

音位標音法的多能性

趙元任

THE NON-UNIQUENESS OF PHONEMIC SOLUTIONS OF PHONETIC SYSTEMS

BY Y. R. CHAO

In reading current discussions on the transcription of sounds by phonemes, one gets the impression of a tacit assumption that given the sounds of one language, there will be one and only one way of reducing them to a system of phonemes which represent the sound-system correctly. Since different writers do not in fact agree in the phonemic treatment of the same language, there arise then frequent controversies over the "correctness" or "incorrectness" in the use of phonemes.

The main purpose of the present paper is to show that given the sounds of a language, there are usually more than one possible way of reducing them to a system of phonemes, and that these different systems or solutions are not simply correct or incorrect, but may be regarded only as being good or bad for various purposes.

I. DEFINITIONS OF A PHONEME.

The most comprehensive discussion of the phoneme and related ideas seems to be that by H. E. Palmer¹, of which we shall now give a brief summary. Palmer begins by quoting at length Jimbo's writing on "The Concrete and Abstract Nature of Sounds." (音の具體性と抽象性) "One concrete sound has one definite quality, one definite pitch, one definite loudness, one definite length," in other words, it corresponds to one particular oscillograph curve or a stretch of the groove of a faithful gramophone record, which is therefore not the usual object of study for phonetics. By collecting examples of actual utterances of what is considered the same word with the same meaning by speakers of the same language of concrete sounds, one arrives at "an abstract sound of the first degree," such as the first sound in the word *army*. By

1. H. E. Palmer, *The Principles of Romanization*, 1931, Tokyo, pp. 52 ff.

comparing *different* words as *army*, *archer*, *art*, *argue*, one concludes, after due examination, that the first sound in these words are "the same," which is then an abstract sound of the second degree.

Taking Palmer's own system, we note that he finds it more convenient to replace the term *abstract speech-sound* by the term *phone*. His system of phones is then as follows:

Phones	{	Monophones	{	(1) Contactual phonemes	
				(2) Free phonemes	
	{	Metaphones		(3) Dynamophones	
				(4) Diaphones	
				(5) Phonogenes	

A Monophone is "any phone of the first or second degree of abstraction of which the concrete members are so similar in point of production and of acoustic effect even when observed by a competent observer, that it may be regarded as a minimal unit of pronunciation (i.e. practically insusceptible of sub-division.)" (We may add: "or of further differentiation.") "Contrasted with *monophones* we have *metaphones*, which we may define as two or more phones which serve jointly as units of meaning within the limits of a given linguistic community."

(1) Palmer goes on to identify Jones's definition of a phoneme with his idea of a contactual phoneme: "A *phoneme* is a group of sounds consisting of an important sound of the language (i.e. the most frequently used member of that group) together with others which take its place in particular sound-groups.....The use of subsidiary members of phonemes is, in most languages, determined by simple principles which can be stated once for all, and which can be taken for granted in reading phonetic texts."

(2) A free phoneme is like a contactual phoneme except that it is impossible to say in what phonetic circumstances one or another of its members will be actually used. We can give the apparently random¹ use of the tip or back of the tongue in the nasal ending of words like 因, 今 in Nanking as an example of free phonemes. This is the same as Jones's *variphone*.

(3) A dynamophone is a metaphone which contains two or more phones differing not only in quality, but also in regard to the intensity or force of the articulation that produces them. Palmer cites the first phone in the word *as* as an example which shades from the first phone

1. That is, determined by psychological or physiological conditions other than those which are usually considered to be phonetic.

of *act* to the obscure sound of the first phone of *about*, and even to zero value.

It would seem convenient also to include under this heading metaphones whose members differ according to conditions of length and intonation (in which case a term wider than *dynamophone* will have to be used). Thus, the vowel in French *bette* and *bête* is a metaphone whose members differ slightly in quality according to the conditions of length. Those who transcribe *eat*, *it* as [iːt], [it] are also considering the vowel in these words as forming one metaphone whose members differ in quality according to conditions of length. Again, the vowel in the Foochow words 哥 and 個 is a metaphone whose members differ in quality according as the intonation belongs to one or the other of two sets of tones.

Before taking up the next two terms, it will be well to examine a later definition of a phoneme given by Jones: "Definition of a phoneme: a family of sounds in a given language which are related in character and are such that no one of them ever occurs in the same surroundings as any other in words. (The term 'language' here means the pronunciation of one individual speaking in a definite style. 'In the same surroundings' means surrounded by the same sounds and in the same condition as regards length, stress and intonation."¹ This definition differs from the earlier one quoted above in that it mentions explicitly "and in the same condition as regards length, stress and intonation." It seems therefore that Jones's conception of a phoneme includes not only Palmer's contactual phonemes, but also some at least of his dynamophones.

(4) The term *diaphone* is used by Palmer following the usage of Jones: "The *diaphone* is a family of sounds heard when we compare the speech of one person with that of another." Jones cites [oː], [ou], [əu], [vu] as members of the diaphone occurring in words like *coat*, *road*, *home*. Similarly, we can cite [ʁu], [ou], [əu], [ɜ], [u] [øy], [ei], [ɪ] as members of the diaphone occurring in words like 歐, 狗, 後.

(5) The *phonogene*, a term also proposed by Jones, is "a given phone together with its ancestral forms," thus the vowel [ou] in *stone*, together with [o], [ə], [ɑ], form a phonogene. Similarly, [ɜ], [əɪ] [ɹ], [ɹi], [zi], [pzi] [pi], [ni] form one phonogene in words like 兒, 耳, 二.

Bloomfield gives no formal definition of a phoneme. He begins by distinguishing the *gross acoustic features* of language (Jimbo's "concrete sounds" or sounds of low degrees of abstraction) and *distinctive* or *significant features*. By comparing the partial identities and

1. *Proceedings of the International Congress of Phonetic Sciences*, 1932, Amsterdam, p. 23.

differences between words like *pin*, *tin*, *tan*, *tack*, he succeeds in analyzing the distinctive features of words like *pin* into indivisible units which cannot be analyzed any further (from the standpoint of the language under investigation): each of these units is "a minimum unit of distinctive sound-feature, a phoneme,"¹ which phrase is the nearest Bloomfield comes to a formal definition of a phoneme.

Differences of quality conditioned by length are grouped by Bloomfield under the same phoneme, as German *Beet* [be:t], *Bett* [bet]. He also writes *hatte* [ˈhate] where the stress on the first syllable indicates sufficiently the weakened and obscure value of the second vowel. Bloomfield's phoneme therefore also includes Palmer's dynamophones.

Bloomfield makes no explicit mention of free phonemes or variphones. In cases like the apparently random use of final [n] and [ŋ] in some Chinese dialects for the same word in the same phonetic surroundings, he would probably consider simple nasality as being the distinctive feature and the place of articulation as among the gross acoustic features. In other words, variphones are also phonemes, except that the choice of the exact shade of the sound used is determined by psychological and physiological factors other than those of phonetic environment. Since, however, whether variation of sounds determined by non-phonetic conditions are wide enough to be called two or more "different" sounds or simply inevitable small "accidental" variations depends upon the degree of narrowness of the phonetician's scale of division, Bloomfield is within his rights in neglecting the existence of variphones.²

From the preceding, it may seem that Bloomfield has a different conception of the phoneme from that of Jones and Palmer. For Jones and Palmer, a phoneme is a group of sounds, while for Bloomfield it is a sound-feature. If, however, we examine the two ideas more closely, we shall find that they amount to the same thing. Take for example the English phoneme [h]. From one point of view, we may say that it is a group of different sounds [h_i], [h_e], [h_a], [h_u] etc., where the subscripts are indication of the tongue and lip positions during the pronunciation of the consonant. But from the other point of view, we may just as well say that the phoneme [h] is simply the feature of voiceless glottal friction and leave the other non-significant features unspecified. There is therefore no real difference in the use of the term phoneme by those writers, so far as this point is concerned.

1. Leonard Bloomfield, *Language*, 1933, New York, p. 79.

2. See however III below on the finiteness of the number of distinguishable speech-sounds.

For the present discussion, we shall group together Palmer's contactual phoneme, free phoneme, and dynamophone, all under the term phoneme, to be defined as follows:

A phoneme is one of an exhaustive list of classes of sounds in a language¹, such that every word in the language can be given as an ordered series of one or more of these classes and such that two different words which are not considered as having the same pronunciation differ in the order or in the constituency of the classes which make up the word.

Observations:

(1) This definition presupposes that it is possible to enumerate exhaustively the total number of phonemes for any given language.

(2) It does not exclude the possibility of the same sound belonging to more than one class (Cf. II 2 (f), (g) below).

(3) It is non-committal as to whether given a language, there is one unique way for grouping its sounds into phonemes or there are other possible ways.

(4) It leaves unspecified the scope of the word "sound" as regards size and kind, i.e. the degree of analysis into successive elements and the degree of differentiation into kinds.

(5) It includes both the cases where given the phonemes in a word and its phonetic environment, it is possible to determine the actual pronunciation of the word by a set of "rules of pronunciation" (i.e. to know which member-sounds of the sound-classes will actually be used) and those cases where a given word in a given phonetic environment may still contain a phoneme of which one or another member may be used. The former would be a contactual phoneme or a dynamophone and the latter a free phoneme. (This remark, however, would be superfluous if we repudiate the validity of descriptive phonetics, with its narrow transcriptions.)

(6) The clause that every word consists of a series of "classes" may sound a little strange. But if, as it is convenient in the study of languages, to speak of recognizable words consisting of recognizable phonemes, then such phonemes are usually *classes* of sounds, which a trained ear would distinguish as different sounds. The statement sounds no more strange than that 1, 2, 3, 4 are a series of "classes", which is what mathematicians define numbers as.

1. Taken in the sense of the pronunciation of a homogeneous speech community, such that members of the same community will find absolutely no "accent" in one another's speech.

(7) If each phoneme is written with one definite symbol, then every word will have a definite form of transcription. Homophones, or different words having the pronunciation, will be transcribed alike. It should be noted, however, that the boundary between a homophone and a word with variations in meaning is often hard to determine.

(8) A phonemic transcription is pronounceable without reference to grammatical or lexical consideration. Thus, the Chinese National Phonetic Script and the National Romanization are phonemic transcriptions in a sense in which English or even German orthography is not.

II. FACTORS WHICH INFLUENCE THE PHONEMIC SOLUTIONS OF PHONETIC SYSTEMS

As the grouping of sounds in a language into phonemes as defined above does not necessarily lead to one unique solution, we shall now consider the various factors which influence the form of the solutions.

1. Size of unit in time.—

(a) *Under-analysis*.—In the early days of phonetic transcription, the slogan was “one sound, one symbol.” In these days of phonemic transcription, this has been changed to “one phoneme, one symbol,” so that it is now permissible to represent more than one sound by one symbol.

