THE PSYCHOLOGICAL REALITY OF PHONEMES*

The concept of the “phoneme” (a functionally significant unit in the rigidly defined pattern or configuration of sounds peculiar to a language), as distinct from that of the “sound” or “phonetic element” as such (an objectively definable entity in the articulated and perceived totality of speech), is becoming more and more familiar to linguists. The difficulty that many still seem to feel in distinguishing between the two must eventually disappear as the realization grows that no entity in human experience can be adequately defined as the mechanical sum or product of its physical properties. These physical properties are needed of course to give us the signal, as it were, for the identification of the given entity as a functionally significant point in a complex system of relatedness; but for any given context it is notorious how many of these physical properties are, or may be, overlooked as irrelevant, how one particular property, possessing for the moment or by social understanding an unusual sign value, may have a determinedness in the definition of the entity that is out of all proportion to its “physical weight.”

As soon, however, as we admit that all significant entities in experience are thus revised from the physically given by passing through the filter of the functionally or relatedly meaningful, as soon as we see that we can never set up a scale of added or changed meanings that is simply congruent to the scale of physical increments, we implicitly make a distinction, whether we know it or not, between the phoneme and the sound in that particular framework of experience which is known as language (actualized as speech). To say that a given phoneme is not sufficiently defined in articulatory or acoustic terms but needs to be fitted into the total system of sound relations peculiar to the language is, at bottom, no more mysterious than to say that a club is not defined for us when it is said to be made of wood and to have such and such a shape and such and such dimensions. We must understand why a roughly similar object, not so different to the eye, is no club at all, and why a third object, of very different color and much longer and heavier than the first, is for all that very much of a club.

Some linguists seem to feel that the phoneme is a useful enough concept in an abstract linguistic discussion—in the theoretical presentation of the form of a language or in the comparison of related languages—but that it has small relevance for the actualities of speech. This point of view seems the reverse of realistic to the present writer. Just as it takes a physicist or philosopher to define an object in terms of such abstract concepts as mass, volume, chemical structure, and location, so it takes very much of a linguistic abstractionist, a phonician pure and simple, to reduce articulate speech to simple physical processes. To the physicist, the three wooden objects are equally distinct from each other, “clubs” are romantic intrusions into the austere continuities of nature. But the naîve human being is much surer of his clubs and poles than of unnamed objects to be hereinafter defined in physical terms. So, in speech, precise phonetic stations can be abstracted only by patient observation and frequently at the expense of a direct flouting of one’s phonic (one should say “phonemic”) intuitions. In the physical world the naîve speaker and hearer are sensitive to sounds, but what they feel themselves to be pronouncing and hearing are “phonemes.” They order the fundamental elements of linguistic experience into functionally and aesthetically determinate shapes, each of which is carved out by its exclusive laws of relationship within the complex total of all possible sound relationships. To the naîve speaker and hearer, sounds (i.e., phonemes) do not differ as five-inch or six-inch entities differ, but as clubs and poles differ. If the phonetician discovers in the flow of actual speech something that is neither “club” nor “pole,” he, as phonetician, has the right to set up a “halfway between club and pole” entity. Functionally, however, such an entity is a fiction, and the naîve speaker or hearer is not only driven by its relational behavior to classify it as a “club” or a “pole,” but actually hears and feels it to be such.

If the phonemic attitude is more basic, psychologically speaking, than the more strictly phonic one, it should be possible to detect it in the unguarded speech judgments of naive speakers who have a complete control of their language in a practical sense but have no rationalized or consciously systematic knowledge of it. “Errors” of analysis, or what the sophisticated onlooker is liable to consider such, may be expected to occur which have the characteristic of being phonetically unsound or inconsistent but which at the same time register a feeling for what is phonemically accurate. Such “errors,” generally overlooked by the practical field linguist, may constitute valuable evidence for the dynamic reality of the phonemic structure of the language.

