
WORD

VOLUME 8

April, 1952

NUMBER 1

FUNCTION, STRUCTURE, AND SOUND CHANGE

ANDRÉ MARTINET

I. INTRODUCTION

Today, just as twenty years ago, many linguists would be tempted to agree unreservedly with Leonard Bloomfield that 'the causes of sound-change are unknown.¹ Not a few would infer from this statement that any research aimed at determining such causes is inevitably doomed to failure. Scholars who regard linguistics, primarily and even exclusively, as a descriptive discipline both on the synchronic plane and in evolutionary matters will most naturally be tempted to favor these views since they afford a justification for their reluctance to go beyond mere statements of directly observable facts, such as 'English *p* and *b* are distinct phonemes' or 'French *u* ([ü]) corresponds to Latin *ū*'. The modern followers of the neogrammarians who are ready to reckon with mutual influence IN THE SPOKEN CHAIN, but who ignore and would deny the fact that the nature of a given phoneme depends on that of its neighbors IN THE PATTERN will find themselves in agreement with those phonemicists who conceive of 'structure' as resulting from combinatory latitudes of distinctive units in the chain rather than as based upon the latter's partial phonic identities and differences. Whoever sees in a phonemic pattern nothing but a convenient way of summarizing the behavior of segments in the utterance is hardly prepared to conceive of it as a dynamic reality. The componential analysis of phonemes which is the first step toward the setting up of a pattern based upon phonic identities and differences is still often looked upon as an amusing but impractical refinement of more traditional methods, and its wide implications are rarely perceived. Yet it is felt by an increasing number of structurally minded linguists² that it paves the way toward a better understanding of phonetic evolution.

What would seem to prevent a general acceptance of diachronic phonemics as a legitimate discipline is not only the wide-spread feeling that linguistics would jeopardize its hard-won scientific character by venturing beyond the limits of pure description, but also an irrational conviction that there should be one and only one answer to the question: Why do sounds change? and that a principle of explanation which can be shown not to account *in toto* for any change chosen at random is *ipso facto* to be rejected as invalid. This conviction

¹ *Language*, New York 1933, 385.

² A bibliography will be found in A. G. Haudricourt & A. G. Juillard, *Essai pour une histoire structurale du phonétisme français*, Paris 1949, 119-120; see, *ibid.*, ix-xiv and 1-13. Later contributions will be cited in the course of this paper.

is so ingrained that, in the case of conditioned sound changes, some linguists would probably reject the view that well-known conditioning factors afford a partial answer to the question. In Old English, intervocalic *s* as in *céosan* is found to have been shifted to [z] whereas *s* has been preserved as [s] elsewhere. We know for sure that intervocalic position was instrumental in the change. But, of course, intervocalic position was not enough, since Old Icelandic has preserved a voiceless *s* in *kjósa*, and hundreds of easily accessible languages show no sign of voicing their intervocalic sibilants. The unknown reason, or reasons, which let Old English *s* become a prey to its vocalic environment are, in the opinion of some, the only elements of the case that deserve the name of 'cause'. And yet, can we not imagine that the voicing in question may have resulted from a concurrence of phonetic circumstances, one of which (intervocalic position) we know, and the others (nature of the accent or various other prosodic features) we are not well enough informed to recognize? All of these would be of a similar nature, and if we should have to distinguish among them, the distinctions would not be made because they are of intrinsically different nature, but exclusively on the basis of what we happen to know about the one or the others. It is methodologically unsound to assume anything about the importance or lack of importance of unidentified factors. Above all we have no right to postulate that these should all be of one and the same type, and that, short of the identification of some sort of ever present *deus ex machina*, any theory of the causes of sound change has to be resolutely brushed aside. Bloomfield's sweeping statement that the causes of sound change are unknown should be replaced by the one that some of the causes are still either unknown or difficult to identify or to verify. This could by no means be interpreted as an invitation to restrict linguistic activities to descriptive practices, but, on the contrary might incite the reader to investigate the possibilities of reducing the domain of the unknown.

So far external factors of sound change, among which interdialectal and interlinguistic influences stand in the foreground, have been the object of much theorizing but of little factual observation. Among internal factors only those that can be found in the spoken chain and account for allophonic differences, have been submitted by phoneticians to a more or less exhaustive examination. It remains to investigate to what extent the coexistence in the pattern of a number of phonemic units can account for their synchronic nature and diachronic comportment. We know that an [s], when placed in a GIVEN CONTEXT in the utterance, may develop in a certain way. We have to determine what we can expect from /s/ when placed in the frame of A GIVEN PATTERN.

The problem of the causes of sound change would remain one of the central problems of linguistic science even if we should refuse to consider linguistics as an historical discipline, because we shall not fully know what language is, and how it works before we have determined why languages change. No one would maintain that morphology, syntax, and lexicon change irrespective of the communicative needs of the speakers. It remains to be seen whether this is true or not of phonology. If it can be shown that phonetic evolution is not as 'blind'

as some of our predecessors meant it to be, we shall be able to discover not all but some of the so-far unknown factors of phonological evolution.

We shall, in what follows, center our attention on 'regular' sound changes, the type whereby all the performances of a given phoneme, everywhere or in a well defined context, are eventually affected. There are sporadic sound changes of many kinds, some of which may be due to causes similar to those which may be adduced or supposed for 'regular' changes. But it is felt that the consideration of sporadic changes would needlessly complicate our research. It should further be stressed once more that we are ultimately concerned here with the behavior of speakers keeping distinct or merging various phonemes of A PATTERN and not at all with what has normally been so far the practically exclusive preoccupation of historical phoneticians, namely the mutual influencing of successive phonemes in THE SPOKEN CHAIN. In phonemic terms, our predecessors were intent upon accounting for the appearance of combinatory variants or allophones. By now, it should be a well established fact that one and the same phoneme when appearing in different contexts may be submitted to divergent treatments, and this should need no further emphasizing. In the frame of the present exposition, it is completely immaterial whether a change affects a phoneme in all contexts or only in phonemically well defined ones, whether what is eventually merged or kept distinct is two phonemes or two combinatory variants of different phonemes. We know that combinatory factors of sound change play a considerable role, but if we want to be able to identify functional and structural factors, we have to concentrate upon them and keep the former out of our field of vision as far as this is practicable. In order to simplify the exposition, it is therefore advisable not to stress at every turn the existence of allophonic deviations, and to establish the following convention: unless otherwise stated, what will be said of phonemes applies equally to those allophones whose phonetic evolution happens to be deviating. In theoretical discussions, it will look as if we were always operating with phonemes whose unity is never endangered. But illustrations will show that allophones are also involved. Allophones will usually be presented in the form of a cluster of phonemes: the front allophones of a /k/ will appear as /ki/, /ke/, or both /ki/ and /ke/. In other words we no longer deal with a single phoneme /k/ but the phoneme clusters /ki/, /ke/. In view of the fact that phoneme clusters often coalesce into single phonemes in the course of phonological evolution and would seem frequently to exert an influence upon the pattern behavior of single phonemes, the use of clusters instead of allophones will actually result in a simplification. We can accordingly reword what we said above about our statements applying to allophones as well as to phonemes: unless otherwise stated, what we say about phonemes applies to larger phonemic units as well.

II. FUNCTION

It is an obvious fact that the pronunciation of a given phoneme in one and the same word by a given speaker varies from one utterance to another. The variation is normally imperceptible, but strictly speaking, no two pronunciations

can be exactly alike. Under certain conditions the variation may be more considerable. In any case, we have to reckon with a range of possible dispersion even in the speech of one person and still more so probably if we consider all the speakers of a given community. The existence of such a range of dispersion is of course obvious if we consider a phoneme with important allophonic variations, i.e. a phonemic unit whose actual performances are largely dependent on the context as is the case for instance with English /k/ whose range of dispersion covers a large part of the palato-velar area, or with Russian /a/ which varies, depending on context, from [æ] to [a]. But what is stressed here is not the dispersion resulting from combinatory variation, but that which may affect a phoneme in a well characterized context.

Some scholars have been tempted to interpret de Saussure's statement that a linguistic 'value' is everything that the other 'values' of the same system are not³ in the sense that this range of dispersion of every phoneme is limited only by those of other phonemes. This is certainly not universally true. It is probably meaningless to try to imagine whether [φeɪl] would, by English speakers, be interpreted as *pale* or as *fail*, because [φ] is a normal rendering of neither /p/ nor /f/, and if, under most unusual circumstances, any one said [φeɪl], the interpretation as *pale*, *pail*, or *fail* would depend on the context. In the frame of a homogeneous speech community it is probable that the normal range of dispersion of every phoneme in a given context will not be contiguous to those of its neighbors, but that there will be a margin of security in the form of a sort of no man's land. We speak here of 'normal' range because it is a well-known fact that, under unusual circumstances such as severe intoxication, neighboring phonemes of the pattern may be completely merged. It is then clear that the minor evil consisting in an impingement upon the margin of security must also occur in 'abnormal' circumstances. Abnormal circumstances of the kind considered here are not likely to affect the articulation of a single phoneme only, but that of most, if not all, of the phonemes of the pattern, and this in itself will be a perceptible mark of their abnormality. Hearers will unconsciously make allowances for it, discount deviations, or rely more heavily on context and situation in their interpretation of what they hear. Among 'abnormal' circumstances we might also include the cases where the language is spoken by a foreigner who has not achieved a complete mastery of the phonetics of the language. Here again allowances will be made.

For a full understanding of what will follow, one should remember that, on all occasions, it is far easier for man not to be than to be accurate; as Jespersen puts it⁴ 'it requires less effort to chip wood than to operate for cataract'; the main difficulty for children in learning to speak; or to write or draw, for that matter, is not to produce sounds, bars, or curves, but to hit upon the right sound, bar, or curve asked for at a given instant by the necessities of communication. This applies to adults' phonemes as well. For each one of them, in a given context at least, there must be an optimum which we might call the center of gravity

³ *Cours de linguistique générale*, Paris 1931, 162.

⁴ *Language, Its Nature, Development, and Origin*, London 1922, 263.

of every range of dispersion, but actual performances will normally fall somewhat off the mark. In the normal practice of speech, some of them are even likely to fall very far off it. If too dangerously near the center of gravity of some other phoneme, they may be corrected, and, in any case, will not be imitated. If unusually aberrant, slightly beyond the normal range of dispersion, but not in a direction where misunderstanding might arise, they would in no way threaten to impair mutual understanding. If not, in themselves, imposing any strain on the organs, they might well end up as establishing a legitimate extension of the acceptable range.