But there are two aspects to the idea of “one sound.” From the point of view of differentiation of quality, “one sound” is one *kind* of sound (一種音), which is what one usually has in mind when using the phrase in discussions about phonemes. But from the point of view of analysis in time, “one sound” is one *piece* of sound (一個音), such that its quality is homogeneous throughout its duration. Discussions about phonemes do not seem to have been very explicit about the change of quality in time which may be included within the scope of one phoneme. We recall that Palmer defines a monophone “as a minimal unit of pronunciation (i.e. practically insusceptible of further subdivision).” All the preceding discussions in the passage quoted have to do with the question of differentiation, but as the words “minimal” and “subdivision” can also be taken in the temporal sense, it would seem that a monophone should be both one *kind* of sound and one *piece* of sound.

Now if it is convenient to group into classes and call phonemes different *kinds* of sounds in a language which go together in a certain way, it would also be convenient to join into compounds successive *pieces* of sounds which act as units in a language. This is by no means new

practice. Our point here is only to make it explicit and put it on a par with the differential aspect of phonemes.

All kinetic speech-sounds, diphthongs, affricates, aspirates, and other sounds with their usual glides are compounds which act as units and can be treated as phonemes. Thus, Bloomfield considers the English affricates [č] and [j] as independent phonemes. The English plosives [p], [t], [k] are treated by all writers as single phonemes, although in initial stressed positions they have a slight aspiration and have a larger size than in unstressed positions or after [s] (in [sp-], [st-], [sk-]). In the former case, the inclusion of [č] and [j] is optional. For these could be resolved into the phonemes [tʃ] and [dʒ] respectively. In cases like *he cheats* [hi:tʃi:ts], *heat sheets* [hi:tʃi:ts], *each eye* [i:tʃ'ai], *eat shy* [i:tʃai], the distinction may either be made by considering [č] and [tʃ] as different phonemes, as with Bloomfield, or simply by the difference in the position of the minimum point, as with most other writers, that is to say, since the [č] in *each eye* and the [tʃ] in *eat shy* never occur under the same conditions as regards stress, [č] need not be considered as a separate phoneme. In many Chinese dialects, the initial [k] always occurs before low front vowels or central or back vowels, and initials of the [tʃ] type always occur before high front vowels. The two may therefore be taken as the same phoneme, although the latter is an affricate. Similarly, the [t] in [ta] ㄊ, the [tʃ] in [tʃei] ㄊㄟ and the [ts] in [tsu] ㄊㄣ in Japanese may be taken as belonging to one phoneme.

Kinetic sounds of the diphthong type need special consideration. While affricates, aspirates and sounds with characteristic glides can usually be analyzed, if desired, into two or three recognizable elements, kinetic vowels and quasi-vowels are sounds with even more gradual change in quality. The usual method of representing these sounds is simply to indicate the two end-positions of the whole movement, as [ei], or to indicate the open position and the extreme close position even though never actually reached, as [ai] for what is actually never wider than [ae]. In the case of movement not by the most direct line, the turning point is indicated by inserting an additional symbol, as [uei], but not [əu], as [au] means [əu] or [əəu].

Now by our definition of a phoneme, there is nothing to prevent us from regarding characteristic kinetic open sounds in a language as independent phonemes, which is in fact the practice of the designers of the Chinese National Phonetic Script, who represent [ai], [ei], [au], [ou] by the single symbols ㄞ ㄟ ㄠ ㄡ, and even [an], [ən], [aŋ], [əŋ] by ㄢ ㄣ ㄥ ㄥ. It may seem unorthodox if we took the National Phonetic Script as serious phonemic transcription, but we should be

less sure of ourselves if we come to cases of narrow-range kinetic sounds. There is a real difference in practice, if not of opinion, between Bloomfield's use of [ij] and [uw] for English and other writers' use of [i:] and [u:] (as contrasted with [i] and [u]); or of [i] and [u] (with implied relative length) (as contrasted with [ɪ] and [ʊ]). Again, in many American dialects, it is a toss whether to write *bet*, *bait* as [bet], [beit] or as [bet], [be:t], or [bet], [bet] (with implied length). The most interesting case of the size of unit question is that of the Foochow dialect, where a whole series of vowels *in the same words* are static or kinetic according to the actual tone in which it is pronounced. Thus, 氣 [k'ei₁] "air", 竹 [tøyk₁] "bamboo," 護 [hou₁] "protect," take on the following sounds when they are pronounced in the following combinations of tonal environment: 氣壓 [k'ei₁ ak₁] "air pressure," 竹節 [ty₁ zai₁] "bamboo section," and 護兵 [hu₁ bi₁] "guards" (protecting soldiers), respectively. We have therefore on our hands the question of choice between (1) admitting phonemes of which some members are static and other members kinetic vowels, or diphthongs, and (2) regarding the static members as forming one phoneme and the corresponding kinetic vowels as two phonemes in succession, thus allowing the same word to have two forms. The presence and absence of the aspiration in English [p], [t], [k] mentioned above is also a similar case, though not so striking.

Another very peculiar case is that of a vowel in a concave circumflex tone in a number of Chinese dialects, such as the *yangshaang* tone of Hwangyan (黃巖陽上), where the valley is so low or simply so narrow that the voice is lost into a glottal stop in the middle of the syllable, so that [ə₁] actually becomes [ə₁ʔ ə₁]. Phonetically, it sounds like three sounds forming two syllables. But phonemically, it is much more natural to consider it as a form of [ə₁] in a certain tone.

On the whole, the usual practice allows a great deal of latitude in taking kinetic consonants as single phonemes, but not so free in giving single symbols for kinetic vowels. Bloomfield gives a list of eight diphthongs and one triphthong for English, and call them "compound primary phonemes," all their elements occurring also as single primary phonemes.

The chief point we wish to emphasize here is that it is not always advisable or convenient to take the smallest static unit of sound analyzable by the trained ear as the unit of phonemic members ("one *piece* sound, one symbol"), and that according as we take a smaller or a larger unit for our phonemic members, we sometimes arrive at different forms of

phonemic pattern for the same language, which are equally valid, though they may not be equally suitable for this or that purpose.

(b) *Over-analysis*.—The principle of “one piece sound, one symbol” has yet to allow a class of exceptions in the opposite direction, namely, one piece sound, two or three piece symbols. Jones and Camilli give the following cases where combinations of letters are permitted to represent single phonemes:¹

- a) The affricates [pf], [bv], [ts], [dz], [tʃ], [dʒ], [tʂ], etc.
- b) The aspirates [ph], [th], [kh], [tʰ], [dʰ], etc. and weak aspirates [pʰ], [tʰ], etc.
- c) The aspirated [s] or [sh].
- d) [t], [d] with lateral explosions or [tɫ], [dɫ].
- e) The voiceless nasals, [hm], [hn], [hɲ], [hŋ], when these are distinct phonemes.
- f) Retroflex vowels, as American [əɻ], or Peiping [uɻ].
- g) Labiovelar consonants as [kp], [gb].

Of these cases, a) and b) are recognizably compound sounds, which we should consider as two or three piece sounds, for which the use of [tʃ], [dʒ], [ph], [th], etc. would be considered as normal and the use of [tʃ], [dʒ], (or [c], [ɟ]), [p], [t], etc. would be considered as cases of under-analysis. c) and d) may be regarded as borderland cases. e), f), and g) are clear cases of over-analysis, that is, cases of one homogeneous sound represented by two or three piece symbols, each of which represent some aspect or aspects of the sound². Thus, [hm] is a [m]-sound which is breathed (i.e. [h]-ized) or a [h]-sound with labio-nasal articulation (i.e. [m]-ized). It is meaningless to ask which is the substantive and which is the adjective, as they are all constituting attributes which together form the sound in question and could be represented by Jespersen's over-analytical alphabetic symbols. Similarly, American [əɻ] is a single vowel formed by the middle of the tongue in the [ə] position with the apex curled back (sometimes transcribed as [ə̠]). The representation of voiceless [w], or [ʍ] by [hw] is another case, which is mentioned by Jones and Camilli under an earlier section in the same pamphlet quoted³.

1. *Fondamenti di Grafia Fonetica*, by Daniels Jones and Amerindo Camilli, 1933, Aube and London, pp. 11-12.

2. G. M. Bolling must have overlooked such cases when he says, “At least I can recall no example of.....a digraph for a non-compound phoneme,” from editorial note on R. G. Kent's review of Bloomfield's *Language* in the journal *Language*, X, 1, 1934, pp. 51-52.

3. *Fondamenti*, p. 11, section 15.

Among the uses of diacritical marks, Jones and Camilli¹ mentions "the saving of a series of new letters," such as adding ~ to [ɑ], [ɔ], [œ], [ɛ] to form [ã], [õ], [œ̃], [ẽ] in French. The reader will recall the great furore which was aroused by Passy's proposal to use [ɑ̃], [ɔ̃], [œ̃], [ẽ] for these French vowels in the first post-war issues of *le Maître Phonétique*. He modestly called it orthographic transcription, but if [ɛ̃] can represent [ẽ], there is no reason why [ɑ̃] cannot represent [ã]. To object that other French dialects or German actually has [ɑ̃] as two successive sounds is beside the point, as we are talking about phonemic transcriptions and our universe of discourse is limited to one dialect or one language, otherwise we should have to go back to narrow phonetic transcriptions. Not that [ɑ̃] is the only right way or even a good way of representing French [ã], but there seems to be nothing wrong, so far as usage in other cases goes, in representing one piece sound by two piece symbols.

Jones and Camilli do another thing along the same line. Without mentioning the saving of a series of modified letters under any of the principles, they also use the device of representing one piece sound by two piece symbols in transcribing the Russian palatalized consonants, where the explanatory note says, "j is used as the sign of palatalization, that is, tj=t̃, nj=ñ, lj=l̃, snj=s̃ñ, tnj=t̃ñ, lnj=l̃ñ, etc."² This [j] is therefore a significant feature, but it does not necessarily occupy any time of its own.

Another important case is that of the "voiced *h*", which plays a very important part in the Wu-dialects in China. These dialects usually have an ordinary [h], which has different values according to the vowel following and may therefore be taken as one phoneme, just as in the case of English or German, so that instead of having $2n$ symbols for $h_1a_1, h_2a_2, \dots, h_na_n$ (where a_1, a_2, \dots, a_n are the vowels which may follow the *h* in the language), we need only $n+1$ symbols for ha_1, ha_2, \dots, ha_n . But in the case of the voiced *h*, not only the vowel quality (or the vowel articulation) begins at the very beginning of the breathing, but the breathiness also lasts till the very last moment of the vowel, so as to form one homogeneous breathy vowel, and there is neither question of order of succession nor question of substantive and adjective. If we must have one piece symbol for one piece sound, we should have to have either a series of different voiced *h* symbols for different vowels, or an extra series of breathy vowels have to be recognized. The only practical thing to do here is to consider voiced *h* as one phoneme and write the vowel

1. *Fondamenti*, p. 4. section 3).

