec- 4(a). Given a long sample of text tran- bed phonemically, we write under each nbol a column of pluses, minuses, and zeros resenting its distinctive features in some ılar order (as in Table A). The horizontal ıences of pluses, minuses, and zeros pro- ed in this way can then be used to measure "continuity" of the various features. The abilities of such sequences may be written m), $p_-(m)$, $p_0(m)$, where $m = 1, 2, 3$, etc. It bvious that such distributions may provide sis for the statistical specification of the nemic differences between one language another.

The statistical analysis of the phonemes their sequences in connected messages t be supplemented by a similar analysis of dictionary, in order to understand the dis- ıtion of phonemes in the lexical code of given language.[12] The comparison of the two sets of data is certain to be most instructive. The statistical analysis of the dictionary permits us to draw conclusions about the phoneme sequences peculiar to different types of morphemes and to words of different grammatical categories.[13] Furthermore, it forms the basis for definitive statements about phoneme combinations with probabilities of 1 and 0; for no phoneme sequence can occur in messages if it is not provided by the code.

Finally, among the problems which remain to be investigated are those transitional probabilities which operate backwards, i.e., which depend not on earlier but on subsequent events, or, in linguistic terms, not on the progressive but on the regressive action of phonemes in a sequence. The comparison of these two sets of statistics is very important, because it is obvious that for different types of sequences the predictability is greater in one direction than in the other. Analysis of such data will provide the most solid basis for setting up a statistical model of the syllable as a recurrent link in the chain of speech.

[12]In R. Carnap's terminology, the occurrences ıonemes, having been studied in the Russian -*events*, are to be investigated in the word- *ns*, just as we have here studied the occurrences stinctive features in the phoneme-*designs;* cf. *duction to Semantics*, 3 (Cambridge, Mass.,). Charles S. Peirce, the founder of modern ıtics would say that besides the application of honemic *legisigns* within the lexical *sinsigns*, ın application must be scrutinized again within lexical *legisigns;* cf. his *Collected Papers*, 2.245–247 (Cambridge, Mass., 1932).

[13]An exhaustive statistical analysis of the phonemic structure of Russian root morphemes has been prepared by Robert Abernathy within the framework of the research program mentioned in footnote 2.

Written jointly in Cambridge, Mass., 1952; published in *Language* 29 (1953).

32

THE STRATEGY OF PHONEMICS*

MORRIS HALLE**

I. INTRODUCTION. DISCRETE OR CONTINUOUS?

From the time of the invention of alphabetic writing until the end of the nineteenth century, students of language accepted without question the view that speech consisted of sequences of discrete sounds which are tokens of a small number of basic types. It was generally held that we speak in a manner quite similar to the way in which we write, and the idea that there might not be any simple, one-to-one correspondence between letters and sounds did not seem as obvious to an earlier generation as it does to us. The learned Baron van Helmont even 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.[1]

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 grave doubts were cast on the validity of practically all the standard notions 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.[2]

Although the view of language as a continuous phenomenon is simple and straightforward from a strictly physical standpoint, it has certain inherent difficulties which make it undesirable as a basis for descriptions, and investigators of language, phoneticians as well as physicists, have usually preferred to describe language as a sequence of discrete events.[3]

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 later state some of the conditions as they apply to speech (see Chapter II). At this point in the argument it is only necessary for us to establish the possibility of segmenting into discrete events the continuous acoustical phenomenon that is speech. A person taking dictation is performing just such an operation. His eardrums receive the continuous acoustical wave, his hand writes (types, if you will) sequences of discrete symbols, the letters. The acoustical wave, therefore, must contain clues which enable human beings to perform this operation. If we could state what these clues are, we could presumably build a machine to perform the same operation. In sum, both the continuous and the discrete representation of speech can—at least in principle—be achieved by a set of

[1]Cf. reproduction of one of his figures in H. Dudley and T. Tarnoczy, "The Speaking Machine of Wolfgang von Kempelen," *Journal of the Acoustical Society of America* 22.151–167 (1950).

[2]P. Menzerath and A. de Lacerda, *Koartikulation, Steuerung und Lautabgrenzung* (Berlin-Bonn, 1933).

