HANDBOOK OF MIDDLE AMERICAN INDIANS

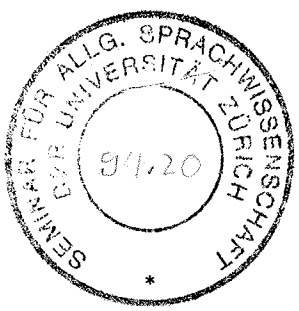
ROBERT WAUCHOPE, General Editor

VOLUME FIVE

Linguistics

NORMAN A. McQUOWN
Volume Editor

UNIVERSITY OF TEXAS PRESS AUSTIN



Published in Great Britain by the
University of Texas Press, Ltd., London

International Standard Book No. 0-292-73665-7
Library of Congress Catalog Card No. 64-10316

Manufactured in the United States of America

Second printing, 1972

The preparation and publication of the
Handbook of Middle American Indians
has been assisted by grants from
the National Science Foundation.

5. Systemic Comparison and Reconstruction

ROBERT LONGACRE

0. INTRODUCTION. The systemic comparison of languages and reconstruction of earlier stages has been pre-eminently successful in the field of Indo-European studies. In spite of limited application to other language groups (noticeably Semitic, Finno-Ugric, Bantu, Sino-Tibetan, and Algonquian), the accomplishment of Indo-European comparative reconstruction still remains unparalleled. Nevertheless, over the past several decades the comparative reconstruction of American Indian languages, especially those of Middle America, has gone forward at an ever accelerating rate. Whorf's Uto-Aztecan reconstruction of 1935 was not only the first piece of systemic reconstruction involving Middle American languages, but remains a model of craftsmanship in respect to its sensitivity to basic

considerations of linguistic structure. Its only defect is its brevity. Other serious comparative studies did not follow until the latter part of the next decade when Swadesh's article on Proto-Zapotecan (1947) and Wonderly's article on Proto-Zoquean (1949) appeared. The Newman and Weitlaner articles on Proto-Otomí and Proto-Otomí-Mazahua (1950) provided an auspicious start for the past decade which saw the appearance of Arana's article on Proto-Totonac-Tepehua (1953), McQuown's tentative sketch of Proto-Mayan (1956), and volumes by Longacre (1957) and Gudschinsky (1959c) on Proto-Mixtecan and Proto-Popolocan. In this decade has appeared a sketch on Proto-Mixtec by Mak and Longacre (1960), an article of Bartholomew's revising Proto-Otomian consonants (1960), and an ethnolinguistic article of Millon and Longacre which incidentally sketches a few features of Proto-Amuzgo-Mixtecan (1961). The present backlog of unpublished studies includes: Arana's sketch of Proto-Amuzgo-Mixtecan; Fernández de Miranda and Weitlaner's volume on Proto-Chiapanec-Manguean (1961); a very brief sketch of Smith and Weitlaner (1957) on Proto-Chinantecan; and an article of Bartholomew's (1959) on Proto-Otomí-Pame. A significant study now in progress is Fernández de Miranda's (1960) revision and amplification of Proto-Zapotec which was first reconstructed by Swadesh almost a quarter of a century ago.[1]

All told, the above makes an imposing tally. Notice that all branches of the so-called 'Mexican Penutian' (excluding Huave whose affinities to Zoquean, Mayan, and Totonac-Tepehua are doubted by both McQuown and Swadesh) have been subjects of some systemic comparison and recon-

[1] In the five years since this article was written, further comparative studies have appeared (not included in the main bibliography of this volume): C. F. Voegelin, F. M. Voegelin, and Kenneth Hale, *Typological and comparative grammar of Uto-Aztecan I* (phonology) (*IJAL*, Memoir 17, 1962); Calvin Rensch, *Proto - Chinantec phonology* (master's thesis, Univ. Pennsylvania, 1963); Terrence Kaufman, *Mixe-Zoque diachronic studies* (unpublished, 1963); Allan Wares, *A comparative study of Yuman consonantism* (doctoral dissertation, Univ. Texas, 1964): Burton W. Bascom, Jr., *Proto-Tepiman* (Tepehuan-Piman) (doctoral dissertation, Univ. Washington, 1966); Paul Kirk, *Proto-Mazatec* (doctoral dissertation, Univ. Washington, 1966); Calvin Rensch, *Proto-Otomanguean phonology*, (doctoral dissertation, Univ. Pennsylvania, 1966).

The following articles have appeared (also not included in the main bibliography of this volume): Maurice Swadesh, "The Oto-Manguean hypothesis and Macro-Mixtecan" (*IJAL*, 26: 79–111, 1960); Longacre, "Swadesh's Macro-Mixtecan hypothesis" (*IJAL*, 27: 9–29, 1961); Wigberto J. Moreno, "Estudios Mixtecos" (reprinted from the introduction to the facsimile edition of *Vocabulario en Lengua Mixteca* of Fray Francisco de Alvarado, Mexico City, 1962); Longacre, "Amplification of Gudschinsky's Proto - Popolocan - Mixtecan" (*IJAL*, 28: 227–42, 1962); Norman Nordell, "On the status of Populuca in Zoque-Mixe" (*IJAL*, 28: 146–49, 1962); Longacre, "Progress in Otomanguean reconstruction" (*Proc. Ninth Int. Cong. Linguistics*, Cambridge, Mass., 1962 [1964]); Ronald Olson, "Mayan affinities with Chipaya of Bolivia I: correspondence" (*IJAL*, 30: 313–24, 1964); "Mayan affinities with Chipaya of Bolivia II: cognates" (*IJAL*, 31: 29–38, 1965); Kaufman, "Materiales lingüísticos para el estudio de las relaciones internas y externas de la familia de idiomas mayanos" (in *Desarrollo cultural de los Mayas*, pp. 81–136, E. Z. Vogt and A. Ruz L., eds., Mexico City, 1964); Longacre, "On linguistic affinites of Amuzgo" (*IJAL*, 32: 46–49, 1965); B. W. Upson and Longacre, "Proto-Chatino phonology" (*IJAL*, 31: 312–22, 1965).

It is evident that the volume of study completed since I wrote this article considerably exceeds the work summarized here. To take account of the recent work would require rewriting the entire article. A recent summary which takes account of most of the above is found in Longacre, "Comparative reconstruction of indigenous languages" (*Current Trends in Linguistics*, vol. 4).

One important shift in my own position is that I now regard Amuzgo as constituting a seventh language family within Otomanguean. This is reflected in the classification embodied in the linguistic map (App. F), but is not reflected elsewhere in this article.

Some other articles of possible relevance here but not mentioned in the body of this article are: Bright, 1956; Ecker, 1939; Fernández de Miranda, 1951; Greenberg and Swadesh, 1953; Hale, 1958; 1959; McQuown, 1942; Mason, 1952; Mayers, 1960; Radin, 1916; Ravicz and Romney (unpublished MS); Swadesh, 1956; Weitlaner, 1942; Wonderly, 1953. To these can be added an unpublished file of Uto-Aztecan cognates prepared by Wick J. Miller.

struction. The time would soon be ripe, it seems, for comparing Proto-Zoquean, Proto-Mayan, and Proto-Totonac-Tepehua, and demonstrating the assumed relationship of the three to each other—if, indeed, it be demonstrable.[1a] The chief difficulty at present would seem to be the lack of extensive published ensembles of cognate sets in each of these language families. A further consideration will then be the possible affinity of the Mexican stock to Penutian of California. Each of the six branches of Otomanguean (Mixtecan, Popolocan, Chiapanec-Manguean, Otomí-Mazahua-Pame, Zapotecan, and Chinantecan) has been reconstructed either in published or unpublished studies. Some of these reconstructions (noticeably Chinantecan) are indeed most sketchy. Two others (Mixtecan and Popolocan) are extensive. For Chiapanec-Mangue we have all we can ever hope for, because the two languages are extinct and exist only in rather fragmentary records. Although Swadesh's Zapotecan sketch does not give a very extensive body of cognate sets, the present studies of Fernández de Miranda should supply what is lacking here. The combined efforts of Newman and Weitlaner plus Bartholomew readies Otomí-Mazahua-Pame for comparison with other branches of Otomanguean. The time is therefore very near when Proto-Otomanguean may itself go onto the drafting-board, if Otomanguean proves to be a valid genetic grouping.[2]

[1a] Terrance Kaufman has assembled evidence (unpublished cognate sets) to substantiate this grouping.

[2] With the completion of Rensch's dissertation (of footnote 1) Otomanguean is now firmly established. In this massive work (427 cognate sets), every phase of comparative Otomanguean phonology (consonants, vowels, laryngeals, tones, consonantal alternations) is carefully considered. Every constituent family of Otomanguean is given chapter length treatment. This study modifies some details of my Proto-Mixtecan reconstruction and replaces entirely my speculations regarding the probable structure of Proto-Otomanguean. In particular, Rensch reduces drastically the number of vowels reconstructed for Proto-Otomanguean and for Proto-Mixtecan as well. The latter reduction is made possible by the reconstruction of two laryngeals for Proto-Mixtecan rather than my one laryngeal (*?). Kirk in Proto-Mazatec and Bartholomew in Proto-Otopamean likewise reconstruct systems of but four vowels. Rensch also rejects the whole system of postposed deictic particles first suggested by Gudschinsky (1959) and further developed in Longacre 1962 and 1964.

Meanwhile Gudschinsky's comparison of Popolocan and Mixtecan (1959c) and the Fernández-Weitlaner common reconstruction of Popolocan-Mixtecan with Chiapanec-Mangue (1961) amount, in effect, to two approximations to Proto-Otomanguean. When reconstructed with detail and care, the latter may give us a piece of Middle American linguistic reconstruction approaching in depth and diversity the accomplishment of Indo-European scholarship, although the work will long be hampered by the skeleton crew manning Middle American comparative studies as compared to the more favorable number of scholars in Indo-European.

In this article we sketch (1) the operation of the comparative method (as illustrated by Mixtecan materials) along with factors obscuring regular sound change, with mention of some important ethnolinguistic by-products of the comparative method; (2) reconstructed phonological systems involving Otomanguean languages with mention of any grammatical features that have also emerged in the course of such work, and with attention to approximations already made to the reconstruction of Proto-Otomanguean itself; (3) reconstructed phonological systems involving other Middle American languages; (4) problems arising from "diffusion" versus "family tree" concepts.

1. THE COMPARATIVE METHOD as worked out on the terrain of Indo-European studies is "comparative" in a very specialized sense of that term. We are not here comparing gross similarity of shape and function (as in, say, comparative anatomy). Linguistic typology in the latter sense has been slow developing, although this may be a fruitful development in this decade. Rather we are concerned with the discovery of systematic

sound correspondences between languages.

In the application of the comparative method several stages may be noted: (a) By systemic comparison regular sound correspondences are noted. (b) An attempt is then made to discover which sets of sound correspondences are in contrast in the same or similar phonological environments, and which sets are in noncontrastive distribution (i.e., occur in mutually exclusive environments or in free variation). In this fashion the number of reconstructed phonemes is ascertained. (c) The reconstructed phonemes are then assigned symbols according to phonetic plausibility, i.e., in terms of the phonetic nature of the reflexes (the present-day phonemes which have developed from the reconstructed sounds), in terms of the geographical distribution of those reflexes, and according to their possible placement in the emerging phonological system. (d) Once the phonological system is tentatively worked out for a proto-language (i.e, a "parent" language), we are in a position to recognize other features (grammatical and lexical innovations) which in part obscure the regularity of appearance of the postulated reflexes. These latter features may have considerable ethnolinguistic import.

These stages in the application of the comparative method are illustrated with some detail in the following section. Illustrative data are drawn from the reconstruction of Proto-Mixtecan (by comparison of Mixtec, Cuicatec, and Trique). But here let it be noted that, beginning in the most pedestrian fashion with a search for regular sound correspondences, systemic reconstruction eventually leads to the uncovering of facts which relate not only to the phonological history of a language family but also to its grammatical and lexical history as well.

1.1. Sound correspondences. For Mixtec (of San Miguel el Grande), Cuicatec (of Concepción Pápalo), and Trique (of San Andrés Chicahuaxtla) the following sound correspondence is observed before front vowels:

(1) M *š* ~ C *d* ~ T *c* as in M *vīšì* 'sweet', *ndùšī* 'honey'; C *nā-dī²ī* 'honey', *čyāādē* 'sugar'; T *zi³ci²i*[43] 'candy', *ga³ci²i*[43] 'honey'.

There is also a further sound correspondence observed before back vowels and *a*:

(2) M *s* ~ C *d* ~ T *t* as in M *tï-sāà*, C *²yáádá*, T *ža³taha*[43] 'bird'.

These two sound correspondence are partially similar in that the Cuicatec reflex is identical in both cases while Mixtec *š* and Trique *c* could be considered to be palatalizations of Mixtec *s* and Trique *t* respectively. Since reflexes Mixtec *š* and Trique *c* occur in environments which may exert palatalizing influence (before front vowels) while Mixtec *s* and Trique *t* occur in nonpalatalizing environments, it seems plausible that we have here not two contrasting reconstructed phonemes but simply one such phoneme with reflexes distributed in mutually exclusive and noncontrasting environments.

However, in contrast to these two sound correspondences the following sound correspondences occur:

(3) M *t* ~ C *t* ~ T *č* as in M *kātā*, C *kāātā*, T *ga³ča*[21] 'to sing'.

(4) M *č* ~ C *t* ~ T *č* as in M (*kōò*) *yúčí* 'alligator', C *²yáátě* 'lizard', T *če*[343] 'scales'.

These latter two sound correspondences are also found to be in complementary distribution relative to one another: (4) is found in ultima syllables before front vowels in all three languages; and in Mixtec before *a* and after *i* or *u*; (3) is found elsewhere.

1.2. Reconstruction. From the above sound correspondences in which (1) and (2) contrast as a pair with (3) and (4) we reconstruct two phonemes. What phonetic values may be assigned to them? For sets (3) and (4) we assume a Proto-Mixtecan voiceless alveolar stop, viz. **t*. For sets (1) and (2) we need to reconstruct another sound which is presumably also alveolar; **s*, **đ* and **θ* are likely choices. The first option **s* is not so likely in view of the Mixtec dialect evidence

summarized in section 4.5. Choice of symbol for our reconstructed Proto-Mixtecan phoneme is arbitrary as to *θ or *$đ$. I have chosen the former in that two other spirants, *x and *x^w, seem to be witnessed to in the three languages. By choosing *θ we obtain a series *θ, *x, *x^w which parallels the *t, *k, *k^w stop series, and the *nd, *^ng, *$^ng^w$ pre-nasalized series.

1.3. APPARENT EXCEPTIONS to the appearance of regularly postulated reflexes may be occasioned by various features. Some such aberrances may indicate the need for reconstructing another phoneme or may indicate a previously postulated phoneme in a hitherto unsuspected position (i.e. forming a consonant cluster, a vowel cluster, or some post-vocalic consonantal element coloring the quality of the preceding vowel). Thus, from the following frequently encountered sets of sound correspondences six Proto-Mixtecan vowels are reconstructed (a number of specially conditioned reflexes are omitted here; but some alternative reflexes are indicated below without statement of conditioning factors):

	M	C	T	PMx
(1)	*i*	~ *i*	~ *i*	**i*
(2)	*i/e*	~ *e*	~ *e*	**e*
(3)	*ï*	~ *i/e*	~ *ï*	**ï*
(4)	*a*	~ *a*	~ *a*	**a*
(5)	*u/o*	~ *u/o*	~ *o*	**o*
(6)	*u/ï*	~ *u/e*	~ *u*	**u*

Nevertheless, having reconstructed these Proto-Mixtecan vowels on the basis of the above sound correspondences, there yet remains a further sound correspondence:

(7) M *u/o* ~ C *u/o* ~ T *a*

While this further sound correspondence is partially similar to numbers (4) to (6), it apparently contrasts with all three.[3] It would seem, therefore, that the only course open to us is that of reconstructing a further vowel phoneme, e.g. *$ɔ$. Such a phoneme was reconstructed in my Proto-Mixtecan study. At the same time I reconstructed a post-vocalic *-*m*, which evidently occurred with considerable frequency, which was the main source for nasalized vowels in the Mixtecan languages, and which survives to this day in closely related Amuzgo as well as in Mixtec dialect forms that have added a final vowel. Apparently this *-*m* patterned as some sort of suffixal element. Consequently the occurrence of a given vowel reflex colored by old post-posed *-*m* versus reflex uncolored by *-*m* is largely unpredictable. Thus in the same language may occur nasalized and unnasalized vowel reflexes in etymological doublets that seem to stem back to the same Proto-Mixtecan root plus or minus *-*m*. I had also realized that on occasion, the *-*m* not only resulted in nasal coloring of the vowel, but in raising and backing of the vowel as well (e.g. PMx **nam* > M *nu*). What I failed to realize was that the raising and backing influence of the post-posed bilabial nasal was general enough to account entirely for the aberrant set of reflexes, M *u/o* ~ C *u/o* ~ T *a*. It now seems plausible that every instance of reconstructed *$ɔ$ vowel may be reconstructed simply as **am*. In brief, instead of reconstructing a further vowel phoneme, we simply expand the distribution of post-vocalic bilabial nasal to a number of sets where its presence was previously unsuspected.

Thus, sets 167 and 222 of my Proto-Mixtecan study are remarkably parallel, but for the former set I reconstructed *$ya^2na^{2?}$ 'network bag', whereas for the latter set I reconstructed *$k^wa^2nɔ^{2?}$ 'to weave'. The latter we now reconstruct simply as *$k^wa^2na^2m^?$. Thus, set 167 represents a Proto-Mixtecan noun built on the root 'to weave' and not characterized by post-posed *-*m*, while the root 'to weave' has the post-posed nasal. Notice that both the Proto-Mixtecan noun and verb correspond as to the vowel of their first syllables, as to tone class *22 and as to presence of final glottal

[3] Swadesh and Arana's attempt to dispose of this as an assimilation of Proto-Mixtec *u/o ... a > u ... u or o ... o ignores a mass of data.

stop. Similarly, my set 223, formerly thought to witness to **nɔ* in its last syllable, is seen to witness to **nam* instead.

Such a restatement typically leads to adjustments all along the line. However, the postulation of **am* rather than **ɔ* not only obeys the law of parsimony in avoiding reconstruction of further phonemes, but constitutes a more adequate explanation for several apparent anomalies of phonological development (see Appendix A).

1.4. PATTERNS OF PHONEMIC SUBSTITUTION IN THE PROTO-LANGUAGE. Another factor apparently interfering with regular appearance of postulated reflexes may consist of patterns of phonemic substitution in the proto structure. Once such patterns are recognized, it is seen that they do not constitute exceptions to regularity of sound development; the developments are regular enough, but a cognate in one language may not hark back to exactly the same form as does a cognate in another language. Granted slightly different starting points, it is not surprising that the end points differ also. Thus, in Indo-European languages vowel gradation has played an important part and results in apparent anomalies of phonological development until its role is understood. In the Mixtecan languages consonant gradation of a systematic sort may be postulated for the parent structure. The patterns of consonant gradation were primarily of two sorts, with Proto-Mixtecan **k*^{w}-, **y-/*w-/*x-*, **k-*, and **^{n}d-/*n-* marking verb paradigms with four tense-aspects (and various allomorphs of the second and fourth aspects), and **t-/*θ-*, **y-*, **^{n}d-*, and **n-* marking nouns. Possibly the **y-* of the noun paradigms indicated *indefinite* (which became Trique unpossessed and Amuzgo plural), whereas **t-/*θ-* marked *definite* (which became Trique possessed and Amuzgo singular). A given noun reconstructs with *either* a paradigm involving **t-* *or* with a paradigm involving **θ-*. In the cognate sets of my Proto-Mixtecan study there is almost no overlap (total of only four cases) of these two paradigms. Such overlap as occurs presumably reflects a mixture of nouns harking back to **t-* versus **θ-* paradigms but built on the same Proto-Mixtecan root.