We shall reckon with a sound shift as soon as the normal range of a phoneme (in a given context—from now on this shall be understood) is being ever so little displaced in one direction or another, whereby the margin of security which separates it from its neighbors increases or decreases. We do not choose to discuss at once the possible causes of such a shift, but rather try to determine how it may affect other phonemic units of the pattern. Let us call A the phoneme whose normal range is being displaced, B the one separated from A by an increasing margin, C the one separated from A by a decreasing margin. The dynamic situation will thus look as follows:

$$B \quad A \rightarrow C$$

If, as the saying goes, 'phonetic laws work blindly' i.e. irrespective of communicative needs, the outcome of this situation will necessarily be a merger of A and C unless, for some mysterious reason, the trend is stopped or reversed. If it is found that B and C begin to shift in the same direction as A, so that the situation becomes

$$B \rightarrow A \rightarrow C \rightarrow$$

it will be assumed that the same unknown reason is affecting the three units equally. As a rule, in such a case, it will be difficult to prove that A actually began to shift before B and C did, and, if this could be shown, one could probably argue that, for some unknown reason, A was more susceptible to being shifted and therefore yielded to the push before the two others. If one is not inclined to be economical, one could of course also assume three different causes for the different shifts.

The basic assumption of functionalists in such matters is that sound shifts do not proceed irrespective of communicative needs, and that one of the factors which may determine their direction and even their appearance is the basic necessity of securing mutual understanding through the preservation of useful phonemic oppositions. Lest we should give the impression that we are dealing with some sort of linguistic providence, we shall have to present a fairly detailed analysis of how we may conceive the working of the various observable phenomena.

Let us revert to the afore-mentioned situation where A is drifting toward C and away from B, and concentrate our attention first on the possible ensuing compartments of B: (1) the normal range of B may remain what it was before

A began to shift, or it may happen to start drifting in any direction but toward A; in this case, we shall assume that the shift of A exerts no influence, or at least no direct influence, on the comportment of B; (2) the normal range of B will begin to shift in the direction of A: thus

$$B \rightarrow A \rightarrow$$

If it can be shown that the shift of A actually preceded that of B, and/or that the shifts of A and B can not be ascribed too well to the same general phonetic trend such as a general tendency toward aperture, closure, or other, so that the situation would more adequately be represented as

$$B \rightarrow \quad A \uparrow$$

functionalists will assume that B has, as it were, 'taken advantage' of the space left vacant by the drifting away of A. In fact, the chances are that B is environed by other phonemes and separated from them by margins of security which we may assume, for simplicity's sake, to have been just as wide as the one which originally separated B from A. At that time, any random deviation of B out of its normal range and in the direction of any one of its neighbors was not likely to be imitated since it would have tended to conflict with communicative needs. When however A started to shift away from B, chance deviations out of the normal range of B and in the direction of the receding A would no longer conflict with communicative needs; from that time on, B was contained on all sides except in the direction of A, and the center of gravity of its range naturally began to shift away from the sections of the field where it could not expand. What will often happen in such cases is that one of B's neighbors will in turn take advantage of the space left vacant by B so that a sort of chain reaction will be set in motion which may eventually affect an important section of the pattern.

At this stage of the exposition, it is not easy to present illustrations taken from actual languages, because every shift considered will involve the play of certain internal factors which have not so far been presented and discussed. Yet, in a few cases, such factors may be temporarily discounted without distorting the facts, although it will become clear at a later stage that the proof of the coherence of the shifts presented can only be administered if all factors are taken into consideration.

A comparison of the phonemic pattern of the Hauteville dialect⁵ with those of vernaculars spoken in the same region shows that, at about the same time, the following shifts must have taken place: (1) /ē/, from Lat. ĪN, > /ĕ/ (lowering); (2) /ĕ/, chiefly from Lat. EN, > /ɛ/ (denasalization); (3) /ɛ/, from Lat. Ī, Ē, > /a/ (lowering and retraction); (4) /a/, chiefly from Lat. A in open syllables, > /ɔ/ (raising and rounding). Since there was previously no /ɔ/ in the pattern, none of these four shifts has resulted in any phonemic merger. Schematically the process can be represented as follows

$$\tilde{e} \rightarrow \bar{e} \rightarrow \varepsilon \rightarrow a \rightarrow \text{ɔ}$$

⁵ As presented in A. Martinet, *Description phonologique du parler franco-provençal d'Hauteville (Savoie)*, *Revue de linguistique romane* 15.1-86; see, in particular; 2-3.

There has been no wholesale lowering in the front series since /i/ and /e/ are intact, no wholesale raising in the back series since /u/ and /o/ have not moved, no wholesale denasalization since /ã/ and /õ/ remain by the side of the new /ẽ/. Therefore the whole shift can not be accounted for as resulting from one and the same general phonetic trend. Taken one by one, each of the first three shifts should have resulted in a merger. As a matter of fact, every one of the four phonemes involved has kept clear of the others. Since the margin of security between the old /a/ and /o/ was twice as wide as every one of the others in the vocalic pattern, it seems most likely that /a/ was the first to start moving. To-day all margins of security have approximately the same width and no shift is in progress. It seems difficult to escape the conclusion that some necessity of preserving existing phonemic distinctions has been at work throughout the process.

Another illustration is afforded by the Portuguese dialect of São Miguel in the Azores⁶. A comparison with standard Portuguese shows that /u/ has been shifted to /ü/, /o/ has passed to /u/, /ɔ/ has been raised toward /o/ without always reaching it, /a/ has assumed a back value 'tending toward open ó'. This description of the shift is not exhaustive; some features of it, which would only assume full significance at a later stage of this study, have been left out. The powerful influence of standard Portuguese has obviously exerted some disturbing influence. It seems clear, however, that /u/ took the lead in its shift toward /ü/, /o/ soon followed, /ɔ/ began its shift with a certain delay, and /a/ was last to move. Schematically the process could be represented as follows:

a → ɔ → o → u →

The shift of /u/ to /ü/ raises a problem which we are not yet ready to tackle⁷. Let it suffice to say that it may have resulted from a pressure exerted upon /u/ by its partners of the back series. It should only be stressed here that if three of the particular shifts involved can be described as raisings, the /u/ > /ü/-shift is of a totally different phonetic nature, and yet a causal connection between the fronting and the raisings can hardly be denied.

Let us now direct our attention to the possible comportment of C, the phoneme toward which the range of dispersion of A is moving for reasons so far unknown. The range of C may well not move away from invading A, and a phonemic confusion will take place. The undeniable frequency of such mergers is sometimes held as a powerful argument against the assumption that the preservation of phonemic distinctions is a factor of phonological evolution. Since phonemes, by definition, serve to distinguish between words and forms, any phonemic merger will inescapably involve confusions detrimental to the normal functioning of the language, and yet mergers do take place.

In a number of cases it might be argued that C is, as it were, at the end of its tether, that its performance represents an extreme phonetic possibility as

⁶ Cf. F. M. Rogers, *Insular Portuguese Pronunciation: Porto Santo and Eastern Azores*, *Hispanic Review* 16.1-32, in particular 13.

⁷ The problem is dealt with by Haudricourt-Juillard, *op. cit.*, 100-113.

when it is an /i/ badly pressed by an invading /e/ with surrounding diphthongs which block all way of escape. An objection would be: how is it that these circumstances have not, from the start, prevented the range of /e/ from moving into the margin of security separating it from /i/? But of course the unknown reasons pushing /e/ upward may simply be more powerful than the functional factors working for conservation. This does not mean that the latter do not exist. It must be stressed over and over again that no one has ever pretended that internal phonemic factors are the only ones or even necessarily the most potent. What we have to show is not that these factors explain all features of phonological evolution, but that there are cases where no understanding can be reached unless they are duly taken into account.

It will be seen that both the articulatory and acoustic nature of the distinctive features involved may be a factor of some importance in the fate of an opposition. But, at this stage, the problem which shall detain us is whether the relative importance of the opposition in the satisfaction of communicative needs plays a role or not in its own elimination or preservation. The question to be answered is whether, everything else being equal, a phonemic opposition which serves to keep distinct hundreds of most frequent and useful words will not offer a more successful resistance to elimination than one which only serves a useful purpose in very few instances. What makes this answer particularly difficult is that we know, thus far, so little that is definite about other factors involved. The first step we have to take in order to bring some clarity into the affair is to investigate whether and how the distinctive importance of a phonemic opposition can be evaluated.

The functional importance of a phonemic opposition is often called its functional yield or burden (Fr. *rendement fonctionnel*, German *funktionelle Belastung*). There is no complete agreement as to what this term is meant to cover. In its simplest somewhat unsophisticated acceptation, it refers to the number of lexical pairs which would be complete homonyms if it were not that one word of the pair presents one member A of the opposition where the other shows the other member B: the pair *pack-back* is part of the functional yield of the /p/-/b/ opposition in English, and so are *repeI-rebel*, *cap-cab* and hosts of others. The number of such /p/-/b/ pairs being considerable, it is said that the functional yield of the /p/-/b/ opposition is high. If we try to do the same with, say the English /θ/-/ð/ opposition, we shall find only a few pairs like *thigh-thy*, *mouth n.-mouth v.*; it will be said that the functional yield of this opposition is low. Provided we consider a given dictionary as fully representative of the lexicon of the language under consideration, it is possible to make exhaustive lists for every one of the phonemic oppositions, although in practice only those would be considered that involve phonemes which componential analysis has shown to be minimally distinct: In English /s/-/ʃ/ and /s/-/z/ would be included in the research, but not /s/-/ʒ/ or /ʃ/-/z/. Thereby vague labelings like 'high', 'medium', 'low' can be advantageously replaced by exact numerical ratings.