[3]The major exception is the work of some communications engineers on long time spectra of speech; e.g., H. K. Dunn and S. D. White, "Statistical Measurements on Conversational Speech," *Journal of the Acoustical Society of America* 11.278–288 (1939), and L. G. Kraft, "Correlation Function Analysis," *ibid.*, 22.762–764 (1950).

*Reprinted from *Word* 10.197–209 (1954) by permission.

**I wish to acknowledge my indebtedness to those of my colleagues and students who have served as none too passive guinea pigs on whom the ideas expressed in this essay were tested at various stages of maturation. In particular I want to thank Roman Jakobson, whose influence ought to be apparent on every page, and Noam Chomsky, 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 Office of Naval Research.

physical operations. We can, therefore, assert that both views are meaningful in an empirical, physical sense.[4]

It is now necessary for us to show why the discrete picture of language is preferable. Our answer is that it enables us to account for many facts which on the assumption of continuity would be extremely difficult, if not impossible, to explain. The grossest of these is the well-known fact that speech is perfectly intelligible in the presence of very great disturbances. We shall call this property of speech "resistance to noise."

In a recent paper B. Mandelbrot[5] has shown that it is impossible to account for the high resistance to noise of speech on the basis of a continuous view. If linguistic messages (utterances) be thought of as continuous, the correction of errors in the reception cannot begin until the entire message is received, which would make correction well nigh impossible, certainly infinitely more difficult than it actually is. On the other hand, if a discrete view be adopted, correction of errors can begin upon receipt of each discrete unit (quantum), since the discrete units in the language are just a small fraction of all possible things that the ear can receive.

Mandelbrot investigated in detail the consequences of the discrete character of language only on one level, that of words. The necessity for discrete units on other levels is implicit in his argument. The words themselves are thus viewed as being composed of discrete components, usually known as *morphemes*, which in turn consist of other discrete units, the *phonemes*.

The reasons for our using the phoneme as the smallest quantum of language in preference to other units that have been suggested, as, e.g., the syllable, are as follows:

(1) Since the larger units (words, morphemes, syllables) can usually be subdivided into phonemes, it follows that the number of different phonemes necessary for describing a given set of messages will be smaller than that of the larger units.

(2) The phonemes are extremely useful (if not the only possible) units when it is necessary to describe other facts of languages, such as rules of grammar, regularities of linguistic change, development of language in children, pathological manifestations of language, poetic devices such as rime, assonance, etc. One can easily convince oneself of this by trying to state such a simple grammatical rule as the distribution of the English past tense suffixes /-d/, /-ɪd/ or /-t/ in terms of syllables or words.

We want to insist on this point in particular. To us the major criterion for the applicability of a certain category to linguistic description is whether or not this category yields simple statements not only on the particular level for which it was introduced, but on all levels which are pertinent to descriptions of a language. It always must satisfy a multiplicity of criteria.

We have chosen to represent language as a sequence of discrete events and thereby to complicate our description on the physical level, because the physically simpler, continuous view would have made impossible other statements about language which we would like to make; e.g., to give a simple description of grammar. The nonlinguist need not follow us there, and as a matter of fact a high-fidelity recording of human speech involves neither phonemes nor words nor any other linguistic quantization.

In sum, then, the traditional view of language as a sequence of discrete events of phoneme length is empirically meaningful and has certain clear advantages not only over the purely physical view of language as consisting of utterances which themselves are continuous, but also over other proposed schemes of quantization (syllable, for instance).

II. THE PHONEME

It is now necessary to sketch a procedure whereby we would be guaranteed to arrive at answers to two questions: (1) How many phonemes does our language possess? (2) What acoustical properties are used for the identification of the phonemes?

Since in our view the most important property of language is that it serves as an

[4] "A term (predicate) is a legitimate scientific term (has cognitive content, is empirically meaningful) if and only if a sentence applying the term to a given instance can possibly be confirmed to at least some degree." R. Carnap, "Truth and Confirmation," in H. Feigl and W. Sellars, *Readings in Philosophical Analysis* (New York, 1949), p. 123.