2. *Fondamenti*, p. 17.

signs *after* it, as [hɑ], [hɛ], [hɔ], etc., although we know that these digraphs represent perfectly homogeneous sounds.

There are also borderland cases where it is open to question whether certain sound-elements are simultaneous or successive. According to ordinary transcriptions, the English word *sway* is transcribed as [swei] while the Chinese word 歲 is transcribed as [suei], from which it would seem that the first two elements in Chinese 歲 would be separated more clearly than in English *sway*. As a matter of fact, the contrary is the case. While the [s] in English *sway* is not at all labialized for most of its duration, the [s] in Chinese 歲 is completely labialized. Moreover, the diphthong [ei] starts almost as soon as the tongue leaves the [s]-position without leaving any appreciable duration for the [u] or [w] to stand alone, so that a narrow transcription might give 歲 as [ɕei] or, as the velar element is rather weak in this type of word, as [œei]. But in similar syllables in other tones or with other initial consonants, there is more independence in the [u]-element. It would be contrary to the spirit of phonemic transcription to write 歲 as [œei] and 對 as [tuei]. Consequently, we must allow as a possible phonemic "solution" the over-analysis of [ɔ] into two phonemes [su] or [sw], and so long as our universe of discourse is Chinese (Mandarin) phonemes, we should not be disturbed by the fact that [sw] in English is a succession of two sounds in which [s] is little or not at all [w]-ized.

From the consideration of these cases of under-analysis and over-analysis, we see the great advantage of Bloomfield's speaking of sound-features instead of sounds. If we consider a sound as made of a number of features, then a phoneme is a combination of certain (simultaneous and/or successive) features, leaving other features unspecified. The English [t]-phoneme, for instance, consists of the features of voicelessness, apico-alveolar articulation of a certain range (*eighth, tea, tray*), and complete stop of breath, while the exact position of articulation, the force of stopping, the nature of on-glides (*heat, hoot*) and off-glides (*tar, star, tea, two, little, button, but*) are left unspecified. The Chinese [u]-phoneme consists of the features of lip-narrowing, a slight velar action, and voice, and as the position of the tip of the tongue is left unspecified, it is perfectly free to form the [s]-articulation while the [u]-articulation is being held, so that we can entertain the idea of two phonemes [s] and [u] being telescoped into one single sound [ɔ] without necessarily considering the sound [ɔ] as one new phoneme or as *one* member of a new phoneme. Similarly, the [ɦ]-phoneme in the Wu-dialects consists of the feature of emitting more air than usual in producing voice, and as it does not specify anything about the oral or nasal features of articulation, the speaker is free to do all kinds of

articulatory tricks at the same time with [ɦ]¹, so that there is an [ɑ] type of [ɦ], and [e] type of [ɦ], etc., and even an [ɱ] type of [ɦ], as [ɦɱ] (“have not”), as contrasted with [ɱ] in [ɱ-ma] (“mother”), and yet all this does not prevent us from considering the [ɦ] and [ɑ] in [ɦɑ] as two theoretically separate phonemes.

(c) *Zero Symbols*².—As limiting cases of the variation of the size of unit, we have the possibility of using zero symbol for sounds or sound-features and of counting absence of sound as a phoneme or as one member of a phoneme.

Where there are several degrees of significant stress, significant length, or kinds of significant intonation, it is the usual practice to represent one of them by zero symbol. Thus, unmarked syllables in polysyllabic English words are understood to have the low degrees of stress. Vowels without length marks are understood to be short. In most systems of tone-marking, the first tone in Chinese is “marked” by not marking it.

In the Chinese syllables [tʂʌ], [tʂʌ̃], [ʂʌ], [ʂʌ̃], [tsʌ], [tsʌ̃], [sʌ]³, there is a vowel which is a vocalized prolongation of the preceding consonant, and is understood to be present when these syllables are written as ㄗ, ㄗ̃, ㄘ, ㄘ̃, ㄘ̣, ㄘ̣̃, ㄘ̣̣ in the National Phonetic Script. This is therefore a way of representing actual sounds by zero symbol.

In German stressed syllables beginning orthographically with a vowel, there is normally a glottal stop. Some writers give the symbol [ʔ] for this sound, but others omit the symbol, and in internal positions as in *verein*, a stress mark suffices to indicate the presence of the [ʔ], as [ferˈain]. It would be perfectly possible, though hardly conventional, for us to favor some other phoneme with the saving of a symbol, say [h], and transcribe *hauch* as [aux] and *auch* as [ʔaux].

Readers of Bloomfield's *Language* who are used to ordinary types of transcriptions of English must have been impressed by forms like these:

1. There is a trick recitation in one of the dialects near Nanking in the form of a story consisting mostly of phrases like 鵝對鴨 [ŋə tuei ŋaʔ] (“goose versus duck”), in which a flapped click is made with the front of the tongue each time [ŋ] is pronounced. The effect is that of beating a pair of clapping boards as an independent rhythmic accompaniment to the recitation. In other words, the [ŋ]-phoneme consists of the features of voice, nasality, and articulation with the back of the tongue. The front of the tongue can do as it pleases.

2. Under this heading, we are not including cases like ancient Hebrew, in which the vowels were not written. For in this system of writing, the vowels cannot be deduced from the phonetic environment alone by any set of phonetic rules. The writing is therefore an orthography and not a transcription.

3. The symbols ʌ and ʌ̃ are Karlgren's.

<i>gentleman</i>	[ˈʤentlmp]	p. 111
<i>atom</i>	[ˈetmp]	112
<i>maintenance</i>	[ˈmejntnps]	”
<i>maintain</i>	[mpˈtejn]	”
<i>stirring</i>	[ˈstriŋ] vs. <i>string</i> [striŋ]	121
<i>pattern</i>	[ˈpetrŋ] vs. <i>patron</i> [ˈpejtrŋ]	”
<i>erring</i>	[ˈriŋ] vs. <i>ring</i> [riŋ]	122
<i>error</i>	[ˈerɾ]	”
<i>butter</i>	[ˈbotɾ] on a par with <i>bottle</i> [ˈbatl]	”
<i>bottom</i>	[ˈbatmp] on a par with <i>button</i> [ˈbotn]	”
<i>anatomy</i>	[eˈnetmpij] vs. <i>met me</i> [met mij]	”

Now Bloomfield systematically avoids the use of the obscure vowel letter [ə] and plays his game admirably well. The *e* in French *le*, he considers as a short variety of [æ] (p. 106), which agrees more or less with the idea of the French themselves. For German, he lets the difference in stress take care of the difference between [e] and [ə]. For American English, he uses the strong forms where there is no following consonant or where the following consonant is not usually considered to be syllable carriers in English, but leaves out the symbol entirely in other cases. Now from the point of view of actual sound, weakened orthographically written vowels either become [ə] or disappear entirely. If we take ordinary deliberate conversation as the style of “language” to consider, we can say, according to the writer’s own observation of Middle Western American speech, that the presence or absence of a vocalized [ə] is about as follows:

[ə] compulsory or preferred.	[ə] optional.	Absence of [ə] compulsory or preferred.
<i>arbor</i> [-bər] vs. <i>club rate</i>	<i>happen</i> [-p(ə)n]	<i>able</i> [-bɪ]
<i>upper</i> [-pər] vs. <i>upright</i>	<i>often</i> [-f(ə)n]	<i>simple</i> [-pɪ]
<i>gentleman</i> [-mən] vs. <i>autumnal</i>	<i>even</i> [-v(ə)n]	<i>dismal</i> [-mɪ]
<i>humor</i> [-mər] vs. <i>am ready</i>	<i>bacon</i> [-k(ə)n]	<i>careful</i> [-fɪ]
<i>kingdom</i> [-dəm] vs. <i>bed-mate</i>	<i>winkum</i> [-k(ə)m]	<i>devil</i> [-vɪ]
<i>London</i> [-ndən] (cf. <i>sudden</i>)	<i>Beauchamp</i> [-č(ə)m]	<i>sudden</i> [-dn]
<i>under</i> [-dər] vs. <i>shad roe</i>	<i>Gresham</i> [-š(ə)m]	<i>middle</i> [-dɪ]
<i>atom</i> [-təm] vs. <i>met me</i>	<i>patron</i> [-tr(ə)n]	<i>colonel</i> [-nɪ]
<i>pattern</i> [-tərŋ] vs. <i>outright</i>	<i>Durham</i> [-r(ə)m]	<i>wiggle</i> [-gɪ]
<i>maintenance</i> [-nəns] vs. <i>main news</i>	<i>Coral</i> [-r(ə)l]	<i>engine</i> [-ʃɪ]
<i>Barnum</i> [-nəm] vs. <i>on me</i>	<i>handsome</i> [-s(ə)m]	<i>cordial</i> [-ʃɪ]
<i>corner</i> [-nər] vs. <i>Henry</i>	<i>bosom</i> [-z(ə)m]	<i>luncheon</i> [-čn]
<i>Helen</i> [-lən] vs. <i>hell no</i>	<i>Bentham</i> [-θ(ə)m]	<i>celestial</i> [-čl]

[ə] compulsory or preferred.	[ə] optional.	Absence of [ə] compulsory or preferred.
<i>alum</i> [-ləm] vs. <i>elm</i> (though [eləm] in some Am. dialects)	<i>fathom</i> [-ð(ə)m]	<i>nation</i> [-ʃn]
<i>Keller</i> [-ləɹ] vs. <i>all right</i>	<i>lengthen</i> [-θ(ə)n]	<i>special</i> [-ʃl]
<i>finger</i> [-gɹ] vs. <i>big row</i>	<i>heathen</i> [-ð(ə)n]	<i>vision</i> [-ʒn]
<i>poker</i> [-kɹ] vs. <i>quick ride</i>		<i>listen</i> [-sn]
<i>singer</i> [-ŋɹ] vs. <i>sing right</i>		<i>tassle</i> [-sl]
<i>ginger</i> [-jɹ] vs. <i>age wrong</i>		<i>dozen</i> [-zn]
<i>teacher</i> [-tʃɹ] vs. <i>teach right</i>		<i>hazel</i> [-zl]
<i>pleasure</i> [-ʒɹ]		<i>ethel</i> [-θl]
<i>error</i> [-rɹ] vs. <i>her right</i>		<i>brothel</i> [-ðl]
<i>tracer</i> [-sɹ] vs. <i>viceroi</i>		
<i>Caesar</i> [-zɹ] vs. <i>phase-rule</i>		
<i>ether</i> [-θɹ] vs. <i>Ruth ran</i>		
<i>father</i> [-ðɹ] vs. <i>with red</i>		

Opinions may differ as to the placing of particular cases under each heading, but there seems to be no doubt as to the presence of [ə] in *gentleman* [-mən] or as to the absence of [ə] in *able* [-bl]. Historically, as the orthography indicates, many of these words had clear vowels. Now some of them have an obscure vowel even in deliberate speech, which does not however entirely disappear in some cases. Since the presence, option, or absence of the [ə]-sound are more or less determined by the nature of the sounds preceding and following, and sometimes by conditions of syllabication, we can regard this as one phoneme of which one member is the obscure vowel [ə], a second member is a variphone (or dynamophone) consisting of [ə] and zero, and a third member is zero. Bloomfield has therefore as much right to represent this phoneme by zero symbol as one has to represent German [ʔ] by zero symbol. Apparent ambiguities as in the case of *string* and *stirring*, may be avoided by marking the syllabication [stɹɪŋ], which will remind us to explode the [t] before the [ɹ], as it is a case of the first member of the phoneme.