In the course of many years of experience in the recording and analysis of unwritten languages, American Indian and African, I have come to the practical realization that what the naîve speaker hears is not phonetic elements but phonemes. The problem reaches the stage of a practical test when one wishes to teach an intelligent native, say one who can

* Published originally in French under the title “La Réalité psychologique des phonèmes,” Journal de Psychologie Normale et Pathologique, 30 (1933): 247–265.
read and write English reasonably well and has some intellectual curiosity besides, how to write his own language. The difficulty of such a task varies, of course, with the intelligence of the native and the intrinsic difficulty of his language, but it varies also with the “phonemic intuitiveness” of the teacher. Many well-meaning linguists have had disappointing experiences in this regard with quite intelligent natives without ever suspecting that the trouble lay, not with the native, but with themselves. It is exceedingly difficult, if not impossible, to teach a native to take account of purely mechanical phonetic variations which have no phonemic reality for him. The teacher who comes prepared with a gamut of absolute phonetic possibilities and who unconsciously, in spite of all his training, tends to project the phonemic valuations of his own language into what he hears and records of the exotic one may easily befuddle a native. The native realizes when what he is taught “clicks” with what his phonological intuitions have already taught him; but he is made uncomfortable when purely phonetic distinctions are pointed out to him which seem real enough when he focuses his attention on them but which are always fading out of his consciousness because their objective reality is not confirmed by these intuitions.

I have selected for brief discussion five examples of phonemic versus phonetic hearing and writing out of many which have come to me in the course of my experience with natives and students. In each of these, it will be observed, we have clear evidence of the unconscious reinterpretation of objective facts because of a disturbing phonological preparedness not precisely adjusted to these facts.

I. When working on the Southern Paiute language of southwestern Utah and northwestern Arizona I spent a little time in trying to teach my native interpreter, a young man of average intelligence, how to write his language phonetically. Southern Paiute is an unusually involved language from the phonological standpoint and, as my point of view at that time stressed phonetic accuracy rather than phonemic adequacy, I doubt if I could have succeeded in teaching him well enough to satisfy my native. The native realizes when what he is taught “clicks” with what his phonological intuitions have already taught him; but he is made uncomfortable when purely phonetic distinctions are pointed out to him which seem real enough when he focuses his attention on them but which are always fading out of his consciousness because their objective reality is not confirmed by these intuitions.

Tony was not “hearing” in terms of the actual sounds (the voiced bilabial \( \beta \) was objectively very different from the initial stop) but in terms of an etymological reconstruction: \( \text{pa}' : \text{“water” plus postposition } \text{*-pa}’ \text{ “at.”} \) The slight pause which intervened after the stem was enough to divert Tony from the phonetically proper form of the postposition to a theoretically real but actually nonexistent form.

To understand Tony’s behavior, which was not in the least due to mere carelessness nor to a tendency of the speakers of this language “to confuse sounds,” to quote the time-worn shibboleth, we must have recourse to the phonology of Southern Paiute. The treatment of the stopped consonants may be summarized in the following table:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Labial</td>
<td>p</td>
<td>( \beta )</td>
<td>( m_p )</td>
<td>( p' )</td>
<td>( p )</td>
</tr>
<tr>
<td>Dental</td>
<td>t</td>
<td>t</td>
<td>( n_t )</td>
<td>t</td>
<td>t</td>
</tr>
<tr>
<td>Guttural</td>
<td>k</td>
<td>k</td>
<td>( n_k )</td>
<td>k</td>
<td>k</td>
</tr>
<tr>
<td>Labialized guttural</td>
<td>( k_w )</td>
<td>( \gamma_w )</td>
<td>( \eta_k )</td>
<td>( k_w )</td>
<td>( k_w )</td>
</tr>
</tbody>
</table>

The postvocalic forms of the stops of types 1, 2, and 3\( a \) are further modified before an unvoiced vowel, the voiced spirants becoming unvoiced spirants (\( t, R, x, xW \)),\(^1\) and the nasalized and geminated stops becoming aspirated (\( m_p', p'', n', t', \eta_k', k', \eta_kW, k'W \)). It is impossible here to give a systematic idea of the phonologic processes which bring about the sound interchanges within a given articulatory series, but it is important to know that the spirantized, nasalized, and geminated stops can occur only in postvocalic position and that they are largely determined by the nature of the element (stem or suffix) which precedes them and which may be said to have an inherently spirantizing, nasalizing, or geminating force. The stem \( \text{pa}' - \) is a spirantizing stem, and the spirantizing of a theoretical \( \text{*-pa}’ \text{ “at”} \) to \( \beta a’ \) is parallel to the spirantizing of \( \text{pa}’ - \) “trail” to \( \beta a’ - \) in such a compound as \( \text{pa}’ \beta a’ - \) “water-trail.” In other words, the language is so patterned that examples of type \( \text{pa}’ - \) lead to the proportion \( \*\text{pa}’ : \*\beta a’ \) and, while \( *\text{pa}’ \text{ “at”} \) does not actually exist as an independent element but must always be actualized in one of the three possible postvocalic forms, its theoretical existence suddenly comes

\(^1\) \( W \) represents voiceless \( W \)

\(^2\) \( W \)
to the light of day when the problem of slowly syllabifying a word is presented to a native speaker for the first time. It then appears that the -\( \beta \alpha \) of speech behavior, as a self-contained syllabic entity without immediately preceding syllable, is actually felt as a phonologic \( pa' \), from which it differs in two important phonetic respects (voiced, not voiceless, consonant; spirant, not stop).