[5] B. Mandelbrot, "Structure formelle des langues et communication," *Word* 10.1–27 (1954).

artifact for the transmission of information, it is evident that we must be able to record those features of the speech wave which by themselves serve to convey information. It is to be stressed that "information" is used here in its technical meaning, which involves a selection of one from an ensemble of possible messages. In other words, we must be able to detect those properties of the speech wave which enable us to tell that a particular utterance is *bill* and not *pill* or *dill* or *gill*, etc.[6] We have placed ourselves hereby in the position of the receiver, and our description is to be made primarily from the receiver's point of view.[7]

In the preceding paragraph we have spoken of "an ensemble of *possible* messages." This means that we must be prepared to deal only with those sound sequences which are possible in our language, and not with all sounds that man might emit; i.e., we need to establish the conditions sufficient for decoding the messages which we might receive, for speakers of a language are in the position of people who communicate by means of a common code. The linguist's role is in part similar to that of the cryptanalyst who must describe the code from observations made on both the messages and the behavior of the users of the code (not excluding his own).[8]

Since, if taken as physical events pure and simple, no two utterances are alike, the decision of whether two utterances are the same or different can only be made on another level, namely, by observing the behavior of the users of the code, including their verbal behavior. It has been shown by Harris[9] that this decision can be established without recourse to "meaning" or "difference of meaning." Two utterances as spoken by the same informant are recorded on tape and a test tape is prepared by re-recording the original two utterances in a random order. The test tape may thus contain fifty or sixty re-recordings of the original utterances. The two original utterances are played to an audience of native speakers who are instructed to make a check on their ballots (or signify in some other manner) whenever they hear the first utterance during the playing of the test tape. If the utterances are "different," i.e., phonemically distinct, the audience will obtain an almost perfect score; if the utterances are homophonous the score will be in the vicinity of 50 percent. Since all linguistic descriptions are based on analyses of a finite number of utterances, it is possible, at least in principle, to subject the entire corpus to this test. This is, however, not necessary since shortcuts can be easily devised and justified. A particular advantage of this method is the neat solution which it provides for the difficulties connected with phonemically distinct utterances having the same meaning, like /ekən'amɪks/ and /ikən'amɪks/.

Having thus established *sameness* and *difference* among the utterances in his corpus, the linguist must next outline a procedure for the identification of the elementary discrete units that compose the utterances; i.e., of the phonemes.

In the past, the problem of identification has been approached essentially from two directions. The first approach is formulated in Potter and Steinberg's article, "Toward the Specification of Speech": "If different speakers are asked to speak one of the vowels, the utterances will, of course, be different. *The problem is to determine those physical properties that are invariant in the several utterances that enable the ear to identify them as a given vowel.*"[10] In other words, these investigators assume that from a study of the physical

[6] Cf. F. de Saussure, *Cours de linguistique générale* (Paris, 1949), p. 163: "Ce qui importe dans le mot, ce n'est pas le son lui-même, mais les différences phoniques qui permettent de distinguer ce mot de tous les autres, car ce sont elles qui portent la signification."

[7] For a detailed discussion of this point see B. Mandelbrot, *op. cit.*

[8] Whether the linguist can "break the code," just by investigating records without recourse to observations of the speakers' behavior, has been the subject of spirited debate. While I doubt the possibility I do not take a definite position in this matter.

[9] Z. S. Harris, *Methods in Structural Linguistics* (Chicago, 1951), pp. 32 ff. It is to be noted here that in Harris' presentation recourse is had to meaning in the instructions given the original speaker, where the two utterances are identified "by translation or otherwise." On the relation between meaning and translation cf. R. Jakobson, *Sound and Meaning* (in preparation).

[10] *Journal of the Acoustical Society of America* 22.807 (1950).

characteristics of a great number of utterances of a given vowel they will be able to isolate the invariant properties which presumably serve the human being as cues for the correct identification. The attention is focused here on the properties which all repetitions of a given phoneme have in common, a sort of common denominator.

The second approach to the problem of identification was already foreshadowed in the quotation from de Saussure (footnote 6); we can add another which reads: "Or ce qui les [the phonemes] caractérise, ce n'est pas, comme on pourrait le croire, leur qualité propre et positive, mais simplement le fait qu'ils ne se confondent pas entre eux." Primary attention is here directed not to the properties which all repetitions of a given phoneme may possess in common, but rather to the properties which differentiate each repetition of a given phoneme from all other possible phonemes which might have been uttered in its stead; i.e., the focus is on the distinctive differences instead of the similarities. Focusing on the distinctive difference presupposes, of course, a knowledge of all the possible judgments which one might be expected to make; it presupposes a knowledge not only of the phoneme under investigation but also of all the phonemes in the language.