It is of course easy to point out the drawbacks of such a method as a tool for determining the actual number of cases where a given phonemically distinctive feature is by itself the sole element which prevents misunderstanding, as the

degree of vocalic aperture would be if *give me a pen!* were uttered in a situation which did not give any clues as to whether a *pen* or a *pin* is wanted. In order to be fully valid, any rating of the functional yield of an opposition should be based upon a frequency rating of such linguistic situations as the one just mentioned. Since such a count is practically impossible, one might be content with a listing of those lexical pairs that could be conceived of as likely ever to give rise to such indeterminacy as illustrated above by *pen* and *pin*. But, in order to avoid subjective decisions, one would probably have to be satisfied with the exclusion of only such pairs as involve words belonging to different parts of speech and therefore not likely ever to appear in the same grammatical context. In that case; among the pairs cited above *pack-back*, *repel-rebel*, *cap-cab* would pass muster, but *thigh-thy* and *mouth n.-mouth v.* would be rejected. Yet, even then, some complications might arise: the minimally distinct Fr. pair *poignée-poignet* will never give rise to conflict as long as the two words are used in the singular since, in that case, the difference of gender will show up one way or another and would tell which is which even if the /e/-/ɛ/ opposition happened to be blurred; but in the plural (*les poignées-les poignets*) the vocalic opposition might have to bear all the distinctive burden. Furthermore this type of evaluation would completely disregard the essential factor of frequency and would give equal rating to *brig-brig* (Thorndike frequency levels 19 and 16 respectively) and *pack-back* (2 and 1). In view of all the difficulties involved, it is perhaps just as indicative in most cases and certainly incomparably simpler to determine the lexical frequency of every phoneme involved, assuming that the more frequent a phoneme is, the more likely it is that it will have to assume clearly distinctive functions. Lexical frequency is probably preferable to actual frequency in texts or utterance because it is not exceptional that a phoneme such as English /ð/, which enters into the minimally differentiated pair /θ/-/ð/ with a very low functional yield, appears very frequently in texts or utterances. There might however be cases where a lexically very frequent phoneme is less frequent in speech than some other one with lower lexical frequency, if for instance the former appears mostly in learned, the latter in everyday lexical items. In such a case, the conclusions derived from a lexical count would have to be tempered by reference to the actual situation in speech. Generally, the method would have to be adapted to the language under consideration.

From what precedes, it is clear that the functional yield of an opposition can only be evaluated with any degree of accuracy if we deal with linguistic stages for which fairly exhaustive word lists are available. This circumstance makes it practically impossible to check the validity of the functional assumption in the case of prehistoric sound shifts. It would seem, for instance, that the merger of *o, and *a in Slavic, Baltic, and Germanic is in some way connected with a relative rarity of *a (from *a or *ǎ) in these languages where *ǎ is dropped in second medial syllables, and vocalic sonants universally develop high vowels. But since we do not know the lexicon of Slavic, Baltic, and Germanic at the time when they merged *o and *a, we can hardly go beyond vague assumptions. Even in the case of early Romance, our fairly exhaustive knowledge of Classical Latin vocabulary gives us an imperfect picture of the lexical resources of the vulgar

language from which we would have to start. In the case of certain mergers taking place in modern cultural languages for which full data are available, the functional yield has been found to be extremely low: the Parisian French merger of / $\bar{\epsilon}$ / and / $\bar{\alpha}$ / which is in full swing, practically never results in any homonymic conflict, and the lexical frequency of / $\bar{\alpha}$ / is of the lowest⁸. The same could be said of the merger of / \bar{n} / and / \bar{n}_i / which seems to be gaining ground,⁹ and of the earlier confusion of / \bar{y} / and / \bar{y} / . In the same language, the old distinction between long and short /i/, / \bar{u} /, /u/, /e/ whose function was practically restricted to distinguishing between masculine and feminine words is now practically eliminated among Parisians¹⁰. Since gender in French is usually expressed in accompanying articles or pronouns, the actual yield of these oppositions was very low, and this circumstance may well have been instrumental in the merger. It is interesting to notice that, to this day, French speakers have not found a universally accepted solution for the irritating problem resulting from the homonymy of *l'amie*, *l'ami*, *mon amie*, *mon ami*¹¹. If low functional yield is accepted as a factor of the merger, we shall have to conclude that even one very useful pair is not enough for preserving a phonemic opposition.

The actual importance of functional yield in the preservation of phonemic oppositions can not be assessed on the basis of the limited information available to date. It will have to be tentatively considered as one of the internal factors of phonological evolution, and the possible extent of its influence will have to be evaluated wherever feasible. The problem will have to be reconsidered when we possess a large body of relevant data. It should however be pointed out immediately that (1) two neighboring phonemes will not necessarily tend to merge, simply because the functional yield of their opposition is practically nil: / \bar{s} / and / \bar{z} / in English are not found to approach each other in spite of the exceptionally low yield of their opposition; (2) semantic extension, word composition, and morphological reshuffling frequently afford easy solutions to the problems which may arise when a functionally important opposition is being threatened by the drifting together of two phonemes: as soon as the margin of security is invaded and danger of misunderstanding arises, speakers will be induced to give preference to such alternative words, phrases, or forms as will remove all ambiguity.

We now revert again to our theoretical example of a phoneme A drifting in the direction of a phoneme C, but this time we shall assume that C, instead of

⁸ Lip-rounding, which distinguishes / $\bar{\alpha}$ / from / $\bar{\epsilon}$ /, is an unstable feature in the case of such very open articulations. The same is true of course for / $\bar{\delta}$ / which we might expect to merge with / $\bar{\alpha}$ / in forms of speech where / $\bar{\alpha}$ / merges with / $\bar{\epsilon}$ / (actually [$\bar{\alpha}$]). But the functional yield of the / $\bar{\delta}$ /—/ $\bar{\alpha}$ / opposition is very high in French, and the merger is only attested in such northern Gallo-Romance dialects (and the corresponding local forms of Standard French) as have kept *en* phonemically distinct from *an* so that the frequency of / $\bar{\alpha}$ / (= *an*) is much lower than in the standard language.

⁹ Cf. A. Martinet, *La prononciation du français contemporain*, Paris, 1945, 170-173.

¹⁰ *Ibid.*, 94-109.

¹¹ The language affords no easy solution by means of composition, such as exists in English: *boy friend*, *girl friend*; most French speakers will pronounce the *-e* of *amie*, which results in a phonemically exceptional combination, phonetically [$a^1mi^1\bar{o}$] or [$a^1mi^1\bar{\alpha}$].

awaiting the impending merger, recedes before the invader preserving all the time a margin of security between A and itself. This type of assumption conflicts of course with the traditional views concerning the 'blindness' of 'phonetic laws.' Yet it is not too difficult to understand how a phoneme can yield under the pressure of one of its neighbors. As soon as the margin of security separating A from C is invaded by the former, any performance of C that falls too close to that margin will incur the danger of being misinterpreted and will therefore be disfavored. Thereby the center of gravity of the range of C will be displaced away from A. If may be that, in so doing, C will exert upon one of its other neighbors the kind of pressure that A is exerting upon it, and that neighbor will in its turn be shifted further, away from invading C. We shall thus observe a chain of reactions similar to the ones we have noticed in the case considered above of A and B.

In practice, it may often be difficult to tell whether we have to do with a $B \rightarrow A \rightarrow$ chain, or drag-chain, or an $A \rightarrow C \rightarrow$ chain, or push-chain. Even in the $B \rightarrow A \rightarrow$ type, there is some amount of pressure from such neighbors of B as are not included in the diagram. In order to simplify the exposition, we have purposely refused to investigate factors acting upon A and determining the drift of its range. But among them may, in the case of $B \rightarrow A \rightarrow$, figure a pressure exerted by B and its neighbors. We have, in the case of Hauteville, suggested the existence of what amounts to a tendency toward equidistance between the phonemes of the same pattern or, in other terms, toward equalization of the mutual pressures. Hauteville's /a/ passing to /ɔ/ would result from this equalizing tendency. São Miguel's /u/ > /ü/ would result from a pressure exerted upon /u/ by the other three back phonemes of a series where margins of security are; by nature, narrower than in the corresponding front series. We thus probably have to reckon with pressure everywhere, so that the suggested distinction between drag and push would often be blurred. We may say that, in some cases, the move of the leading phoneme is one which our phonetic and phonemic experience would lead us to expect, and in others, that it is the move of the last phoneme which would seem to make more sense. What we have called A is the first phoneme in the former case and the last phoneme in the latter. Now A was the phoneme whose move we took for granted all the time, so that our final judgment in such matters will depend on our interpretation of such factors as we have not so far investigated, or as will ultimately remain out of the frame of this study.

The difficulty of deciding which unit is leading the shift may be illustrated by the following example: (1) Italian *qui* is derived from $\text{kw}i$ and generally, in the traditional vocabulary, /kwɪ/ should result from Lat. dissyllabic /ku + i/ or /ko + i/; (2) *chi* /ki/ is from Lat. *qvi*, *qvis*, and generally, /ki/ derives from Lat. /kwi/; (3) *ci-* /çi/ as in *città* is from Lat. *ci-* /ki/ as in *civitatē*; three phonemic units have thus been kept distinct although the articulation of every one has changed. The whole shift can be schematized as follows:

$$kui \rightarrow kwi \rightarrow ki \rightarrow \text{çi}$$

Since palatalization of dorsals before front vowels is a most frequent phenomenon, we might be tempted to call this a drag shift: /ki/ was first palatalized; then /kwi/ could be reduced to /ki/, and /kuí/ could become a monosyllabic /kwi/. But we could also start from /kuí/ and argue that since hiatuses in general were being widely reduced in Imperial Latin, /kuí/ would tend to pass to /kwi/ and thereby exert a pressure on former /kwi/'s. These in turn would press upon /ki/'s with the result that they would be articulated farther forward in the mouth and become palatalized. This whole shift can not have been general in the Romania: qvi must still have been something like /kwi/ pretty late in northern Gallo-Romance since the purely French palatalization, which is found to affect /ki/ in Germanic loans as in *échine* from **skīna*, leaves the dorsal intact in the reflex of qvi; only the palatalization in cr- must have spread out of its original domain to the provinces with the well known exceptions of Sardinia and Dalmatia. This latter account of the shift is highly satisfactory in as much as it ties up neatly with what is universally recognized to have been the fundamental trends of Vulgar Latin phonological evolution: the tendency to eliminate hiatuses obviously resulted from the development of stress accent; the resistance of individual phonemes or clusters must have been negligible in comparison with such a powerful irreversible trend; /kuí/ had to become /kwi/ and actually did everywhere. But was the functional yield of the /kuí/-/kwi/ opposition so important that /kwi/ had to recede before the invader? Many Latin speakers in northern Gaul and elsewhere just let the two groups merge. Could we not think that if the merger did not take place in central Italy, it was because Latin speakers there had already palatalized cr and considerably weakened the /w/ of qvi thus making room for /kuí/? On the other hand, it can not be argued that a push shift is to be discounted here on the ground that there are so many known cases of palatalization of dorsals which certainly do not result from a pressure upon /ki/ exerted by /kwi/ and /kuí/. There is no valid reason for assuming that the ultimate cause of such a palatalization is necessarily the same in all cases. What is needed here, as elsewhere, is a large body of tentative functional and structural explanations for the most varied cases of the type of phenomenon under consideration, and a set purpose never to let one's self be deterred from causal research by the complexity of the problems.