All this has an important bearing on the construction of a maximally correct orthography of Southern Paiute, if by "maximally correct" we mean, not most adequate phonetically, but most true to the sound pattern of the language. As it happens, there is reason to believe from both internal and comparative evidence that the spirantized form of a consonant is its normal or primary form after a vowel and that the nasalized and geminated forms are due to the emergence of old nasal and other consonants that had disappeared in the obsolete form of the preceding element. It follows that the postvocalic \(-\alpha \) is more closely related functionally to a simple initial \( p- \) than is the postvocalic \(-p\) (after unvoiced vowel), which must always be interpreted as a secondary form of \(-p\). These relations are summarized in the following table of theoretical nonfinal forms.

<table>
<thead>
<tr>
<th>Phonetic Orthography</th>
<th>Phonologic Orthography</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. pa-</td>
<td>pa-</td>
</tr>
<tr>
<td>2. pa( \alpha )-</td>
<td>pa( \alpha )-</td>
</tr>
<tr>
<td>3. pa( \alpha )A-</td>
<td>pa( \alpha )A-</td>
</tr>
<tr>
<td>4. pap'( \alpha )-</td>
<td>pap'( \alpha )-</td>
</tr>
<tr>
<td>5. pa( \alpha )A-</td>
<td>pa( \alpha )A-</td>
</tr>
<tr>
<td>6. pap'A-</td>
<td>pap'A-</td>
</tr>
</tbody>
</table>

The phonetic orthography is more complex and, in a sense, more adequate, but it goes against the grain of the language in one important respect, for it identifies the second \( p \) in type 5 with the initial \( p \), which is phonologically unsound. The phonologic orthography, on the other hand, is useless for one who has not mastered the phonology of the language, as it leads, or seems to lead, to incorrect pronunciations which would have the cumulative effect of making the language, so read, entirely unintelligible to a native. To a slightly schooled native, however, there can be no serious ambiguity, for the phonetic forms result from the phonologic only by the application of absolutely mechanical phonetic laws of spiran-

tizing, alternating stresses, and unvoicing. It is not necessary to deal with these laws here but we can indicate their operation by the following table of theoretical final forms:

<table>
<thead>
<tr>
<th>Phonetic Orthography</th>
<th>Phonologic Orthography</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. pa( \beta )A</td>
<td>pa( \beta )A</td>
</tr>
<tr>
<td>2. pa( \beta )'</td>
<td>pa( \beta )'</td>
</tr>
<tr>
<td>3. pa( \beta )A</td>
<td>pa( \beta )A</td>
</tr>
<tr>
<td>4. pa( \beta )'A</td>
<td>pa( \beta )'A</td>
</tr>
<tr>
<td>5. pa( \beta )'</td>
<td>pa( \beta )'</td>
</tr>
<tr>
<td>6. pap'( \beta )</td>
<td>pap'( \beta )</td>
</tr>
</tbody>
</table>

Obviously, in such a language as this, spirants, whether voiced or voiceless, and voiceless vowels are not phonemes but are merely phonetic reflexes of stopped consonants and voiced vowels under fixed dynamic conditions. Long consonants and long vowels are sub-phonemes. The former are the resultants of simple phonemes (stopped consonants) and the operation of certain phonologic (and morphologic) latencies in given syllables, present or formerly present. The latter are phonologically resolvable into short vowel plus short vowel, i.e., into two syllables of unit length (moras), of which the second begins with a zero consonant.

Southern Paiute, then, is a language in which an unusually simple

---

\(^{1}\) They are described in detail in E. Sapir, The Southern Paiute Language, Proceedings of the American Academy of Arts and Sciences, 65 (1930).

\(^{2}\) \( M \) is voiceless \( m \).
phonemic structure is actualized by a more than ordinarily complex phonetic one. Tony’s “error” unconsciously registered this contrast.