The two approaches might be illustrated by the following analogy: It is necessary to find a car in a parking lot. If it is not known what other cars are in the parking lot, then the only sufficient way to specify the wanted car is by giving its total description: make, year, model, color, and such other distinguishing characteristics as bumps, scratches and broken windshield. If on the other hand, it were known that the wanted car is the only red car in the parking lot it would be sufficient to specify this one distinctive property in order to find the wanted car. It is, of course, true that the first specification is sufficient in all cases—however, under the given conditions the second method is doubtless the more economical. In the case of language it is evident that the second method, i.e., concentration upon distinctive differences rather than upon common properties, is a more reasonable approach.

One technique for establishing the number and properties of the phonemes of a language consists in arranging the utterances in *minimally different sets*. These are sets of different utterances (preferably of short words) which differ from one another by only one phoneme but are alike in every other respect; for example, sets like:

(a) *bin, pin, kin, Lynn, tin, din, sin, thin, fin,* etc.
(b) *bin, bean, boon, bun, ban, Ben,* etc.
(c) *bin, bit, big, bill, bib, bid, bing,* etc.

It may be objected that the ability to make such an arrangement necessarily presupposes a knowledge of the solution: one cannot arrange utterances in sets in which each member differs from all others by one phoneme without first knowing the phonemes.

The objection overlooks one other way in which this particular arrangement might be arrived at, namely, by an exhaustive examination of all possible arrangements. Since the number of utterances on which our analysis is based is finite, there is also a finite number of ways in which these can be arranged; i.e., in a finite time we could generate all possible arrangements. If we now possessed criteria for determining which arrangement actually consists of minimally different sets, we might conceivably check through all possible arrangements with the guarantee that in a finite time we should end up with the desired arrangement.

As a matter of fact we possess such criteria (see the following discussion), and it is by means of these criteria that we establish the minimally different sets. The procedure is as follows: each arrangement is provisionally assumed to contain nothing but minimally different sets; i.e., all words in it are assumed to differ from one another by one phoneme only and to be alike in every other respect. In the overwhelming majority of cases the incompatibility of this assumption with the requirements for simple solutions are immediately obvious. Thus, for example, we never consider minimally distinct, sets like *bee, antidisestablishmentarianism,* and *psychic,* for it is immediately apparent that no conceivable manner of segmentation of these words will yield units which even remotely resemble phonemes. (Cf. requirements of physical uniformity stated below.) In more difficult cases, as, e.g., whether the affricate [tʃ] as in *chew* should be considered as one or two phonemes (or in our terminology, whether *chew* belongs in the set *shoe, sue, too, do,* etc., or both in the set *true, tew,* and in the one-member set *chew*), both assumptions are actually tested and the interpretation yielding the simpler over-all description is chosen.[11]

The criteria which phonemes have to meet are the following:[12]

(*1*) They must be of relatively simple physical structure: within a phoneme segment there cannot be (a) turning on and off of the exciting source or sources; (b) a switching on and off of the nasal resonances; (c) sudden severe drops in over-all level; (d) certain changes in the formant positions; (e) sharp changes in the bandwidths of the formants (exact specifications have still to be given here).
(*2*) The segments considered different must be replaceable in the members of the set without destruction of intelligibility; i.e., by careful cutting and splicing it must be possible to transform minimally different utterances like *bill* and *mill* into each other, as has been done by P. Menzerath,[13] Carol Schatz[14] and others.
(*3*) There must be as many phoneme intervals in each utterance as will enable us to distinguish each utterance from each other utterance which is not a repetition of the former, and no more.[15]
(*4*) The identification in terms of the distinctive features (see below) must be possible and show no inconsistencies.

It is to be noted that in the above specifications there are involved both physical and distributional properties. In recent years the former have tended to be under-emphasized with the result that phonemic solutions often have had a very artificial appearance.

By application of criteria 1, 2, and 4, above, we establish where the phoneme boundaries approximately lie. However, we have not as yet identified a single phoneme.