The scheme of Proto-Mixtecan consonantal gradation sketched above is, of course, an end product of comparative reconstruction. It affords a systematic explanation of many apparent exceptions and anomalies of phonological development. Therefore, we summarize and illustrate here some of the problems encountered at this point in the course of reconstructing of Proto-Mixtecan.

In reconstructing **CVCV* forms, both of the vowels as well as the second consonant reconstruct quite regularly—although not without problems of the sort illustrated above in reference to the reconstruction of **am* versus **ɔ*. In the second consonantal position phonemes **θ* and **t* are reconstructed from the sound correspondences already presented and exemplified. In this second consonantal position we also catalogue such additional correspondences as:

	M	C	T	PMx
(1)	*y/ñ*	∼ *y* ∼	*y/zero*	< **y*

(with Mixtec *ñ* reflex conditioned by occurrence of another nasal in the Proto-Mixtec form, and Trique zero reflex conditioned by following *i*).

(2)	*n*	∼ *n* ∼	*n*	< **n*
(3)	*č*(<PM**t*)	∼ *n* ∼	*n*	< **^{n}d*

(Mixtec reflex *č* occurs intervocalic; other conditioned reflexes, including *^{n}d* occur elsewhere.)

But although these reflexes and reconstructions are clearly established in reference to the second consonant of **CVCV* forms, they do not seem directly applicable to the first consonantal position. Furthermore, we experience the same difficulty in reconstructing the consonant of some **CV* forms. Evidently something has disturbed the regularity of anticipated reflex for the first consonant of many forms, whether

disyllabic or monosyllabic. Thus, from such sets as 97, 98, and 171, we can abstract the sound correspondence:

$$\text{M } \tilde{n} \text{ (PM } *y) \sim \text{C } y \sim \text{T } t;$$

from such sets as 277 we can abstract

$$\text{M } \tilde{n} \text{ (PM } *y) \sim \text{C } d \sim \text{T } t;$$

and from set 178 we can abstract

$$\text{M } \tilde{n} \text{ (PM } *y) \sim \text{C } t \sim \text{T } n.$$

Taking these sets at their face value, we could reconstruct three additional Proto-Mixtecan phonemes—albeit phonemes of very limited distribution: t^y, d^y, and $^nd^y$. But positing these three additional phonemes would be but a beginning, in that inspection of further sets would necessitate the reconstruction of more such phonemes of restricted distribution. This would be especially true in regard to sound correspondences abstracted from the first consonantal position of *CVCV* forms. Here, for example, Arana and Swadesh reconstruct $*r$ from M $y \sim$ T d (with inconsistency of C d versus t^y reflex; why not postulate r versus $ɾ$?). In brief, postulation of such further phonemes on the basis of these apparent irregularities in the first consonants of forms rapidly works in the direction of reducing our proto structure to the status of a phonological wastebasket choked with an abundance of phonemes of very restricted distribution. However, consonantal gradation may be posited as lying back of all these apparently irregular and haphazard sound correspondences. It then becomes unnecessary to posit a host of further phonemes of restricted distribution.[4]

1.5. ANALOGICAL CHANGE. A factor which often obscures regular sound correspondences between languages may be analogical change in one language. In such circumstances we suppose the phonological development to have been originally quite regular, but subsequent grammatical and/or lexical developments have obliterated the regular reflex (as "foots" may be substituted for "feet" in a child's language).

In Trique the fortis phonemes t and k are restricted to word-final syllables where they contrast with lenis d and g. In non-word-final syllables this contrast does not occur; rather there occur somewhat colorless alveolar and velar stops (varying to spirants) which we assign to the lenis phonemes. In that word stress typically falls on the word-final syllable, it seems possible that the fortis-lenis contrast in alveolar and velar stops was originally a phonetic distinction correlating with onset of stress or lack of onset of stress. However, some instances of d and g in word-final syllables do occur and the phonemic contrast is not to be doubted in modern Trique. Spanish loans such as *na²du³* 'soldier' are possibly too few to have played a very significant part in the lenis stops becoming phonemic. A few native roots display g in second syllable—perhaps in forms that hark back to Proto-Mixtecan $*x^w$ or $*^ng^w$. But it is striking that the clearest fortis-lenis contrasts in word-final syllables involve a morpheme *d-* (replacive of first consonant) and *g-* (likewise replacive) with the former indicating POSSESSED status on nouns and the latter indicating PUNCTILIAR aspect of verbs. Note, e.g. the following minimal pairs:

tą³ ni³ 'their blood'
dą³ ni³ 'their palm-trees'

(with *yą³* 'an unpossessed palm-tree')

ką³ 'squash'

gą³ 'she dug (it)'

(with *wą³* CONTINUATIVE ASPECT, 'she's digging it')

[4] Telltale traces of the old alternation survive in the form of etymological doublets in all three languages. Thus we find Cuicatec *nąʔąʔą* 'firewood' side by side with *yąʔąʔą* 'stick' as are Mixtec *vīšì* 'sweet', and *ⁿdùšī* 'honey'; Cuicatec *dà-kààčī* 'to boil' and *ʔyūūcī* 'to fry'; Cuicatec (*dāīyā*) *déénó* 'granddaughter' and (*dāīyā*) *hīínó* 'grandson'; Mixtec *ʔīnī* 'afternoon, early evening', and *šínī* 'supper'; Trique *ga³nąʔh* 'to weave' and *ži³nąʔ¹* 'loom'.

In this first pair, presumably Proto-Mixtecan *θom lies back of the Trique form POSSESSED PALM. However, the regular reflex of Proto-Mixtecan *θ in ultimate syllables is not *d*, but *t*. Nevertheless, since bisyllabic stems with initial *y-* alternate to *d-* in the possessed forms of such nouns, it seems plausible here to believe that (a) regular development of Proto-Mixtecan *θ to Trique *t* took place in an earlier stage of Trique; but (b) analogical extension of *d-* SIGN OF POSSESSED to monosyllabic nouns has obliterated this earlier *t* reflex. Similarly, in respect to the second pair, we assume that regular development of Proto-Mixtecan *k to Trique *k* first took place, but that this regular reflex was obliterated by later extension of *g-* morpheme for PUNCTILIAR aspect.[5]

1.6. LEXICAL BORROWING. Other apparent exceptions to regular sound development may be attributed to lexical borrowing at a time when certain sound developments had already taken place and were no longer operative. Thus, as we have stated, M *s/š* ~ T *t* < PMx *θ, but there is no regular correspondence M *s* ~ T *s*. Nevertheless, a set occurs exhibiting this apparently anomalous correspondence: M *sõ²õ* 'ear', *só²ó* 'deaf' ~ T *zo³²o³* 'deaf'. In that the Mixtec and Trique forms are so very similar (T has *z* as lenis phoneme in penultimas, and *z* versus fortis *s* only in ultimas) we suspect here a Mixtec loan word in Trique. This suspicion is reinforced by noting another set: M *ⁿdūčī* 'beans', 'eyes' *ⁿdúčí* 'blind' ~ regularly to T *ru³ne⁴³* 'beans', *ru³ne⁴ yą⁴³* 'eyes' (<PMx **ⁿdu³ⁿdi⁴*, set 16). But there is a Trique form *du³či³* 'blind' which is phonologically quite different from the inherited item *ru³ne⁴³* 'beans', 'eyes'. Again it seems best to assume that *du³či³* 'blind' is a Mixtec loan word in Trique. We note that both these items presumed to be Mixtec loan words refer to pathology. Knowing that the Mixtecs were culturally dominant over the Triques before and after the conquest, we might wonder if these two borrowings (and a few other items) do not reflect a cultural situation involving Mixtec shamans and Trique clients.[6]

1.7. WÖRTER-UND-SACHEN ANALYSIS OF THE RECONSTRUCTED CORPUS. In the above sections we have illustrated the nature of systemic comparison and reconstruction by use of Mixtecan data. We have noted that apparent exceptions to regular phonological developments do not lead us to abandon the assumption of regular sound change. On the contrary, by taking account of these exceptions we bring into better focus the reconstructed phonological system with possible discovery of some facts about the grammar as well. We may also uncover some lexical borrowings that can be recognized as such only by the application of systemic reconstruction. Nevertheless, analysis of such borrowings may have considerable ethnolinguistic value.

[5] This assumption is not wholly gratuitous. An old Proto-Mixtecan root meaning 'hard' and later 'metallic' is seen in Trique *zi⁵ka⁵²* 'hard' and *zi³-ka³⁴h* 'rattle (of rattlesnake)'. But in a stressless syllable following a syllable with phrase stress we find forms with *k* weakened to *g*: *du³kwa²-ga²a⁴³* 'jail (='house of iron')' and *žu³kwa²¹-ga¹²* 'rattlesnake'.

[6] In this respect it is interesting to note two other Trique lexical items which may belong to the same lexical domain: *la³kwą³* 'lame' from my set 142 and *la³kwe³h* 'mucus' from set 127. These items involve the only occurrences in our cognate sets of penultimate *l-* in Trique forms; penultimate *l-* is somewhat more common in Mixtec dialects. Furthermore, we can in set 127 match the Trique form *la³kwe³h* 'mucus' with Mixtec *lākʷā* 'pus'—although we have no Mixtec form with initial *l-* to match the Trique word *ʷla³kwą³* 'lame' (and, in fact, the Mixtec cognate *yakʷa* simply means 'crooked'). However, it is not impossible that these two Trique terms might also be Mixtec loans. If so, then Trique has borrowed them from Mixtec at a comparatively early period—earlier than the Trique shift of *(kʷ)a(h)* to *e*. Moreover, we have evidence that the Triques—who are rather conservative about linguistic borrowing—borrowed the Spanish word cojo 'lame' in the form *ga³čo²* some time in the past when they (or their Mixtec overlords) were in contact with a Spanish dialect in which *j* was pronounced as an alveopalatal sibilant (no such Spanish dialect exists in the region today).

One further possible ethnolinguistic by-product of systemic reconstruction consists in the application of a *Wörter-und-Sachen* analysis to the reconstructed corpus. Thieme (1958) has done some work of this sort in Indo-European studies. Millon and Longacre (1961) have made such a cultural analysis of the reconstructed Proto-Mixtecan and Proto-Amuzgo-Mixtecan vocabularies. In attempting to guard against some of the pitfalls inherent in a *Wörter-und-Sachen* approach, we sought for evidence of the existence of whole cultural complexes rather than merely isolated traits. Thus, on the Proto-Mixtecan horizon (some 3000 years ago?) there seems to be clear evidence for the following cultural complexes: maize complex, masa preparation complex, agricultural complex, maguey complex, and weaving complex.

2. PHONOLOGICAL SYSTEMS RECONSTRUCTED FOR BRANCHES OF OTOMANGUEAN. In this section we present in summary form phonological systems reconstructed for the six language families comprising what used to be called 'Macro-Otomanguean' but what we will refer to here simply as Otomanguean. These six families are Mixtecan, Popolocan, Chiapanec-Manguean, Otopamean, Zapotecan, and Chinantecan.

2.1. PROTO-MIXTECAN, used illustratively in the preceding sections, has the following phonological system:

Consonants

*t	*k	*k^w	*ʔ
*θ	*x	*x^w	
*nd	*^ng	*$^ng^w$	
*n		*m	
*y		*w	

Vowels

*i	*$\ddot{i}$	*u
*e	*a	*o

Tones

*1 high register (restricted to tone sandhi variants)
*2 mid-high register
*3 mid-low register
*4 low-register

Of the consonant phonemes only *ʔ and *m occur final in forms; *mʔ could occur final as well. The glottal stop *ʔ could occur in medial or initial cluster with possibly any consonant. Aside from these considerations the canonical forms are *$CVCV$ and *CV with no other consonant clusters, and no vowel clusters.

The semi-vowel *y apparently had a variety of allophones. A lateral allophone *[l] occurred before *i and *u but its occurrence is partly conditioned there by distribution of *ʔ. It is possible that a phoneme *l of limited distribution (as I posited in 1957) should be retained.[7] A further nasalized allophone of *y, i.e. *[$\tilde{y}$] occurred at least in the Proto-Mixtecan dialects that became Mixtec and Cuicatec; occurrence of the allophone was contingent on occurrence of a nasal consonant (often final *-m) somewhere in the form.

Of the last vertical column of consonants, only *k^w and *w occurred in penultimate syllables. Furthermore, *m is extremely rare in pre-vocalic (but frequent in post-vocalic) except in the cluster *ʔm. Of all the consonants the alveolars *θ, *nd, and *n are the least restricted in distribution.

Of the vowels, *e and *o do not occur in penultimate syllables nor before post-posed *-m.

In regard to the tones: One tone or two non-identical tones occurred on *CV forms, while two tones, identical or non-identical, occurred on *$CVCV$ forms. Basic tone pattern *$4(4)$ did not occur, while tone patterns involving tone *1 were restricted to sandhi variants. The occurring basic tone patterns and tone sandhi variants are here listed:

$$\left.\begin{matrix}*22\\ *24\end{matrix}\right\} \sim *21$$

[7] At any rate *l was apparently phonemic in Proto-Mixtec from which Trique possibly borrowed such l- initial forms as are mentioned in 1.6.

*33
*23 ~ *11
*34 ~ *14
*43 ~ *13
*42 ~ *41
*32 ~ *31 (in the dialect which became Mixtec)
*32 ~ *12 (in the dialects which became Cuicatec and Trique)

2.2. PROTO-POPOLOCAN, as reconstructed by Gudschinsky (1959c) has the following phonological system:

Consonants

*t	*t^y	*k	*k^w	*ʔ
*c	*č			
*s	*š	*h	*h^w	
*n	*ñ		*m	
*y	*l (?)		*w	

Vowels

*i *u
*e *o
*a

Tones

*1 high register (restricted to tone sandhi variants)
*2 mid-high register
*3 mid-low register
*4 low register

Clusters of two and three consonants occurred, but no vowel clusters. Since these clusters occurred initial in forms, and display some phonological variety, no sequence of nasal plus stop need be interpreted as a single prenasalized phoneme (as is, by contrast, necessary in Proto-Mixtecan). Gudschinsky lists (1959c, p. 21) *l as a Proto-Popolocan phoneme but this phoneme does not appear on any subsequent chart nor come in for comment in any of her expository paragraphs—although her statement to the effect that Proto-Popolocan has one liquid is probably a reference to *l.[8]

For Proto-Popolocan, unlike Proto-Mixtecan, no regular patterns of consonantal alternation may be posited in reconstructed nouns. Rather, various elements occur which Gudschinsky considers remnants of a former system very similar to that found in Proto-Mixtecan. The various preposed elements in both nouns and verbs are, on the whole, similar to preposed elements (consonantal alternates) in Proto-Mixtecan with the striking exception of Proto-Popolocan *t-, *tV-, and *c- which are alveolar elements of a sort not found in Proto-Mixtecan verb paradigms (although t- occurs in verb paradigms of Amuzgo).

Regarding the reconstructed Proto-Popolocan tones, Gudschinsky remarks, "*1 was almost certainly always a sandhi variant rather than a 'basic' tone," and "a disproportionately large number of items are reconstructed with *4 as one alternant, possibly because *4 was also a sandhi variant."[9]

Apparently the labiovelar and labial consonants *k^w, *h^w, *m and *w, did not occur before *o and *u. Among the vowels *o is apparently rare (reconstructed in three sets on the basis of the correspondence *u ~ o ~ u ~ o* and in three other sets on the uniform *o* reflex, Gudschinsky, 1959c, p. 38).

In monosyllabic reconstructed items Gudschinsky reconstructs forms on all four contrasting pitch levels. But while Proto-Popolocan items on pitch level *2 reconstruct solidly enough, such items do not

[8] In her set 248 Gudschinsky reconstructs *la as a further alternate along with *na, *nta, *ya 'tree', 'stick', 'wood'; the *la form is witnessed to by Chocho only. In set 354 she reconstructs *lihi, *nt^{v}ihi 'pasture', 'grass' with only the Huautla dialect of Mazatec witnessing to *lihi. In set 355 she reconstructs *laʔa 'dancer, singer' with Ixcatec and Huautla Mazatec witnessing to the initial *l.

[9] Gudschinsky cautions: 'A tentative reconstruction of PPn tone structure is postulated in spite of severe difficulties' (p. 38). These difficulties arise from the fact that only for two dialects of Mazatec have definitive tone analyses been published. Tones are not recorded on the Chocho forms and tone analyses of Ixcatec and Popoloca are 'in tentative phonemicization' (p. 38).

seem to have been involved in tone sandhi variation. Furthermore, pitch levels *1 and *4 are suspect of being only tone sandhi variants. I therefore venture to suggest that in Proto-Popolocan all monosyllables sort essentially into two broad tone classes: (1) those with invariable tone *2; and (2) those with basic tone *3 which varied to *4 and/or to *1. However, the latter broad tone class may possibly have had three subclasses: (a) those varying only to *4; (b) those varying only to *1; and (c) those varying to both *4 and *1 (Gudschinsky, 1959c, pp. 39–40).

In disyllabic reconstructed items Gudschinsky reconstructs five basic patterns (including *44) with sandhi variants (p. 40). Actually, I believe that the underlying picture is simpler than her summary might indicate. I therefore venture to restate Proto-Popolocan tone in disyllabics as follows: (1) Basic tones *42 and *32 are witnessed to by one set each with *42 ~ *44 ~ *14, and *32 without sandhi variation. (2) Aside from these rare reconstructed patterns there are four tone classes determined by types of possible sandhi variation to higher and lower pitch levels. Common to all these tone classes are *33 and *44:

(3) Class A: *33 ~ *34 ~ *44 ~ *13
(4) Class B: *33 ~ *34 ~ *44 ~ *31 ~ *11
(5) Class C: *33 ~ *43 ~ *44 ~ *13
(6) Class D: *33 ~ *43 ~ *44 ~ *31 ~ *41

Conspicuously lacking in the above are *21, *22, *23, and *24, all of which are possible in a four-level system (cf. 2.7. below).

2.3. Proto-Chiapanec-Manguean, as reconstructed by María Teresa Fernández de Miranda and R. Weitlaner on the basis of written records of these two extinct languages, has the following phonological system:

Consonants

*p		*t		*k		*ʔ
		*s		*h	*h^w	
*^{m}b		*^{n}d			*^{n}g	
*m		*n	*ñ(?)			*M
*w		*y				
		*l(?)				
		*r				

Vowels

*i	*ɨ	*u	
*e	*a		

Vowel Clusters

*ii		*uu
*ee	*aa	
*ai		*au

Proto-Chiapanec-Manguean *ñ and *l are reconstructed by Fernández de Miranda and Weitlaner with some hesitation. Reflexes of *ñ overlap with reflexes of *n; the former may be posited only on the strength of four sets. Reflexes of *l overlap with reflexes of *r in such a fashion so as to suggest that these two were the same phoneme at an earlier stage. Fernández de Miranda and Weitlaner also characterize the vowel clusters as being somewhat dubious. The consonants *t and *^{n}d had palatalized allophones *[ts] and *[nz] before *u, while *s is considered to have had an allophone *t before *u also.[10] The velar stop *k had a palatalized allophone before *i and *e. The phoneme *M is posited rather than consonant group *hm. Presence of phonemic pitch seems probable but the sources do not mark this feature well enough to permit even the most general sort of speculation. Nasalization is marked in some of the sources but was probably conditioned by contiguous nasal consonant.