It should be noted that our discussion here is to find a methodological justification for Bloomfield's use of zero symbol for an actual sound. There are other considerations from which this avoidance of the symbol [ə] seems rather inconvenient. Thus, when there is no final consonant like [l], [n], etc., to act as a syllable carrier, as in *America*, suppose, *jealous*, he is obliged to use exclusively strong forms like [e'merike] or [e'merika], [so'pouz], [jelos], which are rarely heard even in deliberate speech (understanding of course that [o] is the "short u"). The definite

article *the* will have to be either [ðij] or [ð], with no middle ground. Those who favor Bloomfield's system for English will find that he is simply carrying the omission of [ə] to its logical conclusion. Those who do not will consider forms like [ˈstrɪŋ], [ˈmeɪtɪŋs], [eɪˈmerɪkə] a *reductio ad absurdum*.

Under cases of under-analysis, we considered the representation of affricates, aspirates, and narrow-range diphthongs by single symbols. Now if the symbol used is obviously *one* of the elements in the compounds, as [p] for [pʰ], [c] (instead of [ç]) for [cɸ], [ɹ] (instead of [ɹ̥]) for [ɹɸ], or [o] for [ou], then we can regard that element which is understood but not represented as having zero symbol. For instance, in the Soochow dialect, labials go with [ɸ], velars and dentals go with [əu], and alveolars go with an apical vowel with protruding lips, for which the writer has proposed the symbol [ɰ]¹, as 布 [pɸ], 故 [kəu], 註 [tɰ]. All these can be considered as members of one phoneme [u], in which case the [ə] in [əu] would be a sound with zero symbol. Again, in the Foochow vowels [u]: [ou], [i]: [ei], [y]: [øɣ] according to tone, as cited above, it is common practice to consider the first tone, which goes with [i], [u], [y], as more primary or representative, and for certain reasons, it may be more convenient to write these phonemes as [i], [u], [y], in which case a tone mark would suffice to remind one of the addition of [e-], [o-], [ø-] (by no means weak and parasitic), though these elements still have no symbol to themselves except as implied by the tone.

(d) *Zero Sound*.—In the cases of over-analysis, as in [fɛ], we had two features representing separate phonemes which together make one single sound. But if we take the series [ɸ], [əu], [ɰ] in Soochow and consider them as varieties of [əu], of which the [ə] is absent after labials and alveolars, then under the latter conditions, the phoneme [ə] will have zero as a member. Similarly, if we write in the symbol [ə] for *maintenance* [-nəns], *happen* [-pən], *button* [-tən], all alike, then the [ə] will be a symbol for a phoneme, of which one member (in words of the type in the third column in the preceding table) has the value zero. Again, Bloomfield's use of [ij] and [ow] in unstressed positions may be regarded as cases of [j] and [w] with zero sound. In Passy's "orthographic" notation referred to above, he spelt out the "mute e" as [ə] in all cases, letting the "rule of three consonants" take care of the presence or absence of the actual sound. From our point of view, [ə] would then be a phoneme with zero as a possible member. In the system of ancient Chinese initials, there are two called *yǐng* (影) and *yuh* (喻) which have been reconstructed by Karlgren as [ʔ] and smooth vowel

1. A combination of Karlgren's [ɿ] and [ɰ].

respectively. Those are of course only the *names* of the initials. But Jang Tayyan (章太炎) has devised an *alphabet* with a symbol for each of the 36 initials, so that his symbol for *yuh* would be a symbol with zero value, very much like the “’”-symbol for the smooth ingress of vowels in Greek.

In the theory of *sheh* (攝) or “rim-emes” in traditional Chinese phonology, the use of a symbol for zero is extremely useful. Taking again the National Phonetic Script, which is constructed very much in the spirit of traditional phonology, we have the rimemes ㄨ, ㄩ, ㄣ, ㄥ, which, like the other rimemes, may be preceded by the medials ㄟ, ㄨ, or ㄩ, so as to form the following complete finals (i.e. syllables minus initial consonant, if any) which actually occur in words:

Without medial :	ㄨ	ㄩ	ㄣ	ㄥ
with medial ㄟ :	ㄟ	ㄟ	ㄟ	ㄟ
with medial ㄨ :	ㄨ	ㄨ	ㄨ	ㄨ
With medial ㄩ :	ㄩ	ㄩ	ㄩ	ㄩ

A simple phonemic transcription in the IPA would be:

əi	əu	ən	əŋ
	iəu	iən	iəŋ
uəi		uən	uəŋ
		yən	yəŋ

In these twelve finals, the [ə] in [iən], [iəŋ], and [yən] always has zero value (in [yən], [y] is broken up into an intermediate value between [iu] and [yu]), just like the [e] in [ba:den] for German *baden*¹. In the case of [uəi] and [iəu], the [ə] has zero sound in the first and second tones and has some sound in the third and fourth tones, except that in [uəi] not preceded by an initial consonant, [ə] does not entirely disappear in any tone. In [uən], the [ə] has zero sound in the first and second tones when there is an initial consonant, is fully sounded when there is no initial, and is very weak in other cases. With [uəŋ], the [ə] is sounded only when there is no initial consonant. With [yən], the [ə] is sounded (with a value [ɪ]) when there is a palatal initial or no initial, but has zero sound with other initials. With such a complicated group of facts, where each case is a law unto itself, we should still fail to attain perfect phonetic accuracy by writing something like:

ei	ou	en	ɛŋ
	iu	in	iŋ
uei		un	uŋ
		yn	iuŋ,

1. Bloomfield, *Language*, p. 113.

although this may be a useful form of transcription for certain purposes. The paradoxical appearance of a symbol with widely different values, including zero, would disappear if we stuck to the National Phonetic Script or used some non-committal symbol as "ø" for the phoneme in question, thus:

øi	øu	øn	øŋ
	iøu	iøn	iøŋ
uøi		uøn	uøŋ
		yøn	yøŋ

This is of course not the only or even the best phonemic treatment of these finals, but by allowing the possibility of zero members of phonemes, we do gain a number of advantages.¹

(e) *Phonemic Treatment of Conditional End-consonants*.—In ordinary transcription of French, cases of liaison and elision are spelt as they sound. The word *pas* then has two forms [pa] and [paz], *le* has [lə] and [l], and by the "rule of three consonants," the word *demandeur* has the two forms (vous) [dmãde] and (pour) [dãmãde]. Similarly, Southern English *sore* has the two forms [sɔː]² (throat) and [sɔːr] (eyes). The presence or absence of the sound in question is not distinctive, so that it and zero may be considered as members of the same phoneme. But the difference between *saw* [sɔː] and *sore* [sɔːr] is distinctive, and for the phoneme with the conditional [r], the symbol "*" has been used in dictionaries, though the writer has never seen it used in texts, probably because ordinary transcriptions are not phonemic. From arguments with unsophisticated Frenchmen, who insisted that *point* did not have the same pronunciation as *poing*, the writer would think that a special phonemic symbol for these optional sounds would be welcomed by the French, say something like [paz], [pwɛt], so as to avoid the pitfalls of the "[patakes]" business³. Better symbols than these may be devised. Our interest here is in the obvious phonemic nature of these groups [z]: zero, [t]: zero, etc. It may not be necessary to outlaw the writing of two alternate forms for one word. But it would be an advantage not to have to do so.⁴

1. In this article, we are limiting ourselves to the discussion of phonemes of single languages. If we extend our universe of discourse to diaphones, say about 100 miles south of Peiping, the advantage of the above form will be enormously increased.

2. One type of Southern English.

3. "Puisque ce n'est pat à moi et n'est pions à vous, je ne sais pat à qu'est-ce." From Passy's *Chresthomathie*.

4. The case of English *a*: *an* is somewhat doubtful. If English never had a system of writing, or if its orthography had come to writing *for uncle*: *fo mother*, just like *an uncle*: *a mother*, we might then be inclined to treat the indefinite article as one word (as it was) and provide a special phoneme [-n] as its second element, a phoneme which occurs only in one word. Cf. II(e) below on word identity.

In this connection, we may mention the so-called "aspirated *h*" in French as a consonant phoneme which always has zero sound, but has a very definite "feature" of its own, and may be conveniently symbolized as [*h*]. The great advantage in regarding this as a consonant phoneme lies in that it greatly simplifies the description of the behavior of other phonemes. We can then say that [-*t*] (liaison *t*) has the sound [*t*] before vowels, and zero sound before consonants or in end-position. If we refuse existential status to [*h*], we should have to say that [-*t*] has the sound [*t*] before vowels, except before the following exhaustive list of words: [a₂z₁r] [ɛ₂z₁], etc., etc., which is no way of stating the "rule of pronunciation" for phonemes.

In many Chinese dialects, final consonants like [-*n*], [-*k*], [-ʔ] are pronounced very clearly at the end of phrases, but become weakened or disappear entirely when followed immediately by another word. The [-ʔ] in Foochow or the Wu-dialects is a phoneme which has zero value before another word. Thus, Soochow 八 [poʔ]. "eight," 八百 [popoʔ] "eight hundred," 八百八 [popapoʔ], "eight hundred eight(y)." The vowel is not even lengthened (as it is in Foochow under certain conditions) to make up for the time of the original [ʔ]. If we write phonemically, we can represent this phoneme with [ʔ] and zero sound as its two members either by (1) zero symbol (and let the symbol for the entering tone, with which it is always associated in these dialects, indicate its presence), or (2) the symbol [-ʔ] or [-ʔ] in all cases, whether the glottal stop is articulated or not.

It is not our purpose here to propose purely for the pleasure of perversity either to under-analyze two or more piece sounds and treat them as single phonemes or to over-analyze one piece sounds and treat them as successions of phonemes; nor purposely to write something where there is nothing to write or to write nothing where there is something to write. We wish only to indicate that all such tricks are actually being done in current transcriptions, and that according to the way in which we treat the time unit of phonemes in a language, we may arrive at one or another of various possible solutions for that language.