II. When working on Sarcee, an Athabaskan language of Alberta, Canada, I was concerned with the problem of deciding whether certain words that seemed homonymous were actually so or differed in some subtle phonetic respect that was not immediately obvious. One such homonymous, or apparently homonymous, pair of words was dins (“this one”) and dinta “it makes a sound.” In the early stage of our work I asked my interpreter, John Whitney, whether the two words sounded alike to him and he answered without hesitation that they were quite different. This statement, however, did not prove that he was objectively correct, as it is possible for perfectly homonymous words to give the speaker the illusion of phonetic difference because of the different contexts in which they appear or because of the different positions they occupy in their respective form systems.\(^8\) When I asked him what the difference was, he found it difficult to say, and the more often he pronounced the words over to himself the more confused he became as to their phonetic difference. Yet all the time he seemed perfectly sure that there was a difference. At various moments I thought I could catch a slight phonetic difference, for instance, (1) that the -ni of “this one” was on a slightly lower tone than the -ni of “it makes a sound”; (2) that there was a slight stress on the d- of “this one” (analysis: stem d- “this” plus suffix -ni “person”) and a similarly slight stress on the -ni of “it makes a sound” (analysis: prefix d- plus verb stem -ni); (3) that the -ni of “this one” ended in a pure vowel with little or no breath release, while the -ni of “it makes a sound” had a more audible breath release, was properly -ni. These suggestions were considered and halfheartedly accepted at various times by John, but it was easy to see that he was not intuitively convinced. The one tangible suggestion that he himself made was obviously incorrect, namely, that the -ni of “it makes a sound” ended in a “t.”

1 The grave accent represents a low tone, the acute accent a high one. Sarcee is a tone language.

4 Thus, in English, the word led (e.g., “I led him away”) is felt as having a vowel which has been deflected from the vowel of lead (e.g., “I lead him away”) and is therefore not psychologically homonymous with the word for a metal, lead, in which the vowel is felt to be primary, not deflected (cf. further, “the leading of the windowpane,” “the leaded glass,” “the different leads now recognized by chemists”). The homonymy of led and lead (metal) is therefore of a different psychological order from the homonymy of yard (“He plays in my yard”) and yard (“I want a yard of silk”), for the last two words enter into roughly parallel form systems (e.g., “Their yards were too small to play in”: “I want two yards of silk”; “yard upon yard of railroad tracks”; “yard upon yard of lovely fabrics”). It is probably easier for the naı́ve speaker, who does not know how to spell either led or lead (metal), to convince himself that there is a phonetic difference between these two words than between the two words yard.

9 The lack of a tone mark indicates that this syllable is pronounced on the middle tone. [d̥] is an over-long a, consisting of a long a followed by a weak rearticulated a. Syllables of this type result in Sarcee from contraction of old final vowels with following suffixed vowels. The change in quality from [d̥] to [d̥] is due to historical factors. -ni “person” is an old *-nd (with pept vowel), relative -i is old *-d; two pept vowels contract to long open *-i, as Athabaskan becomes Sarcee a, this older *-* passes into Sarcee -a.

11 ft is voiceless spirant [f], as in Welsh ll.

We see at once that dinta “this one” behaves like a word ending in a smooth vowel (witness contraction of *f + a to an over-long vowel and

\(8\) The lack of a tone mark indicates that this syllable is pronounced on the middle tone. [d̥] is an over-long a, consisting of a long a followed by a weak rearticulated a. Syllables of this type result in Sarcee from contraction of old final vowels with following suffixed vowels. The change in quality from [d̥] to [d̥] is due to historical factors. -ni “person” is an old *-nd (with pept vowel), relative -i is old *-d; two pept vowels contract to long open *-i, as Athabaskan becomes Sarcee a, this older *-* passes into Sarcee -a.

\(11\) ft is voiceless spirant [f], as in Welsh ll.
unaffected of *la), while dint “it makes a sound” acts as though the final vowel had a voiceless consonantal latency, which registers partly as *-a (‘-la passing, as always, to -la), partly as -e.

