

## I THE CONCEPT OF THE DISTINCTIVE FEATURE

### 1.1 RESOLVING SPEECH INTO ULTIMATE UNITS.

In a typical test of the intelligibility of speech, an English speaking announcer pronounces isolated root words (bill, put, fig, etc.), and an English speaking listener endeavors to recognize them correctly. For the listener this situation is in one sense simpler than normal speech communication because the word samples with which he deals cannot be broken up into shorter meaningful entities and are not grouped into higher units. Thus the division of sentences into words and of words into their grammatical components does not concern this listener. Nor need he account for the interrelation of words within a sentence and of various grammatical components within a complex word (ex-port-s, im-port-ed, re-port-ing, mid-night).

In another sense, however, this test is more complicated than normal speech communication. Neither the context nor the situation aids the listener in the task of discrimination. If the word bill were to appear in the sequence one dollar bill or as a single word said to a waiter after a meal, the listener would be able to predict its appearance. In such a situation, the sounds which compose this word are redundant to a high degree, since they "could have been inferred a priori"(1). If, however, the word is deprived of any prompting context, either verbal or non-verbal, it can be recognized by the listener only through its sound-shape. Consequently, in this situation the speech sounds convey the maximum amount of information.

The question arises: how many significant units, i.e., units relevant for the discrimination of the samples, do the sound-shapes of the samples contain? Upon perceiving syllables such as bill and pull, the listener recognizes them as two different words distinguishable by their initial part /bi/ and /pu/ respectively. This distinctive fraction, however, may be decomposed in turn. The listener, and any member of the English speech community, has in his vocabulary words such as pill and bull. On the one hand, identical means are employed for distinguishing bill from pill and bull from pull. On the other hand, the distinction between bill and bull is the same as that between pill and pull. Thus to distinguish between bill and pull a double operation is necessary. The fraction /bi/ in bill proves capable of being split into two segments /b/ and /i/, the first exemplified by the pair bill - pill and the second by bill - bull.

Each of the two segments derived serves to distinguish the word bill from a whole series of vocables, all other things being equal.\* For each of them a set of other segments can be substituted. This substitution of one segment by others is called commutation.

\* Henceforth we shall use the more condensed Latin equivalent of this formula: ceteris paribus.

We can list one whole commutation set. Commuting the first segment we obtain the series bill - pill - vill - fill - mill - dill - till - thill - sill - nil - gill /gil/ - kill - gill /zil/ - chill - hill - ill - rill - will. A closer examination of such a series permits certain inferences.

For some pairs of words in this set the discriminatory minimum is identical; hence one is warranted in saying that bill is to pill, as vill is to fill, or dill to till, or gill to kill, etc. or, for the sake of a more graphic presentation: bill:pill  $\simeq$  vill:fill  $\simeq$  dill:till  $\simeq$  gill:kill etc.

By the same token,

- 1) bill : vill  $\simeq$  pill : fill  $\simeq$  till : sill etc.
- 2) bill : mill  $\simeq$  dill : nil etc.
- 3) bill : dill  $\simeq$  pill : till  $\simeq$  fill : sill  $\simeq$  mill : nil etc.

A distinction is called minimal if it cannot be resolved into further distinctions which are used to differentiate words in a given language. We owe this term to Daniel Jones, from whom we also borrow the following definition\*: "Wider differences may be termed duple, triple, etc., according to the number of minimal distinctions of which the total difference is composed. Duple distinctions are the result of two minimal distinctions." (2)

The distinctions between bill and pill, or bill and vill or bill and dill are minimal distinctions since they cannot be resolved into simpler discriminations, which are, in turn, capable of differentiating English words. On the other hand, the relation of bill to till is a duple distinction, composed of two minimal distinctions: 1) bill - dill (which is equivalent to the distinction pill - till) and 2) bill - pill (equivalent of dill - till). The relation of bill to sill is a triple distinction: in addition to the two minimal distinctions cited, it includes a third one: bill - vill (equivalent to pill - fill and to till - sill).

The discrimination between the words bill and fell implies a duple distinction in their initial segments (/b/ - /f/), and a minimal one in the middle segments (/i/ - /e/). To discriminate between words such as bit and said, we need a triple distinction in their first segment and one minimal distinction in each of the two others.

Without further examples, it becomes clear that the listener of a speech sample is faced with a series of two-choice selections. To identify the message bill, he must decide for the non-vocalic inception against the vocalic and for the consonantal against the non-consonantal. By this double operation, vowels, liquids and glides are eliminated because if the word had begun with a vowel,

\* We, alone, are responsible for the way in which these concepts are hereafter applied to the empirical material.

the inception would have been identified as vocalic and non-consonantal; if with a liquid, as both vocalic and consonantal; and if with a glide, as neither vocalic nor consonantal. (For the interpretation of these distinctions see Sec. 2.2).