2. The Grouping of Sounds into Phonemes.—

So long as we confine ourselves to the consideration of stock examples like *keep*, *call*, *cool*, our construction of phonemic systems is smooth-sailing. We need only to disregard slight variations of what is generally regarded as "the same sound" and call it a phoneme. But on many questions of the identification of sounds in a language, we are not favored with such general consensus of opinion. Is the second element of the English "long *i*" to be identified with the first element in *yes* (Bloomfield's [a_j]), or with the first element in *it* ([a_i] by many

writers), or with the final element in *very* (Palmer's [aɪ]), or with the undistinguished [i] in *it* [it], *eat* [i:t], *very* [ˈveri] ([aɪ] by many writers), or with the first element in *eight* ([ae] in certain "narrow" transcriptions)? Is the palatal series [tɕ], [tɕʰ], [ɕ] in words like 家, 青, 下 (occurring only before high front vowels) to be identified with the velar series [k], [kʰ], [x] or with the retroflex series [tʂ], [tʂʰ], [ʂ] (none of either series ever occurring before high front vowels)? According as we emphasize this or that motive, we should arrive at a different system of organization of elements into phonemes. We may desire to have (a) phonetic accuracy, or smallness of range of phonemes, (b) simplicity or symmetry of phonetic pattern for the whole language, (c) parsimony in the total number of phonemes, (d) regard for the feeling of the native speaker, (e) regard for etymology, (f) mutual exclusiveness between phonemes, (g) symbolic reversibility, and these motives are often conflicting.

(a) *A minimum degree of phonetic accuracy* is provided for by the "similar in character" clause contained in Jones's later definition. By our purely logical definition, we should have the possibility of regarding English [h] and [ŋ] as members of one phoneme, which never occur in the same phonetic environment, and we could write forms like [hæt], [biŋeiv], [soŋ], [ˈsiŋə*] for *hat*, *behave*, *song*, *singer*, and learn very quickly when to say [h] and when to say [ŋ]. Such practice, however, would not be favored by either the phonetician or the philologist. Now the automaticity of variation within a phoneme has two senses. (1) The variation of [h] of the shades [h_e], [h_a], [h_o], [h_o]¹, etc. according to the following vowel is automatic practically in all languages which have these sounds. So is the variation of the [t] in [ts] and [tʃ] in all languages which have these affricates, that is, if we take affricates as successions of two phonemes. But such cases are much rarer than we are inclined to think. (2) In most cases, the automaticity of variation holds only for the particular language in question, although familiarity with the language may give one the impression of its universality. Thus, speakers of one language, e.g., Japanese, would find the change of [h] into [ç] before [i] so natural as to be something inherent in the nature of speech sounds, while in another language, e.g., German, [h] can be followed by [i] without becoming [ç], which belongs to another phoneme. The variation of Foochow [a] and [ɛ] 能, "to be able to," according to tonal environment is so natural to the native speaker, that he refuses to admit that he is not pronouncing it always in one and the same way, while in many languages these are widely different phonemes. Since,

1. Not to include cases of high vowels, which involve other questions.

therefore, the automaticity of variation is mostly of conditional nature, we shall have to allow a good deal of latitude in the interpretation of the "similar in character" clause. For the sake of phonetic accuracy, it would be an advantage to construct our phonemes with as narrow ranges of variation as possible (though it is never desirable to limit ourselves to universally automatic groups of the type (1) mentioned above), but this one desideratum may have to be sacrificed to some extent for other motives.

(b) *Simplicity or symmetry of phonetic pattern* is a factor which greatly influences our organization of phonemes. Bloomfield wishes to say that there are no long vowels in English, a statement which, from our standpoint, is neither true or false, but may be estimated as methodologically desirable or not desirable. He has eight vowels:

i	u
e	o
ɛ	ɔ
a	ɑ

and eight diphthongs or triphthong:

aj	oj ¹	ej	ij	juw
aw	ow		uw	

It would seem that he could gain phonetic accuracy by writing [ai], [oi], [ou], even without the addition of special symbols like [i:] and [u:], but then he could not very well go on and write [ii], [uu], and if he indicated the diphthongal character of these vowels by [ij], [uw], the system would look much less symmetrical. The table would also look less symmetrical if he wrote [i:] [u:], with the American narrow-range [e:] and [o:] lurking around for recognition, while [aj], [aw], and [oj] must still remain as diphthongs. The use of the nonce phoneme "o" for Chinese (see I (d) above) with zero as a possible member of the phoneme, gives great symmetry to the system. Again, the series ㄣ, | ㄣ, Xㄣ, | ㄣ may be symmetrically rendered as [an], [ian], [uan], [yan] instead of the usual [an], [ien]², [uan], [yan], which is phonetically more accurate, but by no means necessary. When symmetry runs parallel to structural or etymological considerations, so that the phonemes also agree with diaphones or phonogenes, its claim for consideration will of course be greatly increased.

1. Regard for "similarity in character" probably prompted him to identify the first element of *oil* with the first element in *or*, rather than the first element in *up*. He would gain still more symmetry if he wrote [oj], [ow], or still better [oj], [ow], as the first element in *own* is much nearer the first element in *or* [ɔr] than the first element in *up* [op] in American English.

2. Considering | ㄣ as the nasal ending counterpart of | ㄣ [iɛ].

(c) *Parsimony of entities* in the spirit of "Occam's razor" is of course the hobby of symbolologists. We already noted the admission of digraphs for single sounds for the saving of a whole series of new letters. The use of [ij] and [uw] or introduction of a symbol for length saves the use of the letters [ɪ], [ʊ], and [ɒ] for English. Palmer deplores this "exaggerated compliance with the principle of symbol economy,"¹ because, among other reasons, the symbol for length, e.g. in *although* [ɔɪlðou] does not necessarily indicate length. The writer can recognize the usefulness of the letters [ɪ], [ʊ], and [ɒ] from motives of phonetic accuracy, but the objection to the length mark does not seem to be fatal, for the symbol [ɔɪ] may also be taken phonemically in such a way that it is long in stressed positions, less long before voiceless consonants, and short (without change of quality) in unstressed positions, while [ʊ] can still be considered as a separate phoneme. Bloomfield's avoidance of [ə] and his identification of the vowel in *son* with the first vowel in *own* (instead of writing the former [ʌ] or [ɜ]) also effects a saving of "queer symbols."

The extent to which one could go in the parsimony of symbols can best be illustrated by Liu Fu's numerical code for the Peiping syllables.² He used only six symbols in six positions (or *plus* six positions, if we count positions as part of the symbols) as follows:

Position Figure	I. HEAD: Place of articulation	II. FACE: Manner of articulation	III. NECK: "Medial"	IV. ABDOMEN: Principal vowel	V. TAIL: Final vowel or consonant	VI. EXPRESSION: Tone
0	zero		zero	zero	zero	—
1	labial	unaspirated	i	ə	i	1st
2	dental	aspirated	u	ɑ	u	2nd
3	velar or palatal	nasal	y		n	3rd
4	retroflex	voiceless continuant			ŋ	4th
5	dental advanced	voiced continuant				

Thus, 光 [kuɑŋ] would be 312241, where 31 stands for [k], 224 is [uɑŋ] and the last figure 1 means the first tone. 000042 would be the nasal interjection meaning "What did you say?" This system is extremely symmetrical in structure, economical in the number of *kinds* of symbols used, and very illuminating as to the phonetic pattern of the language,

1. H. E. Palmer, *Principles of Romanization*, pp. 68-69.

2. "A Table of the Analytical Numbers of the Beiping Dialect," *The Kwoshyue Jihkan*, III, 3, 1932, pp. 533 ff.

but it can hardly be used as a system of transcription and was never intended to be. It may be noted here that his "Abdomen No. 1" includes [ɤ], [ʌ], [ɛ], [ʌ], [ɪ], [ə], and zero as members, and corresponds to our "o". In the body of the table, he gave also a somewhat narrow transcription of all the syllables.

(d) *The feeling of the native speaker* is a factor which is greatly emphasized by Sapir. Where the feeling comes from obvious misconceptions, arising often from orthographic considerations, such as the idea that *principal* and *principle* have different pronunciations¹, or that $ng=n+g^2$, we need not take it very seriously. But when there is no question of misconception, but one of preference of choice between alternate manners of organization of phonemes, then the feeling of the native should be given due consideration, though it need not be taken as the deciding factor. Thus, while the phonetician would write Chinese ㄢ, ㄣ, ㄤ, ㄥ as [an], [ien], [uan], [yan], the speaker of the dialect of Peiping feels that they all belong to the same rimeme with different medials. This is further supported by the fact that when the [-n] is dropped when the syllable is amalgamated with a following retroflex vowel, [ien] does not become [ier], but [iar], as in 一點兒 [i tien ər] > [itiar], "a little." The speaker of the Foochow dialect feels that among the vowels in the following words,

音	in ₁	詠	ein _Λ
鶯	ein ₁	限	ain _Λ
溫	un ₁	問	oun _Λ
恩	oun ₁	筭	oun _Λ

those in the same row are tonal variations of the same vowel, while refusing to recognize that the vowels in 詠 [ein] and 鶯 [ein] or those in 問 [oun] and 恩 [oun] are the same. As there are very definite rules for the diphthongization of single vowels (or opening of close vowels, as [ɛ]: [a]), it is quite possible to arrange the Foochow vowel phonemes according to the native conception as an alternate and for some reasons a better way of grouping the phonemes. On the ambiguity of the phonemic membership of Peiping ㄣ, ㄥ, ㄤ, the native speaker will also have something to say. The distribution or patterning of these sounds and related sounds are as follows:

1. Except when the former is pronounced [prinsi / pæl], which is merely an abbreviated way of saying "the word which ends in -p-a-l."

2. Even this is open to question, if we take a broader linguistic (as contrasted with phonetic or phonemic) point of view. Cf. Sapir's discussion on this point in "Sound Patterns in Language," *Language*, I 2 (1925), p. 49.