It is clear that, while John was phonetically amateurish, he was phonologically subtle and accurate. His response amounted to an index of the feeling that dint “this one” = dint, that dint “it makes a sound” = dint, and that this -nt = -nt. John’s certainty of difference innow objective identity is quite parallel to the feeling that the average Englishman would have that such words as saved and soared are not phonetically identical. It is true that both saved and soared can be phonetically represented as so’d,12 but the -ing forms of the two verbs (saving, soaring), phonetically so-ing and so-ing, and such sentence sandhi forms as “Saw on, my boy!” and “Soar into the sky!” combine to produce the feeling that the so’d of saved = so’d but that the so’d of soared = so’d. In the one case zero = zero, in the other case zero = r. Among educated but linguistically untrained people who discuss such matters differences of orthography are always held responsible for these differences of feeling. This is undoubtedly a false assumption, at least for the great mass of people, and puts the cart before the horse. Were English not a written language, the configuratively determined phonologic difference between such doublets as saved and soared would still be “heard,” as a collective illusion, as a true phonetic difference.

III. The most successful American Indian pupil that I have had in practical phonetics is Alex Thomas, who writes his native language, Nootka,13 with the utmost fluency and with admirable accuracy. Alex’s orthography, as is natural, is phonologic in spirit throughout and it is largely from a study of his texts that I have learned to estimate at its true value the psychological difference between a sound and a phoneme. Anyone who knows the phonetic mechanisms of Nootka can easily analyze his orthography. Thus, hi14 phonologically parallel to si or ni, is actually pronounced he, with a vowel which is much nearer to the e of English met than to that of sit. This is due to the peculiar nature of the laryngeal consonants, which favor an a-timbre and cause the following vowels i and u to drop to e and a respectively. The orthographies hi and hu are entirely unambiguous because there can be no phonologically distinct syllables of type he and ho.

Another mechanical peculiarity of Nootka is the lengthening of consonants after a short vowel when followed by a vowel. This purely mechanical length has no morphological or phonological significance and is ignored in Alex’s orthography. His hisk and hina are, then, to be normally pronounced his-k and hina. It sometimes happens, however, that a long consonant, particularly s and š, arises from the meeting of two morphologically distinct consonants (e.g., š + s > s or š + š > š, or, less frequently, š + s or š + š > š). In such cases the long consonant is not felt to be a mechanical lengthening of the simple consonant but as a cluster of two identical consonants, and so we find Alex writing, for example, tsik’sitlassatin15 “we went there only to speak,” to be analyzed into tsik’sit’lassa-sa-(a)tl-ni. The s of ‘as “to go in order to” and the s of -sa “just, only” keep their phonologic independence and the normal intervocalic -s- of -’assall is interpreted as -ss-. Similarly, kwis-sita “to do differently,” to be analyzed into kwis-si-ta. It does not, however, that there is an actual phonetic difference between the -s- (phonologically -s-) of such words as kwas’alt “the stick takes an upright position on the beach” (= kwa-salt), pronounced kwas’alt, and the -s’- of -’assall above. Here again we have objectively identical phonetic phenomena which receive different phonologic interpretations.

IV. In the earlier system of orthography, which Alex was taught, the glottalized stops and affricatives were treated differently from the glottalized nasals and semivowels. The former were symbolized as pl, fl, kl, kw, ql, qwl, tl, tc! (= tkl), and Ll (= tl); the latter as m, n, y, and w. The reason for this was traditional. The glottalized stops and affricatives, as a distinctive type of consonants, had been early recognized by Dr. F. Boas in many American Indian languages and described as “fortes,” that is, as stops and affricatives “pronounced with increased stress of articulation.” The type m, n, l, y, and w was not recognized by Dr. Boas until much later, first in Kwakiutl, and described as consisting of nasal, voiced lateral, or semivowel immediately preceded by a glottal closure. The orthography for these consonants (later discovered in Tsimshian, Nootka, Haida, and a number of other languages, but not as widely distributed as the so-called “fortes”) suggested their manner of formation, but the orthography for the glottalized stops and affricatives was purely conventional and did not in any way analyze their formation except to suggest that more energy was needed for their pronunciation.16 As a pure matter of phonetics, while the Nootka glottalized stops and affricatives seem to be somewhat more energetic in articulation than the corresponding unglottalized consonants, in others there is no noticeable

13 These remarks apply to British, not to normal American, usage.
14 This is spoken on the west coast of Vancouver Island, B. C.
15 h is a voiceless laryngeal spirant, almost identical with the Arabic h.
stops and affricatives are roughly parallel in formation with the glottalized sonant consonants, they are not and cannot be entirely so. In a glottalized p, for instance, our present p and former pl, there is a synchronous closure of lips and glottal cords, a closed air chamber is thus produced between the two, there is a sudden release of the lip closure, a moment of pause, and then the release of the glottal closure. It is the release of the lip (or other oral) closure in advance of the glottal closure that gives consonants of this type their superficial “click-like” character. On the other hand, in a glottalized m, our ’m, while the lip closure and glottal closure are synchronous as before, the glottal closure must be released at the point of initial sonancy of the m. Roughly speaking, therefore, p may be analyzed into p + ', while ’m may be analyzed into ’ + m. Such an orthographic difference as pl versus ’m, therefore, which I had inherited from the Americanist tradition, was not unjustified on purely phonetic grounds.