2.4. Proto-Otopamean as such has not been reconstructed at the present time. Instead we have: (1) Proto-Otomí, first reconstructed by Newman and Weitlaner (1950) and recently revised by Doris Bartholomew. (2) Proto - Otomí - Mazahua ('Primitive Central Otomian') by Newman and Weitlaner (1950); (3) Proto-Otomí-Mazahua tone tentatively reconstructed in an unpublished paper of Stewart (1957);

[10] The seeming overlap between phonemes *t and *s is not serious in that *[tV] (not *u) contrasted with *[sV] (not *u) while *[tsu] contrasted with *[tu].

and (4) Proto-Otomí-Pame (not taking into account Mazahua) by Bartholomew.

2.4.1. PROTO-OTOMIAN. Bartholomew's revision of Proto-Otomí consonants[11] gives us the following system; postulated phonetic qualities of consonants are indicated in brackets in that phonetic features cannot be readily inferred from the phonemic symbols:

*p [hp]	*t [ht]	*c [hts]
*b [p]	*d [t]	*z [ts]
*m [^{m}b]	*n [^{n}d]	*y

Proto-Otomí vowels are reconstructed as follows according to Newman and Weitlaner:

*i	*ə	*u	*į	*ų
*e	*a	*o	*ę	*ǫ
*ɛ	*a	*ɔ	*ą	

2.4.2. PROTO - OTOMÍ - MAZAHUA. Mazahua and Otomí are closely related. It may therefore be expected that most or all of the revisions of the sort that Bartholomew made in the Newman-Weitlaner inventory of Proto-Otomí consonants might hold as well for Otomí-Mazahua. We omit here the Newman-Weitlaner line-up of Otomí-Mazahua consonants pending re-examination by Bartholomew or someone else. The Proto-Otomí-Mazahua vowel inventory is identical with that reconstructed for Proto-Otomí.

Newman and Weitlaner also reconstruct a considerable number of diphthongs for Proto-Otomí-Mazahua. Some of these diphthongs they reconstruct simply as a device to handle certain Otomí-Mazahuan vowel correspondences. In respect to these 'diphthongal formulas' they comment: 'In view of the progressive assimilatory influences that have operated upon Mazahua it is likely that this type of diphthong formula represents an assimilatory factor, giving a high or front vowel reflex in PMz, rather than a genuine diphthong in the parent language. But, because the conditioning factor cannot be determined from the data ... we must symbolize these special Otomí-Mazahua correspondences as distinctive vowel combinations at the PCO level.' (p. 78)[12]

*k [hk]	*k^{w} [hk^{w}]	*ʔ
*g [k]	*w [k^{w}]	
*s		*h

For another sort of diphthong more historical reality is claimed by Newman and Weitlaner. This latter sort of diphthong is posited on the basis of such correspondences as POt $V_1V_2 \sim$ PMz V_1mV_2, e.g. *ęi is posited as the basis of POt *ęi ~ PMz *ęme; *ɔi, on the basis of POt *ɔi ~ PMz *ɔmə; and *ųi, on the basis of POt *ųi ~ PMz *umə. Similar are cases with *h interposed between the two members of the posited cluster: *ąhi is reconstructed from POt *ąhi ~ PMz *ahma. Commenting on these cases Newman and Weitlaner write: 'Since most of these diphthongs have a nasal vowel as the first element in PCO and POt, the PMz *m in such cases is obviously a reflex of the vowel nasalization' (1950, p. 77). I suggest here an alternative hypothesis, viz. that the Mazahua -m- is not an innovation but an archaism, and that the nasal vowel quality in corresponding Otomí forms is a vestige of the older post-posed nasal. It seems quite plausible to take the

[11] Consonant clusters are of four sorts (Bartholomew, p. 327); (a) *p, *t, *c, *k, *k^{w}, *m, *n, *y, or *w in cluster with *h, which follows the stops, but precedes the other consonants. (b) The same consonants listed above in cluster with *ʔ, which follows the stops, but precedes the other consonants. (c) *m or *n preceding any other consonant (but not in cluster with each other); *m preceded *p and *b, while *n preceded all others. (d) Clusters of three consonants involving combinations of (a) or (b) with (c), such as *mph, *nhy, *nkʔ, *nʔn and an additional two clusters *ʔtʔ and *ʔcʔ. Bartholomew discusses some systematic features in the patterning of these clusters in reference to the concept of phonetic rank (Eunice Pike, 1954) and simulfixation in Otomí (Wallis, 1956).

[12] Newman and Weitlaner abbreviate Proto-Mazahua as PMz, and Proto-Otomí-Mazahua—which they term 'Primitive Central Otomian'—as PCO.

Mazahua witness here as basic and explain the Otomí form by loss of **m* with resultant development of diphthong. Furthermore, such a suggestion is in line with the role of **m* in other branches of Otomanguean (cf. 2.10.). If my explanation of the Otomí-Mazahuan reflexes is correct we can probably assume (a) no vowel clusters in Proto-Otomí; and (b) Proto-Otomí-Mazahua postposed **-mV* and **-hmV* elements similar to certain postposed deictics suggested by Gudschinsky for Proto-Popolocan-Mixtecan.

Proto-Otomí-Mazahua tones as reconstructed by Stewart on the basis of three Otomí dialects and one Mazahua dialect include high, low, rising, and falling tones (Appendix B).

2.4.3. PROTO - OTOMÍ - PAME. Bartholomew's unpublished reconstruction of Proto-Otomí-Pame presumably has to do with an earlier layering than does the reconstruction of Otomí-Mazahua. As such, Otomí-Pame probably is more indicative of the structure of Otopamean as a whole—although definitive statements regarding the latter must await the common reconstruction of Otomí-Mazahua, Matlatzinca-Ocuilteco, Chichimeca, and Pame.

The inventory of consonant phonemes for Otomí-Pame is a small one (noticeably absent are semi-vowels).

**p*	**t*	**c*	**k*	**ʔ*
**m*	**n*	**s*		**h*

But these nine consonants occur as geminate clusters and in clusters with **h*, **m*, and **ʔ*:[13]

**pp*	**tt*	**cc*	**kk*	**ʔʔ*
**mm*	**nn*	**ss*		**hh*

[13] Pame has other consonant clusters not indicated here; some such clusters occur word-final as well. Some of these clusters correspond to *CV* elements in Otomí. Thus Otomí *-di* or *-de* ~ Pame *-dn* in several sets while Otomí *gi* ~ Pame *-gy* in two sets. Bartholomew gives no source for Pame *d* or *g*; I conjecture that these correspondences may hark back to Proto-Otomí-Pame **-tVn* and **-kVn* elements.

**ph*	**th*	**ch*	**kh*
**mh*	**nh*	**lh*[14]	
	**tʔ*	**cʔ*	
**ʔm*	**ʔn*	**lʔ*	
**mʔm*	**nʔn*		

Proto-Otomí-Pame vowels are likewise fewer in number than Proto-Otomí-Mazahua vowels: **i*, **e*, **ɛ*, **a*, **o* with the corresponding nasalized phonemes. Bartholomew comments: 'A number of cognate sets indicate proto vowel clusters. These clusters are the probable source of the Otomí extra vowels. Though the data enable us to make this basic assumption, the actual correspondences do not permit us to trace the precise development from the proto clusters to Otomí' (Bartholomew, 1959, p. 30).

Pame has three phonemic tones which occur on stressed syllables: high, low, falling. It is evident here that we have a tone system which invites comparison with the four tones reconstructed by Stewart for Otomí-Mazahua.

2.5. PROTO-ZAPOTECAN as reconstructed by Swadesh (1947) and currently under modification by Fernández de Miranda has the following inventory of consonants:

**p*	**t*		**k*	**kʷ*	**ʔ*
	**c*	**č*			
	**s*	**š*			
**m*	**n*				
	**l*				
	**r*				
**w*	**y*				

These consonants may appear geminated with the exception of **kʷ*, **ʔ*, **r*, **w*, and **y*.

Vowels, as reconstructed by Swadesh

[14] Notice that a further consonant **l* is reconstructed as occurring in clusters **lh* and **lʔ*. In that **sh* and **sʔ* do not occur, **l* and **s* are in complementary distribution. Underlying the **lh* reconstruction is Otomí *th* ~ Pame *lh*; and underlying the **lʔ* reconstruction is Otomí *tʔ* ~ Pame *lʔ*. The possibility of considering **s* and **l* to be but one Otomí-Pame phoneme should be considered.

are:

*i		*u
*e	*a	*o

The pairs of contrasting phonemic entities reconstructed here as single consonants versus geminates could alternatively be reconstructed as lenis versus fortis consonants. But Swadesh's arguments in favor of the former reconstruction appear on the whole to be valid. He argues that there is some evidence that other clusters existed of a non-geminate variety in Proto-Zapotecan. In addition, some present-day dialects have phonetically geminated versus phonetically single consonants. In the materials of Fernández de Miranda there is considerable phoneme substitution between a reconstructed single consonant and its corresponding geminate (for a possible historical explanation, see §2.10).

Swadesh does not reconstruct *m and *mm for Proto-Zapotecan. In view of the fact that apparent reflexes of bilabial nasal are very rare, he states: 'Since we have no evidence that necessarily points to an *m in Proto-Zapotec, we have to assume a secondary development of mm from *np in cases like 40 (*cunpi), even though mm or m is found in all the dialects.' (p. 223). Fernández de Miranda reconstructs *mm and *m as rare Proto-Zapotecan phonemes.

2.6. Proto - Chinantecan. The preliminary reconstruction of Proto-Chinantecan made by Smith and Weitlaner (1957) exhibits the following inventory:

Consonants

*p	*t	*k	*ʔ
	*c		
*b	*d	*g	
	*s		*h
	*l		
	*r		
*m	*n	*ŋ	

Vowels

*i	*ï	*u
*e	*a	*o
	*ą	

(with all corresponding nasalized vowels as well).

Noteworthy in the above is the reconstruction of *ŋ, in that a velar nasal is reconstructed for no other branch of Otomanguean.

Consonant clusters with *ʔ and *h are posited: *hg, *ʔg, *hl, *ʔl, *hm, *ʔm, *hn, *ʔn, *hŋ, *ʔŋ. The restriction of these clusters to those with *g, *l, or a nasal consonant as a second member is a curious feature here; it leads one to suspect that Proto-Chinantec *g in such clusters (where no other stop occurs) is a reflex of some Proto-Otomanguean semi-vowel such as *y or *w.

A quantity of vowel clusters is posited. In Proto-Chinantec, the absence of semi-vowels together with the presence of many vowel clusters—the bulk of which involve *i or *u (whether nasal or oral)—make it probable that Proto-Chinantec vowel clusters arose from *VyV and *VwV sequences in Otomanguean.

2.7. Proto - Otomanguean. A first approximation to Proto-Otomanguean has been made by Gudschinsky (1959c) in her reconstruction of 'Proto-Popotecan,' i.e. Proto - Popolocan - Mixtecan. She reconstructs the following phonemes[15] for Proto-Popolocan-Mixtecan (abbreviated PPMx):

Consonants

*t	*tʸ	*k	*kʷ	*ʔ
*θ		*x	*xʷ	

[15] I have modified Gudschinsky's reconstructed system slightly here by subtracting *ɔ—which was reconstructed for Proto-Popolocan-Mixtecan largely on the supposed strength of my reconstruction of such a vowel in Proto-Mixtecan. Proto-Popolocan-Mixtecan *tʸ is rare; Gudschinsky, in fact, reconstructs it from but three sets in which PPn *tʸ ~ PMx *t (her sets 34, 36, and 42). She mentions in a footnote (p. 44) a number of residue sets in which PPn *č or *tʸ corresponds to PMx *θ. It is possible that both Proto-Popolocan-Mixtecan *tʸ and *θʸ should be reconstructed. Gudschinsky is uncertain about the status of *ny as a phoneme or cluster in Proto-Popolocan-Mixtecan.

*n	*n^{y}		*m
*l	*y		*w

Vowels

*i	*ï	*u
*e		*o
*ë	*a	

Some of the more noteworthy phonological developments from Proto-Popolocan-Mixtecan to Proto-Popolocan are (1) PPMx *yk or *ky > PPn *t^{y} (before PPMx *ë), thus augmenting distribution of this presumably rare phoneme; while before other reconstructed vowels the PPn reflex was *č. (2) A three-way split of PPMx *θ into PPn *s, *š, and *c. It seems to me to be plausible that *s and *š were freely varying allophones in early Proto-Popolocan, but that contrast arose in late Proto-Popolocan when considerable dialect difference had already arisen. The third Proto-Popolocan reflex of Proto-Popolocan-Mixtecan *θ, viz. Proto-Popolocan *c, occurred in cluster with *h or *n—but these latter two elements, especially *n, have been in many instances lost in all Popolocan languages except Mazatec. Gudschinsky posited *č as another Proto-Popolocan reflex of Proto-Popolocan-Mixtecan *θ but possibly Proto-Popolocan-Mixtecan *θ^{y} might be reconstructed here instead. (3) Proto-Popolocan-Mixtecan *mn cluster was reduced to Proto-Popolocan *m by loss of second member (unpublished Popolocan-Mixtecan sets).

Some noteworthy phonological developments from Proto-Popolocan-Mixtecan to Proto-Mixtecan are: (1) Unconditional merger of Proto-Popolocan-Mixtecan *t and *t^{y} into Proto-Mixtecan *t. (2) Development of a Proto-Mixtecan prenasalized series *^{n}d, *^{n}g, *^{n}g^{w} from fusion of stop and/or spirant with preposed nasal. This development came about by reduction of other Proto-Popolocan-Mixtecan clusters in initial position (exclusive of *$^{?}$C), so that the three surviving clusters were reinterpreted as unit phonemes. (3) PPMx *mw > PMx *m.

In tone systems Proto-Popolocan and Proto-Mixtecan are very similar. Both are four-level systems with *1 restricted to sandhi variants. Either tone pattern *44 did not occur at all in Proto-Popolocan-Mixtecan (it is clearly witnessed to in Proto-Popolocan but absent in Proto-Mixtecan), or it was possibly restricted to tone sandhi variants. A portion of Gudschinsky's summary of this and related considerations follows:[16] 'For PPtn, I postulate the occurrence of all the possible sequences, with loss of *44 in PMx, and the augmenting of the frequency of *44 in PPn by special sandhi development. The sequences with *2 on the initial syllable may have been lost in PPn, but it is more likely that they are hidden by the inadequate analyses of the living languages. This hypothesis is supported by the fact that PPn has a much larger inventory of alternations and variants than PMx, and it seems probable that some of these should rather be reconstructed as basic sets with *2. For example, PMx has *43 ~ *13 and *42 ~ *41. PPn has *43 ~ *13 in some sets, and *43 ~ *41 in others. The obvious conclusion is that *43 ~ *41 reflects PPtn *42 ~ *41. Unfortunately it has not been possible to match the majority of the PMx reconstructions in this fashion, and exact correspondences are too few for any definite conclusions' (Gudschinsky 48-9).[17] It also

[16] Gudschinsky labels what I here term Proto-Popolocan-Mixtecan as 'Proto-Popotecan' and abbreviates to PPtn.

[17] But just as Gudschinsky conjectures that one class of Proto-Popolocan *43 stems could actually reflect Proto-Popolocan-Mixtecan *42, it could similarly be suggested that one class of Proto-Popolocan *34 could be Proto-Popolocan-Mixtecan *24 (PPn *34 ~ *13 versus PPn *34 ~ *31; see Section 2.2.). Similarly there are classes of Proto-Popolocan *33 stems which may possibly mask *22 versus *23. Gudschinsky reconstructs Proto-Popolocan *32 in one set with no witness to sandhi variant. Proto Mixtecan *32 is peculiar as well, in that it has two sandhi variants: *31 in the late Proto-Mixtecan dialect that became Mixtec, and

seems plausible, by way of more general observation, that Proto - Popolocan - Mix - tecan had sandhi variation both to higher and lower tone patterns. The latter sort of sandhi variation was dropped in the transition to Proto-Mixtecan (with concomitant disappearance of pattern *44 in Proto-Mixtecan as well).

Gudschinsky reconstructs for Proto-Popolocan-Mixtecan two noun declensions very similar to those I posit for Proto-Mixtecan:

*t declension: *t, *y, *n, *nt, *x
*θ declension: *θ, *y, *ny, *nθ

These noun declensions with some modifications survive as declensions in Proto-Mixtecan, but only as fragments in Proto-Popolocan. Possibly the three-way split of Proto-Popolocan-Mixtecan *θ into late Proto-Popolocan *s, *š, and *c (not to mention *č)—which shattered the *θ declension into bits—had something to do with the disappearance of neatly delineated noun declensions in Proto-Popolocan. On the Mixtecan side, where *θ does not undergo split and merger with other phonemes, the declensions were retained.

Items preposed to verbs are more varied on the Proto-Popolocan-Mixtecan horizon than on the Proto-Mixtecan horizon. Even on the Proto-Popolocan horizon we do not find anything as neat and regular as the four tense-aspects of Proto-Mixtecan. Noticeably, *t and *θ figure in Proto-Popolocan-Mixtecan and Proto-Popolocan verb forms, but not in Proto-Mixtecan verb forms (Gudscinsky p. 58).

Gudschinsky reconstructs postposed deictics for Proto-Popolocan-Mixtecan. An outer layer deictic has survived only as Proto-Mazatec *-V^3/*-V^4 marking end of noun phrase and slightly deictic in function, and as Proto-Trique *-a^3 end of noun phrase. But she postulates that this deictic particle is a Proto-Popolocan-Mixtecan development antedated by an older deictic layer which was already losing deictic significance and fusing with the Proto-Popolocan-Mixtecan stems. These 'inner layer' deictics she reconstructs as: *xmV, *$ʔmV$, *xV, and *$ʔV$. It seems likely that these reduce essentially to two forms plus or minus *m, i.e. either to *$xV(m)$ and *$ʔV(m)$ or to *$x(m)V$ and *$ʔ(m)V$.[18]

As a first approximation to Proto-Otomanguean, Gudschinsky's 'Proto-Popotecan' may prove to be (a) actually on the Proto-Otomanguean horizon of reconstruction, as a reconstruction based upon but two language families and needing to be amplified by taking account of further language families; or (b) slightly more recent than the Proto-Otomanguean horizon of reconstruction, in that these two families may prove to constitute a group more closely related to each other than to other Otomanguean languages.

2.8. Proto-Popo-Manguean. A second approximation to Proto-Otomanguean has been made by Fernández de Miranda and Weitlaner in their reconstruction of 'Proto-Popo-Mangue' on the basis of Popolocan, Mixtecan, and Chiapanec-Manguean (1961).

The following phonological system is

*12 in the late Proto-Mixtecan dialects that became Cuicatec and Trique. It is therefore possible that Proto-Popolocan-Mixtecan *32 was rare and had no sandhi variant, and that such variants developed only in later Proto-Mixtecan dialects.