- | | | | | | |
|-----|-------|----|-----|---|---------------------------|
| (1) | ㄔ ㄨ ㄒ | tɕ | tɕ' | ɕ | always before [i] or [y] |
| (2) | ㄍ ㄎ ㄌ | k | k' | x | } never before [i] or [y] |
| (3) | ㄒ ㄗ ㄘ | ts | ts' | s | |
| (4) | ㄗ ㄘ ㄙ | tɕ | tɕ' | ɕ | |

It is therefore possible to identify the series (1) phonemically with any one of the other three series. Wade identifies it *partially* with (4): he writes *ch, ch, hs* for (1), and *ch, ch, sh* for (4). The National Romanization identifies (1) with (4) completely by writing *j, ch, sh* for both. The French system of romanization for Chinese has (2) or (3) according to etymology, which was what (1) came from, and over-zealous adopters of the French system identify (1) with (3) completely, and write forms like *Sien Sien* for 獻縣, although both belonged to series (2). Now as to the feeling of the native, the favoured series is (2). For he feels [kə, tɕi, ku, tɕy] or [xə, ɕi, xu, ɕy] to be alliterative series with only different vowels. Moreover, in the system of a secret language which breaks every syllable with initial-final *I+F* into *Iai+kF*, (e.g. 北 [pei]: [pai-kei])¹, the [k] becomes [tɕ] when the final begins with a high front vowel, as 米 [mi]: [mei-tɕi].

(e) *Regard for etymology* is properly not within the scope of our present study, which is concerned only with the descriptive study of one language of one period. But in the very frequent case of possibility of alternate phonemic treatment, we should certainly be allowed to steal a squint towards extrinsic factors. As a matter of fact, consideration of etymology does have a great weight with many writers. The identification of [tɕ], [tɕ'], [ɕ] with [k], [k'], [x] is etymologically preferable, if only partially, to identifying them with [tɕ], [tɕ'], [ɕ]. It would, however, cease to be strictly phonemic transcription of the Peiping dialect if we split [tɕ], [tɕ'], [ɕ] into a velar and a dental series according to derivation, as [xi] for 希 and [si] for 西, for then no rule of phonemic membership short of lexical enumeration could tell us when it is [xi] and when it is [si].

It is also of etymological interest to try to secure identity of words by giving them constant phonemic forms. Thus, we can write [sɜ:ʔ] or [sɜ:r] for *sir* and let the phonetic environment decide when it is to be pronounced [sɜ:r], [sɜ:], [sɜr], or [sɜʔ]. Again, by writing [ɛ:trə] for *être*, instead of [ɛ:tr] before vowels, [ɛ:trə] before consonants, and [ɛ:trʔ] at the end of phrases, Passy gives the word a constant form, the value of the phoneme [ə], which may be written in italics if desired, to be determined by the "rule of three consonants," etc. The Foochow

1. Y. R. Chao, "Eight Varieties of Secret Language," Based on the Principle of Fanch'ieh, *This Bulletin*, II, No. 3, 1931, pp. 320 ff.

word 能 “to be able to,” may be given the constant form [a], or a compromise form [æ], and the choice of values between [ɛ] and [a] may be determined by a very simple tonal rule. The identity-of-word interest, however, must not go so far as to cover grammatical considerations, where the rule of pronunciation would have to contain other than purely phonetic conditions. Thus, while we can write French *en* as [an], understanding that it is to be pronounced [ã] before consonants (*s'en va*) and [an] before vowels (*s'an aller*), we cannot write *fin* as [fin] in order to provide for the pronunciation of the feminine form [fɛn]. In such cases, we shall have to consider *fin* [fɛ̃] (or [fɛn]) and *fine* [fin] as two separate words, as much as *fil*s and *fil*le.¹

(f) *Mutual exclusiveness between phonemes* is another desideratum we wish to consider, that is, the list of phonemes shall not only be exhaustive for the language, but other things being equal, we should try to make the membership of the classes mutually exclusive. Other things, however, are never equal, and we have in fact already allowed the possibility of over-lapping of membership between phonemes in cases like the Foochow:

one phoneme	i	ei
another phoneme	ei	ai
one phoneme	u	ou
another phoneme	ou	u,

and in cases of different phonemes each of which containing zero as a member. The treatment of affricates as independent phonemes where their occlusive and fricative element can easily be identified with other phonemes in the same language, such as Bloomfield's [ʃ] and [č] for what many other writers give as [dʒ] and [tʃ], may also be considered as a case of overlapping of membership. Palmer calls this “multiple identity,”² under which he cites a number of examples from Japanese and English. We should note, however, that “the same sound” which belongs to two or more phonemes may be taken in two senses. In a conditional sense, “the same sound” never occurs under the same conditions as to contiguous sounds or as to conditions of stress, length, and intonation. The [ei] in the Foochow [i]:[ei] phoneme occurs always in the tones [˥], [˨˥], [˨˨], while the [ei] in the [ei]:[ai] phoneme occurs always in the tones [˥], [˨˥], [˨˨], [˨˨˥].³ The English [ɛ] and [j] also occurs under different

1. Cf. Bloomfield's distinction between *phonetic alternation* and *formal alternation*, “A Set of Postulates for the Science of Language,” (the journal) *Language*, II, 3, 1926, p. 160.

2. *The Principles of Romanization*, p. 151.

3. A pure phonetician would therefore prefer to take [ei] as one phoneme (or succession of two phonemes) in seven tones, although this would be against the “feeling of the native.”

conditions of stress from combinations like *heat sheets* and *and Jeanne*. In an absolute sense, two phonemes will have in common one member identical in all respects. Thus, there is absolutely no difference between the initial in 希 [ɕi] and the initial in 西 [ɕi], discussed above under (e). We could, if we liked, put both into the [x] phoneme or both under the [s] phoneme, but if we write 希 [xi] and 西 [si], then the identical [ɕ] would belong to two phonemes under the same conditions. This treatment brings up the question of.

(g) *Symbolic Reversibility*. The use of symbols has two aspects, the aspect of reading, or the determination of the object from the given symbol, and the aspect of writing, or the determination of the symbol from the object. The reading aspect of phonemic symbols is always determinate with respect to the language in question. Given a phonemic symbol, the range of sounds is determined, and the choice within the range is usually further determined by phonetic conditions. It would also be a desirable thing to make this reversible, so as to include the aspect of writing, that is, given any sound in the language, its phonemic symbol is also determined. If phonemes do not overlap, this is obvious. If they overlap, and the common members occur under different phonetic conditions, the reversibility still obtains. For instance, although 電報 is normally pronounced [tiempau], so that the *m* sounds exactly like the *m* in 門 [mən], yet we can tell that it is only a member of the phoneme *n*, as the phoneme *m* never occurs in this position in standard Chinese. Again, in the dialect of Foochow, if we had symbol *A* for the [i]:[ei] phoneme and symbol *B* for the [ei]:[ai] phoneme, we could still tell whether a given case of the sound [ei] is to be written *A* or *B* according to the tone.¹ But if the identity of a common member between phonemes is unconditional, as the distinction of 希 [xi] and 西 [si] for the Peiping dialect, then it would be impossible to go from the sound to the symbol even for the native speaker. Strictly, a non-reversible symbolization of sounds based on etymological or other considerations becomes an orthography and ceases to be a transcription, and the French system of romanization of Chinese, which distinguishes 基, 欺, 希 *ki*, *k'i*, *hi* from 躋, 妻, 西 *tsi*, *ts'i*, *si* (also favored by Bernhard Karlgren) is a case of this kind. In other words, homonyms should not have different transcriptions. There is, however, a class of intermediate cases, where the common member between two phonemes occurs sometimes under exactly the same phonetic conditions, but at other times becomes differentiated in some way under other sets of identical conditions. Thus, the same [ə] which occurs in *mica* [maikə],

1. This is not as complicated as the description looks on paper. The native speaker is not even aware of the vocalic identity or similarity of the [ei] in the two sets of tones.

and in *poker* [ˈpoukə] before consonants become differentiated into [ə] and [ər] respectively before vowels. If we write the former as [ə] and the latter as [ə*] or as [ər], then it will be possible to go from sound to symbol only when the sound in question is followed by a vowel, but not when followed by a consonant. The reversibility is therefore only partial. Usage is by no means uniform in such cases. Sometimes, symbolic reversibility is secured at the expense of word identity, the same word *poker* appearing in two forms [ˈpoukə] and [ˈpoukər], considered as different sets of phonemes. At other times, identity of word form is secured at the expense of reversibility, the same word Fr. *espece* always appearing as [ɛspɛs], where the final [s] is pronounced [z] when followed by a voiced consonant, so that given the final sound [z], one cannot tell whether it is a member of the [s]-phoneme or a member of the [z]-phoneme.

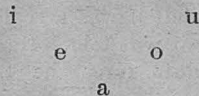
3. Choice of Symbols.—

It is one problem to group the sounds of a language into such and such phonemes and another thing to assign such and such symbols or letters to these phonemes. As a phonemic transcription has reference to one language, there is a great degree of freedom in our use of symbols. The freedom, however, is not so unlimited as in the case of mathematics, where the same symbol changes value not only from problem to problem, but also within the same problem. From purely logical considerations, it would seem that once the phonemes themselves are agreed upon, it is only a "matter of form" as to the symbols used for them, "What's in a letter?" Who ever heard of one mathematician writing *l, m, n* and another insisting that the same items shall be written as *p, q, r*? In phonetic symbols, however, there is a certain tradition, or rather, what is more unfortunate, a number of conflicting traditions in the use of symbols. Consequently, there arise frequent controversies with as much vehemence as about the use of words. We shall feel the importance of the use of symbols when we realize that it often has an influence on our actual organization of phonemes. Some of the factors which influence our choice of symbols run parallel to those which influence the organization of phonemes. Thus, symmetry and simplicity of phonetic pattern corresponds to a certain degree of symmetry and simplicity in the symbols. Parsimony in the number of phonemes implies also parsimony in the number of symbols. The feeling of the native as to sound will also apply to the choice of the symbol if the language already has an alphabet, although this is often less dependable than his feeling for the pattern in the abstract. In addition to these, we have the following questions specially concerned with the choice of symbols.

(a) *The desire to keep within the limits of the ordinary 26 letters of the roman alphabet* is such a powerful one that transcribers yield to it at great cost to other considerations. Thus, if a language has [ɑ], [ɔ], or [ɛ] but no [a], [o], or [e], then the latter symbols will be used as a rule.¹ If a language has only [ɐ], but no [r], then [r] would be used, although phonetically it would be taking as much liberty as writing [n] for [ɳ]. Bloomfield's use of [o] in the phoneme [o] and the diphthong [ow] avowedly comes from the desire to avoid "queer symbols." So far as parsimony of number of phonemes and symbols is concerned, [ɐ] would do just as well as [o], but would be even more appropriate, as it is more natural to say that the [ɐ]-phoneme is rounded in the diphthong [ɐw] on account of the labial [w], than to say that the vowel [o] in American English is an unrounded vowel except in the diphthong [ow]. This avoidance of queer letters means that while theoretical phonetics tells us that there are such and such sounds, or at least advises us to recognize conveniently such and such distinguishable sounds in the main, yet we feel inclined to identify the phonemes of a language with those sounds which happen to be favored with "lower case" letters.