We now come to the intuitive phonologic test whether p and ’m are consonants of the same type or not. Alex learned to write consonants of type p and pl very readily (our earlier pl and pl), e.g., papi “ear” (earlier plapi, pli), lsa’ak “stream” (earlier lsa’ak). To my surprise Alex volunteered mi in such words as ‘ma’ti’qe’u “the older [brother or sister],” which he wrote mlami’i’qe’u. In other words, we had valuable evidence here for the phonologic reality of a glottalized class of consonants which included both type p (with prior release of oral closure) and type ’m (with prior release of glottal closure). A phonologically consistent orthography would require p and ml (or pl and ml). Once more, a naive native’s phonetic “ignorance” proved phonologically more accurate than the scientist’s “knowledge.” The phonologic justification for Alex’s difference so far as “stress of articulation” is concerned. In the Athabaskan languages that I have heard (Sarcee, Kutchin, Hupa, Navaho) the aspirated voiceless stops and affricatives (of type t, k, ts) are far more “fortis” in character than the corresponding glottalized consonants (e.g., t, k, ts). There is no necessary correlation between laryngeal type of articulation (voiced, voiceless, glottalized; or any of these with aspiration) and force of articulation (fortis, lenis). So far as Nootka is concerned, it did not seem to me that the glottalized stops and affricatives (Boas’ “fortes”, “fortis”) were significantly different in emphasis from the ordinary stops and affricatives. In such languages as recognize a phonologic difference of emphatic and nonemphatic and, at the same time, possess glottalized consonants, there is no reason why the glottalized consonants may not appear in both emphatic and nonemphatic form. As Prince Trubetzkoy has shown, some of the North Caucasian languages, as a matter of fact, possess both emphatic and nonemphatic glottalized stops and affricatives.

These consonants are apparently identical with the “ejectives” of Daniel Jones. There is another, apparently less common, type of glottalized stop or affricative in which the oral and glottal releases are synchronous.

“error” is not difficult. Consonants of type p are entirely analogous to consonants of type ’m for the following reasons.

1. Each occurs at the beginning of a syllable and, since no word can begin with a cluster of consonants, both p and ’m are felt by Nootka speakers to be unanalyzable phonologic units. In other words, the glottal stop can no more easily be abstracted from ’m than from p. Similarly, the affricatives and glottalized affricatives are phonologically unanalyzable units.

2. All consonants can occur at the end of a syllable except glottalized stops and affricatives, glottalized sonant consonants (’m, ’n, ’y, ’w), semivowels (y, w), nasals (m, n), the glottal stop (‘), and h. This rule throws consonants of type ’m more definitely together with consonants of type p.

3. Many suffixes which begin with a vowel have the effect of “hardening” the preceding consonant, in other words, of glottally affecting it. Under the influence of this “hardening” process p, t, k become p, t, k, while m and n become ’m and ’n. For example, just as the suffixes ’a’au (’a’au “on the rocks” and ’a’ah “in a receptacle” change the stem wi’nap “to stay, dwell” to wi’nap (e.g., wi’nap’a “so stay on the rocks”) and wi’k “to be not” to wi’k (e.g., wi’kah “to be not in a receptacle”) and wi’k “to be not” to wi’k, so t’lu’ma” (alternating with t’l’ma”) “to be hot” becomes t’lu’m- (alternating with t’lu’m-) “to be hot” becomes t’lu’m- (alternating with t’lu’m-) “to be hot” becomes t’lu’m- (alternating with t’lu’m-) “to be hot” becomes t’lu’m- (alternating with t’lu’m-) “to be hot” becomes t’lu’m- (alternating with t’lu’m-) “to be hot” becomes t’lu’m- (alternating with t’lu’m-) “to be hot” becomes t’lu’m- (alternating with t’lu’m-) “to be hot” becomes t’lu’m- (alternating with t’lu’m-) “to be hot” becomes t’lu’m- (alternating with t’lu’m-) “to be hot” becomes t’lu’m- (alternating with t’lu’m-) “to be hot” becomes t’lu’m- “summer, hot season” = parallel t’lu- + t’i-”tch “season” and kan- “to kneel” (e.g., canil “to kneel in the house”) becomes ka’n- (e.g., ka’n’aba “to kneel in a canoe”). As there seem to be no stems ending in h or h, the group ’m, ’n, ’y, ’w, y is left over as functionally related to the group m, n, w, y in the same sense as the group exemplified by p. Morphology, in other words, convincingly supports the phonologic proportion p: p = m: m. It is maintained that it was this underlying phonologic configuration that made Alex hear ’m as sufficiently similar to p to justify its being written for the following reasons.