III. STRUCTURE

We have, in what precedes, been generally considering the problem of sound change as if every phoneme were characterized by one specific articulatory feature, entirely different from that of every other phoneme of the language. In fact this would seem to be the exception rather than the rule. The articulation of the majority of consonants in most languages implies the combination of two or more characteristic features, every one of which is to be found in some other phoneme or phonemes of the language. These features may be defined in articulatory or acoustic terms. We shall here as a rule operate with articulatory data, since they are more readily available and better known. A feature is said to be characteristic in this connection if it is phonemically distinctive. In a language like English, the lungs play a role in the production of every single

phoneme, and practically the same role; therefore the pulmonic articulation is never characteristic and never distinctive. On the contrary, the bilabial articulation characterizes three phonemes /p/, /b/, and /m/ and is distinctive since it keeps these phonemes apart from e.g. /t/, /d/, and /n/ respectively. The occlusive nature of these bilabials is not characteristic or distinctive since English bilabials are always articulated as stops. Instrumental research may show that the bilabial articulation is not quite the same for /p/, /b/, and /m/, nor is the apical articulation quite of the same type for /t/, /d/, and /n/, but whatever difference may be found could, in the case of English, easily be shown to result from such concomitant (glottal or nasal) articulations as distinguish /p/ from /b/ or from /m/, /t/ from /d/ or from /n/, and so forth. We have thus to do with an automatic deviation with no distinctive significance and comparable to the one which makes the /k/ of /ki/ different from that of /ka/. A number of consonantal phonemes characterized by one and the same articulation will be said to form a 'series' if their other characteristic articulations can be located at different points along the air channel. Thus in English /p/, /t/, /ð/, /k/, all characterized by the same glottal articulation but distinguished by the region where the stoppage takes place, will form a series, and so will /b/, /d/, /ǰ/, /g/.

A number of phonemes characterized by one and the same articulation at a given point of the air channel, but distinguished from one another by some other distinctive articulation will be said to form an 'order'. Thus in English /p/, /b/, /m/, will form a labial order, /t/, /d/, /n/ an apical order, and so forth. In regard to vowels, it seems more advantageous to label as 'series' a number of phonemes characterized by the same type of resonance cavities, but distinguished by different degrees of oral aperture, and as 'order' a number of phonemes characterized by the same degree of aperture but distinguished by different types of resonance cavities. In English, /i/, /e/, /æ/ form a front series; /i/, /u/ a high order. In such matters /y/ and /w/ are often advantageously grouped with the vowels and may form a special order if they are phonemically distinct from /i/ and /u/.

It should be pointed out that both series and orders are oppositional in nature just like any other phonemic entity. Just as a phoneme as such presupposes other phonemes, a series presupposes one or more other series, an order, one or more other orders. A language whose consonantal inventory was restricted to /p/, /t/, /t/, /c/, /k/, /q/ would not present any consonantal series because its six phonemes would have no distinctive features in common. A language with /p/, /t/, /k/, /m/, /n/, /ŋ/ would present two series, one of non-nasals and one of nasals, and three orders, labial, apical, and dorsal. Series and orders presuppose a larger unit grouping them into a whole, namely, the 'correlation', which includes two parallel series and a number of coupled phonemes belonging to the same orders. The six phonemes of our second theoretical example would form the following correlation:

p	t	k
m	n	ŋ

Strictly speaking, a phoneme which phonetically would seem to belong to one series, is actually no member of that series and of the correlation to which that series belongs if it has no correspondent in the other series: if a language had only the five consonants /p/, /t/, /k/, /m/, /n/—and no /ŋ/—/k/, in the theory, would not belong to the non-nasal series comprising /p/ and /t/ since the absence of nasality is not distinctive in combination with dorsal articulation. In a language where there is only one lateral phoneme /l/ articulated with the tongue tip in the same position as that of /t/, /d/, /n/, it could not be said to belong to the apical order because the apical articulation is not distinctive in combination with laterality. In diachronic phonemic practice, it will however be found convenient to include a phoneme in a series even when it has no counterpart in the parallel series, or in an order even when it has no counterpart in parallel orders, if its general phonic behavior (allophonic deviations, distribution, etc.) is similar to that of the phonemes of that series or of that order. In the case presented above of a language with /p/, /t/, /k/, /m/, /n/, it would probably be advisable to include /k/ in the non-nasal series. But there might not exist the same reasons in our second theoretical example for placing /l/ in the apical order.

The relationships existing between phonemes of the same order are usually rather different from those between phonemes of the same series. In the former case they would seem generally to be bilateral, whereas in a series they would be multilateral. In other words, phonemes of the same order would form a binary opposition or, if there are more than two of them, a complex of binary oppositions. On the contrary, all phonemes of a series would stand in the same relation to one another. If a language has, among other phonemes, /p/, /b/, /m/, /t/, /d/, /n/, they will form three series, and an order of labials, an order of apicals, and so forth. Thence:

p	t
b	d
m	n

The /m/ and /n/ phonemes are likely to be normally voiced, but occasionally unvoiced without losing their identity; /p/ will be defined as unvoiced (in opposition to /b/), non-nasal (in opposition to /m/), labial (in opposition to /t/ and others); /b/ will be defined as voiced (in opposition to /p/), non-nasal (in opposition to /m/), labial (in opposition to /d/ and others); /m/ will be defined as nasal (in opposition to /p/ and /b/) and labial (in opposition to /n/ and others). It is clear that /p/ and /b/ have two characteristics in common, non-nasality and labiality, which they are the only ones to share. They are said to form a bilateral opposition, and, as one unit, they enter into another bilateral opposition with /m/. One can also say that /p/ and /b/ stand in exclusive relation, since they are the only phonemes to share the distinctive features of labiality and non-nasality.

The relation between the different phonemes of a series (or the different pairs of a correlation) seems to be of a different nature. Theoretically at least, every one of them is opposed exactly in the same way to any one of the others.

This will explain why a correlation, the simplest coherent partial pattern, consists of an indefinite number of orders but of only two series of phonemes standing in a one-to-one exclusive relation, the same for all pairs.

Two or more parallel correlations form what is called a 'bundle'. A bundle can be made up of three series, as in the case of

p	t	.	.	.
b	d	.	.	.
m	n	.	.	.

presented above; of four series grouped in various ways, as for instance in a language combining phonemically voice and aspiration and presenting e.g. the four labials /p/, /b/, /ph/, /bh/; of five series or more.

In practice however, there would seem to be exceptions to this clear-cut opposition between bilaterality inside orders, and multilaterality characteristic of series: a labial order consisting of /p', /p/, and /p'/ might be more naturally conceived of as a triad than as a combination of two binary oppositions. In many languages two orders of hissing and hushing sibilants seem to stand in particularly close relation since they appear in partial complementary distribution¹². In the case of vocalic patterns, three vowels of the same order such as /i/, /ü/, /u/ form a triad, and to present them in the frame of two binary oppositions would certainly distort reality. On the other hand; phonemes of the same series such as /i/ and /e/ or /e/ and /ε/ are found in certain languages to be in partial complementary distribution.

These facts and a number of theoretical considerations have induced some scholars to attempt a reduction of all phonemic oppositions to the type we have seen to prevail inside orders¹³. It has for instance been suggested that the phonemes of consonantal series actually form a more closely knit pattern than the one which is suggested by a linear presentation. The oppositions in such a series should result from combinations of acute or grave quality with two different degrees of 'compactness'. Thereby a close parallelism could be established with vocalic patterns, and a considerable reduction in the number of distinctive features would be achieved.

We can not enter here into a discussion of the advantages or disadvantages of such a method in synchronic studies. In diachronic matters it would seem so far that not too much is gained by departing from a linear conception of the relations between the consonantal phonemes of varying degrees of articulatory depth. In a pattern with the four phonemes /p/, /t/, /c/, and /k/, /p/ would

¹² As for instance in German.

¹³ Roman Jakobson was the first scholar to advocate such a reduction: see *Proceedings of the Third Intern. Congress of Phon. Sciences*, Ghent 1939, 34-41, and *Kindersprache, Aphasie und allgemeine Lautgesetze, Språkvetenskapliga sällskapets förhandlingar* 1940-1942, Uppsala 1941, 52-77. It was applied by J. P. Soffietti in his *Phonemic Analysis of the Word in Turinese*, New York 1949, and by Roman Jakobson and J. Lotz in Notes on the French Phonemic Pattern, *Word* 5.151-158. The latest exposition of the procedure is to be found in *Preliminaries to Speech Analysis, The Distinctive Features and their Correlates*, Technical Report No. 13, January 1952, Acoustics Laboratory, Massachusetts Institute of Technology, by Roman Jakobson, C. G. M. Fant, and M. Halle.

share with /k/ the distinctive feature of graveness, and /p/-/k/ would be parallel to /t/-/c/. Yet it is found that, diachronically, passages from the /c/ type to the /t/ type are quite frequent, and so are shifts from /k/ to /c/; but /p/ and /k/ are kept well apart. The frequent shift of [k^w] to [p] can not be adduced to support a close kinship of /p/ and /k/, because [k^w] combines a dorsal and a labial articulation and [p] can only result from a hardening of the latter and a release of the former. Generally, a diachronic approach requires a greater concern with phonetic reality than is possible when we are bent upon reducing the number of distinctive features to a minimum. Even in a language like French where /k/ has no exact fricative counterpart /x/, and /š/ no occlusive partner /č/, /š/ can not be said to be the fricative or continuant counterpart of /k/ because it can not be maintained that a velar fricative normally tends toward a hushing articulation as a result of its fricative nature¹⁴. Two phonemes can only be said to belong to the same order if they both present the local characteristic articulation in exactly the same form or in forms which deviate from each other only through features which can be fully accounted for as due to the synchronic influence of a concomitant articulation: in Arabic the tongue-tip articulation of 'emphatic' /t/ takes place much farther back than that of 'non-emphatic' /t/, but the two phonemes still belong to the same order, because the retracted articulation of /t/ is readily accounted for as resulting from the concomitant velar or pharyngeal articulation which is the permanent characteristic of modern Arabic 'emphasis'. On the contrary, in a language where /t/ has the normal apical articulation, and /s/ is predorsal, we have no right to include the two of them in the same 'dental' order, because we do not see why a tense fricative counterpart of /t/ should have a predorsal and not the same apical articulation. It may, in certain cases, be difficult to decide whether two phonemes belong to the same order or not, and we have in practice to reckon with borderline cases, but phonologists should be warned against identifying orders as defined above with the traditional loose grouping of the phonemes of a language into the ready-made classes of labials, labiodentals, dentals, palatals, and velars.