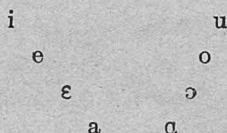
(b) *Of those symbols which are not the ordinary letters of the alphabet, some are considered less "queer" than others*, either on account of old standing or on account of the importance of their position in the scheme of general phonetics. Thus, [ŋ], [ʃ], [ð], [ø], [ɔ] are usually considered much less queer, and less effort is usually made in avoiding them than in the case of symbols like [ɣ], [β], [ɹ], [ɻ]. Again, in the abstract scheme of cardinal vowels, a special symbol for the part between [ɛ] and [a] would be of less importance than the eight main positions. And since it is possible to group all the [e]-[ɛ] region sounds in English under the phoneme [e], the symbol [ɛ] is thus left free for indicating the phoneme between cardinal [ɛ] and [a], which is what Bloomfield does: using the less queer symbol [ɛ] instead of the symbol [æ], which is "queer" in that it occupies a less strategic position.

(c) *The scale of division* into which a variable range of sounds is supposed to be divided will have a great influence on the choice of the symbols. Thus, the traditional triangular scale

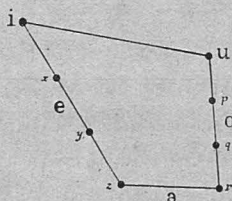


1. Jones and Camilli, *Fondamenti*, p. 3.

and the cardinal scale



differ in the number of intervals into which the vowels are divided. The difference would be less confusing if we had non-conflicting symbols in the new scale, something like:



As a matter of fact, one does have a partiality for using [e] for [ε] and [o] for [ɔ] (Cf. (a) above), and, less frequently, [a] for [ɑ], which shows the influence of prestige of the i-e-a-o-u system. Every transcriber feels that somehow [ε] is a variety of [e] and not a variety of [a], [ɔ] is a variety of [o] and not a variety of [ɑ]. If we took our scheme of cardinal vowels seriously, we ought not to have such feelings.¹

In Karlgren's scheme of vowels, using Lundell's dialect alphabet, the 3-point 2-interval high vowel scale of [i(y)-i(ɨ)-u(u)] of the IPA is given as a 2-point 1-interval scale of *l(y)-u(u)*. (More accurately speaking, *u* is placed by Karlgren a little to the front of *u*.)² The Russian *и*, which on the 3-point scale is nearest to [i], is therefore given as [i] in the IPA³, but as *u* by Karlgren, as it is nearer to the back vowel than the front *i* on his 2-point scale.

(d) *The avoidance of diacritical marks*, which are now reserved for modifiers, also influences our choice of letters. We already noted that rather than writing [ɕ] for the single sound in Chinese 歲, we allowed the modifier to be written separately, thus: [suei]. Again, if a language has only two series of voiceless plosives, one unaspirated and one aspirated, but no voiced plosives, then either [p, t, k; ph, th, kh] or [b, d, g; p, t, k] would be preferable to [p, t, k; p', t', k'] or [p̥, t̥, k̥; p, t, k].

1. The writer once heard a piece of music and interpreted it as being here in major and there in minor and its notes as being *do, re, mi*, etc., only slightly "off," but subsequently learned to his surprise that it was a scale of seven equal steps in the octave. The illusion persisted even after it was told. He had forced his own intervals into the new scale, just as we all tend to force the 4-step *i-e-a-o-u* scale into the 7-step cardinal scale.

2. Bernhard Karlgren, *Études sur la Phonologie Chinoise*, p. 316.

3. As for instance by Daniel Jones.

(e) *Consistency with phonemic transcriptions of other languages* is a thing that one may keep in mind, but which one must not go out of one's way to obtain. Where our phonemes are of narrow ranges and the symbols given them are the nearest phonetic letter we happen to have, the resulting transcription is not likely to conflict seriously with other transcriptions. But if for one reason or other, our phonemes vary within very wide ranges, and if, further, we wish to secure certain symbolic advantages by departing somewhat from the usual range of values of the letters, then the chance of conflict with other transcriptions will be greater.

III. PHONETIC AND PHONEMIC TRANSCRIPTIONS.

It is usual practice to distinguish between phonetic, or narrow, transcriptions and phonemic, or broad, transcriptions. The former express the actual sounds, [raɪt], [tɹaɪ], [lʌvɪ], [eɪt], [gɛt], [ðɛə], [æt], while the latter only indicate the distinctive classes of sounds, [raɪt], [tɹaɪ], [lʌvɪ], [eɪt], [gɛt], [ðɛə], [æt] (or [ɛt]). From the previous discussions, however, we have seen that there is no such thing as *the* correct phonemic transcription for any given language. According as we emphasize one or another factor in the size of the unit, method of phonemic grouping, and choice of symbols, we arrive at one or another form of phonemic solution. There is nothing in our definition of a phoneme or any other of the definitions quoted that can decide for us, for example, whether the Chinese [ɕ] shall be a member of [x] or [ɣ] or [s], or how the [ɪ] in [aɪ], the [j] in [ij], the [ɪ] in [ɪt], and the [j] in [jɛs] should be grouped into phonemic classes. The definition *permits* us to devise ways and means of grouping together distinguishable sounds that are not distinctive with respect to the particular system of phonemic grouping. It also implies that certain sounds in a language are never distinctive in that language by any reasonable manner of symbolic juggling, e.g., the difference between the [k]'s in *keep*, *call*, *coo*, etc. and the [h]'s in *heap*, *hall*, *who*, etc. can never be considered as being distinctive, unless we should do the very unnatural thing of considering all the vowels [i:] [ɔ:], [u:] etc. as non-distinctive members of one vowel phoneme X, the value to be determined by the nature of the preceding consonant k_1 , k_2 , k_3 , etc., h_1 , h_2 , h_3 , etc., or zero₁, zero₂, zero₃, etc. (i.e., in words like *eat*, *all*, *ooze*). But many sounds in a language are neither distinctive nor non-distinctive *per se*, but depend upon our particular manner of phonemic treatment. Thus, by writing *up*, *owe*, *oil* as [op], [ow], [əj], Bloomfield considers the difference between the first elements

in *up* and *owe* as non-distinctive and the difference between the first elements in *owe* and *oil* as distinctive. But precisely the reverse thing will have to be said if we treat the same sounds as [o], [ow], [oj], a modification which would do no damage to Bloomfield's system as a whole either by way of compromising the parsimony of letters, or by way of introducing queer symbols. Again, in most of the Wu-dialects, in words of the type [tɕ¹a], [ɕ¹a] [ɲ¹a], etc., as against [ka], [xa], [ŋa] the ['] is so short that it can be considered as a glide of the preceding consonant and can be left out in the transcription, in which case the difference between [k], [x], [ŋ] and [tɕ] [ɕ], [ɲ] would be considered distinctive. On the other hand, if we write the ['] on the line, then we could consider the [tɕ]-series as members of the [k]-series phonemes: [ki], [xi] [ŋi], and it is now the difference between [a] and [ia] that is distinctive. In practice, no phonetic transcription is so narrow and concrete as to distinguish between the [h]'s in [he], [hɛ], [tə] in any language, and no phonemic transcription is so broad and so purely abstract as to group English [h] and [ŋ] under the same phoneme [h]. Between these extremes, there are all intermediate proportions of phoneticity and phonemicity. On the whole, we may say that a phonetic transcription is one which makes use of all the usual distinctions which the majority of phoneticians are expected to be familiar with, irrespective of their distinctiveness in words, and that a phonemic transcription is one which, given a particular set of directions of approach, makes only such distinctions as are necessary in distinguishing words from that particular set of directions.

The reader will notice the unsatisfactory nature of the phrase "the usual distinctions which the majority of phoneticians are expected to be familiar with." This comes from the unsatisfactory nature of the actual state of affairs. In the field of descriptive phonetics, there is nothing like the near unanimity of opinion which exists among physicists, either as to the organization of facts or as to the use of symbols for referring to them. Thus, Bloomfield says, "The phonetician's equipment is personal and accidental, he hears those acoustic features which are discriminated in the languages he has observed. . . . He should remember that his hearing of non-distinctive features depends upon the accident of his personal equipment, and that the most elaborate account cannot remotely approach the value of a mechanical record."¹ This is all true to a great extent, but in the opinion of the writer, Bloomfield is going too far in saying further: "Only two kinds of linguistic records are scientifically relevant. One is a mechanical record of the gross acoustic features, such as is produced in the phonetic laboratory. The other is a record in terms of

1. Bloomfield, *Language*, pp. 84-85.

phonemes, ignoring all features that are not distinctive in the language. Until our knowledge of acoustics has progressed far beyond its present state, only the latter kind of record can be used for any study that takes into consideration the meaning of what is spoken." We need not, however, be worried if we cannot read or copy the grooves of a phonograph record. The phonograph record is at best an *icon*, or a picture, not a *symbol* in the usual sense of something that we can "read" and "write." Nor need we be worried that the number of sounds in human speech is infinite. The number of distinguishable sounds in human speech is relatively small, limited by the condition of *oral-auditory transmission of phonemic distinctions from one generation to the next*. When the average actual difference falls below a certain finite limen, the distinction becomes unstable, and the two phonemes soon coalesce into one later phonogenic member. We cannot say, as Bloomfield seems to imply, that phonetic transcriptions are mostly subjective and that phonemic transcription are mostly objective. We already saw how phonemic transcriptions are not unique and to that extent subjective. On the other hand, there is also a certain degree of practical agreement as to the non-phonemic use of symbols in general phonetics. For purposes of (1) citation of forms where a feature which is non-distinctive in the language cited is relevant to the point under discussion, (2) giving forms of words or sounds in comparative dialectology, (3) noting incipient or vestigial traces of sound-change, (4) impartial consideration of the gross features of a language before a good phonemic system has been worked out for it, and (5) as a less worthy purpose, for pedagogical use,—for all these, a narrow phonetic transcription is sometimes very useful and sometimes quite indispensable. One should not do the worst of narrow transcriptions all the time, but one should be prepared for the worst at any time. The dialect alphabet of Lundell, used by Karlgren in his *Phonologie Chinoise*, both in his main discussions and in the appended dialect dictionary, is a very narrow and non-phonemic transcription. The writer has nevertheless found the system thoroughly usable and understandable, and although for typographical reasons, he has changed it into IPA form in the Chinese translation,¹ he has been able to equate the symbols of the system with relatively few additions and few doubtful points of classification arising from the number of scale-steps problem. In the writer's own experience in the recording of Chinese dialects, he found that besides the matching and comparison of words with related sounds, a very important procedure is to give a reasonably narrow phonetic transcription at the start, so that we have materials to base our decisions upon when we come to questions of choice among alternate treatments.

1 中國音韻學研究，趙元任，羅常培，李方桂合譯，Shanghai, 1934.