[18] Furthermore it seems plausible to me that it will also eventually be necessary to reconstruct either *-m or *-mV (possibly *-Vm) in sets that possibly show influence of a bilabial nasal but no evidence of *x or *$ʔ$. Thus, *$ɔ$ on the Proto-Mixtecan horizon has been eliminated in favor of *am. But Proto-Popolocan a versus u reflexes of my formerly postulated *$ɔ$ are similar to a versus u reflexes of Proto-Popolocan-Mixtecan *am versus *am. Furthermore, Gudschinsky has expressly formulated that '*a has reflex *u before *m or *w in some *dialects* of PPn' (p. 46). If we eliminate *$ɔ$ on both Proto-Mixtecan and Proto-Popolocan-Mixtecan horizons, then it seems simplest to attribute Proto-Popolocan reflex *u of reputed *$ɔ$ to *a plus postposed bilabial nasal which disappeared early and without trace (except vowel nasalization?) in Proto-Popolocan. Gudschinsky attributes reflex *u of reputed *$ɔ$ to a preceding nasal.

reconstructed for Proto-Popo-Manguean:

*t	*t^y	*k	*k^w	*ʔ
*c		*h	*h^w	
*n			*m	
*l	*y		*w	

Vowels

*i	*ï		*u
*e			*o
*ë	*a		

The reconstructed consonants comprise a system which resembles that projected for Proto-Popolocan-Mixtecan in respect to absence of *p and presence of series *c, *h, and *h^w which resembles the Proto-Popolocan-Mixtecan series of spirants (*θ, *x, *x^w). However, it is within the realm of possibility that Proto-Popo-Manguean *p could be reconstructed rather than *k^w. The latter is attested to in but one set where on the Mixtecan side presence of consonantal alternation (PMx *w/*^{n}g^w) further weakens the witness to the postulated *k^w. Furthermore, on cursory examination it appears to me to be possible that some of the Fernández-Weitlaner reconstructions with Chiapanec-Manguean *p may prove to be cognate with sets with Proto-Mixtecan *k^w.

Fernández de Miranda and Weitlaner reconstruct consonant groups *kh and *kh^w in contrast with *k and *k^w on the one hand and with *h and *h^w on the other hand. Leaving aside the more dubious reconstruction *k^w just discussed, the pertinent sound correspondences are:

(1)	PChM *k	~	PPMx *k	<	PPM *k
(2)	PChM *h	~	PPMx *x	<	PPM *h
(3)	PChM *h^w	~	PPMx *x^w	<	PPM *h^w
(4)	PChM *p/^{m}b	~	PPMx *x^w	<	PPM *h^w(ï)
(5)	PChM *h	~	PPMx *k		
(6)	PChM *h^w	~	PPn *k/zero	~	PMx *k^w

For correspondences (5) and (6) Fernández de Miranda and Weitlaner reconstruct *kh and *kh^w respectively. This is a plausible reconstruction in the light of Gudschinsky's Popolocan-Mixtecan reconstructions which indicate a considerable number of consonant clusters. However, a possibility suggested by Swadesh (1960a, p. 96) should be investigated here, viz. the possible presence of front versus back velars. For the moment the presumption is against this suggestion in view of the fact that clear evidence has not yet been presented to substantiate presence of this sort of phonemic contrast in Otomanguean. Nevertheless, the presence of two *k*-positions in Mayan and Totonacan should alert us to the fact that contrasts of this sort are found in Middle America as a possible areal feature of the sort sometimes found to transcend genetic groupings.

Besides the clusters *kh and *kh^w just discussed, Fernández de Miranda and Weitlaner posit *nt, *nt^y, *nk, and *hm. It seems to me that their data also suggest *mw, *mn, *mp, and/or *nk^w. Furthermore, there is some evidence that the forms with a preposed nasal alternated grammatically with forms without a preposed nasal. Perhaps Proto-Popo-Manguean *hm harks back to an earlier *mh. The following Pre-Popo-Manguean scheme might be suggested:

*t/*nt
*t^y/*nt^y
*k/*nk
*c/*nc
*n/*mn
*w/*mw

To this we could probably add *y/*ny (as underlying PChM *ñ) and either *p/*mp or *k^w/*nk^w (or both?). To this general pattern of consonant alternating with nasal plus consonant, might be added a more specialized alternation *t^y/*y of which there appear to be traces.

Certain preposed elements in Proto-Popo-Manguean verbs are similar to those posited for Proto-Popolocan-Mixtecan verbs (especially if PChM *p ~ PMx *k^w). Preposed *ti-/*tu- for Proto-Chiapanec-Manguean past tense and *ta- for future tense have similar *tV- elements in Proto-Popolocan

and in Amuzgo, but no parallels in Proto-Mixtecan. For discussion of some postposed Proto-Popo-Manguean elements see Appendix C.

2.9. MACRO-MIXTECAN. Swadesh has recently made several suggestions concerning what I continue here to call Otomanguean phonology in spite of his attempt to expunge Chiapanec and Manguean from this group and to rename the whole Macro-Mixtecan. One suggestion of his—that possibly Mixtec is the most typical language of this family—must await evaluation until we have a better picture of Otomanguean structure. However, two suggestions of Swadesh seem to me to be worthy of special mention: (a) Both **p* and **kʷ* may need to be reconstructed for Proto-Otomanguean with these both falling together as **kʷ* in Proto-Mixtecan and Proto-Popolocan. Note the possibility that these both fell together as **p* in Proto-Chiapanec-Manguean. (b) It may eventually prove necessary to reconstruct front versus back velar positions.

2.10. PROTO-OTOMANGUEAN **m*. One of the most interesting features of Otomanguean structure may well prove to be the consonant **m*.[19] While possibly somewhat rare in stems, this phoneme was apparently quite frequent as a preposed and postposed element (possibly with an associated vowel). (1) In Proto-Mixtecan stems, **m* is rare but **ʔm/*m* is fairly frequent. It is possible that ultima **-ʔmV* syllables in Proto-Mixtecan may hark back to a Proto-Otomanguean postposed deictic **-ʔmV* which eventually became part of the stem (cf. various **-ʔV* and **-hV* elements in Proto-Popolocan). (2) In the Popolocan languages, incidence of *m* in stems has been augmented by the development of older **mn* cluster (from preposed **m-* plus initial **n-*) to **m* in Proto-Popolocan. Similarly, some cases of Proto-Mixtecan **m* in stems hark back to older cluster **mw* (from preposed **m-* plus initial **w-*). (3) Proto-Mixtecan and Proto - Chiapanec - Manguean prenasalized stops originated from such clusters as **mt*, **mθ*, **mtʸ*, **mk*, **mx*, **mkʷ*, and **mxʷ* which fell together in various ways with concomitant reduction of other **-mC-* or **mC-* clusters. (4) As postposed **-mV* and **-ʔmV* elements that fused into the stem (with loss of **m* in Otomí leading to development of vowel clusters), the phoneme **m* played an important role in Otomí-Mazahua phonology and possibly in general Otopamean phonology as well. (5) In the development of Proto-Zapotecan from Proto-Otomanguean it is possible that **m*—which all but disappeared in that branch—played a crucial role. I offer here somewhat diffidently a hypothesis which is outlined below.

Many Proto-Otomanguean **mC* clusters became Proto-Zapotecan **CC* (geminate) by assimilation of the **m* to the following consonant. Thus, the Proto-Zapotecan geminate clusters developed from what seems to be a prevailing type of Proto-Otomanguean non-geminate cluster, viz. **mC*. The fact that the 'fortis and lenis' (i.e. geminate versus simple) are often somewhat confused in Zapotecan reflects Proto-Otomanguean **mC* versus **C*, pursuant to a frequent pattern of consonantal alternation testified to in such an early layer as Proto-Popo-Mangue (cf. under 2.8). Thus, such forms as **CVCV* and **CVmCV*, as well as **CV* and **mCV* existed side by side for many Proto-Otomanguean roots; haphazard survival of forms with or without **m-* accounts for a certain inconsistency in 'fortis versus lenis' (i.e. geminated consonant versus single consonant) reflexes in Proto-Zapotecan. The fact that **m* is very scarce in Proto-Zapotecan is due precisely to this wholesale assimilation of that consonant.

A detail of Isthmus Zapotec morphophonemics is possibly explained by such a hypothesis, viz. *Ci-* or *Ca-* prefix before a lenis consonant alternates to *Cu-* prefix

[19] Cf. Appendix D.

before a fortis consonant in causative paradigms. We may assume here that the *Ci-* and *Ca-* basic forms of the prefixes hark back to Proto-Otomanguean **CV-* syllables not followed by **m-* of the succeeding syllable, while the *Cu-* forms of these same prefixes hark back to Proto-Otomanguean **CV-* syllables followed by **m-* in the succeeding syllable. The bilabial **m* presumably leaves a trace in labializing *i* and *a* to *u*.

Note the following Isthmus Zapotec data:[20]

giǰu?uni	'It's going to be fried'
guču?uni	'Fry it'
rizi?ɬdé	'I learn'
rusi?ɬdé	'I teach'
nizɬ?idikabe	'they did not learn'
nusɬ?idu	'you do not teach'
riǰi?icibě	'she gets angry'
ruči?ime (na?a)	'he makes (me) angry'
nižale?	'this did not come open'
nušalétú	'you (pl.) did not open it'
zado?oni	'it will sell'
zuto?olu	'you will sell (something)'

In the last two examples we suggest the following development: Proto - Otomanguean **θat-* > PZ **sat* > IZ *zad-*; Proto-Otomanguean **θamt* > PZ **sutt* > IZ *zut.*

We hasten to add that it is by no means true, either internally in Isthmus Zapotec or in Proto-Zapotec, that instances of fortis consonants are regularly preceded by vowel *u*. Nevertheless, we have cited a tendency in this direction. Our observation offers an explanation of occurrences of *i* and *a* alternating with *u* where observed; it does not attempt to explain parallel cases where such a development might be expected but does not occur. For further Zapotec data of possible relevance here see appendix D.

3. PHONOLOGICAL SYSTEMS RECONSTRUCTED FOR OTHER MIDDLE AMERICAN LANGUAGE FAMILIES. With various systems reconstructed for branches of Otomanguean and with Proto-Otomanguean itself beginning to take shape, I now turn to the summary of reconstruction in other Middle American linguistic stocks. There first follows a sketch of Proto-Uto-Aztecan of the Aztec-Tanoan phylum; then follow Proto-Zoquean, Proto-Mayan, and Proto-Totonacan, which are usually considered to be related to each other and have sometimes been termed 'Mexican Penutian'. We do not have materials available for Proto-Hokaltecan (or for Hokaltecan plus Jicaque and Yurumangui).[21]

3.1. PROTO - UTO - AZTECAN. In a pithy little nine-page article written in 1935 Whorf gave us an excellent sketch of Proto-Uto-Aztecan phonology with some attention to structural features of the reconstructed language as well. In Whorf's words: 'The comparative linguistics of the Uto-Aztecan stock is now entering upon its second stage.... This first stage may be described as finding out what we had to deal with. One of its leading ideas was classification of the stock into sub-groups and attempts to grade degrees of linguistic kinship.... If the key word of the first stage was classification, that of the second stage has become structure. In the first stage we compared words willy-nilly, looking for resemblances to appear. On the basis of much rough and quick comparison we attempted to arrange our hodge-podge of facts into something like order. Eventually a feeling for the basic structure of Uto-Aztecan speech emerged; these languages had a characteristic structure of word and stem, as Semitic and Bantu have' (p. 600). The above quotation describes well the genesis and growth of a comparative reconstruction project, which begins with desultory comparisons and ends with systemic comparisons, reconstruction of proto-forms, and some awareness of the structure of the reconstructed layer.

[20] From a problem of Velma Pickett's included in Nida's Morphology (1949, p. 262 ff.).

[21] Cf. Greenberg and Swadesh, 1954; Bright, 1956.

Whorf gives the following inventory of Proto-Uto-Aztecan phonemes:

Consonants

*p	*t	*c	*k̲	*k	*kʷ	*ʔ
*m	*n	*ⁿs	*[ñ]	*ŋ	*ŋʷ	
*v	*r	*s				*h
*w	*l	*y				

Vowels

*i			*u	
*e			*o	
	*a	*a·		

In setting up the Proto-Uto-Aztecan consonants Whorf postulated *p/*v, i.e. *p as a stop phoneme and *v as a spirant overlapping distributionally with the stop to such an extent that it is very difficult to separate with certainty reflexes of the one versus reflexes of the other. Whorf similarly postulated a relationship of *m/*ŋʷ ('ŋʷ, related in some similar but not quite the same way to *m*' p. 606). Regarding *ñ Whorf commented 'Possibly *ñ* should be added to the list' (p. 606).

Several morphophonemes are reconstructed which can be better summarized by first sketching briefly Whorf's picture of Proto-Uto-Aztecan structure. Stems are primarily of *CVCV* structure with a small minority of *CV* stems. *CVCV* stems occurred in patterns of derivation and inflection that involved suffixes, a limited number of prefixes, 'changes in and apocopation of the second stem vowel', and reduplication. In addition to these grammatical devices, Proto-Uto-Aztecan was also characterized by the alternations *p/*v and *m/*ŋʷ.

Whorf appears to have felt that alternation between *p and *v in Proto-Uto-Aztecan stems involved interchange of full phonemes harking back to an earlier period when *p had a stop allophone *[v] occurring in intervocalic position. But apparently Proto-Uto-Aztecan stops in general had spirant allophones in intervocalic position; in many of the daughter languages the conditioning factor or factors disappeared with the result that the stop/spirant alternation became phonemic. There are, however, certain Uto-Aztecan stems with medial consonant which appear to be resistant to spirantization. Whorf shrewdly reasoned that such consonants were not originally intervocalic, but that another consonant occurred in cluster before the consonants that proved resistant to spirantization—so that no spirant allophone occurred in these Proto-Uto-Aztecan stems. Whorf further noticed that there was a nasalizing influence associated with some instances of occurrence of consonants resistant to spirantization. Putting all these and other data together, Whorf formulated the hypothesis that the medial clusters in which spirant allophones did not occur consisted of stop preceded by nasal consonant or by a limited class of other consonants 'perhaps *l, *r, possibly no more.' (p. 606). His 'spirant-proof unnasalized set' *ˣp, *ˣt, *ˣc, *ˣk̲, *ˣk, and *ˣkʷ—presumably are clusters consisting of stop preceded by *l or *r; while his nasalized set *ᵐp, *ⁿt, *ⁿc, *ᵑk̲, *ᵑk, *ᵑkʷ are to be interpreted simply as nasal plus stop clusters.

Furthermore, Whorf seems to indicate that *l, *r, and nasals could also appear as final consonants in a *CVCVC* sub-type of the broad *CVCV* pattern. In short, syllable patterns *CV* and *CVC* occurred; when *CVC* occurred the final *C* was *l, *r, or a nasal. Combinations of these two syllable patterns give **CVCV*, **CVCVC*, and **CVCCVC*.

In regard to the relative chronology of the above, Whorf is somewhat vague. He seemed to hint that these clusters existed in Pre-Proto-Uto-Aztecan, in that he reconstructed Proto-Uto-Aztecan morphophonemes with antispirantizing and nasalizing tendency rather than the consonant clusters as such—which he may have wanted to relegate to an earlier stage. But this hardly seems possible. Either the clusters still

existed in Proto-Uto-Aztecan, and spirant allophones were in complementary distribution with stops; or the clusters had already disappeared from Proto-Uto-Aztecan, and the spirant allophones were already separate phonemes. There scarcely seems to be room for an intermediate stage when 'anti-spirantizing and nasalizing influences' existed, while the clusters had nevertheless disappeared and the spirant allophones were not yet phonemic.

Whorf seems to feel that vowel length characterized Proto-Uto-Aztecan, but reconstructs only *a versus *a· with the comment 'Varied reflexes of other vowels may also be a function of length but this has not yet been worked out' (p. 606). Here we probably would do well to trust Whorf's appraisal of the situation. Nevertheless, as his reconstruction of vowels now stands, it could be restated as follows: Converting Whorf's *e to *ï (by far the commonest reflex is ï), and interpreting his *a as *ä and his *a· simply as *a we obtain:

*i	*ï	*u
*ä	*a	*o

The only regrettable feature of Whorf's study is its brevity. We could wish for a catalogue of cognate sets appended to this most useful and provocative article, but this apparently was beyond the scope of Whorf's intentions.

3.2. PROTO-ZOQUEAN. Wonderly has reconstructed Proto-Zoquean by the comparison of Zoque (4 dialects), Tapachulteca, Sierra Popoluca, and Mixe (1949). In regard to Tapachulteca, Wonderly comments: 'Materials on the extinct or nearly extinct Tapachulteca are scanty and not too reliable, but are included where possible' (p. 2). He postulates the following reconstructed system:

Consonants[22]

*p	*t	*c	*k	*k^w	*ʔ
	*s				*h
*m	*n		*ŋ		
*w	*y				*W

Vowels

*i	*a	*u
*e		*o
*ä		*a

Of the above phonemes, *k^w is set up on the basis of the correspondence Zoque and Popoloca $k^w \sim$ Tapachulteca and Mixe p in one set; it is thus not well attested. The vowel *ä is set up on the basis of the correspondence Mixe-Zoque-Popoluca $a \sim$ Tapachulteca e. Although its reconstruction might be questioned in view of the poor quality of the Tapachulteca data, it can probably be regarded as solid.

Wonderly reconstructs a morphophoneme *W which deserves special comment. It is reconstructed on the basis of Zoque morphophoneme $X \sim$ Sierra Popoluca morphophoneme V. In Central Zoque, X has the following phonological characteristics: (1) $-hC + X > -CX > C$ since X itself is phonologically zero. (2) $-V\eta + X > -Vw$. (3) $-Vh + X > -Vy$ (if V is a, e, i, or $ä$) and $-Vw$ (if V is u). In Northern Zoque, morphophoneme X does not include $h > y$ when preceded by a, e, i, or $ä$ as in (3) above, but includes the other morphophonemic changes. In Sierra Popoluca—in suffixes cognate with Zoque suffixes containing morphophoneme X—there occurs a morphophoneme $\boldsymbol{V}$ which is actualized phonologically as follows:

(1) *CVC-* stem + morphophoneme $\boldsymbol{V} >$ *CV·C-*

(2) *CVʔC-* stem + morphophoneme $\boldsymbol{V} >$ *CVʔVC-*

[22] Wonderly does not reconstruct an alveopalatal series *t^y, *č, *š, *ñ for Proto-Zoquean (in spite of its well-nigh universal presence) in that he feels there is ample cause to believe that these phonemes have arisen 'from an alveolar phoneme palatalized by a contiguous y or i'. It seems safe to infer that palatalized allophones were present in Proto-Zoquean, and have independently become phonemic in the various languages and dialects by break-downs in the pattern of complementary distribution.