18 m and n may be followed by a murmured vowel of i-timbre which is a reduced form of a, u, or i. Syllables or half-syllables of type m or n are preceded by an assimilated product of a, u or i; ’m and ’n result, therefore, in part, from by i, an assimilated product of a, u or i; ’m and ’n result, therefore, in part, from.

19 “the symbol” indicates the “hardening” effect of a suffix.

20 The pronunciation of ’m, ’n, ’y, and ’w as a simple sequence of glottal stop (‘) plus m, n, w, and y is rejected by the Nootka ear as incorrect.
obtain between 'm and 'p would have a significantly different psychologic weighting.

V. In a course in practical phonetics which I have been giving for a number of years I have so often remarked the following illusion of hearing on the part of students that there seems no way of avoiding a general phonologic theory to explain it. I find that, after the students have been taught to recognize the glottal stop as a phonetic unit, many of them tend to hear it after a word ending in an accented short vowel of clear timbre (e.g., a, e, ε, i). This illusion does not seem to apply so often to words ending in a long vowel or an obscure vowel of relatively undefined quality (a) or an unaccented vowel. Thus, a dictated nonsense word like *smé* or *pild* would occasionally be misheard and written as *smé* and *pild* but there seems far less tendency to hear a final glottal stop in words like *pila* or *pild*. What is the reason for this singular type of "overhearing"? Is it enough to say that students who have learned a new sound would have a significantly different psychologic weighting?

It must be remembered that the language of my students is English. We may therefore suspect that the illusion of a final glottal stop is due to some feature in the phonologic structure of English. But English has no glottal stop. How, then, can English phonology explain the overhearing of a consonant which is alien to its genius to begin with? Nevertheless, I believe that the students who projected a final glottal stop into the dictated words were handling an exotic phonetic element, the glottal stop, according to a firmly established but quite unconscious phonologic pattern. It requires both the learning process, with its consequent alert preparedness to recognize what has been learned, and English phonology to explain the illusion. If we study the kinds of syllables in English which may normally constitute an accented monosyllabic word or an accented (or secondarily accented) final syllable of a word, we find that they may be classified into three types:

A. Words ending in a long vowel or diphthong, e.g., sea, flow, shoe, review, apply.
B. Words ending in a long vowel or diphthong plus one or more consonants, e.g., ball, cease, dream, alone, amount.
C. Words ending in a short vowel plus one or more consonants, e.g., back, fill, come, remit, object.

The theoretically possible fourth class:

D. Words ending in a short vowel, e.g., French *fait*, *ami*; Russian *zarádo* does not exist in English. English-speaking people tend to pronounce words of type D in a "drawing" fashion which transfers them to type A (e.g., *ami* for *ami*). Observe that the apparently inconsistent possibility of a nonfinal accented syllable ending in a short vowel (e.g., *fiddle*, *butter*, *double*, *pheasant*) is justified by the English theory of syllabification, which feels the point of the syllabic division to lie in the following consonant (d, t, b, z, in the examples cited), so that the accented syllables of these words really belong phonologically to type C, not to type D. Inter-vocalic consonants like the *d* of *fiddle* or *z* of *pheasant*, in spite of the fact that they are not phonetically long, are phonologically "flanking" or two-faced, in that they at one and the same time complete one syllable and begin another. Should the point of syllabic division shift back of the consonant, the preceding vowel at once lengthens in spite of its "short" quality (type A), and we thus get dialectic American pronunciations of words like *fiddle* and *pheasant* in which the accented vowel keeps its original quality but has been lengthened to the unit length of "long vowels" of type *feebly*, *reason*, and *ladle*.