III. DISTINCTIVE FEATURES

It has already been stated that the method of identification to be adopted here will concentrate on the differences existing between the phonemes, and not upon the properties common to all utterances of a given phoneme. Our first step is to characterize the differences between the members of a minimally different set: we intend to state how *bin* differs from *pin*, from *tin*, from *din*, etc.

Given n entities there may be as many as $n(n-1)/2$ differences between them. This would be the case if none of them had any properties in common: e.g., if we assumed that the following eight phonemes /p/, /t/, /f/, /s/, /b/, /d/, /v/, /z/ had no properties in common, then in order to characterize each one of them by the method of differences adopted here, we would need a total of $8 \times 7/2 = 28$ statements. We would have to be prepared to answer questions of the form "Is the phoneme under consideration /p/ or /t/, /p/ or /b/, etc.?" Once we have discovered, however, that some of the phonemes possess common properties which others lack, we can pose much more "perceptive questions"; e.g., "Is the phoneme under consideration voiced? continuant?"

Since to each of the questions two answers (yes or no) are possible, we can identify by means of three questions eight different entities corresponding to the eight phonemes

[11] The necessity of setting up a special set containing just the phoneme in question (since stops cannot precede fricatives in English unless the affricate be considered a cluster) would incline us towards a monophonematic interpretation. Further complications which arise in other parts of the description would cause us to prefer this interpretation even more.

[12] The first two requirements are of a physical nature and are introduced specifically to account for two facts: (a) that a fairly rough segmentation, the so-called phonetic transcription, can be made just by listening to samples of the language, and (b) the remarkable agreement of various phoneticians on how to segment a sequence. On this point see the discussion in K. L. Pike, *Phonetics* (Ann Arbor, 1943), pp. 42–55.

[13] "Neue Untersuchungen zur Lautabgrenzung und Wortsynthese mit Hilfe von Tonfilmaufnahmen," *Mélanges de linguistique et de philologie offerts à Jacques van Ginneken* (Paris, 1937), pp. 35–41.

[14] Verbal communication.

[15] This is a slightly paraphrased version of a statement by Z. S. Harris, *op. cit.*, p. 43. On homophones see *ibid.*, pp. 32 ff.

of our sample. Add a ninth phoneme to our example—/t͡s/, for instance—and the three questions no longer suffice. We now have a choice to complicate our description in one of two ways, either by admitting more than two answers to some or all of our questions (e.g., instead of asking "continuant? yes? or no?" we may now pose the question in the form "continuant? stop? or something in between?" thereby accepting three answers) or by increasing the number of questions (e.g., by introducing an additional "yes or no" question, like "Is it strident?"). 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 requirement that our terms be empirically meaningful, the "questions" in the preceding paragraph are of a kind to which answers can be provided by physical measurement; i.e., they are questions regarding the presence of certain definite physical properties. Thus if we ask whether a certain phoneme is voiced we have in mind definite acoustical measurement procedures to determine it.[16]

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 n phonemes is $\log_2 n$. In the language of communications engineering such a description would be said to possess minimal redundancy.[17] In a natural language we should 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 not to have excessive redundancy, for a very redundant language is an inefficient language.

The physical properties whose discovery is the purpose of the "questions" of the preceding paragraphs have received a detailed discussion in another place.[18] They are referred to there by the term *distinctive features,* which will also be used in the remaining part of this essay.

It may justifiably be asked: "How does the linguist know what 'questions' to pose?" The answer is much like the one we gave when we discussed the establishment of phoneme boundaries. As yet, we have no procedure for arriving at the correct set of "questions"; we can only guess. We do, however, possess a method for checking whether or not any proposed set fits our requirements for a simple description of language on all levels.

Many distinctive features, though under different names, have long been used in phonetics. They are implicit in the work of the Hindu grammarians as well as in the works of Western phoneticians. In spite of their apparently independent development these two major traditions show certain striking similarities. Both traditions agree in assigning great importance to the "point" of articulation, the place in the vocal tract where the stricture is narrowest. They also agree to some extent in their treatment of the "manner" of articulation and of nasality. There is, however, one striking difference between the two traditions. The Hindus used the "point" of articulation for the classification of both the vowels and the consonants, while in the West the vowels were classified according to the so-called "vowel triangle" (a two-dimensional classification).