(3) *CVCC-* or *CVCCC-* + morphophoneme *V* > no change

Wonderly's choice of symbol **W* is based on the following argument: (a) The Proto-Zoquean phonological feature here symbolized seems to have been a semi-vowel. Thus, PZoq **W* > SP *V* which acts like a vowel in that Sierra Popoluca *V* > *V·* while *V?* > *V?V*; but acts like a consonant in that, given a *CC* or *CCC* cluster in Sierra Popoluca, this *V* morphophoneme is not actualized at all (thus *V* figured as a consonant which is eliminated in reduction of certain clusters). On the Zoque side, Proto-Zoquean **W* > *X* which acts like a semi-vowel in that it conditions *h* to *w* apparently in most if not all Zoque dialects; and conditions *h* to *y* as well after certain vowels in Central Zoque. But, again, Zoque *X* acts like a consonant in that its presence results in loss of *h* from final *-hC* clusters. (b) But if this Proto-Zoquean morphophoneme were a semi-vowel, a bilabial semi-vowel seems indicated. Crucial here is the fact that in Central Zoque and in Sierra Popoluca *w* 'does not appear in word-final position, and appears in syllable-final position in a very limited number of intervocalic *-CC-* clusters most of which are Spanish loans.' But **W* as set up in Proto-Zoquean is apparently syllable-final or at least post-vocalic. Wonderly suggests that possible his Proto-Zoquean **w* and **W* were in complementary distribution. His only hesitation seems to be as to whether Proto-Zoquean **w* was actually lacking in syllable final. Nevertheless, in the data of his article, post-vocalic **w* is clearly lacking, and Wonderly apparently knew of no exceptions or he would have included them. Furthermore, **h*, **?*, **s*, **y*, **m*, **n*, and **ŋ* all appear in syllable final, i.e. every non-stop phoneme except **w*; while in postvocalic position preceding a final *C*, both **h* and **y* occur but not **w*. The apparent phonetic similarity of **w* and **W*, their situation in complementary distribution in available reconstructions, and the curiously defective distribution of **w* as just mentioned all combine to make it seem very plausible that **W* was simply an allophone of Proto-Zoquean **w*.

3.3. PROTO-MAYAN, as reconstructed by McQuown (1955) has the following phonological system:

Consonants

*p	*t	*c	*č	*k	*ḳ	*?
*'p	*'t	*'c	*'č	*'k	*'ḳ	
		*s	*š	*x		*h
*m	*n			*ŋ		
*w			*y			
	*l					
	*r					

Vowels

*i		*u		*i·		*u·
*e	*a	*o		*e·	*a·	*o·

A slightly more recent work of McQuown's (1956) posits not five but six Proto-Mayan vowel positions, posits presence of two contrasting pitch-accents, and adds **ç* to the reconstructed consonants.

The consonant system exhibits symmetrical series of stops and glottalized stops including the affricates **c* and **č* (and **'c* and **'č*) as well as fronted and backed velars **k* and **ḳ* (and **'k* and **'ḳ*). As less extensive series occur spirants, nasals, semi-vowels, and liquids. Consonant clusters occurred as follows: C_1C_2, with C_2 consisting of the semi-vowels **y* and **w*; and C_1 consisting of any consonant except: the bilabials **p*, **'p*, **m*, the liquids **l*, **r*, and the semi-vowels themselves.

3.4. PROTO-TOTONACAN, as reconstructed by Arana (1953) on the basis of three Totonac dialects and one Tepehua dialect, has the following phonological system:

Consonants

*p	*t	*c	*č	*tɬ	*k	*ḳ	*?
		*s	*š	*ɬ	*x		
*m	*n						
*w			*y				
	*l						

Vowels

*i	*u	*i·	*u·
*a		*a·	

Initial clusters of spirant plus stop occurred: *sk, *šk, *st, *šk, *lk, *lt. In medial position, syllable-final consonants (e.g. *t, *k, *s, *n, and *y) followed by syllable-initial consonants or clusters constituted clusters of two or three members.

Proto-Totonac was characterized by consonantal alternation in the following patterns: *k/*k̇; *c/*č/*tl; *s/*š/*l. There is also some suggestion of vowel alternation.

In the list of phonemes, *ʔ is included in the stop series, perhaps for want of a better place. Arana actually reconstructs glottalized vowels in preference to either a separate glottal stop phoneme or a series of glottalized consonants.

As a whole, phonological developments from Proto-Totonacan to Totonac and Tepehua are rather regular and pedestrian. Evidently the dialects have not diverged far from each other. There is clear evidence in the Totonac dialect of San Pedro Petlacotla of considerable lexical borrowing from adjacent Tepehua.

4. Diffusion phenomena and linguistic family trees. Against the background of the above summary of reconstructed systems in Middle America I now discuss certain matters having to do with diffusion of features versus 'family tree' relationships, and with classification into subgroups within a language family.

The position here taken is that an intelligible and useful concept of language relationship must retain the notion of linguistic 'family' in the sense of (a) an internal relationship of peculiar relevance for the languages included within a family; (b) external relationships of less relevance with languages of related families; and (c) related versus non-related languages and language families. This we explain in a hypothetical situation below:

Linguistic groups X, Y, Z, and W each contain a small number of related languages which may be symbolized as l, m, n, and o respectively. The l languages which comprise group X are related primarily to each other, as are the m languages of group Y, the n languages of group Z, and the o languages of group W. The linguistic groups X, Y, and Z (but not including group W) comprise a broader grouping, and are mutually related in a fashion analogous to that in which l languages of Y are related (or the languages of group Y, Z, or W). There is a given language α among the l languages of group X, and a given language π among the m languages of group Y. Relationship of α to π is by virtue of the broader relationship of group X to group Y rather than by direct relationship of α to π as such.

In keeping with the above assumptions, we entertain the following questions: (1) Since in group X, languages α, β, γ, and δ are found, what are the internal relationships of the four languages? Are they four coordinate off-shoots of one parent stock, or do two or three of the languages constitute a sub-group? (2) Is a language ϵ—reputedly of group X—actually found to belong to group Y or Z? Similarly in keeping with our assumptions, we consider the following to be non-fruitful lines of inquiry: (3) Which language, α or β, of group X is the most closely related to language π of group Y? (4) May not the linguistic divergence between α and γ of group X be greater than the divergence between γ of group X and π of group Y? To answer the last question affirmatively seems to this writer to lead either to combining groups X and Y, or to scrapping the very concept of linguistic family. Therefore, if we grant the validity of group X versus group Y this question may not be answered affirmatively.[23]

[23] The above attempt to make explicit what we mean by linguistic 'family' is not simply a theoretical construct. It attempts to give weight to some known facts of language history and development. For example, we know that the Romance languages have developed from a somewhat

However, the family-tree concept may be criticized on the grounds that it does not permit us to give adequate weight to resemblances between certain languages. What if, e.g., a given language σ assigned to group R does as a matter of fact appear to be very resemblant to a language of group T? What of resemblances between apparently unrelated languages? Or to ask another question of a slightly different sort: Why are family-tree genealogical schemes *within* a language family or among major branches of a stock often so indeterminate?

Resemblances between languages can be from one of three sources: (1) genetic relationship; (2) chance parallelism of development; and (3) diffusion from one language or language group to another. In respect to the third consideration, several conclusions seem possible. First of all, it seems that borrowing is on the whole most frequent and uninhibited between two mutually intelligible dialects of the same language. But borrowing may take place between any two languages granted a situation of contact with some bilingualism. Finally, languages in the same geographical area tend to become similar (e.g. note common grammatical features in the Balkans; French-like phonemic systems in the Breton dialects of France; and presence of domal stops in the Indo-Pakistan subcontinent).

On finding a language σ of genetic group R which is resemblant to another language τ of genetic group T, we can assume one of two things (other than fortuitous coincidence) and maintain the concept of language family: (a) We can assume a strong diffusion of σ to τ, e.g. influence of French on English; or (b) we may re-examine the scheme of genetic relationship itself. We may, e.g., find that language σ forms a one-language group Q which is coordinate with groups R and T. Or we may find in structural resemblances between supposedly unrelated languages significant clues to unsuspected genetic affinities which are nevertheless demonstrable by the comparative method. We may, finally, find that the resemblance is simply an areal feature of no relevance to genetic grouping.

Some Middle American problems of this general nature (i.e. problems in sub-grouping and considerations of diffusion) are now discussed here.

4.1. MAYAN SUB-FAMILIES. Reconstructing Proto-Mayan on the basis of Mam, Huastec, and Yucatec, McQuown (1956) classified the Mayan languages on the basis of shared retention of seven features assumed to be Proto-Mayan.

McQuown points out that Huastec (isolated from the rest of Mayan and lying to the northwest), Yucatec (in the Yucatan peninsula), and the Highland Guatemalan languages form three points of a triangle which 'may be thought of as long established differentiated blocks' to which may be added the somewhat central-western Mayan languages of Tabasco and Chiapas (Chontal, Chol, Tzeltal-Tzotzil), Chicomuceltec and Motocintlec (towards the Pacific Coast) and Chorti (southeast of the Highland Group). He points out that there are special connections between Huastec and Chicomuceltec (which share particular

---

homogeneous vernacular Latin spoken in the first few centuries of the Christian Era. We also have documents to substantiate earlier stages of Greek, Indic, Germanic, and Slavic—so that we can with considerable assurance speak of Germanic, Slavic, and Indic (or Indo-Iranian) as language families. But these families and others are considered to be related qua families in the Indo-European stock. We entertain questions concerning internal relationships within, say, Germanic or Romance. We entertain similar questions about mutual relationship among the branches of IE. We may also entertain questions concerning the linguistic family into which some poorly attested language fits. But we have not entertained questions such as: Which Germanic language is most closely related to French? Nor are we accustomed to considering that the distance between two Germanic languages could be greater than the distance between some Germanic language and some Slavic language. We do not attend to the last two considerations simply because we believe in the historical validity of language families within Indo-European.

lexical items found only in them), between Mam and Chuh (which share survival of *η although they differ sharply in most other aspects), and between Chol and Chorti (which are separated phonologically only by the line between a six-vowel and a five-vowel system).

Against this general schematized background McQuown notes: (1) Certain isolated archaisms survive at the points of the triangle: (a) lexical tones (high and low) in Yucatec; (b) k versus k^w in Huastec; and (c) k versus $k̓$ in the Guatemalan highlands. In addition, $č$ versus $č̓$ survives in some of the Highland languages (Mam, Aguacatec, Ixil, and Kanhobal). On the western edge of the Highlands n versus η survives 'in a fully exploited form' only in Chuh but is also present word-final in Mam. In the central-western belt the old six-vowel system survives in Chol and Chontal—which are in this respect western outliers in reference to the languages of Highland Guatemala and Yucatec. To this picture of scattered archaisms surviving at the fringes of the Mayan region may be added the common survival of vowel quantity in Huastec, Yucatec, and Highland Guatemala. Thus, in terms of dialect geography, the general picture of surviving archaisms is a plausible one, with the Guatemalan highlands (especially Mam, Aguacatec, and Ixil) emerging as the most conservative block of languages.

The languages found along the northwest edge of the Guatemalan highlands and in the Chiapas highlands extending on down to the coast (i.e. Chuh, Toholabal, Tzotzil, Tzeltal, Chol, Chontal) are the more innovating. To these could be added Chorti—which has a special affinity with Chol although found at the eastern fringe of the Mayan region. Kanhobal shares some characteristics both of the conservative block (retains k versus $k̓$ and $č$ versus $č̓$) and also of the innovating block (loss of quantity).

Combining the above insights with 'an impressionistic sampling of the lexical material' (p. 194), McQuown emerges with the following 'cautious sub-grouping of the Mayan languages' (our summary here omits a few details): (1) Huastecan; (2) Cholan (Chontal, Chol, Chorti); (3) Tzeltalan (Tzeltal, Tzotzil, Toholabal); (4) Chuh; (5) Kanhobalan; (6) Motocintlec; (7) Mamean (Mam, Aguacatec, Ixil); (8) Quichean (Rabinal, Uspantec, Quiché, Cakchiquel, Tzutuhil); (9) Kekchian (Kekchí and Pocomam - Poconchi); (10) Mayan proper (Yucatec, Lacandone). He mentions that by 'less cautious sub-grouping' we could combine Cholan and Tzeltalan, Chuh and Kanhobalan, Motocintlec and Mamean, Quichean and Kekchian.

4.2 POPOLOCAN SUB-GROUPING. Using at first the Proto-Popolocan reconstruction of Fernández de Miranda and later Gudschinsky's more extensive treatment of the same, Hamp has attempted to clear up certain problems in the mutual degrees of affinity within the Popolocan family. Hamp at the outset of his first treatment of the subject (1958) makes very clear the principle on which he feels genetic proximity should be based: 'The only criterion for genetic proximity consists in the recognition of a decisive set, whether in number or in structural placement, of shared structural innovations; and these must be innovations by addition or replacement, rather than by loss.' (p. 150). On these grounds he rejects both Fernández de Miranda's belief that Popoloca and Ixcatec are especially closely related and Gudschinsky's belief that Popoloca, Chocho, and Ixcatec are coordinate, in favor of a belief that Chocho and Popoloca are especially closely related. (See fig. 1, which reproduces these alternative stemmata from Hamp, 1958.)

In his more recent treatment of the problem Hamp (1960) abandons five of the seven significant Chocho-Popoloca common innovations that he at first posited. However, he adds two more shared innovations involving these two languages; and further

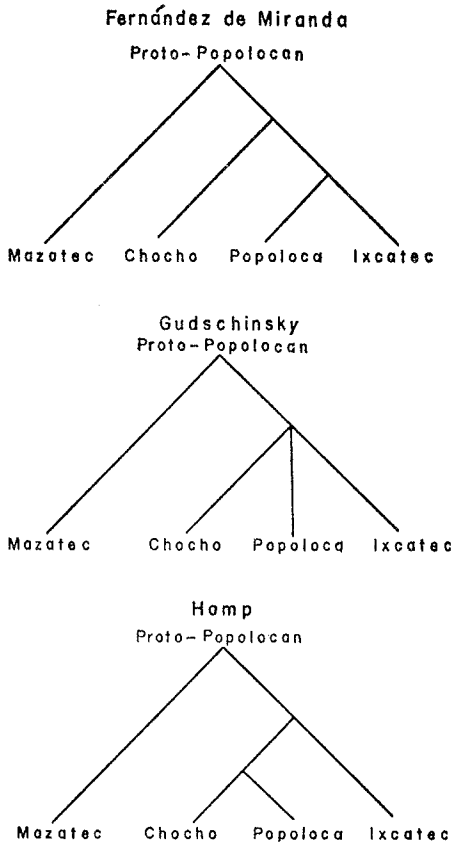


FIG. 1—THREE PROPOSALS REGARDING AFFINITY IN POPOLOCAN. (From Hamp, 1958.)

clarifies and strengthens one of the shared innovations at first posited. By comparison with these Chocho-Popoloca shared innovations, the Ixcatec-Popoloca and Chocho-Ixcatec agreements do not seem to Hamp to carry so much import in that these latter agreements are fewer and not so strategically placed structurally.

The Chocho-Popoloca shared innovations that Hamp considers of abiding importance are: (1) loss of *s and *š from clusters *št, *šk, *škh, šh, and *sk (as opposed to retention of reflexes of *s and *š in Ixcatec). (2) parallel treatment of allophones of *c including a three-way split into Chocho and Popoloca c, č, and r. (3) parallel merger of *tʸ before *h with *t in Chocho-Popoloca (and with r as well in Chocho). (4) parallel treatment of *š before back vowels (> Popoloca š and Chocho š and r).

Hamp closes with the caution: 'Finally, it is worth pointing out how misleading it would be if one were to classify languages on the basis of mere appearances; a glance at the superficial agreements between Ixcatec and Chocho in Gudschinsky's chart 11 should satisfy anyone on the futility of such an approach' (p. 62).

4.3 UTO-AZTECAN SUB-GROUPING. In concluding his sketch of Proto-Uto-Aztecan, Whorf was pessimistic about the validity of sub-grouping within that family: 'In time past some of us hoped that the stock could be classified in such a way that we could summarize the situation by stating generalized reflexes for sub-groups such as "Shoshonean", "Piman", "Sonoran", from which the reflexes of the individual tongues in these groups could be derived as a second step. This hope is vain. No such groups exist. The nearest thing to one is Piman, but it is unsafe to generalize about even so small and compact a group of dialects as Piman.' (p. 606).

In such a classification as McQuown's (1955), Uto-Aztecan languages of Mexico are divided into three groups: (1) Taracahitian (including Mayo-Yaqui, Tarahumara, and a host of extinct languages); (2) Aztecoid (including Aztec, Cora-Huichol, Pipil of Guatemala, and a number of extinct Mexican and Central American languages); (3) Piman (Papago, Tepecano, Northern Tepehuan, Southern Tepehuan, and a number of extinct languages). The first group may be largely a grouping of geographic convenience; the large number of extinct languages in it makes extensive investigation of the group qua group impossible. But it would be worthwhile to inspect Mayo-Yaqui and Tarahumara to see what, if any, common innovations are shared by them. In terms of Whorf's reconstruction of Uto-Aztecan along with his listed reflexes in eight languages, it is difficult to make out a good case for special Cora-Huichol and

Aztec affinities (Cora and Huichol admittedly are close but the question here is relation of the two to Aztec). Thus, if Whorf's Proto-Uto-Aztecan *e be reconstructed as *i (which seems probable) we note that Cora and Aztec share a common innovation (> A e, Cora e, i). But, unfortunately, Tarahumara also shares reflex e so we do not have here an exclusively shared Aztec-Cora innovation. Slightly better is Aztec-Cora treatment of *u which goes to i only in these two languages in the eight languages for which Whorf lists reflexes. It perhaps does not detract too greatly from the evidential value of this that Cora also has a reflex v (of *u) shared with Southern Paiute. Lastly, Aztec and Cora possibly share a similar handling of the bilabial phonemes *p and *v: PUA *p > Aztec p or *zero*, and to Cora p and *h/zero*; while PUA *v > Aztec w or *zero*, and to Cora w, v, or *zero*. Whorf does not give us the data to discover whether the *p/zero* and *w/zero* reflexes are phonemic splits under parallel conditions in Aztec and Cora.

For Piman we get somewhat better results—as Whorf's comments might lead us to believe in spite of his pessimism regarding sub-grouping. Thus, for the two Piman languages (Tepecano and Papago) whose reflexes are listed by Whorf we note that: (1) PUA *c > T,P *s/š* (split under same conditions?). (2) PUA *k^w > T,P *b*. (3) PUA *$^n s$ > T,P *h/ʔ*. (4) PUA *s > T,P *h/ʔ* (thus merging with *$^n s$). (5) PUA *y > T,P *d*. (6) PUA *w > T,P *g*. Note that peculiar reflexes of *k^w, *y, and *w apparently yield a Proto-Piman voiced stop series *b, *d, *g not found in Proto-Uto-Aztecan (but data should be checked from other Piman languages).

4.4. ZOQUEAN SUB-CLASSIFICATION. In the last section of his Proto-Zoquean study Wonderly (1949) draws some Zoquean isoglosses that deserve extended comment. Zoquean is not internally as diversified as, e.g. Mixtecan. Indicative of this is the fact that phonological developments from language to language are rather pedestrian. For example, Wonderly writes, "PZoq stops in word initial position remain unchanged in all dialects, except that *c > č before *i in Popoluca, and before *i, *ä in Tapachulteca" (p. 3). The relative closeness of Zoquean languages is indicative of the fact that they are relatively less removed from the period when they were mutually intelligible dialects than are languages of families that show greater diversity. Two things therefore may be expected: (1) Extensive dialect borrowing in late Proto-Zoquean and in early stages of each language may leave its trace in a web of mutually inconsistent isoglosses that may even cut across language boundaries. (2) Many loans—even of fairly recent date—may be indistinguishable from inherited stock in view of the rather non-distinctive nature of Zoquean phonological developments.

Considering the last point first, it is instructive to note a comment of Wonderly's: "In SP, M,[24] there is a tendency toward intervocalic voicing of stops, with at least traces of 'Verner's law' in M. In NeZ, the stops *-*p*-, *-*t*- become -*b*-, -*d*- (-*r*- in Chapultenango) in most words; there is a residue of apparent exceptions which may be due to dialect borrowing or other factors" (p. 3). In other words, while -*b*- and -*d*- are reflexes of *-*p*- and *-*t*- in Northeast Zoque, borrowings from adjacent dialects with reflexes -*p*- and -*t*- may be one factor obscuring this picture.