The next decision to be made is between bill and gill /gil/ - diffuse or compact (see 2.41), between bill and dill - grave or acute (see 2.42), and finally, between bill and mill - non-nasalized or nasalized (see 2.44). A decision in favor of the latter of the two alternatives would leave no further selections, since /m/ is the only combination of grave and nasal in English. But the opposite choice being made, there inevitably follows the selection between bill and pill - weak or strong (in more general terms, lax or tense: see 2.43), and, finally, the selection between bill and vill - stop or constrictive (in more general terms, interrupted or continuant: see 2.311). An analogous sequence of operations treats the two succeeding segments of the sample /i/ and /l/. The set of selections to be made is, however, more restricted than for the initial segment. For example, when a sequence begins with a stop, as bill does, the option for vocalic is obligatory, since in English the initial stop may be followed only by vowels or liquids.

Any minimal distinction carried by the message confronts the listener with a two-choice situation. Within a given language each of these oppositions has a specific property which differentiates it from all the others. The listener is obliged to choose either between two polar qualities of the same category, such as grave vs. acute, compact vs. diffuse, or between the presence and absence of a certain quality, such as voiced vs. unvoiced, nasalized vs. non-nasalized, sharpened vs. non-sharpened (plain). The choice between the two opposites may be termed distinctive feature. The distinctive features are the ultimate distinctive entities of language since no one of them can be broken down into smaller linguistic units. The distinctive features combined into one simultaneous or, as Twaddell aptly suggests, concurrent bundle form a phoneme.

For example, the word bill is comprised of three consecutive bundles of distinctive features: the phonemes /b/, /i/ and /l/. The first segment of the word bill is the phoneme /b/ consisting of the following features: 1) non-vocalic, 2) consonantal, 3) diffuse, 4) grave, 5) non-nasalized (oral), 6) lax, 7) interrupted. Since in English 7) implies both 1) and 2), the latter two features are redundant. Similarly 3) is redundant as it is implied by 4).

A speech message carries information in two dimensions. On the one hand, distinctive features are superposed upon each other, i.e., act concurrently (lumped into phonemes), and, on the other, they succeed each other in a time series. Of these two arrangements the superposition is the primary because it can function without the sequence; the sequence is the secondary since it implies the primary. For example, the French words où /u/ "where", eu /y/ "had" (participle), y /i/ "there", eau /o/ "water", œufs /ø/ "eggs", et /ɛ/ "and", aié /e/ "have!", un /ɔ̃/ "one", an /ɑ̃/ "year", etc., each contains a single phoneme.

The difference between the distinctive features of contiguous bundles permits the division of a sequence into phonemes. This difference may be either complete, as between the last two phonemes /i/ and /y/ in the word wing (which have no distinctive features in common) or partial, as between the last two phonemes of the word apt - /p/ and /t/ all of whose distinctive features are the same except one: /p/ is grave and /t/ is acute.

This suprasegmental extension of certain features such as interruptedness, diffuseness or non-nasality is selective: cf. such sequences as asp (continuant and interrupted), act (compact and diffuse) and ant (nasal and oral). On the other hand, strong (tense) and weak (lax) consonants cannot follow each other within a simple English word: cf. nabs/nabz/, nabbed/nabd/, and naps/naps/, napped/napt/. That is to say, in consonant sequences the tenseness and laxness features are suprasegmental.

Any one language code has a finite set of distinctive features and a finite set of rules for grouping them into phonemes and also for grouping the latter into sequences; this multiple set is termed phonemic pattern.

Any bundle of features (phoneme) used in a speech message at a given place in a given sequence is a selection from among a set of commutable bundles. Thus by commuting one feature in the first phoneme of the sequence pat we obtain a series bat - fat - mat - tat - cat. Any given sequence of phonemes is a selection from among a set of permutable sequences: e.g. pat - apt - tap. However, /tp'a/ not only does not, but could not exist as an English word, for it has an initial stop sequence and a single final vowel under stress, both of which are inadmissible according to the coding rules of contemporary English.

## 1.2 INVARIANCE AND REDUNDANT VARIATIONS

The consonants are quite different in the English coo and key or in the French coup and qui. In both languages a more backward (velar) articulation is used before /u/ and a more forward (palatal) articulation before /i/. The formants of the consonant are closely adapted to those of the following vowels, so that the frequency spectrum of /k/ before /u/ has a lower center of area and is closer to that of /p/ than is the case before /i/, where it has a higher center of area and is closer to that of /t/. Both in English and French, /p/ and /t/ are separate phonemes opposed to each other as grave and acute, whereas the two varieties of /k/ represent but a single phoneme. This seeming discrepancy is due to the fact that the opposition of /p/ and /t/ is autonomous, i.e. both /p/ and /t/ occur in identical contexts (pool - tool; pea - tea), while the difference between the two k-sounds is induced by the following vowel: it is a contextual variation. The retracted articulation and the low frequencies of one of these k-sounds and the more advanced articulation and high frequencies of the other are not distinctive but redundant features, since the distinction is carried by the subsequent vowels. In Roumanian, both k-sounds in question occur in one and the same context (e.g. before /u/: cu "with", with a backward articulation, and chiu "cry", with a more forward articulation) and, therefore, they represent two different phonemes.