All this does not mean that a componential analysis of phonemes that strives at maximal reduction of the number of distinctive features and eventually reveals unheeded connections between the seemingly most remote sections of the pattern, may not have to play a role in diachronic considerations: /k/ and /a/, for instance—which are described as 'compact' as opposed to /t/ and /p/, /i/ and /u/—will often evince parallel evolutionary trends, as when they tend toward [c] and [æ] i.e. a more 'acute' pronunciation, a phenomenon which we find for instance in Anglo-Frisian and Old French. This might mean that, in such cases, the palatalization of dorsals before front vowels is not entirely con-

¹⁴ In such a case, it would of course be redundant to state that such an opposition as /š/-/k/ is one of hush-friction vs. velarity-plosion, and descriptive economy is achieved by reducing it to friction vs. plosion (or continuant vs. interrupted; cf. *Preliminaries to Speech Analysis*, 6 and 21) because the hush-velarity opposition can thus be eliminated. But it should be clear that descriptive economy is achieved here through blurring the actual synchronic relationship between two phonemic units. Descriptive economy does not necessarily do full justice to functional and structural reality.

ditioned by the quality of the following vowel as usually assumed, but also by a general fronting of all dorsal consonantal articulations whereby post-velars become velars, velars become post-palatals, and so forth. This 'acutization' would of course still have to be explained; but it is scientifically preferable to operate with one unknown cause than with several, one for each of the individual changes. It is a fact however that a presentation of the pattern in terms of orders, series, correlations, and bundles, with its concomitant insistence on the details of phonetic reality, is as a rule more revealing of evolutionary probabilities.

Since most phonemes actually result from combinations of distinctive articulations, we may expect that in many cases a change in the performance of a phoneme will result from a modification of only one of these articulations. If a /t/, characterized by a certain apical and a certain glottal articulation, is found to change, it may be that only the nature of its apical articulation is affected, or only that of its glottal distinctive feature. A change in the apical articulation, as for instance a retraction of the tip of the tongue from the upper teeth toward the alveolas, if it is not in some way connected with the glottal behavior characteristic of /t/ and the other phonemes of the same series, will probably affect not only /t/, but all the other phonemes of the apical order, e.g. /d/ and /n/. Similarly, a change in the glottal articulation which characterizes /t/ as opposed to /d/ will affect not only /t/ but all the phonemes of the voiceless series, e.g. /p/ and /k/. In other words, it may be expected that every distinctive articulation will change irrespective of the other articulations with which it may combine in order to form individual phonemes. This is what we actually find in the most varied languages: as a rule, when in a given language /t/ is being 'aspirated', it is found that other phonemes of the voiceless stop series are also being 'aspirated', which means that the glottal articulation is shifted irrespective of the oral articulations with which it combines. If /d/ is being unvoiced, /b/ and /g/ will probably be unvoiced too. If /k/ is palatalized in certain conditions, /g/ is likely to be palatalized in the same conditions, and the difference in glottal articulation between /k/ and /g/ will not determine a different treatment. All this is of course well-known and it is felt that what would need investigation are the cases in which one phoneme of a given series shows a specific treatment of its glottal articulation, or one phoneme of a given order presents a shift of its local oral articulation which is not being shared by the other phonemes of the order.

All this has obviously an important bearing upon our present research. If, as we have assumed, the functional yield of an opposition is one of the factors in its preservation or elimination, it is clear that the opposition of two articulatory features which serve to keep distinct not merely two isolated phonemes, but two large series or orders will, everything else being equal, be far more resistant. We have seen that the actual yield of the English /θ/-/ð/ opposition is extremely low. But this is not what really counts: the feature of voice supplemented by concomitant differences in articulatory strength, which distin-

guishes /θ/ from /θ/ is also the one which distinguishes /v/ from /f/, /z/ from /s/, /ž/ from /š/, /ɕ/ from /ç/ and helps to keep /b/ apart from /p/, /d/ apart from /t/, /g/ apart from /k/. The functional yield of the opposition of voice to its absence is in English tremendous, and contributes to the stability of a large section of the consonantal pattern. All this does not mean of course that the phonetic nature of such an opposition is not likely to change in the course of time, but that if a change takes place, it is less likely to result in a merger than if the opposition were limited to a single pair.

Apart from the stabilizing influence exerted by the high functional yield of correlated oppositions, we probably have to reckon with a further factor of stability resulting from the mere frequency of the articulations characteristic of series and orders. Linguistic features which recur frequently in the chain are likely to be learned earlier and remembered better than those which appear less often. This is obvious in the case of morphological and lexical elements and syntactic patterns, and should apply to phonemic items as well. Although we still lack a large body of scientifically observed data relating to the acquisition by children of the most varied phonemic patterns, it would seem that in general correlated oppositions are acquired earlier than non-correlated ones. Here again, stability does not mean resistance to change, since perfect imitation on the part of the child should not prevent sound change from taking place, and should only prevent mergers.

If it is true that such oppositions as are integrated in a correlation or a bundle of correlations are *ipso facto* more stable than the ones between non-correlated phonemes or between a correlated phoneme and a non-correlated one, it will mean that phonemes outside of the integrated pattern will vary much more freely. If for simplicity's sake we assume complete fixity for correlated phonemes, and incessant erratic wanderings for non-correlated ones, we shall come to the conclusion that, at some time or other, every one of the latter will, just by mere chance, assume a phonetic shape which will make it the correlative partner of some other. Let us, for instance, assume the following correlation:

f	s	š
v	z	ž

plus a /x/, theoretically no part of the correlation since it has no voiced partner, but behaving exactly like /f/, /s/, and /š/. There is in the same language a trilled phoneme /r/, normally voiced, which is not integrated because it has no voiceless counterpart, whose articulation has been shifting around, and whose range of dispersion includes some non-trilled performances. A day may come when it will assume a post-velar fricative articulation which will make it the voiced partner of /x/. It will be integrated in the correlation which will henceforward appear as

f	s	š	x
v	z	ž	γ

and that will be the end of its erratic wanderings. As a matter of fact, there certainly is more to this than pure chance, and we have to reckon with some

amount of attraction on the part of the integrated pattern. Let us assume that the /r/ phoneme was at some time a uvular¹⁵trill. Pure least effort would probably result in weakening certain of its performances to sheer friction. But friction at the uvular level would not be so very different, both articulatorily and acoustically, from the post-velar friction characteristic of /x/. Since the performances of /r/ are normally voiced, there is no functional resistance against a merger of the two fricative articulations. The oral articulation of /x/ will exert an attraction on that of /r/ or maybe the reverse. This means that, at a certain point of time, speakers will no longer take the trouble to keep apart two minimally distinct articulations whose distinction does not serve any useful purpose. Attraction thus amounts to confusion of two neighboring articulations that have been allowed to drift closer and closer because their difference is never distinctive, since they characterize only such phonemes as are sufficiently distinguished by means of other features.

This attraction exerted by a closely knit pattern on marginal phonemes has been referred to as the filling of 'holes in the pattern'¹⁵ (Fr. 'cases vides'¹⁶, Sp. 'casillas vacías.'¹⁷) This phrase is undoubtedly picturesque, but it is apt to deter linguists from a painstaking analysis of the successive processes involved. 'Paper phonetics' has been severely and justly criticized. Juggling with the symbols of phonemic charts would be equally dangerous and reprehensible. Isolated phonemes do not rush into structural gaps unless they are close enough to be attracted, and whether they are attracted depends on a variety of factors which always deserve careful investigation. Furthermore, we shall see below that what looks like a hole on the chart does not necessarily correspond to a linguistically favorable combination of articulations. Yet it can not be denied that phonemes in groups tend to impose their articulatory types upon isolated phonemes¹⁸.

In dealing with pattern attraction, it is often tempting to oppose integrated to non-integrated phonemes, but it is more accurate to work with various degrees of structural integration. We have first of all to take into consideration phonemes whose phonic make-up and general behavior are that of an existing series, but which lack the partners that would integrate them in a correlation, e.g. /k/ in a language with /p/, /t/, /k/, /m/, /n/, but no /ŋ/. In such a case we might say that /k/ is ready for integration through the filling of the [ŋ] gap. In a pattern like

p	t	k
b	d	
m	n	ŋ

¹⁵ The term is found in K. L. Pike's *Phonemics*, Ann Arbor 1947, 117b.

¹⁶ Probably used for the first time by this author in *La phonologie synchronique et diachronique, Conférences de l'Institut de linguistique de l'Université de Paris* (1938) 6.53.

¹⁷ See Alarcos Llorach, *Fonología española*, Madrid 1950, 80-81.

¹⁸ For a detailed analysis of a clear case of pattern attraction, see A. Martinet, *The Unvoicing of Old Spanish Sibilants, Romance Philology* 5.139. In his pioneering article *Phonetic and Phonemic Change, Language* 12.15-22, A. A. Hill uses the term 'phonemic attraction' for a different phenomenon resulting in partial or total phonemic confusion; cf. 21.

/k/ is undoubtedly integrated, but less so than /p/ or /t/. It is clear of course that we could not speak of 'holes in the pattern' unless we reckoned with /k/, in the two preceding examples, as somehow integrated. In a language where an apical /l/ is the only lateral, it may be both theoretically and practically advisable not to place it in the same order as /t/, /d/, or /n/; and to consider it as non-integrated. But that language may present geminate consonants whose frequency is comparable to that of their simple partners in intervocalic position. Although, in a descriptive study, these geminates would still be analyzed as successions of two single consonants—so that geminated *t* would be /tt/, geminated *l* /ll/, and so forth—they would, on account of their frequency, play a functional role similar to that of single phonemes. We would be justified in speaking of a correlation opposing a series of single and one of geminated consonants; /l/ and /ll/ would thus be integrated into a correlation just as /t/ and /tt/ and /n/ and /nn/. But of course /t/ and /n/ would remain more fully integrated than /l/ because they would belong to other correlations than only that of gemination.