Notice that this tendency towards voicing of intervocalic stops is found at three of the extremities of the Zoquean area, viz. Sierra Popoluca at the north, Mixe on the west, and Northeast Zoque on the east; its distribution is that of an archaism rather than that of an innovation. Probably, therefore, Proto-Zoquean stops had voiced allophones in intervocalic and the loss of these voiced

[24] Abbreviations: SP: Sierra Popoluca; M: Mixe; CZ: Central Zoque; NZ: Northern Zoque; SZ: Southern Zoque; WZ: West Zoque; NeZ: Northeast Zoque; T: Tapachulteca.

allophones in the rest of Zoque and in Tapachulteca is an innovation.[25]

A similar Zoquean phenomenon is voicing of stops after nasals; Wonderly summarizes this as follows: 'In SP and WZ, stops remain voiceless after nasals. In CZ, NZ, NeZ, stops are voiced after nasals. In M, stops are voiced after nasals except when the cluster is word final. In T and SZ, stops are voiced after nasals in some words and voiceless in others. (In some of the dialects, the voicing is probably allophonic rather than phonemic.)' (p. 6). Since voiced allophones of intervocalic stops seem to have been a Proto-Zoquean feature, it seems probable that voiced allophones of stop after nasal were also Proto-Zoquean and that Sierra Popoluca and West Zoque reflect an innovation by which voiced allophones were eliminated. This same innovation has spread to South Zoque and Tapachulteca but has affected only certain morphemes; as such this may be a result of dialect borrowing. Thus South Zoque *cintak* 'pine forest' exists side by side with *cindak* which may be a borrowing from Central Zoque. Mixe shares the unvoicing innovation to the extent that word-final consonants preceded by nasals are voiceless. Figure 2 summarizes unvoicing of old stop allophones.

One type of palatalization sketched by

[25] Another probable archaism found only in Northeast Zoque is indicated by some puzzling reflexes of the nasals. Wonderly says 'In intervocalic position the picture is somewhat confused: in the NeZ dialects, **m*, **n*, **ŋ* appear in some words as nasals, in others as voiced stops, in still others as voiceless stop + nasal; the conditioning factors can probably be found only after more work is done on the structure of the NeZ dialects, and in some of these cases the PZoq form will probably have to be set up with an intervocalic consonant cluster' (p. 5). In reference to the same phonemes in word-final Wonderly continues: 'In certain words in the NeZ dialects, they correspond to *-p*, *-t*, *-k*; in certain other words of the same dialect they correspond to *-m*, *-n*, *-ŋ*. The correspondences are not all the same for all NeZ dialects, but the tendency is for the correspondences with *-p*, *-t*, *-k*.' (p. 5). These peculiar correspondences of Northeast Zoque are as a whole unparalleled in any of the other languages and dialects.

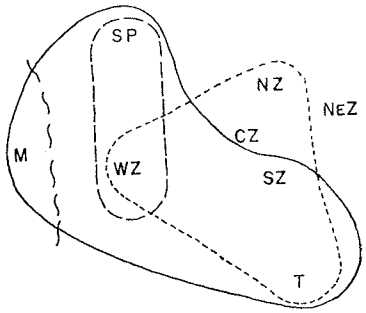


FIG. 2—MAP OF UNVOICING OF OLD VOICED STOP ALLOPHONES. Broken line (short dashes): area where voiced stop allophones in intervocalic are completely eliminated. Broken line (long dashes): area where there is complete elimination of voiced allophones after nasals. Solid line: area where there is partial or complete elimination of voiced stop allophones after nasals. Line of ʃ's: area where unvoicing occurs only following nasals in word final.

Wonderly is entitled: 'Metathesis of **y* with a following consonant' (p. 7). That the metathesis of **yC* > *Cy* is in reality an end-product of palatalization is seen in the varied Zoquean developments: PZoq **yC* > M *yCy* when word-medial and *Cy* when word-initial. In CZ, NZ, NeZ, **yC* > *Cy* but traces of an earlier **yCy* occur in CZ where **yt* > *yty*. I assume, therefore, that PZoq **yC* was phonetically **[yCy]*, that in many areas the palatalized allophones became phonemic, and that in some areas the preceding **y* was entirely lost. In still other areas the palatalized allophone after **y* did not become phonemic (and possibly disappeared).[26] This is summarized in figure 3.

Loss of *h* in word-final cluster is described as follows by Wonderly: "Word-final PZoq **-hC* always loses the *h* in CZ, NZ, SZ, WZ,

[26] Palatalization of **c* to *č* before *i* takes place in Sierra Popoluca and in Tapachulteca but nowhere else in the Zoquean area. In Sierra Popoluca this type of palatalization extends to other alveolar phonemes. In Tapachulteca palatalization of **c* occurs not only before **i* but also before **e*. Wonderly correctly infers that the two developments are independent of each other. However, it is possible that Proto-Zoquean **c* had a palatalized allophone before front vowels and that this allophone has become phonemic in Sierra Popoluca and Tapachulteca independently and in different manners.

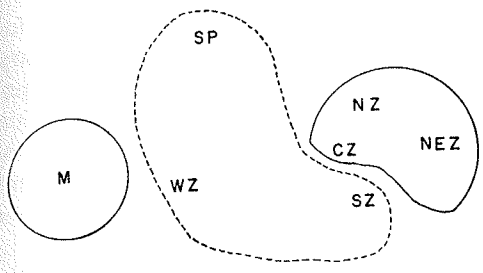


FIG. 3—Solid line: areas where yC^v and/or C^v are found. Broken line: area where *$[yC^v]$ remains phonemically yC with loss of *$[C^v]$ allophone.

T, SP. In certain words the *h is also lost in M. In certain other words the *h is also lost in the Chapultenango dialect of NeZ . . . Map IV leaves part of the Popoluca area outside the main isogloss, as preliminary data collected by Elson indicates that the *h is not lost in the Texistepec dialect of Popoluca" (p. 8).

In accounting for the loss of *h the following scheme seems plausible: (1) Proto-Zoquean *h was not lost from word-final clusters in such late Proto-Zoquean fringe dialects as those which eventually became Northeast Zoque, and in a part of the dialects which became Sierra Popoluca. (2) Proto-Zoquean *h was lost from word-final clusters in central late Proto-Zoquean dialects (including what became West Zoque, Central Zoque, North Zoque, South Zoque, and Tapachulteca as well as part of what later became Sierra Popoluca)—but probably only in certain words at first. (3) Loss of *-h from such clusters was universalized in the central dialects which now were split, along with the contiguous fringe dialects, into early stages of the four languages. The progressive elimination of *h could now proceed somewhat independently in the four proto-languages following the drift that set in late Proto-Zoquean—but loans from language to language could still have figured in the process. (4) Elimination of *h from final clusters then spread sporadically into Mixe and into Northeast Zoque as well. In that the words affected are different in Mixe and in Northeast Zoque, it

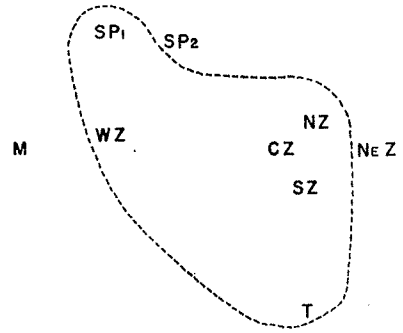


FIG. 4—STAGE I. Broken line incloses central area where late PZoq *h was lost from *-hC in at least some words.

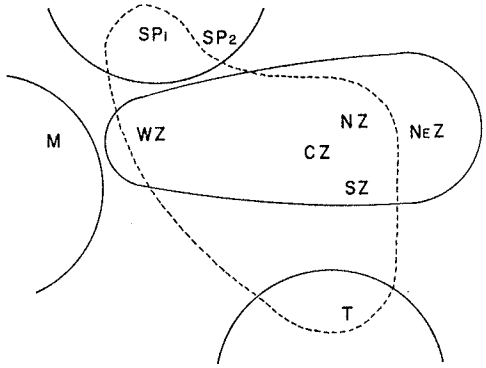


FIG. 5—STAGE II. Broken line incloses area where loss of *h from *-hC clusters continued to be extended until all such clusters were affected. Solid line indicates bundles of isoglosses which came to distinguish early stage of SP, M, Z, and T. These latter isoglosses are at cross purposes to the earlier isogloss, which represents an early dialect boundary which later became of less importance.

seems that the two developments do not represent archaisms (as their fringe position might indicate) but rather a late spread from the contral innovating area. This development is sketched in three stages in figures 4–6.

Plotting all the above features on the same map-like figure we obtain figure 7. which represents the late Proto-Zoquean dialect situation in reference to the central innovating area (involving loss of voiced allophones, loss of palatalized allophones, and loss of h) versus east and west fringe conservative areas.

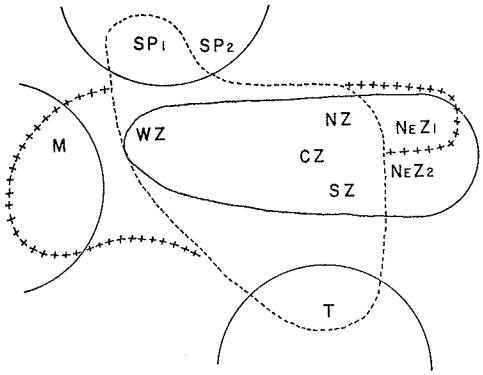


FIG. 6—STAGE III. Line of crosses: later extensions of area in which *h was lost from *-hC in some words. Broken line and solid line as in Stage II (fig. 5).

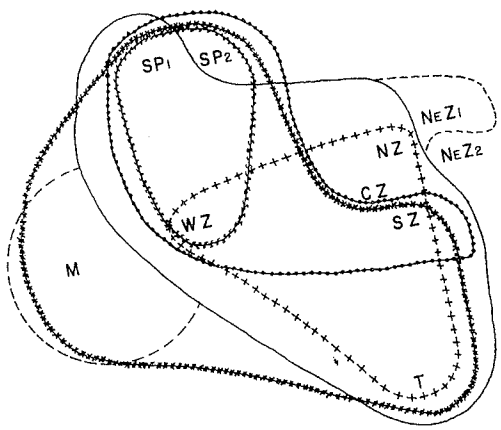


FIG. 7—Line of crosses: 1, area in which intervocalic voiced stop allophones are completely eliminated. Line of ʃ's: 2, area in which voiced stop allophones after nasals are completely eliminated. Line of X's: 2A, area in which voiced stop allophones after nasals are completely or partly eliminated. Dotted line: 3, area where *[yCv] remained phonemically yC with loss of *[Cv] allophone. Solid line: 4, area where *h is completely lost from *-hC clusters. Broken line: 4A, later extensions of the area in which *-hC was lost (but only in certain words).

In terms of the isoglosses plotted in the accompanying figure, we note the primary character of isoglosses 1, 2, 3, 4, and the secondary character of 2A and 4A which represent later extensions of 2 and 4 respectively. We can obtain a coefficient of innovation by scoring each dialect as follows: (a) 2 points for participation in any isogloss of primary character; (b) one point for participation in an isogloss of secondary character. We will not score both participation in a primary isogloss and participation in its corresponding secondary isogloss for the same language, since this would amount to scoring the same feature twice. Scores follow:

WZ participates in 1, 2, 3, 4 and scores 8.
SZ participates in 1, 2A, 3, 4 and scores 7.
SP_1 participates in 2, 3, 4 and scores 6.
T participates in 1, 2A, 4 and scores 5.
SP_2 participates in 2, 3 and scores 4.
CZ participates in 1, 4 and scores 4.
NZ participates in 1, 4 and scores 4.
M participates in 2A, 4A and scores 2.
NeZ_1 participates in 4A and scores 1.
NeZ_2 participates in no innovations and hence scores zero.

West Zoque seems to represent the center of innovation in late Proto-Zoquean. South Zoque represents an area very close to the center of innovation, with Sierra Popoluca (exclusive of Texistepec) only slightly less removed and Tapachulteca still further removed. Sierra Popoluca (Texistepec), Central Zoque, and North Zoque are about equally innovating, but even further removed from the center of innovation than Tapachulteca. The conservative areas are: Mixe (on the west) and Northeast Zoque on the east, with the latter almost completely insulated from the innovation center. Notice that in terms of the features here considered Central Zoque and North Zoque are one dialect area.

The late Proto-Zoquean dialect situation is represented in figure 8. in which the concentric circles represent units on an eight-point scale which marks off the coefficient of innovation.

4.5. MAZATEC DIALECTS. Gudschinsky (1958a) has given us a careful study of Mazatec dialect history. She constructs a plausible family tree for Mazatec dialects. Dividing High Mazatec (HM) from Low

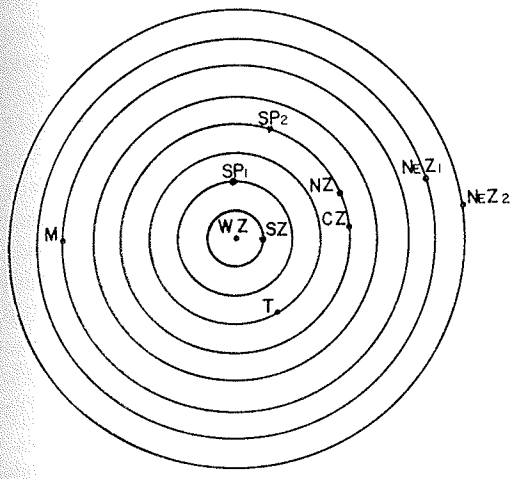


FIG. 8—SCHEMATIC REPRESENTATION OF LATE PZoq DIALECT SITUATION IN REGARD TO LINGUISTIC DISTANCE FROM CENTER OF INNOVATION IN WEST ZOQUE (WZ).

Mazatec (LM) (with common HM innovation **tʸk* > *šk* and common LM innovation **tʸ(i/e)* > *t*) she then further subdivides HM into Western (common innovation: **tk* > *sk*) and Eastern (common innovation: **tk* > *hk*); while LM divides into San Miguel and Valley Mazatec (common innovation of Valley Mazatec: **Vhna/u* > *Vhña/u*). Valley Mazatec in turn divides into Southern (common innovation: **tʸk* > *tk*) and Northern (common innovation: **čha* > *c̣ha*). Further subdivisions occur: (1) eastern High Mazatec divides into Huautla (**tʸhṼ* > *šṼ*) and Jiotes (**kʷ* > *k*). (2) Low Southern Valley Mazatec divides into Mazatlán (**tk* > *hk*) and Jalapa (**hñá/ú* > *hya/u*). (3) Low Northern Valley Mazatec divides into Soyaltepec and Ixcatlán (**tʸk* > *tik; *tk* > *tuk*).

The picture thus obtained is a plausible one both in terms of present day geographical distribution (valley versus highlands, and dissection of area by rivers) and in terms of what is known of Mazatec history. In respect to the latter, Gudschinsky points out that from around A.D. 1300 on, the Mazatec area—either under native rulers or under Mixtecs or Aztecs—was divided into Lowland and Highland 'kingdoms'. Furthermore, at an earlier period (ca. 890–1170) there was a single Mazatec nation concentrated apparently in what is now Valley Mazatec—which would account well for the unity of Valley Mazatec.

Nevertheless, as Gudschinsky points out, there was some spread of innovations across dialect boundary with occasional consequent shift of dialect boundary. Noticeable here is the spread of features from High Mazatec into contiguous portions of Low Mazatec. Thus, the dialect of San Miguel can be classified as Low Mazatec in that it shares the typical **tʸ(i/e)* > *t* innovation of that dialect group. But the dialect of San Miguel is also characterized by the **tʸk* > *šk* innovation which is diagnostic of High Mazatec. Here we have an apparent option; in that either isogloss can be considered to be the older development, it could be argued (1) that the Mazatec of San Miguel is a High dialect that was invaded by spread of an innovation from Low Mazatec; or (2) that it is a Low dialect that was invaded by spread of an innovation from High Mazatec. Gudschinsky adopts the latter alternative in that there are other clues (lexical isoglosses) which seem to indicate spread from such High Mazatec dialects as Huautla to San Miguel. Furthermore, this is supported by what is known of the present cultural situation (Huautla is a dominant center) and past history (formerly the Valley dialects were culturally dominant; later two centers of cultural dominance emerge in Low and High Mazatec respectively). The Mazatec dialect of Mazatlán reflects a similar spread of a High Eastern Mazatec feature (**tk* > *hk*) into a Low Southern Valley dialect.

Gudschinsky also traces a number of lexical isoglosses. She writes: 'A number of the lexical isoglosses closely approximate the dialect boundaries established by the phonological isoglosses, except that there is more diffusion of lexical items especially from the High dialect to Mazatlán and San Miguel' (p. 474). The reference to diffusion of High

Mazatec lexical items to Mazatlán and San Miguel correlates with shifting of dialect boundaries noticed above whereby each of these dialects came to share a phonological isogloss with High Mazatec (or at least with Eastern High Mazatec).

There is another phonological isogloss which Gudschinsky mentions but does not put on her maps: '*sh* in the dialects of Huautla and Jiotes (of the High dialect) and of San Miguel, Jalapa, and Ixcatlán (of the Low Dialect) corresponds to *ch* in the dialects of Tecoatl, Eloxochitlán, and San Mateo (of the High dialect) and of Mazatlán and Soyaltepec (of the Low dialect).' (p. 474). Gudschinsky comments that this isogloss 'seems to contradict the family tree just outlined'. She mentions that both *sh* and *ch* are reflexes of Proto-Popolocan **ch* but does not describe for us the situation in Proto-Mazatec as such. But, considering the geographical distribution of **c(h)* (on the Western and Northeast fringes of Mazatec speaking territory) she recognizes this as an archaism as opposed to the Central area where passage of **c* to *s(h)* is an innovation 'spreading from a central area and leaving conservative islands on the margins'.[27] Figure 9 illustrates this central innovating area as well as the shifting of the boundary of High Mazatec as illustrated above.

[27] Gudschinsky's projected course of development for Proto-Popolocan **c* as posited in her Proto-Popotecan does not agree with the development as sketched in her Mazatec Dialect History. On chart 15 (p. 28) of Proto-Popotecan she traces Proto-Mazatec reflexes as follows: PPn **c* > PMaz **s(h)* and to PMaz **c* elsewhere. On charts 3, 4, 5 (pp. 8–10), she projects the following developments in Jiotes, Mazatlán, and Soyaltepec: PMaz **s(h)* > *c*—thus making PPn **c(h)* develop to PMaz **s(h)* and then, in turn, to *c(h)* once more. Furthermore, the projected PMaz **s(h)* reflex of Proto-Popolocan **c(h)* is, as we have seen, found only in central dialects where it looks like an innovation rather than an archaism. I, therefore, suggest the following reformulation: (1) PPn **c* > PMaz **c* everywhere (including clusters with *h*). Gudschinsky's chart 15 needs to be modified by striking out the special reflex PMaz **s(h)*. (2) PMaz **c* remained *c* in certain archaizing fringe dialects, and (3) PMaz **c(h)* went to *s(h)* in the central dialects.