Bloomfield observes rightly that phonetic transcriptions are often inconsistent as to what features to include and what features to neglect. This difficulty can be met in two ways. In the first place, we can lay down as a principle of symbolology that the position of a symbol in its context may be considered to be one constituent of the symbol. Thus, there is no inconsistency in the figure 1 meaning 1×10 and 7 meaning 7×1 in the form "17," as the symbol 1 is not just "1," but "1 in the second position." Similarly, there is no inconsistency in the symbol $>$ meaning "greater than" in $19 > 17$ and meaning "changes into" in $p > f$, or even between the two uses of $>$ in $a > o$ according as the formula occurs in an article on phonetics or in one on mathematics.¹ So in discussions on diphthongs, we may need to mention forms like [čaj], [čæ], [öæ], etc., while in discussions on affricates, we may have reference to [tšai], [tšai], [tšai], [tšai], etc., just as Bolling finds it perfectly in order to write *Enroughity is coming.—The Enroughities are coming*, so long as the discussion is about the ending of the plural.² But if our discussion should turn on the forms of the indefinite article, it would then be necessary to write: [ɛn eg] but [ej ɪdaɪbi] (the correct pronunciation of the name *Enroughity* according to Bolling), as it would not bring out the point at all if we wrote: *an egg*, but *a Enroughity*.

For avoiding too much inconsistency in the citation of forms, both Karlgren and users of the IPA have resorted to the distinction between broad and narrow transcriptions apart from considerations of significant distinction. Karlgren's practice, as carried out in his *Phonologie* (pp. 260ff.), is very consistent. He has a set of bold-faced letters for a broad transcription, under each of which he puts a number of the Lundell's letters, which are always in italics. Thus, what correspond to the [ɛ] and [æ] in the IPA are both grouped under ä, what correspond to [ɛ], [æ], [ɛ] in the IPA are grouped under [š], and so on. There are a few cases of overlapping groups, but on the whole the groups are mutually exclusive. The relation between the two sets is therefore very much like that between phonemes and its members except that no reference is made to word distinction. A similar tendency is noticeable among users of the IPA, but no systematic division has ever been made between a narrow and a broad transcription. Nevertheless, there are certain unsystematic traditions among phoneticians which are based, on the whole, on the identity of the letters in the roman alphabet.

1. In discussions like the present, where there may be a call for "narrow symbols," one could use \rightarrow for "changes into" and $>$ for "greater than," thus making peace among mathematics, phonetics, and chemistry.

2. From editorial note on R. G. Kent's review of Bloomfield's *Language* in the journal *Language*, X, 1, 1934, pp. 50.

Thus, *r* is somehow recognized as a broad form covering [r] and [R], whereas [t] and [q] are not covered by any broad form. Similarly, [e] and [ɛ] are felt as members of a group of the *e*-type in a way that [i] and [e] do not seem to be. All this points to a conception which no one consciously recognizes, but which seems to be assumed by many, that there are such things as phonemes in general, apart from reference to any particular language, and that all we need to do either for the study of one language or for comparative work is to use one consistent phonemic transcription for all languages. This would of course be recognized by any one as an impossible illusion as soon as the situation is thus made explicit, as we may be called upon at any time to make phonemic distinctions between shades of sounds whose differentiation we never anticipated in either our narrow or broad system of phonetic symbols. The existence of the tradition of usage, however, is real. It is true that the existence of only one common letter *r* for [r] and [R] but two common letters *t* and *k* for [t] and [k] (or [q]) is a matter of historical accident. But we shall see the significance of this accident when we note that as a matter of fact most of the languages which phoneticians, or at least European phoneticians, have studied, do take [t] and [k] as separate phonemes, while [r] and [R] rarely, if ever, occur as separate phonemes. The idea of general phonemes, which we have just proposed and condemned in the same breath, is therefore not entirely baseless. Without entertaining the idea of general phonemes as such, the writer wishes to propose the term *typical phoneme*, to be defined as those groups of sounds which very often go together to form phonemes in many of the major languages studied by phoneticians. This definition of course makes the idea of a typical phoneme depend again on historical accident, the fact that most contemporary phoneticians are speakers and writers of the Germanic and Romance languages. Thus, for a broad transcription using typical phonemes, a European would group [p] and [p'] under one typical phoneme as against [b]¹, while an unsophisticated Chinese phonetician would most likely group [p] and [b] under one typical phoneme as against [p'].

The troublesome part of the transcription problem comes from the inconsistency in using the same symbol sometimes in a general and sometimes in a particular sense. In the citations in this article, the writer has found it hard to do better, and has tried to manipulate the context (taken as part of the symbolic system) in such a way as to eliminate ambiguity. But there is always the danger of slips. When we refer to the English [i], one may not know whether it is narrow [i]

1. Except speakers of certain German dialects.

or [ɪ] that is meant.¹ This is very similar to the old practice of referring to the ancient Chinese initials 照, 穿, 牀, 審 in this way:

General names: 照穿牀審

For the apical series: 莊初牀山 [tʂ], [tʂʰ], [dʒʰ], [ʂ]

For the dorsal series: 照穿乘審 [tɕ], [tɕʰ], [dʒʰ], [ɕ]

so that when 照 is mentioned, one is at a loss as to whether it is 照 in general (including both [tʂ] and [tɕ]) that is meant, or only 照 [tɕ] as against 莊 [tʂ]. He has therefore proposed the following names for the differentiated series, reserving the traditional names for the general sense, incidentally also using an inclusive broad transcription for the general series, thus:

General names: 照穿牀審 [ɕ], [ɕʰ], [j], [ʂ]

For the apical series: 莊初崇生 [tʂ], [tʂʰ], [dʒʰ], [ʂ]

For the dorsal series: 章昌乘書 [tɕ], [tɕʰ], [dʒʰ], [ɕ]

Karlgren's use of a special series of bold-faced types is based on the same principle. Symbols may be as general and inclusive as we may have use for, but must not be vague and ambiguous. An approach to this method of having both general and particular use of symbols is made in connection with the usage of a few symbols in the IPA. Thus, the symbol [ə] is usually understood to be a general form for [ɜ] (half-close) and [ɐ] (half-open). [] and [ʒ] may be used either for [ʂ] and [ʒ] or for [ɕ] and [ʒ] respectively. This latter, however, is less satisfactory, as in the dialect of Lintzy (臨淄), Shandong, [ʂ], [ʃ], [ɕ]: all three exist as separate phonemes, in which the [ʃ] series is intermediate between apical and dorsal articulations of the tongue and is identical with English [ʃ] except that there is no protrusion of the lips. [ʂ] and [ʒ] would be better general symbols, though they are not properly IPA letters.

SUMMARY.

We have proposed a new definition of a phoneme and have endeavored to show that given a language, there is not necessarily one unique solution for the problem of reducing its sounds into elements. We have considered what factors can influence, and have influenced, the phonemic treatment of languages: the variability of the size of the phonemic unit, including the admission of zero symbols and zero sounds, the grouping of phonemic membership, and the choice of actual symbols. Because phonemic solutions are not unique, it is necessary, before arriving at solutions, to have recourse to considerations of descriptive phonetics

1. On the principle of non-uniqueness of phonemic transcriptions, we cannot prohibit the writing of the vowels in *eat*, *it* as [i], [ɪ] and insist on the writing of [iː], [i] or of [ij], [i].

and the use of phonetic transcriptions. These are also necessary for other purposes, such as the comparative study of dialects. We have also noted that there is a tendency among phoneticians to group together sounds under broad symbols, which form phonemes in a number of languages, and we have called them "typical phonemes," although there is no consistency in the use of symbols for these. It is hoped that a more consistent system of symbols be devised for indicating both narrow shades of sounds and typical phonemes for the purpose of phonetic and phonemic transcriptions, but for the time being, we have to let the context serve as part of the symbol to inform us as to shade (if particular) or scope (if general). It is not necessary to take serious exception to anyone's transcription so long as it is self-consistent and its interpretation is clear to the extent it is meant for, and so long as it does not claim unique correctness to the exclusion of other possible treatments. Usage may in time become unified, but problems will always vary. Our motto must be: Write, and let write!

中文撮要

近年提倡音位標音法的各著作家，在他們的言論中，往往像已經假定每一個語言只有一種可能的音位化法。本文就是要證明音位標音法，對於任何語言，不是單答案性的，乃是有多少可能方式的答案的。從事實上人家用的各種標音法，可以尋出許多影響答案的因子出來。

1. 單位的尺寸問題。——平常說“一個音一個符號”，其實常常有多音一號或一音幾號的標音法。在極端的例就發生零符號零音問題。——比方德文元音起頭字必有[ʔ]音，但不寫也可以知道它的存在，這是零符號代表音。“陰平無號”也是零號的例，[ba:dn]寫作[ba:den]就是有[e]號而無音的例。

2. 組類問題。——哪些哪些音歸為一個音位，這問題是跟着許多因子變的。(a)音質準確度，(b)全系統簡單或對稱的要求，(c)音位總數減少的要求，(d)本地人對於音類的見解，(e)字源的顧及，(f)音位與音位間局部重複的避免，(g)讀音知位，見位得音互指可能的要求。這些要求往往互相衝突，對這上對那上輕重的不同，就會得不同的答案。

3. 符號的選擇。——影響組類的各因子有些也影響符號的選擇。此外有(a)用普通羅馬字的傾向，(b)用較常見的語音符號的傾向，(c)分度的數目的不同，(d)增加‘麻子’符號的避免，(e)跟別的音位標音法衝突的避免。

近年有人主張只有音位標音法是有價值的，所謂嚴式標音法或音質標音法是主觀的，沒有意義的。本文從音位標音法多能性的原則證明在音位答案未擬定以前，非得用純粹語音學跟它所用的嚴式音標，才能夠作擬音位答案的初步的工作。從本人方言調查的經驗，深感到這種音質標音法的必要。此外還有稱述某某語言中某某字的音，其中所指的音素在音位標音法所不必標者，而稱述時要提到者；比較各相近的方音；注意音變的初兆或殘跡；為教學時辨別雖同音位而音質相差很遠的音，——以上各種事情都是用得着音質標音法的。至於說各作家所用嚴式不同，但也不是完全不同，不能因此就作因噎廢食的主張。

有些近似的音組，因為它們在一般學者所遇見的語言當中常常合成音位，這種音組可以稱它們作“常見音位”，它們雖常常成為音位，但有時並不。標這種常見音位，往往用其中各音值之一的符號，例如用[r]式表[r]或[R]，那就有如分照母為莊：照，提起“照”字就不知道是指照總類還是與莊對待的照。用音標也應該用作者所提議分照母為莊：章的辦法為原則。但在國際音標沒有改一致化，系統化以前，我們只可以拿上下文來定所用各符號的功用，如本文所引各例標音法就只好靠上下文來定它的確實的意義。