Thus the picture which the classification of the sounds presented was a mixed one: most features were two-valued (for example, sounds were said to be either nasal or nonnasal; voiced or unvoiced; aspirated or unaspirated, etc.); a few others were multivalued. Further complications were introduced in the course of the nineteenth and twentieth centuries when it was discovered that sounds which had been thought to differ only in their manner of articulation differed also in their place of articulation. There were even suggestions in the literature to make the place of articulation the primary variable and to consider the other features as redundant. This led,

[16]For details of the acoustical measurement procedures see M. Halle and L. G. Jones, *The Russian Phonemes* (in preparation). [See now M. Halle, *The Sound Pattern of Russian* (1959)—VBM.]

[17]On the concept of redundancy see C. Shannon and W. Weaver, *The Mathematical Theory of Communication* (Urbana, 1949), and R. Jakobson, C. G. M. Fant, and M. Halle, *Preliminaries to Speech Analysis*, Technical Report No. 13, Acoustics Laboratory, M.I.T., May 1952, pp. 4–8.

[18]R. Jakobson, C. G. M. Fant, and M. Halle, *loc. cit.*

however, to complications, once it was actually put into practice. For example, in French it may be said that the difference between /g/ and /ɲ/ lies in their different points of articulation: /g/ is velar and /ɲ/ prepalatal. It was thought that by this increase in the number of points of articulation, the introduction of an additional dimension, i.e., nasality, would be avoided. French also possesses the phonemes /b/ and /m/, which have identical points of articulation. The dimension of nasality, therefore, must be introduced anyway to distinguish between /b/ and /m/ and might as well be used in the case of /g/ and /ɲ/, for it is patently uneconomical to describe the labial consonants differently from the velars, especially if it is possible to avoid this.

These difficulties were resolved by Roman Jakobson.[19] In making descriptions one usually has the choice between using few dimensions with many significant values and using many dimensions with few significant values. In other words one can trade accuracy of measurement for dimensionality. Reduced to its simplest terms, Jakobson's fundamental argument was that the most satisfactory description of a language would be obtained by using as many dimensions (features) as necessary, but decreasing the accuracy of measurement, i.e., the number of significant decisions which have to be made with regard to each dimension. The dichotomous scale, which underlies the distinctive features, has minimal accuracy of measurement: it is in this sense that it is the simplest possible.[20]

Jakobson suggested that the consonants be subdivided into "strident" vs. "mellow." He thus obtained a fourfold division of consonants where formerly there was only a twofold one: the stops and the continuants were each subdivided into strident and mellow. The increase in dimensionality brought with it a corresponding decrease in the number of "significant" points of articulation. (As a matter of fact it was shown that four such points sufficed for all languages.)

Then, as if taking his cue from the old Hindu grammarians but turning them upside down, Jakobson ordered both vowels and consonants according to a single principle, which, in conformity with his fundamental demand for a substitution of dimensionality for accuracy, he found in the two-dimensional vowel triangle rather than in the one-dimensional, multivalued "points of articulation" parameter.[21]

Finally Jakobson demanded an acoustical classification of the sounds of speech. *Preliminaries* is a step in the implementation of this program, for there the attempt is made to state all the dimensions of phonetic descriptions in terms of acoustical as well as articulatory criteria.

Traditional phonetics was primarily an articulatory phonetics. From time to time attempts were made to translate the articulatory terms into acoustical. Since results were not always convincing, doubts were voiced as to whether this could be done at all.[22] There are two arguments in favor of the possibility of translating into acoustical terms uniformities which have been observed most clearly and first on the articulatory level. The first argument is quite simple: given a certain geometrical configuration and excitation of a resonator, its acoustical output is entirely predictable. Hence, all other things being equal, any uniformity on the articulatory side *must* have a statable acoustical counterpart. The second argument appeals to the fact that the phoneticians who picked a particular articulatory uniformity as a distinctive feature over a whole series of others which they might have chosen were guided by the observation that these functioned as perceptually distinctive marks, and hence must also have existence on the acoustical level.

The distinctive features in the formulation which is given in *Preliminaries* are, with one

[19]R. Jakobson, "Observations sur le classement phonologique des consonnes" [28], *Proceedings of the Third International Congress of Phonetic Sciences* (Ghent, 1939), pp. 34–41, and *Kindersprache, Aphasie und allgemeine Lautgesetze, Språkvetenskapliga sällskapets i Uppsala förhandligar* (1940–1942).