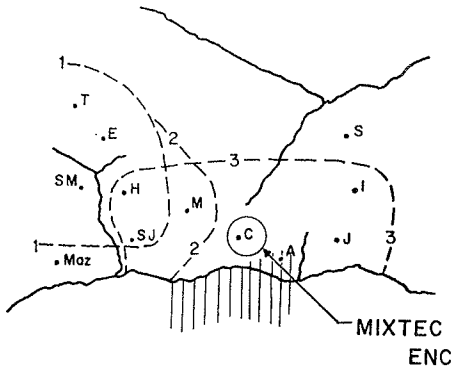


FIG. 9—T (Tecoatl); E (Eloxochitlán); SM (San Mateo); H (Huautla); SJ (Jiotes); Maz (Mazatlán); M (San Miguel); S (Soyaltepec); I (Ixcatlán); J (Jalapa). 1____1 Boundaries of High Mazatec according to primary dialect differentiation. 2____2 Later extension of certain High Mazatec features (but Mazatlán shares more features with Low Mazatec than does San Miguel). 3____3 Central innovating area where PMaz **c* (*h*) > *s* (*h*). Vertical lines: Mazatec-speaking region from which no data are available.

Among the lexical isoglosses other instances occur which, in common with the above phonological isogloss, show survival of archaisms on margins as opposed to an innovating central area. Thus, Gudschinsky (p. 477) notes the following restricted distributions of lexical items which are presumably archaic survivals:

(1) Mazatlán, Soyaltepec, and Ixcatlán
(2) Mazatlán, Tecoatl and Soyaltepec
(3) Mazatlán and Jalapa
(4) Tecoatl and Eloxochitlán
(5) Tecoatl, Eloxochitlán, and Mazatlán
(6) Tecoatl, Eloxochitlán and Ixcatlán
(7) Soyaltepec
(8) Tecoatl and Eloxochitlán

These dialects may be arranged as follows in order of decreasing frequency of mention in the above list: Tecoatl, Mazatlán, Eloxochitlán, Soyaltepec, Ixcatlán, Jalapa. The four most frequently mentioned dialects are also within the archaizing area defined by retention of **c*. With less frequency do Ixcatlán and Jalapa align themselves with the western archaizing fringe along with Soyaltepec. An eastern fringe including Tecoatl, Mazatlán and Eloxochitlán is ap-

parently a more extensive and better defined archaizing area.

In summary, Gudschinsky's study of Mazatec dialects is an excellent example of the manner in which insights of the comparative method may be combined with those of dialect geography. We need not oppose the 'family tree' concept of linguistic relationship to 'wave' phenomena of dialect diffusion. We need not deny the second in order to establish the first, nor need we abandon the first in order to do justice to the second.

4.6. MIXTECAN SUB-GROUPING. In reference to shared innovations, the isoglosses plotted for the Mixtecan family (exclusive of Amuzgo whose precise placement is still doubtful)[28] pattern as follows: (a) strong Mixtec-Cuicatec isoglosses; (b) strong Cuicatec-Trique isoglosses; (c) weak Mixtec-Trique isoglosses. Nevertheless, in spite of the relative weakness of the Mixtec-Trique isoglosses, they are important as vestiges of a former dialect division not correlating well with the subsequent (and more relevant) separation into the three languages.

The shared Mixtec-Cuicatec innovations are: (1) $*x^wa > vaa$. (2) $*{}^ng^wa > vaa$. (3) $*k^wV > ku$ (with loss of vowel and vocalization of labial component) under varying conditions. (4) Mixtec and Cuicatec eliminate *w* from present-day verb paradigms but preserve reflex of pardigmatic $*w$ in fossilized fragments (while Trique preserves $*w$ in present-day verb paradigms). (5) Mixtec and Cuicatec merge Proto-Mixtecan tones $*2$ and $*3$—thus becoming three-level tone systems. (6) Proto-Mixtecan $*o$ > Mixtec and Cuicatec *o/u* (versus consistent reflex *o* in Trique). The first four isoglosses group together; thus Mixtec and Cuicatec show parallelism in the treatment of labialized phonemes, as well as in the elimination of bilabial $*w$ from verb paradigms.

The shared Cuicatec-Trique innovations are: (1) After $*{}^nd$, PMx $*i$ > C *e/i*, T *e*. (2) After $*n$, PMx $*i$ > C *o/u*, T *ï* (C has no *ï* phoneme). (3) After $*n$, PMx $*e$ > C, T *i*. (4) After $*w$, PMx $*e$ > C *a*, T *i* (also after $*x^w$ and $*{}^nd$, $*e$ > C *a*). (5) Before labials, PMx $*i$ > C, T *u*. (6) It is of some interest also that Cuicatec and Trique share a reconstructed $*12$ sandhi variant of basic tone pattern $*32$, while Mixtec reflexes seem to indicate a $*31$ sandhi variant.[29] Notice that Cuicatec-Trique isoglosses (1)–(5) indicate a common Cuicatec-Trique treatment of front vowels, just as the Mixtec-Cuicatec isoglosses (1)–(4) indicate a common Mixtec-Cuicatec treatment of labial consonants.

By contrast the shared Mixtec-Trique innovations are but two: (1) Proto-Mixtecan $*t$ is palatalized in Mixtec and Trique (> T *č* uniformly; > M dialect reflexes, such as t^y, *č*, *c*, *š* or modification of following vowel $*a$ to *e*—with a few Mixtec dialects showing none of these developments). (2) Proto-Mixtecan $*\theta$ splits into two reflexes conditioned by front versus back vowels (T everywhere; M in all but fringe dialects). Both the above reflexes are peculiar in that each of them divides the Mixtec-Trique area in two in such a way so as to include Trique with some Mixtec dialects but not with others. This is reminiscent of the Zoquean isoglosses presented above in that the isoglosses fail to conform neatly to the language boundaries. However, we need to examine in detail the Mixtec dialect evidence:[30] (1) Proto-Mixtec $*t(i)$ is palatalized all over the Mixtec region except for a block of coastal dialects in the southwest corner of

[28] Swadesh (1960) has argued in favor of retaining Amuzgo as Mixtecan. I have raised the possibility that Amuzgo may form along with Mixtecan proper a broader 'Amuzgo-Mixtecan' grouping (Longacre, 1961a; Longacre-Millon, 1961). But, cf. note 1.

[29] The Popolocan evidence suggests that in Proto - Popolocan - Mixtecan (= Otomanguean also?) basic form $*32$ had no sandhi variants. We suggested above in Sec. 2.7. that possibly $*32$ also had no sandhi variants in early Proto-Mixtecan, but that these variants developed in late Proto-Mixtecan dialects. If so, then the Cuicatec-Trique sharing of variant $*12$ of basic $*32$ is another shared Cuicatec-Trique innovation.

[30] As summarized by Mak and Longacre, 1960.

Oaxaca. (2) Proto-Mixtec *$t(i)$ is palatalized only on the northern and northeastern fringes (dialects of Puebla, and Mixtec enclaves in Mazatec) plus the somewhat northwestern dialect of Juxtlahuaca, Oax. (3) Proto-Mixtec *$t(a)$ remains unpalatalized in the southwest coastal block noted above; goes to *te* (with *t unaffected but with palatal influence on the vowel) in northern and northeastern fringes and in a sizeable block of north-central dialects; and goes to *ča* or *tya* in the intervening region. (4) Proto-Mixtec *θ apparently goes to *d* in fringe dialects on the north, northeast, and east as well as in one dialect of the southwest coastal block. Elsewhere, apparently *θ split into *s*/*š* according to back versus front vowels.

To summarize, note that (a) There is a southwest coastal block enclosed by isoglosses (1) and (3) above with one dialect of this block involved in isogloss (4) as well. (b) There is a northern innovating area as seen in palatalization of *$t(i)$, in *$ta > te$ (as distinct from *t itself being palatalized), and in palatalization of *$t(i)$. In that the palatalization of *$t(i)$ occurs in the northern and northeast fringes plus Juxtlahuaca, while *ta becomes *te* in these fringes plus part of the central area, and palatalization of *$t(i)$ occurs everywhere except in the southwest conservative coastal block, these isoglosses show a common tendency in that they inclose the northern and northeast fringes plus varying amounts of the central area. (c) There is a central innovating area in which *$t(a)$ is palatalized to *ty* or *ča* (and in which take place other varied developments beyond the scope of this paper). Development of *s*/*š* reflexes of Proto-Mixtecan and Proto-Mixtec *θ patterns as one of these central innovations which has spread over much of the area and leaves reflex *d* only in northern, northeastern, eastern, and southern fringes.

Trique, which palatalizes all *t to *č* and splits *θ to *t*/*c* (before back and front vowels respectively in a manner quite parallel to Mixtec) shares the first innovation with the Mixtec northern innovating area and the second innovation with the Mixtec central innovating area. It seems plausible to assume that in late Proto-Mixtecan there were two isoglosses delimiting areas where Proto-Mixtecan *t and *θ had palatalized allophones (possibly at first only before front vowels). The isogloss for *$[t^y]$ included what became Trique and some adjacent northern Mixtec dialects (assuming that Mixtec did not extend as far north as at present). From this view, Mixtec palatalization of *ta to *ča* or *tya* in the central innovating area is more recent than *$ta > te$, which is more characteristically northern. The isogloss for *$[\theta^y]$ included late Proto-Mixtecan dialects that became Trique and certain adjacent western and central Mixtec dialects (assuming that Mixtec did not extend westward into Guerrero as at present).

It is probably best to assume that Proto-Mixtecan *θ was phonetically $[\theta]$ and $[\theta^y]$ in late western Proto-Mixtecan dialects that became Trique and west-central Mixtec, but was phonetically $[\eth]$ in the late eastern Proto-Mixtecan dialects that became Cuicatec and the present Mixtec non-western fringes. On this assumption the phonological developments in the various languages and dialects are most easily explained. Thus, western Proto-Mixtecan *$[\theta]$ gives Trique *t* and Mixtec *s*, while its palatalized allophone *$[\theta^y]$ gives Trique *c* and Mixtec *š*. Eastern Proto-Mixtecan *$[\eth]$ gives Mixtec and Cuicatec *d*. It seems clear that but one Proto-Mixtecan phoneme is involved, while it also seems plausible that it varied phonetically from west to east in late Proto-Mixtecan dialect areas.

The above postulated development divides late Proto-Mixtecan into two dialect areas (West PMx versus East PMx) delineated according to distribution of variants of *θ. Division according to distribution of variants of *t yields further dialect areas in which Pre-Proto-Cuicatec and a portion of southern Pre-Proto-Mixtec pattern as fringe areas while Pre-Proto-Trique and

northern Pre-Proto-Mixtec pattern as an innovating central region. Both isoglosses draw a line of separation between dialects that eventually became Cuicatec versus those that eventually became Trique. But within what eventually became Mixtec, the dialects are variously grouped, some with what eventually became Cuicatec, others with what eventually became Trique.

It is postulated here that the two innovations common to Trique and some Mixtec dialects reflect an earlier period of dialect divergence than do the Mixtec-Cuicatec and Cuicatec-Trique shared innovations, which appear to be more numerous and carry greater weight. In that the Mixtec-Trique shared innovations are few and do not correlate well with the present boundary between the two languages it is assumed that (a) Trique, while maintaining contact with Cuicatec, separated geographically from Mixtec at a relatively early period—possibly at a time when early Proto-Trique and early Proto-Mixtec were still mutually intelligible dialects. (b) This separation created a Trique-Cuicatec-Mixtec dialect chain. Consequently, shared innovations between Trique-Cuicatec and Cuicatec-Mixtec continued to develop while shared Mixtec-Trique innovations ceased to develop. It is necessary to carry out the above scheme further by taking into account present-day separation of Trique from Cuicatec, as well as the present-day position of Trique as a language island surrounded by Mixtec. In doing this I assume (c) that Trique and Cuicatec lost mutual contact some time *after* they had ceased to be mutually intelligible dialects while Mixtec and Cuicatec continued to maintain contact in spite of the continued eastward drift of Cuicatec; and (d) that Mixtec, in its period of expansion finally flowed around the Trique-speaking area and engulfed it on all sides. In reference to (c) it may be noted that development of an effectual language barrier between Cuicatec and Mixtec finally checked further development of common

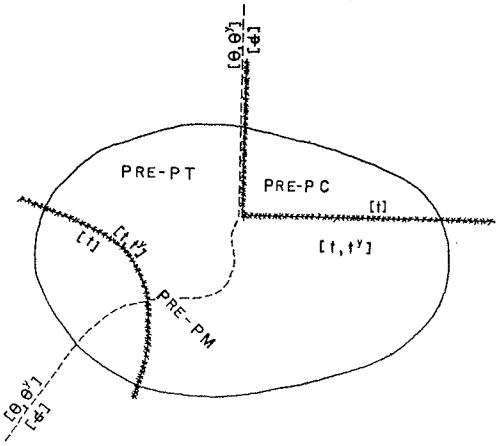


FIG. 10—STAGE 1. The late Proto-Mixtecan dialect situation shows no Mixtec-Trique dialect boundary according to the two isoglosses plotted, but delineates Cuicatec from Trique. Line of X's: presence of allophones *[t] and *[t^{y}] of *t versus presence of only allophone *[t]. Broken line: presence of allophones *[θ] and *[θ^{y}] of *θ versus presence of allophone *[d̪].

shared innovations—which do not appear to be of any more consequence than the Cuicatec-Trique shared innovations. Nevertheless, the continued Mixtec-Cuicatec contact may be reflected in the Mixtec-Cuicatec lexicostatistic count which is apparently consistently higher than that for Mixtec-Trique or for Cuicatec-Trique.[31] In reference to (d) it may be noted that the belated renewal of Mixtec-Trique contact did not result in further shared phonological isoglosses, but is nevertheless reflected in presence of Mixtec loans in Trique.

These four hypothetical stages of development are illustrated in figures 10–14.

5. SUMMARY AND IMPLICATIONS. Systemic comparison and reconstruction of Middle American Indian languages has been carried forward to the point where there is promise of considerable pay-off in the near future. Part of this pay-off will consist in the evolution of

[31] For lexicostatistic counts involving Mixtec-Cuicatec, Cuicatec-Trique, and Mixtec-Trique, see Swadesh, 1960 (which evaluates previous counts by Gudschinsky and Arana); Longacre 1961a; Millon and Longacre, 1961.

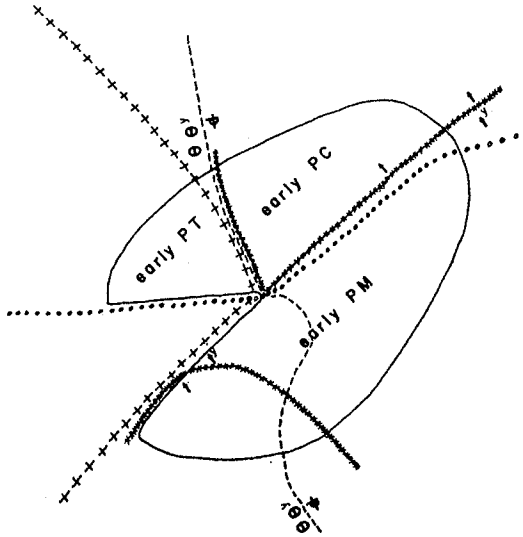


FIG. 11—STAGE 2. Mixtec-Trique isoglosses cease to develop after Trique pulled away from Mixtec. Mixtec-Cuicatec and Cuicatec-Trique isoglosses continue to develop. Line of crosses: separates early Proto-Trique from early Proto-Mixtec and early Proto-Cuicatec which share a common treatment of labials. Line of circles: separates early Proto-Mixtec from early Proto-Cuicatec and early Proto-Trique which share a common treatment of front vowels. Line of X's and broken line: as in fig. 10.

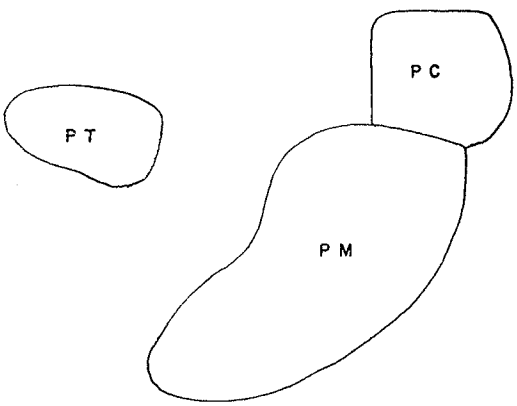


FIG. 12—STAGE 3. Effectual contact between Proto-Trique and Proto-Cuicatec is severed. Proto-Cuicatec and Proto-Mixtec are now mutually unintelligible; most further isoglosses developed between the two are lexical. Proto-Cuicatec is pulling off eastward.

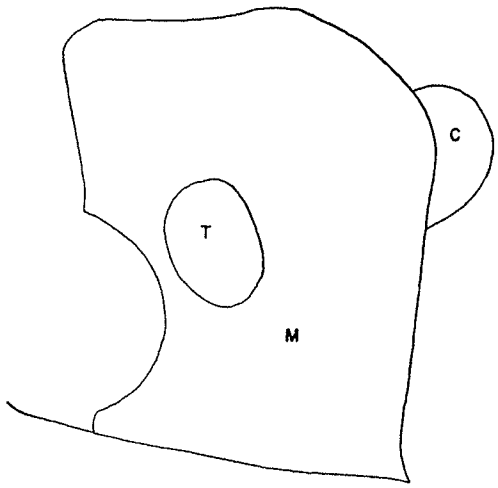


FIG. 13—STAGE 4. Period of Mixtec expansion results in Trique being engulfed by Mixtec on all sides with consequent cultural domination (late Mixtec loans in Trique). Cuicatec is off to the east but still in contact with Mixtec.

a more satisfactory classification of Middle American Indian languages. For example, careful reconstruction of Proto-Otomanguean will make possible some careful judgments as to relative degrees of affinity between Mixtecan, Popolocan, Chiapanec-Manguean, Zapotecan, Chinantecan, and Otopamean; for some preliminary speculations along this line see appendix F. Meanwhile, a somewhat agnostic classification of these languages that simply lists them as related families is preferable to one that at-

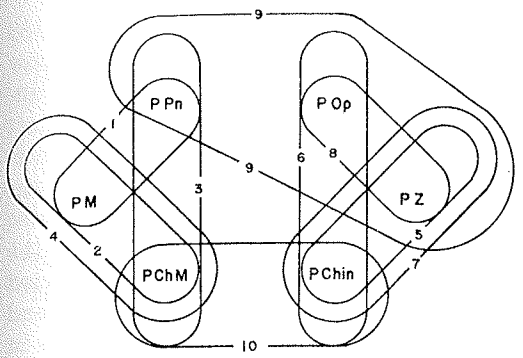


FIG. 14—INFERRED SHARED INNOVATIONS IN OTOMANGUEAN AS NUMBERED IN THE TEXT. Abbreviations: PPn (Proto-Popolocan); PMx (Proto-Mixtecan); PChM (Proto-Chiapanec-Manguean); POp (Proto-Otopamean); PZ (Proto-Zapotecan); PChin (Proto-Chinantecan).

tempts to sub-group them—as, e.g. the dubious sub-grouping 'Otomanguean' (Popolocan, Mixtecan, Otopamean, Chiapanec-Manguean) versus 'Macro-Otomanguean' ('Otomanguean' plus Zapotecan and Chinantecan). Such a classification as Lamb's (1959) is, therefore, quite satisfactory in this respect. Within each family comprising Otomanguean, systemic comparison will eventually elucidate the relationships in terms of such schemes as those here sketched for Mixtec, Cuicatec, and Trique in 4.5. above. Some detailed problems which should be cleared up in the near future are: (a) precise position of Amuzgo (within Mixtecan proper or in some more remote relationship to Mixtecan within Otomanguean?); (b) sub-groupings within Otopamean; (c) position of Chatino (assumed to be Zapotecan).