The theory of pattern attraction could accordingly be summarized by stating that the phonemes of a pattern tend to be as fully integrated as conflicting factors make it possible. This means that filling of holes may involve phonemes which already had some degree of integration, but which, through the process, will emerge as more fully integrated. Let us revert to the above-sketched Hauteville shift as a good illustration of this kind of process.

Both before and after the shift, the normal length vocalic phonemes of Hauteville can be ordered into three series characterized as front-retracted (/i/ type), front-rounded (/ü/ type), back-rounded (/u/ type), with four orders (or degrees of aperture) which we can designate as 1, 2, 3, and 4. The fourth order presents only one phoneme, /a/, in which front-back and retracted-rounded oppositions are neutralized. All these phonemes further enter a correlation composed of one nasalized and one non-nasalized series. The phonemes of the nasalized series are fewer than those of the non-nasalized one, which is frequently the case in similar patterns. Only the more open orders present nasal phonemes. This results from the fact that nasal articulation is detrimental to the clarity of the concomitant oral articulation since it implies that part of the air escapes through the nose and is thus lost for the oral cavity proper. Yet the wider the oral aperture, the more air will flow through it, so that open nasal vowels are likely to be more distinct than close ones. This may account for a frequently observed tendency for nasal vowel phonemes to become more and more open.

Before the shift, the two patterns, oral and nasal, must have been

1	i	ü	u		
2	e	ö	o	ē	
3		ɛ		ē	õ
4		a			ã

There are two gaps in oral order no. 3, to wit [œ] and [ɔ]. The [œ] gap is not rare in such patterns and may be easily accounted for: for a relatively large degree of aperture, it is more difficult to distinguish between retraction and

protrusion of the lips. The comparative rarity of /œ/ as a distinct phoneme is thus due to the same articulatory and acoustic factors as those that determine the frequency of a single phoneme for order no. 4. There are thus only two oral phonemes for the whole of the two most open orders. Since the opposition of /ɛ/ to /a/ is one not only of aperture but also of depth, it is understandable that speakers should have tended to neglect the difference between apertures 3 and 4, which was irrelevant in the rest of the oral pattern, and to stress the difference between front and back articulation, which was largely supported elsewhere. In the process /a/ passed from middle to back. The result was first an oral pattern with only three degrees of aperture

1	i	ü	u
2	e	ö	o
3	æ	a	

where the margin of security was wider between 2 and 3 than between 1 and 2.

The original nasal pattern had one more phoneme at the front than at the back; among nasal phonemes, /ē/ was less fully integrated than /ē/ or /ō/ since it was the only unit to combine nasality with aperture no. 2. We have seen that speakers would tend to open nasal vowels, and therefore /ē/ was exerting a pressure downward. In the frame of the nasal pattern, /ē/ could not become more open without threatening to impinge upon the domain of /ā/, which in its turn could hardly shift toward the back because of the proximity of /ō/; /ē/ was thus squeezed between the gradually opening /ē/ and the resistance of its more open and back congeners. Random weakly nasalized deviations of /ē/ were apt to be favored since there no longer was any /ɛ/ in the pattern. Eventually /ē/ was totally denasalized, and /ē/ could occupy its former position. The resulting situation is actually attested in dialects spoken a few miles from Hauteville where the pattern may be represented as follows:

1	i	ü	u			
2	e	ö	o			
3	(ē>) ε		[ɔ]	(ē>)	ē	ō
4		æ	a			ā

Here, a fourth degree of aperture has reassumed phonemic relevance, but only in the front series; at the back, aperture no. 3 is only represented by contextual variants of the /a/ phoneme. At Hauteville, all the allophones of /a/ have passed to [ɔ], and /æ/ has shifted back to middle position, hence:

1	i	ü	u			
2	e	ö	o			
3		ε	ɔ		ē	ō
4		a				ā

a pattern which shows much more complete integration than the original one.

A few very natural objections to the structural approach could be raised at this point: How is it that after so many millennia of uninterrupted speech prac-

tice, patterns should still be in need of structural integration? What has been called the original Hauteville pattern was of course 'original' only in the sense that we chose to make it the starting point of our research. But, just like any other Romance pattern, it was nothing but one of the numerous avatars of the Latin vocalic pattern, a pattern which may have enjoyed at some period a fair degree of integration. We have of course to assume that the trend toward structural integration is at work all the time. But how can we explain that there should always be grist for its mill? Why could phonemic patterns not reach perfect stability? Or do we mean that the beautifully balanced modern Hauteville pattern has reached such a stage of perfection that it would last forever if the dialect itself were not doomed to disappear in the course of the next sixty years?

These are many questions which require separate answers. First, what we have presented of the modern Hauteville pattern looks perfectly harmonious, but so much harmony may actually involve some strain on the physiological latitudes: the usage of certain speakers would seem to indicate that the back series, with its four phonemes, is somewhat too crowded, and this could be a germ of instability¹⁹. Second, we have left out the short vowel phonemes whose pattern shows clear signs of disintegration²⁰, and if the dialect were to live, we or our successors might probably witness a total reshuffling of the vocalic pattern which might be necessitated by a dephonemicization of quantitative differences. Completely harmonious patterns are probably never reached, and even if one were found which would seem to stand close to structural perfection, it would be at the service of a language which, like all languages, would be used for the expression of changing needs. These needs, acting through syntax, lexicon, morphology, tempo, intonation, and others, would ultimately manage to destroy the beautiful phonological balance. Third, languages do not evolve in ivory towers. The Hauteville dialect for instance has, for centuries, been spoken by an increasing number of bilinguals whose medium of inter-regional communication and intellectual expression is French. Before that time, it was a local variety of a larger dialectal unit whose most prominent and prestige-endowed users were bilinguals, also with French as a medium of wider communication. Even before French was actually spoken in the region by the leaders of the community, a number of linguistic features of all sorts, phonological as well as others, must have seeped through chains of contiguous forms of speech all the way from Northern France, politically dominant since the rise of the Frankish empire. Dominant cultural languages do not necessarily preserve the integrity of their patterns better than local patois when they spread over large heterogeneous areas and become the linguistic mediums of whole nations.

All this accounts for the never-ceasing phonological fermentation that can be observed practically everywhere. There will always be holes in patterns and phonemes moving in to fill them. New series and new orders will appear, result-

¹⁹ Description phonologique . . . 36 and 38.

²⁰ Cf., *ibid.*, what is said, pp. 44 and 56, about a tendency toward making ξ the phonemic equivalent of zero.

ing either from general reshufflings or from the coalescence of successive phonemes of the spoken chain, the result of new accentual conditions, articulatory imitations, etc. These new series and orders will not always be complete from the start; for some time there will remain gaps which ensuing generations may fill through either sound change or borrowings.

The creation of a hushing order in early Castilian affords an interesting illustration of the ways through which a new phonemic type can expand by convergence of the most varied elements²¹: the first hushing units must have resulted from the coalescence of apical articulations with neighboring newly evolved Romance [j], hence word-medial /č/ and /š/; the corresponding holes in the word-initial pattern must have been filled mainly by borrowings from neighboring dialects. Word-initially the voiced hushing phoneme (probably [ǰ]) was normally a reflex of Vulgar Latin *yod*, but intervocalic *yod* was never modified, and the corresponding hole in the word-medial pattern was filled by early Romance /lʲ/ passing to [ʒ]. This rather startling treatment can be understood only if we keep in mind that geminated (at that time probably just strong) *l* was tending toward its modern [lʲ] reflex, and was exerting a pressure upon earlier /lʲ/.

This Castilian process further affords a welcome illustration of what we might call the action of a phonemic catalyst. We have so far assumed that functional yield, even if it were practically nil, would act as a deterrent against merger. But if the opposition in question is between a fully integrated phoneme and one that is not, or upon which some phonemic pressure is being exerted, a minimal functional yield will not act as a deterrent and, on the contrary, an articulatory attraction is likely to take place. In simpler, less technical terms, if a well-integrated phoneme is extremely rare, it may attract a not so well integrated neighboring unit. In the case of early Castilian there must have been a few words in which the [zi] cluster was preserved, having escaped the metathesis whereby *BASIUM* became *beso*; these rare [zi] clusters naturally yielded [ʒ] as in *frijuelo* from *PHASEOLUM*²². This new /ʒ/ phoneme occurring word medially was well integrated in an order which presented, further, word medial /č/ and /š/. But the instances of this phoneme were so few as to exclude any homonymic conflict if what had been /lʲ/ merged with it; /lʲ/, which was being unintegrated by the pressure of a former /ll/, must have been attracted by /ʒ/. Attraction, as we have presented it before, results from the confusion of two articulations when concomitant articulations suffice to preserve phonemic identity. In the case of a catalyst we have the confusion of two characteristic articulations when this does not actually result in confusion of words and forms. Functionally the two phenomena are quite parallel. They both result in articulatory economy without any impairing of communication.

IV. INERTIA AND ASYMMETRY

The most serious resistance to phonemic integration stems from the limitations set up by human physiology to the combination of the most varied articu-

²¹ See *The Unvoicing of Old Spanish Sibilants*, 135-136, 140-141.

²² See Vicente García de Diego, *Gramática histórica española*, Madrid 1951, 103.

lations. The articulations themselves may conflict if they involve neighboring organs. But, more often, the incompatibility will be acoustic, i.e. hearers will find it difficult to perceive a difference between various combinations of the same type, at least in ordinary speech conditions.

We have, in what precedes, indicated in several occasions how some physiological necessities may counteract phonemic integration. We have pointed out that vocalic correlations are quite generally much better represented in the higher than in the lower orders: /æ/ as a phoneme is probably rarer than /ø/ or /ü/; patterns with three series (e.g. of the /i/, /ü/, and /u/ types) practically never keep these three series distinct for the lower order, which is easily accounted for by pointing out that, with maximally open jaws, the lips will be automatically retracted, and that it will become difficult to distinguish between a front and a back oral cavity. The difference of aperture between [o] and [u] will be smaller than that between [e] and [i], although it corresponds to the same maxillary angle. From the point of view of the speaker who has to control the play of his muscles, the proportion [o]:[u] = [e]:[i] will be correct; but acoustically the distinction between [e] and [i] will be clearer than that between [o] and [u]. For the same number of phonemes in the front and in the back series the margins of security will be narrower at the back than at the front, and this may partially account for diverging compartments of the two series. We have also seen that concomitant nasalization affects the clarity of vocalic articulations, which means that there are articulatory combinations which are acoustically good, and others which are not so good.