[20]I. Pollack's recent work in *Journal of the Acoustical Society of America* 25 (1953) and 26 (1954) (see footnote 24 below) provides interesting evidence relevant to this problem.

[21]For details see *Preliminaries* and R. Jakobson's forthcoming *Sound and Meaning*.

[22]Cf. E. Fischer-Jørgensen, "The Phonetic Basis for Identification of Phonemic Elements," *Journal of the Acoustical Society of America* 24. 614–615 (1952).

exception, binary in structure. This is to be understood as an empirical proposition: If the differentiating phonemes in a minimally different set were to be measured for a certain distinctive feature (e.g., degree of voicing), the results would cluster about two values: one for the voiced and the other for the unvoiced consonants. E. Zwirner's experiments with German vowels[23] do not constitute a counter-example that proves the incorrectness of the hypothesis of binarity, because Zwirner did not compare vowels in identical context only (he plotted all vowels in his sample on a single graph) and the above proposition applies only to minimally different sets, i.e., to contexts identical by definition.

The success of the distinctive features as a framework for linguistic descriptions in terms of which a host of difficult linguistic problems can be readily explained is the major reason for their adoption. Recent work in psycho-acoustics has provided further support for the distinctive feature model. Studies by Pollack and others on the transmission of information by multidimensional auditory stimuli show that best results "are obtained when each dimension [is] crudely subdivided into two alternative states. Finer subdivision of each dimension does not produce a proportional gain in information transmission with the display."[24] It is only reasonable to assume that natural languages are constructed in a way fairly closely approximating optimal conditions for auditory transmission of information among human beings.

To justify still further the use of the distinctive features as descriptive parameters we have given in *Preliminaries* under the subheading "Occurrence" examples from the most diverse languages in which the distinctive features provide a convenient framework for the expression of observations which otherwise would require much more complicated statements.

Finally comparisons have been made between the economy of a code utilizing the distinctive features and that of a theoretical code having minimal redundancy. The distinctive-feature code is, as expected, somewhat lower in economy than the optimal code, but not very far below it.[25]

IV. PROCEDURE OF ANALYSIS

The procedure of analysis by means of the distinctive features is as follows: The segments which were established as signalizing 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 fortis (aspirated)? . . . ." through the entire list of features. 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 phoneme[26] 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 the given question has in view. Thus, for example, we shall find that in the set *bin, pin, din, tin*, etc., all voiceless stops are fortis (aspirated), while in the set *spin, skin*, etc., none of the voiceless stops is 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, which is to establish the minimal conditions for identification.

On the other hand, if in a given minimally different set a certain distinction is not represented, it cannot be disregarded if it functions distinctively in another context: e.g., in English there is no *stin* opposed to *spin* although

[23] E. Zwirner, "Phonologische und phonometrische Probleme der Quantität," *Proceedings of the Third International Congress of Phonetic Sciences* (Ghent, 1939), pp. 57–66.

[24] I. Pollack and H. Ficks, "The Information of Elementary Multidimensional Auditory Displays," *Program of Meeting of the Acoustical Society of America, B, 2*, May 7–9, 1953, Philadelphia. This paper has now been published in full in *Journal of the Acoustical Society of America* 26.155–158 (1954).

[25] Cf. E. C. Cherry, M. Halle, and R. Jakobson, "Toward the Logical Description of Languages in Their Phonemic Aspect" [31], *Language* 29.41 [328] (1953).

[26] This constitutes a definition of "phonetic similarity."

the difference between /p/ and /t/ functions in many other contexts, and hence a statement about the quality of gravity (distinctive feature characterizing the difference between /p/ and /t/) must be made.[27]

The method of minimally different sets avoids the difficulties connected with allophones and eliminates the need for reference to "phonetic identity" as was already pointed out in 1935 by Twaddell.[28] Since the members of minimally different sets are by definition identical contexts, we are always comparing things which are otherwise the same, and the question of identifying phonemes purely by complementary distribution is entirely eliminated. The much discussed problem of whether we should consider the English [h] and [ŋ] as one or two phonemes does not arise at all, because the answers from analyzing *hill* in a minimally different set like *hill, pill, till, bill*, etc., differ completely from those obtained by analyzing *sing*, in a set like *sing, sin, sick, sit*, etc.