We are now prepared to understand the illusion we started with. Such words as *smé* and *pild* are unconsciously tested as possible members of class A or class C. Two illusions are possible, if the hearer is to be a victim of his phonologic system. Inasmuch as a final accented short vowel is an unfamiliar entity, it can be "legitimized" either by projecting length into it (misheard *smé* and *pild* fall into class A) or by projecting a final consonant after it (class C). We shall call this imaginary consonant "z" and write *smex* and *pildz*. Now the fact that one has added the glottal stop to his kit of consonantal tools leads often to the temptation to solve the phonologic problem symbolized as *smex* and *pildz* in terms of the glottal stop and to hear *smé* and *pild*'. The glottal stop is the most unreal of consonants to an English or American ear and is admirably fitted, once its existence has been discovered, to serve as the projected actualization of a phonologically required final consonant of minimum sonority. The illusion of the final glottal stop is essentially the illusion of a generalized final consonant ("z") needed to classify the dictated words into a known category (type C). Or, to speak more analytically, English phonology creates the groundwork ("x") of the synthetic illusion, while the learning process colors it to the shape of "z". The error of hearing a glottal stop where there is none in words of type D, is fundamentally a more sophisticated form of the same error as hearing a dictated final glottal
stop as \( p \) or \( t \) or \( k \), which occurs frequently in an earlier stage of the acquiring of a phonetic technique.

The danger of hearing a glottal stop when the dictated word ends in a long vowel or diphthong is of course rendered very unlikely by the fact that such words conform to a common English pattern (type A). The reason why the error does not so easily occur in hearing dictated words ending in an unaccented short vowel (e.g., \( o\-n\-e \), \( u\-i\-l \)) is that such words, too, conform to an English pattern, though the range of the qualities allowed a vowel in this position is not as great as when the vowel is covered by a following consonant (e.g., \( i\-d\-e \), \( v\-e\-r \), \( f\-o\-l \)).

A STUDY IN PHONETIC SYMBOLISM*

The symbolism of language is, or may be, twofold. By far the greater portion of its recognized content and structure is symbolic in a purely referential sense; in other words, the meaningful combinations of vowels and consonants (words, significant parts of words, and word groupings) derive their functional significance from the arbitrary associations between them and their meanings established by various societies in the course of an uncontrollably long period of historical development. That these associations are essentially arbitrary or conventional may be seen at once by considering such a proportion as

\[
\text{phonetic entity } '\text{boy}' : \text{idea (or reference) } '\text{boy}' = \text{phonetic entity } '\text{man}' : \text{idea (or reference) } '\text{man}' .
\]

In passing from the notion of 'boy' to that of 'man' we experience a definite feeling of relationship between the two notions, that of increase in size and age. But the purely phonetic relationship of 'boy' : 'man' takes no account of this. So far as the referential symbolism of language is concerned, the words 'boy' and 'man' are discrete, incomparable phonetic entities, the sound-group \( b-o-y \) having no more to do with the sound-group \( m-a-n \), in a possible scale of evaluated phonetic variants, than any randomly selected pair of sound-groups, say 'run' and 'bad,' have to do with each other.

This completely dissociated type of symbolism is of course familiar; it is of the very essence of linguistic form. But there are other types of linguistic expression that suggest a more fundamental, a psychologically primary, sort of symbolism.1 As examples may be given the interrogative tone in such a spoken sentence as “You say he’s dead?” in comparison with the simple declarative tone of the corresponding “You say he’s dead”; further, the emphatically diminutive \( 'c \) of \( t \) as contrasted with the normal \( i \) of \( t \). In both of these examples the phonetic difference is undoubtedly felt as somehow directly expressive of the difference of meaning in a sense in which the contrast between say ‘boy’ and ‘man’ is not. We may call this type of symbolism ‘expressive’ as contrasted with the merely ‘referential’ symbolism which was first spoken of. It

* Journal of Experimental Psychology, 12 (1929): 225-239. Publication of the Behavior Research Fund, the Institute for Juvenile Research, Chicago (Herman M. Adler, Director), Ser. B, No. 132. For valuable suggestions in the preparation of this paper I am indebted to Professor H. A. Carr, University of Chicago.

1 For the two symbolic layers in speech, as in all expression, see Edward Sapir, “Language as a Form of Human Behavior,” English Journal, 16 (1927): 421-433.