If phonemes were not of phonic nature but resulted e.g. from combinations of flags, if /p/ for instance, instead of being, say, voiceless and bilabial, was performed by stringing the Stars and Stripes and the Union Jack along the same line, /t/ by combining the Stars and Stripes with the French Tricolor, /d/ by adding to the latter the Danish Dannebrog, and so forth, any combination of two flags would be just as good as any other. We can not combine voiced and voiceless articulation, but we could combine the Stars and Stripes with the Dannebrog. Furthermore, if in order to make morphemes or words, we should produce a succession of different flag combinations, any combination could follow any other, so that a word could easily be composed of /ptd/ if the respective units involved were performed as described above. Not so of course with distinctive units performed as sounds. The vowel-consonant dichotomy, with its syllabic corollary is imposed upon us by the nature of the so-called speech organs. The vocalic and consonantal patterns may overlap in certain languages, but they are always organized according to two different models. Even if we should agree with Jakobson that, in human speech generally, the coordinates are the same for vowels and consonants, we would find, in individual languages, no constant parallelism between the two patterns: Czech would have a quadrangular consonantal system and a triangular vocalic one:

t	p	i	u
c	k	a	

and Finnish just the reverse:

t	p	i	u
k		æ	ɑ

The necessity of alternating, in the spoken chain, between closed and open articulatory complexes, which naturally result from the combination of different types of articulation, opposes the integration of all the phonemes of a language into one closely-knit pattern: the opposition of voice to its absence plays a great role in consonantal matters; with vowels, on the contrary, voice is almost indispensable and therefore phonemically irrelevant. Some distinctive features can be found to characterize both vowels and consonants, but not too easily in the same language. Palatalization of consonants and front vowel articulation may be conceived as the same feature, with whatever actual difference that may exist being determined by concomitant vocalic or consonantal features; but where, as in Russian, we might believe that the two coexist, more careful observation will often show that if consonants enter a correlation of palatalization, the vowel series will actually be distinguished by the play of the lips: Russian /i/ is frequently pronounced far back in the mouth, and /u/ may, in certain contexts, be performed as [ü].

The case of nasality is interesting since it will combine with both consonantal and vocalic articulations, but not equally favorably with all consonants and all vowels: most languages distinguish /b/ and/or /p/ from /m/, /d/ and/or /t/ from /n/; /ñ/ is probably about as frequent as its non-nasal counterparts /ɲ/ and /ç/, but /ŋ/ as a distinct phoneme is rarer than /g/ and /k/, which may be due to a tendency of the two velar articulations to conflict. Nasal fricatives as distinct phonemes are extremely rare since friction requires a pressure which can not be obtained if the air is allowed to flow unhindered through the nose. The rarity of liquid nasals can be accounted for in very much the same way. We have already seen that, for similar reasons, nasality combines better with open than with closed vocalic articulations; but, in any case, nasal vowels are never so clear as oral ones, and this should account for their relative infrequency and instability as phonemic units. With stop articulations, experience shows that nasalization as such is easily perceived, but, unless [m] and [n] are clearly exploded, the place of oral occlusion can be identified only with difficulty, as shown by the frequent neutralization of nasal consonants in syllable final position where they are assimilated to the following consonant and, if word final, merged into [n] or [ŋ]²³.

All this means of course that there will be gaps in patterns which are not likely ever to be filled, or if they are, only as a result of a fairly exceptional concurrence of circumstances. The phonemes characterized by acoustically or articulatorily unfavorable combinations will, everything else being equal, be less stable than others combining features with a high degree of compatibility. Orders and

²³ For a general survey of the restrictions imposed upon the expansion of correlations by the inertia and assymetry of speech organs, see A. Martinet, *Rôle de la corrélation dans la phonologie diachronique*, *TCLP* 8.273-288.

correlations will tend to expand as far as human physiology, and certain conditions inherent to the specific language, will permit. To an original /m, n/ nasal series a /ñ/ is added in Vulgar Latin when a palatal order is formed. In Germanic a tendency to simplify the clusters of nasal and homorganic oral stop (cf. the discrepancy between the spelling and the pronunciation of Eng. *lamb, comb*; Danish has gone farthest with its mute *d* in *land*) has resulted in the phonemicization of the dorsal nasal [ŋ]; the same phenomenon is found in Sanskrit as a result of the reduction of all final consonant clusters to one phoneme, hence /n/+dorsal > /ŋ/. A general weakening of implosion is likely to yield a whole pattern of nasal vowels. In all these cases, we can hardly speak of filling of holes since the appearance of the new phoneme or phonemes is obviously determined by trends which have nothing to do with pattern attraction. In most of these shifts the main factor probably was least effort, which was allowed to play in certain domains where communicative needs offered little resistance and in the frame of certain prosodic situations involving a specific nature of accent and a given pattern of syllabification.

Linguistic evolution in general can be conceived of as regulated by the permanent antinomy between the expressive needs of man and his tendency to reduce his mental and physical exertions to a minimum. On the plane of words and signs, every language community will have to strike a balance between an expressive trend toward more numerous, more specific, and less frequent units, and natural inertia which favors fewer, more general, and frequent ones. Inertia will be there all the time, but expressive needs will change, and the nature of the balance will vary in the course of time. Uneconomical expansion, i.e. one which would entail more exertion than the community would deem worth while, will be checked. Inertia, when felt to be excessive, i.e. detrimental to what is felt to be the legitimate interests of the community, will be censored and punished. Linguistic behavior will thus be regulated by what Zipf has called the 'Principle of least effort'²⁴, a phrase which we would rather replace by the simple word 'economy'²⁵.

Linguistic economy is ultimately responsible for the very existence of phonemic articulation: the inertia of the organs involved in the production and reception of speech phenomena makes it impossible for any normal human vocal language to afford a specific homogeneous and distinctive phonic product for every linguistic sign. Yet communication requires distinct expression for each. A satisfactory balance is reached by limiting to a few dozens the number of specific and distinctive expressive units, the phonemes, and by combining them successively into distinct signifiers. Economy is further achieved by making these units result from combinations of non-successive phonic features, but of course only such combinations as will best serve communicative purposes. We find here an antinomy between what we have called the trend toward phonemic integration and the inertia and asymmetry of the organs opposing the inclusion of all phonemes into a theoretically perfect, immutable pattern. When a vocalic

²⁴ *Human Behavior and the Principle of Least Effort*, Cambridge, Mass., 1949, 56-133.

²⁵ La double articulation linguistique, *TCLC* 5.34.

pattern presents four phonemically relevant degrees of aperture in the front series, phonemic integration will tend to preserve or produce four relevant degrees of aperture in the back series. When the vocalic structure of Classical Latin with its three relevant degrees of aperture eked out by a quantitative distinction was, by elimination of the latter, reorganized into a four order pattern, in most of the Romania the reshuffling followed the same procedure in the front and at the back. But for the same number of phonemes in the two series the asymmetry of the organs entailed narrower margins of security in the back series. The mutual pressure could be relieved by diphthongization of the phonemes of one order or more. But in such a case diphthongization resulting from a gradual increase or decrease of the maxillary angle in the course of the articulation of the vowel would not be restricted to the back series but would extend to all the vowels of the same order: where *o* became *uo*, *e* would become *ie*, and pattern symmetry would be preserved. A tendency to merge orders 2 and 3 might also affect equally the front and back series. If on the contrary the pressure was relieved by gradual fronting of /u/, the back series would from then on only distinguish between three orders and the pattern remain asymmetrical²⁶.

The effect of the asymmetry of speech organs is also clear in the case of consonant patterns, not only when we think of the obvious restrictions to the spread of various correlations, but also in the course of certain wholesale mutations when it is found that some orders proceed more rapidly than others. An articulatory weakening will as a rule decisively affect bilabials before the other orders: /p/ is frequently weakened to [h] or zero where /t/ and /k/ are preserved. A shift affecting strongly articulated consonants is likely to act more rapidly on the phonemes of the apical order, as seen for instance in Modern Danish where, of the three energetically articulated aspirates, /t/ is the first to show unmistakable signs of affrication²⁷.

A full awareness of the existence of this asymmetry will, in many cases, help to account for such changes as seem only to affect a single phoneme. It will make clear that this seemingly isolated change is, in fact, the outcome of a general trend, variously warped by specific conditions, hastened if these are favorable, delayed if they are not: a general articulatory strengthening acting upon a series of spirants may change a [p] into [t], but will never change an [f] into a stop, and will simply make the articulation of [s] more energetic; [ɣ] may remain a spirant when [ð] is made an occlusive, whereas [b] becomes a stop in strong (e.g. initial) positions but remains a spirant elsewhere²⁸.

In combination with the various factors considered above, asymmetry should go a long way toward accounting for most of the phenomena which could be described as local modifications or reorganizations of phonemic patterns. When, in such matters, we find two sections of the same language community striking out into different phonological paths, we shall find, as a rule, that both courses

²⁶ See Haudricourt-Juilland, 17-58, 98-113.

²⁷ Cf. Henrik Abraham, *Etudes phonétiques sur les tendances évolutives des occlusives germaniques*, Aarhus 1949, 108.

²⁸ Cf., e.g. A. Martinet, Some Problems of Italic Consonantism, *Word* 6.35-41.

had their functional, structural, and physiological justifications, and that the factors responsible for the divergence must have been such imponderables as will always escape scientific scrutiny when human behavior is involved.

V. PROSODY AND NON-PHONEMIC PRESSURE

Still largely unexplained remain such changes as affect whole orders, series, correlations, bundles, and even the system, both vocalic and consonantal, in its entirety. When investigating these, all the previously discussed factors should be kept in mind. It will be found that the direction and amplitude of every one of them is, as a rule, largely dependent on the nature and diachronic comportment of the structural environment. The behavior of orders and series will be reminiscent of the one we have ascertained above in the case of our three phonemes A, B, and C, and this is easily understandable since the whole of an order or of a series differs from the whole of another order or series in exactly the same way as we assumed A differed from B or from C. Orders and series will merge, just like phonemes, but in most well-documented cases we shall find that, for classes as well as for individual phonemes, a relatively low functional yield may have played a role. But if we can thus account for some of the modalities of these changes, we have so far no way of telling what started them.