V. IDENTIFICATION OF PHONEMES

The analysis just described permits us to establish the inventory 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 necessary for its identification.

The identification of phonemes can then be thought to proceed as follows:

(*1*) The speech is segmented according to principles outlined above.
(*2*) Each segment is analysed in terms of distinctive features. In order to establish the correct answer it may often be necessary to refer to adjacent segments.
(*3*) Each segment is identified by reference to the matrix (see table) with the following instructions: 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 noughts in the table).[29] 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 model of language which has been presented here has the following properties:

(*1*) Language consists of discrete units of short duration which meet certain 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, binary.
(*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.

[27] In cases where a feature is neutralized in a certain context, as, e.g., in English in the above example, the question arises: To which phoneme, /p/ or /t/, is the above consonant to be assigned? I believe that this is to be decided by physical measurement.

[28] W. F. Twaddell, *On Defining the Phoneme*, Language Monograph No. 16 (1935). See also R. Jakobson, "On the Identification of Phonemic Entities" [30], *Travaux du Cercle Linguistique de Copenhague* 5.205–213 (1949).

[29] This provision does two things: it establishes the status of the noughts in our matrix tables and makes allowance for what is known as "free variation," where a feature, being nondistinctive, may or may not be present in various utterances of a given phoneme, e.g., nasality 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	f̂	n	t	d	s	z	ŝ	k	g	x	ʃ	r	l	u	o	a	ü	ö	i	e	æ	h
Vocalic vs. non-vocalic	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+	+	+	+	+	+	+	+	+	+	−
Consonantal vs. non-consonantal	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	−	−	−	−	−	−	−	−	−
Compact vs. diffuse	0	−	−	−	−	−	0	−	−	−	−	−	+	+	+	+	0	0	−	±	+	−	±	−	±	+	0
Grave vs. acute	+	+	+	+	+	+	−	−	−	−	−	−	+	+	+	−	0	0	+	+	+	−	−	−	−	−	0
Flat vs. plain	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	+	+	0	+	+	−	−	0	0
Nasal vs. non-nasal	+	−	−	−	−	−	+	−	−	−	−	−	−	−	−	−	0	0	0	0	0	0	0	0	0	0	0
Continuant vs. interrupted	0	−	−	+	+	−	0	−	−	+	+	−	−	−	+	0	−	+	0	0	0	0	0	0	0	0	0
Strident vs. mellow	0	−	0	0	0	+	0	−	0	0	0	+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Tense vs. lax	0	+	−	+	−	0	0	+	−	+	−	0	+	−	0	0	0	0	0	0	0	0	0	0	0	0	0

The symbol ± indicates an intermediate degree of compactness, for the feature of compactness is ternary for vowels.

The vowels are further differentiated into long vs. short. Among short vowels the distinction between /e/ and /æ/ is non-phonemic.

33

REVIEW OF *FUNDAMENTALS OF LANGUAGE**

BY ROMAN JAKOBSON AND MORRIS HALLE.
('S-Gravenhage: Mouton and Co. [Janua Linguarum, No. 1], 1956. ix + 87 pp.)

NOAM CHOMSKY

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[1] and an individual contribution of Jakobson's that ranges widely over many problems of language and pathological disturbances, literature, and general symbolic behavior. Both essays are written in a rather picturesque and inexplicit 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 unconvincing. 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 it is, concentrating on 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[2] 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 condition is 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 go on to investigate the nature of the physical features and the principles by which segments are assigned to the same phoneme. These are of course interrelated 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

[1] Cf., particularly Jakobson, Fant, and Halle, *Preliminaries to Speech Analysis*, M.I.T. Acoustics Laboratory, Technical Report, No. 13, 1952; Halle, "The Strategy of Phonemics" [32], *Word* 10.197–209 (1954); and Cherry, Halle, and Jakobson, "Toward the Logical Description of Languages in Their Phonemic Aspect" [31], *Language* 29.34–46 (1953).

[2] I.e., nonrepetitions. The notion "phonemically distinct" does not presuppose the notion "phoneme."

*Reprinted from *International Journal of American Linguistics* 23.234–242 (1957) by permission.