Among the possible factors in such changes we shall distinguish between internal and external ones. By internal factors, we mean here the influences exerted upon the pattern of phonemes by those sections of the linguistic structure we have not so far considered, namely the complex of prosodic, 'suprasegmental', features, and the system or systems of meaningful elements. Among external factors, we should distinguish between the influences exerted by other linguistic structures, those of other dialects of the same language or of other languages, and the fairly mysterious non-linguistic factors whose importance may well have been grossly exaggerated by our predecessors. We can not deal here with the influence of other linguistic structures²⁹. Let it suffice to say that they would deserve to be taken into consideration far more than has generally been done so far.

One would have a right to object to our separating, in the present survey, prosodic from phonematic features. It can indeed not be denied that, at least in some languages, suprasegmental features can be arranged into patterns very similar to those we have established for phonemes. For instance, the tones of many south-eastern Asiatic languages could be grouped into orders and series just like the phonemes of the same languages. Most of what we have said above about the comportment of phoneme patterns would apply just as well to such tone patterns. The function of these tones is distinctive like that of phonemes, and the mutual diachronic relations they entertain with the phoneme patterns must be of the same type as those between different sections of such patterns. Accent, which can be defined as prominence given to one syllable in the word, or whatever meaningful unit has prosodic relevance, may at times assume some

²⁹ Cf. The Unvoicing of Old Spanish Sibilants, generally, and 152-156 in particular.

distinctive function, particularly when it is found to appear in two or more phonemically distinct types. These two or more distinct types may pattern in very much the same way as tones proper, their difference being generally one of pitch or melody. In so far as they exert distinctive function, prosodic features have, from a diachronic standpoint, to be considered together with the purely distinctive features we find combined into phonemes. They form with vowels and consonants three natural classes of distinctive units. We have, in what precedes, pointed to a definite tendency to avoid the use, in the same language, of certain features for both consonants and vowels. In a similar way, it is found that, for instance, prosodic intensity, if distinctive, usually excludes phonemically relevant vocalic intensity manifested under the form of quantity³⁰.

Yet the basic function of accent as such is not distinctive. When its place in the unit it characterizes is not predetermined by the phonematic make-up of that unit, it may occasionally evince some sort of distinctive power (cf., e.g. Sp. *córtés-cortés*). But this is normally a by-product. Accent is really there to characterize and localize the word (or a certain type of morpheme or phrase) in the spoken context. If the localization is approximate, its function has been called culminative. If it is accurate, its function is demarcative. The true function of accent is less clear in languages where it is a traditional feature than where a new accentual pattern is being developed at the expense of tradition. In contemporary French, for instance, the weak traditional phrase-final accent seems to be increasingly overshadowed by what has been called the accent of insistence. This accent has two concurrent forms, an emotional and an intellectual one. The former is usually characterized by stressing and lengthening of the first consonant of the word, the latter by some prominence given to the first syllable. Formally, the two varieties are clearly distinct only in words beginning with a vowel (*im'possible* vs. *'impossible*), and, even then, emotional accent can be heard on the first syllable (*c'est 'impossible*)³¹. In what could be dubbed 'didactic style', first syllable prominence is widely prevalent and its function is clearly demarcative; by setting individual words apart from the context, it gives a bolder relief to the successive articulations of thought. The functional difference between phonemes and such an accent is obvious: phonemes contribute only indirectly to the expression of the semantic contents of language, and therefore the expressive needs of man will as a rule affect phonemes only through devious channels; this still optional accent is an immediate reflex of these expressive needs; its intensity will vary from one utterance to another and mirror exactly the communicative purposes of the speaker. Such an accent is linguistic in the narrow sense of the word because its existence is determined by an inherited convention, but its arbitrariness is highly limited in the sense that, once it has been accepted in principle by a community, the details of its actual use will be regulated by what we may call psychological factors. Even when such an accent

³⁰ See, e.g. N. S. Trubetzkoy, *Grundzüge der Phonologie*, Prague 1939, 180, or *Principes de phonologie*, Paris 1949, 215.

³¹ See J. Marouzeau, *Quelques aspects du relief dans l'énoncé*, *Le français moderne* 13.165-168, with references to former contributions by the same author.

has been stabilized and has ceased to be optional, it will be liable to various degrees of intensity whereby meaning will be conveyed directly from speaker to hearer.

What has just been said about accent applies largely to all the prosodic features grouped under the heading of intonation. In many languages, an utterance final melodic rise is the functional equivalent of an interrogative morpheme such as Fr. *est-ce que*, Russian *li*, Lat. *-ne*. Once tone, as automatic accompaniment of every syllable or mora, and occasional phonemically distinctive uses of accents have been discounted, prosodic features may be said to belong to the same linguistic plane as meaningful units, and, just like them, to be liable to be directly affected by communicative needs. But the physical nature of their performances is such as to exert a deep influence on those of the phonematic units of the spoken chain.

This is of course clear in the case of stress accent, which, when particularly strong, is known to play havoc with inflexional endings³² if it does not happen to bear on them, and which must also be ultimately responsible for the most revolutionary reshufflings of phoneme patterns. The umlaut phenomenon illustrates most clearly how processes that originally affect phonemic units in the spoken chain, lead eventually to a reorganization of the phoneme pattern. The prominence of a syllable can only be achieved at the expense of the other syllables of the word; as stress increases in one part of the word, the other parts become more weakly articulated, hence blurring of vocalic distinction and, frequently, syncope. This, in itself, does not affect the inventory of the phonemes, only their distribution. In many cases, this would be expected to result in the elimination of a large number of useful distinctions; OHG *scōno* and *scōni* would have merged into **scōn*, if speakers had not unconsciously favored in the second form such deviations of /ō/ as were determined by a tendency to anticipate the front articulation of /i/; hence, of course, the eventual split of /ō/ into /o/ and /ō/ attested in German *schon* and *schön*. This shows how the phoneme pattern can be made to expand when the average number of phonemes per word is diminishing, and how, more generally, demarcative needs can enlarge the phoneme inventory and lead to a reshaping of the system.

³² It is clear of course that the energy with which stressed syllables are pronounced varies greatly from one language to another. Stress is, for instance, decidedly weaker in Spanish than in Italian, and probably weaker in standard Italian than in standard German. Some languages, like German, show close contact of stressed short vowels and following consonants; others have loose contact in such cases. When Bloomfield writes, *Language*, 385: "many languages with strong word stress do not weaken the unstressed vowels" and cites among them Italian, Spanish, Czech, and Polish, he obviously wants to convince his readers that stress as such can not be held entirely responsible for vowel blurring. But his examples do not carry conviction: neither Czech nor Polish accent can be said to be particularly energetic; in standard Castilian, accent is uncommonly weak. In such matters, it is particularly important to distinguish between the successive stages of the same language. It is commonly assumed that 'Germanic accent' is vowel-blurring. But it remains to be proved that, e.g. in contemporary English and German, absence of stress is actually conducive to the blurring of vocalic distinctions. For Standard German, at least, this seems highly doubtful.

But of course the well-known umlaut process is only one way, among many others, whereby prosodic non-distinctive features can affect the phoneme pattern. This, a central problem of diachronic linguistics, should receive far more attention than has been granted to it so far, since prosodic features are the most normal channel through which the varying communicative needs of speech communities can influence the pattern of distinctive features. As factors of phonological changes, prosodic features are extremely powerful precisely because they are immediate responses to the needs of expression. The speaker of Modern French who makes an extensive use of optional initial accent is prompted to do so by a desire to make his statements as clear and convincing as possible. How could he imagine that he may be paving the way toward the establishment of initial stress as an automatic feature of the language, which may eventually result in blurrings, mergers, and synopes? We do not mean hereby that the phonetic trends launched by, say, a strong stress accent will develop blindly, ruthlessly destroying any piece of linguistic machinery that happens to be in their way. If the vocalism of unstressed endings plays in the economy of the language too important a role to be wiped out, speakers may be induced to save some of their articulatory energy for the final syllable of every word. When, as it seems, prehistoric Latin developed a word initial stress, wide-spread blurring of vocalic timbres took place in medial syllables, but final syllables, in which lay the expression of most morphological categories, were hardly affected. At a much later period, when the language had extensively weakened its adverbs into mere grammatical tools, a new onslaught of prosodic intensity resulted in the wholesale massacre of a declensional pattern which by that time must have become a burden rather than a real help.

In this rapid survey of the ways through which communicative needs may influence the phoneme pattern, we should of course mention again the assumed role of the functional yield of oppositions. But, by the side of this conservative action, it would remain to be seen whether the necessities of expression could not, in some cases, be directly instrumental in enlarging some sections of the phoneme pattern. If some phonemically relevant feature, say, glottalization, happened to be the frequent mark of a morphological or lexical category, could it not be imagined that speakers would be tempted to combine it with new articulations and extend it beyond those sections of the phoneme pattern where we would normally expect to find it? The example we have to offer does not illustrate exactly this type of action, but rather the extension of a correlation beyond its expected range under the pressure of an all-pervading pattern of morphophonemic alternations; most Breton consonants may be grouped into two series of strong and weak units; weak stops are voiced and their strong counterparts are generally voiceless; the strong member of phonetically voiced pairs often evinces a tendency to devoicing, a tendency fully developed in Welsh. This strong-weak opposition frequently coincides with the morphophonemic pattern of 'lenited' versus 'non-lenited' consonants; /b/ for instance is the phonemic weak counterpart of /p/ and also, in the morphology, the 'lenited' equivalent of it. Since, however, [b] could also be the 'non-lenited' counterpart of /v/, a

tendency has been at work to distinguish phonemically between morphologically 'strong' [b] alternating with /v/ and morphologically weak [b] alternating with /p/. As a result, we have today to reckon with two *b* phonemes: a strong *b* transcribed /bb/ which Breton speakers tend to unvoice and which they have in consequence some difficulty in keeping apart from /p/, and a weak fully voiced /b/. The distinction is widely neutralized, or rather it has only got a foothold in a very specific position, namely word initially within an utterance if the last phoneme of the preceding word is a vowel³³.

This last illustration and the preceding suggestions should by no means be conceived of as exhausting all possibilities of direct diachronic influence of one of the linguistic planes upon the other. Neither in this final section nor in the previous parts of this study has there been any attempt at being exhaustive. A complete functional and structural theory of phonological evolution will have to be based upon a much larger body of structurally sifted material than is available to date. The justification for the sketchy and tentative outline which has been presented here is that it may incite diachronically inclined linguists to utilize in their investigations some of the conclusions arrived at after two decades of phonemic research.

Columbia University

³³ See F. Falc'hun, *Le système consonantique du breton*, Rennes 1951, 63-65.