For criticism of existing classifications within Uto-Aztecan, see Section 4.3. above. In view of the perennial interest in this language family, it is hoped that someone will soon take up Whorf's work where he left off and give us a basis for clarifying problems that he raised.

'Mexican Penutian' is still largely in the speculative stage. While it is generally assumed that Zoquean, Mayan, and Totonac-Tepehua go together, little systemic reconstruction has yet been offered to validate this.[32] Nevertheless, if this grouping be valid such work is entirely possible and needs to be undertaken.

Two languages—Huave and Tarascan—remain mavericks. Hauve was classified by Radin (1916) with Zoquean-Mayan-Totonac, but this relationship is now separately questioned by McQuown and Swadesh. Swadesh (1960) has in turn suggested that Huave fits into his 'Macro-Mixtecan' (i.e. Otomanguean minus Chiapanec-Mangue). But, here again, his evidence, like Radin's for Mayan-Huave affinities, is not conclusive. Swadesh also cites resemblant forms between Tarascan and other American Indian languages involved in vague, quasi-universal relationships.

The above considerations are embodied in a linguistic map of Middle America which is presented in Appendix F. On this map Hokaltecan, Uto-Aztecan, Mayan-Zoquean-Totonacan, and Otomanguean are posited as broad groupings, each involving a number of languages. The reality of the Uto-Aztecan grouping has been demonstrated. Current work in various branches of Otomanguean brings ever nearer the day when the reality of this grouping (plus or minus a few languages or language families) will be demonstrated. The other two groupings—Hokaltecan and Mayan-Zoquean-Totonacan—may be regarded simply as probable. Tarascan, Huave, and Xinca-Lenca are considered to constitute three further groups of limited membership. Extinct languages are included in the classification whenever identification of the affinities of such a language seems to be probable; otherwise, extinct languages are plotted geographically without inclusion in the classification.

Systemic reconstruction, vigorously pur-

[32] Swadesh's 'proto-Penutioid' sketch (1956) is on much too broad a front (his 'Penutioid' includes, among other languages and families, Zuni, Totonac, Huave, Tarascan, and Quechua-Aymara) to afford systemic and detailed validation of such a less ambitious (but still problematical) grouping as 'Mexican Penutian.'

sued, will be successful in demonstrating some relationships, but will have to leave other relationships unproven or doubtful, i.e. thresholds will be encountered beyond which systemic reconstruction does not yield results. But, far from being a drawback, this will constitute a criterion for identifying relevant and demonstrable relationships. Within the proper sphere of application the comparative method will lay a basis for sub-grouping and proper evaluation of the results of such methods as glottochronology which are in danger of being run into the ground by free-lance practitioners. Ethnolinguistic implications in terms of related languages, in terms of identification of the nature and direction of diffusions versus inherited items, and in terms of *Wörter-und-Sachen* interpretation of reconstructed vocabularies, will probably be forthcoming as well. Some such preliminary results are already in (Longacre-Millon, 1961). In brief, systemic comparison and reconstruction of Middle American Indian languages is about to put into orbit a satellite or two with considerable pay-load. This is, therefore, no time to discourage, or discontinue, research.

APPENDICES

A. Some apparent anomalies of phonological development resolved by positing Proto-Mixtecan **am* rather than **ɔ*.

I had postulated a special Trique reflex *e/i* of **a* after **w* (giving sound correspondence M *a* ~C *a* ~ T *e/i*) in my sets 179–87. I had also postulated a special Cuicatec reflex *i* of **ɔ* after **w* (thus giving sound correspondence M *a* ~ C *i* ~ T *a*) in sets 226–8. These sets may now all be considered to be reflexes of Proto-Mixtecan **wa(m)*. It is not at all strange that reflexes of **wam* are rare in that there is a tendency in this language family to avoid two bilabials in the same form. Therefore we have more sets witnessing to **wa* (sets 179–87) than to **wam* (sets 226–8). Furthermore, in set 183 I had a troublesome residue which is now seen to be perfectly regular. This set apparently was to be reconstructed as a set with ultimate **wa* in that the Mixtec vowel reflex was *a* and the Trique vowel reflex was *i*. Nevertheless, there was a possible Cuicatec cognate with vowel reflex *i*. Leaving aside the Trique reflex, I could have reconstructed **ɔ* on basis of M *a* ~ C *i*. The Cuicatec and Trique reflexes offered contrary witness in regard to **a* versus **ɔ*, and yet I had no evidence anywhere in Mixtecan of vowel gradation in general, nor of vowel gradation **a/*ɔ* in particular. Now, however, by recognizing that **wa(m)* underlies this set we may attribute the Cuicatec form to **wam* and the Trique form to **wa* without the least difficulty, in that presence versus absence of **-m* is a familiar phenomenon. Furthermore, I had another such anomaly in my set 188 where apparently the correspondence M *a* ~ C *a* ~ T *a* appeared after **w* but where the Trique reflex 'should' have been *e/i*. Here we now attribute the Trique reflex to **am* which regularly gives Trique a after **w*.

B. Otomí-Mazahua tone correspondences and Proto-Otomí-Mazahua tones (as reconstructed by Donald Stewart).

(1) Two Otomí dialects cited have tone systems with high pitch, low pitch and rising pitch; a third Otomí dialect has high, low, and falling pitches. (2) These reconstruct as follows:

O_1	O_2	O_3	POt
high	high	rising	*high (disyllabic)
high	high	low	*high (monosyllabic)
low	low	falling	*low
rising	rising	low	*rising

(3) Mazahua has high, low, rising and falling pitches. (4) These Mazahua pitches reconstruct with the Proto-Otomian pitches as follows:

POt	Mz	POMz
*high	high	*high (before Mz h or vl. cons.)
*high	low	*high (elsewhere in Mz)
*low	low	*low
*rising	high	*rising (before Mz h or vl. cons.)
*rising	falling	*rising (elsewhere in Mz)
*high	falling	*falling

C. Possible Proto-Chiapanec-Manguean reflexes of 'deictics' posited by Gudschinsky for Pre-Proto-Popolocan-Mixtecan.

Postposed elements *-*me* and *-*mu*, which appear with great frequency in Proto-Chiapanec-Manguean form an interesting parallel with Proto-Popolocan-Mixtecan *-*m* or **mV* (see above, Sec. 2.7.); and may constitute archaic survivals of a similar early feature in Proto-Otomanguean. Gudschinsky's Pre-Proto-Popolocan-Mixtecan deictics *-*xV* and *-*ʔV* (plus or minus *-*m*) seem witnessed to as well. Thus, postposed *-hV* elements are fairly common in Chiapanec-Mangue where their archaic nature is seen in the fact that they have often come to function as stems or parts of stems. Such a Proto-Chiapanec-Manguean form as **ai* 'sick, pains' seems to be cognate with Proto-Mixtecan **(k)aʔi* (Mak-Longacre, p. 38). The state of the Chiapanec-Mangue recorded data is such that nothing can be inferred as to presence or absence of Chiapanec and Mangue *ʔ*. At any rate, whether or not *ʔ* was found in Chiapanec and Mangue forms for 'sick', it is possible that Proto-Chiapanec-Mangue **ai* witnesses to an earlier **a* plus postposed *-*ʔi* (Gudschinsky's *-*ʔV* deictic?). In one of the sets of Fernández de Miranda and Weitlaner there is a postposed Proto-Chiapanec-Manguean *-*Mi* (PPM **hmi*) while in other sets this element seems to have become part of the stem. Putting together these various scattered clues, it appears possible that Chiapanec-Mangue witnesses to *-*mV*, *-*hV*, *-*ʔV* and **hmV* postposed elements all of which figure prominently in Popolocan-Mixtecan as well.

D. Further Zapotec data of possible corroborative value in reference to the posited development: Proto-Otomanguean **mC* cluster > Proto-Zapotec **CC* (geminate) cluster.

Another datum further points in the direction of supposing that Zapotec geminated consonants came from Pre-Proto-Zapotec (presumably Proto-Otomanguean) **mC* clusters. Fernández de Miranda reconstructs **ll* and **l* for Proto-Zapotecan. It may be significant, however, that while there is a uniform reflex *l* for **l* in all dialects, there is the following set of reflexes for **ll* (abbreviations A, VA, M, Cu, Co, R, and I are for Zapotec dialects cited in Fernandez's study; I represents Isthmus dialect of Zapotec): *ll* in A, VA, M corresponds to *l* in Cu, Co, and R; and corresponds to *nd* in I. (e.g. 'stinky': A *rillàaʔ*, VA *ǰlla-*, M *lla-*, *te-laʔ*, Co *laaʔ*, R *rila̋ʔ*, I *rinda̋ʔ*; and 'snake' A *bèllà*, VA *bèll*, M *bäll*, Cu *mbelʔ*, Co *ᵐbälʔ*, R *bála*, I *beʔenda̋ʔ*. Here it seems plausible to suppose that the Proto-Zapotecan opposition was not **ll* versus **l* (the latter is mentioned above without illustrative cognate sets), but rather **nl* versus **l*. Swadesh (1947) posited cluster **nl* as well as **ll*, but both do not seem to be necessary in the light of further data. Thus, while most Otomanguean **mC* clusters were reduced to geminates in the transition to Proto-Zapotecan, at least **nl* (originally **ml*) was not so reduced, and remains as an indication of the original state of affairs.

Another isolated datum occurs to lend plausibility to the above hypothesis: Thus, Fernández de Miranda reconstructs Proto-Zapotecan */ttuttu ca* ~ **ttatta wica* 'daily' on basis of A *tutu cá*, VA *toto zá*, and R *tutuǰa* versus Cu *atata wiz* and Co *tä wižʔ*. It seems clear from other Otomanguean language families that there is a definite tendency toward dissimilation of labial consonants in

the same form. Supposing, therefore, a final *-*m* on the Proto-Otomanguean form lying behind Proto-Zapotecan **ttatta* (probably itself a reduplication) we obtain **ttuttu ca* from a Proto-Otomanguean form with final *-*m* and **ttatta wica* from a Proto-Otomanguean form with final *-*m* reduced to zero before following **w*.

E. Some possible isoglosses conjectured for Proto-Otomanguean.[33]

It is not possible yet to draw definitive isoglosses for Otomanguean as a whole in that shared innovations can not unambiguously be distinguished from shared retentions until Proto-Otomanguean itself is reconstructed in some detail. Nevertheless, we here suggest a few possible shared innovations obtained on comparing qua systems, the six phonological systems reconstructed for the various families of Otomanguean. Consequently, these inferred shared innovations are not based on systemic reconstruction of Proto-Otomanguean itself, but on extrapolation from reconstruction in less depth. As such, the suggestions of this section involve considerable speculation and conjecture.

Some inferred Otomanguean shared innovations are:

(1) Disappearance of *p in Proto-Mixtecan and Proto-Popolocan (by merger with *k^w?).

(2) Development of a prenasalized series of consonants in Proto-Mixtecan and Proto-Chiapanec-Manguean (by reduction of other *mC and/or *nC clusters).

(3) Development of *ny or *$\tilde{n}$ in Proto-Popolocan and Proto - Chiapanec - Manguean (may be simply a shared retention?).

(4) Merger of *t^y and *t in Proto-Mixtecan and Proto-Chiapanec-Manguean.

(5) Disappearance of *x^w in Proto-Zapotecan and Proto-Chinantecan (cf. also the disappearance of *x in Proto-Zapotecan).

(6) Disappearance of *w and *y in Proto-Chinantecan and Proto-Otopamean with consequent development of a considerable number of vowel clusters.

(7) Development of *r (in contrast with both *l and *y, or with a phoneme of which *[l] and *[y] are allophones) in Proto-Zapotecan and Proto-Chinantecan. (Swadesh, 1960, reconstructs Proto-Otomanguean *r on what seem to me to be insecure data.)

(8) Development of contrasting geminated consonant versus single consonant (by assimilation of *mC > *CC?) in Proto - Zapotecan and Proto - Otopamean. Possibly parallel is development of voiced versus voiceless consonant series in Proto-Chinantecan. (Swadesh, 1960, posits Proto - Otomanguean geminated consonants; I do not believe that this will prove to be valid.)

(9) Disappearance of *$\tilde{i}$ (assuming its existence in Proto-Otomanguean) in Proto-Popolocan, Proto-Zapotecan, and Proto-Otopamean.

(10) Disappearance of *k^w in Proto-Chiapanec-Manguean and Proto-Chinantecan. (Swadesh, 1947, does not reconstruct *k^w in Proto-Zapotecan, but this phoneme is reconstructed by Fernández de Miranda.)

It is interesting to note that inferred Proto-Otomanguean isoglosses (1)–(4) apparently group Proto-Mixtecan, Proto-Popolocan, and Proto-Chiapanec-Manguean while (5)- (8) apparently group Proto-Zapotecan, Proto-Chinantecan, and Proto-Otopamean. Number (9), however, groups Proto - Popolocan with Proto - Zapotecan and Proto-Otopamean; while number (10) groups Proto - Chiapanec - Manguean and Proto-Chinantecan. This preliminary exploration suggests the possibility that Proto - Popolocan, Proto - Mixtecan, and Proto-Chiapanec-Manguean may form a group as over against Proto-Zapotecan,

[33] cf. note 1.

Proto-Chinantecan, and Proto-Otopamean. But the latter two isoglosses listed, if reinforcable by other similar features, could result in a pattern of mutually conflicting isoglosses which would make any such dichotomy artificial. The isolation of Chiapanec-Manguean from the rest of Otomanguean within historic times, i.e. at least from 1525 on, is at all events, not to be considered as indicative of any necessary remoteness of relationship within Otomanguean. Otopamean lies to the northwest of the mass of Otomanguean languages. It is therefore plausible to expect that Otopamean is more closely related to the Mixtecan and Popolocan languages which lie geographically nearer, than to Chinantecan and Zapotecan which lie more to the east (with Mixtecan and Popolocan intervening). Nevertheless, we must follow the evidence wherever it leads—when that evidence is finally in. It may eventually be established that the placement of Otopamean is with the eastern rather than with the western languages. Resolution of this problem must await detailed reconstruction of Otomanguean. We have gone about as far as we can go until this further step be taken. At any rate, we must be prepared for some surprises. Thus, while I would personally be inclined to consider Zapotecan closer to Mixtecan and Popolocan than to Chinantecan or Otopamean, nevertheless it may be that the extensive cultural contact of Mixtec and Zapotec has resulted in certain diffusions which obscure the original picture of dialect split and development.

F. Linguistic map of Middle America. The accompanying map indicates the probable distribution of speakers of various Indian languages as spoken in Middle America at the time of the conquest. It is based on the previous linguistic maps of Mendizábal and Jiménez (1936, 1941), Frederick Johnson (1940), and McQuown (1955). Johnson's map is accompanied by a helpful discussion of the sources for such a map and of the problems encountered in drawing it. McQuown's maps treat not only Mexico and Middle America but South America as well; like Johnson's map they are accompanied by a helpful discussion. All three maps embody linguistic classifications which agree in broad outline and diverge in details.

The map here presented includes most languages referred to in any source as spoken in the area at the time of the Conquest. It includes more languages than are found on either the Mendizábal-Jiménez map or the Johnson map. McQuown lists along with his map some additional extinct languages not included here, while Arana and Swadesh (1959) in their series of linguistic maps (based on glottochronology as well as on former classifications) list still more languages than McQuown. The sources contain, in fact, a welter of names and labels for various peoples and languages. It is by no means uncommon to find the same language or dialect referred to by several names or to find the same name apparently applied to quite unrelated languages. In letting the map reflect a level of medium complexity we have hoped to avoid the diminishing returns which might result from further cluttering with names of doubtful linguistic import. In deciding which 'languages' to include we have followed the rule of requiring inclusion on two of the previous maps on which the present map is based. Practically, this has meant inclusion of any 'language' mentioned by either Mendizábal-Jiménez or Johnson since McQuown's rather full listing scarcely ever omits what either of the other maps mention.

The geographical limits of the map are meant to correlate roughly with the geographical delimitation of Mesoamerica as a cultural entity. A broken line is given on the map to indicate the northern frontier of Mesoamerica as indicated some time ago by Kirchhoff (1943). On the south, the frontier of Mesoamerica seems to be at once a linguistic and a cultural line of demarcation. Beyond the Maya, the Lenca, and the southern outliers of Otomanguean are groups

which seem to have South American affinities both linguistically and culturally. Beyond this line the only linguistic group which we indicate on the map is the Jicaque which we include because they seem to be related linguistically (but not culturally) to Hokaltecan. On the north, the cultural line delimiting Mesoamerica cuts across linguistic lines. Thus, while most of Otomanguean lies in Mesoamerica, Otopamean splits with Pame and Chichimeca-Jonaz lying outside Mesoamerica. Each of the main divisions of Uto-Aztecan (Piman, Taracahitian, and Aztecoidan) are found on both sides of the boundary. Of the Hokaltecan languages, Yuman, Serian, and Coahuiltecan are found outside the northern boundary of Mesoamerica, while Tequistlatecan and Supanecan are found well to the south within Mesoamerica. Because of the linguistic relations involved we have, therefore, included on the map the area up to the Mexican-United States border.

In plotting the languages geographically we have followed Mendizábal-Jiménez more closely than the other two previous maps but have occasionally modified Mendizábal-Jiménez in the direction of distributions indicated by the other two cartographers. Methodologically, the present map is a compromise between Mendizábal-Jiménez and McQuown. The former attempt to mark boundaries not only between language families and stocks but between the individual languages as well; the latter indicates no boundary lines as such but simply plots languages on the map by means of code numbers. In the present map I use hatchings, lines, and shadings to mark off languages of different stocks and families but use only code numbers to indicate individual languages within a family (parentheses inclose code numbers referring to extinct languages). It is hoped that the boundaries shown—even if often arbitrarily drawn—and groupings indicated will contribute to the intelligibility and usefulness of the map; while it is also hoped that absence of boundary lines between individual languages of each family will not only spare my making further arbitrary decisions but will amount to an overall simplification.

A few specific points in which the present map differs from one or more of the previous maps are here mentioned: (1) Classification-wise I leave Guaycuran unclassified; refer to Mayan-Zoquean-Totonacan as Macro-Mayan rather than as 'Mexican Penutian' or as 'Macro-Penutian'; reject the sub-grouping 'Otomanguean' within 'Macro-Otomanguean' in favor of a more agnostic grouping of the six families[34] concerned into a stock termed simply Otomanguean; retain Coca and Tecuexe as Aztecoidan (Mason and Johnson) in spite of the doubts of Mendizábal-Jiménez and McQuown; do not attempt to set up sub-groupings within Mayan; classify Jicaque as Hokan; consider the 'Mazatec' of Jalisco, Guerrero, and Tabasco to be three unclassified languages which possibly share only a name in common with Mazatec within Otomanguean; consider the 'Chontal' of Guerrero to be an unknown language not necessarily related to the Chontal of Oaxaca (Tequistlatecan); consider Tapuchultec I and Aguacatec II to be Zoquean; follow Mendizábal-Jiménez in leaving Pison-Janambre unclassified rather than as Coahuiltecan (McQuown and Johnson, but see Mason's accompanying discussion of the problem, 1940, pp. 62–63); reject Mendizábal-Jiménez's classification of Guachichil as Coahuiltecan but follow Johnson and McQuown in classifying it as Uto-Aztecan (Aztecoidan). (2) Apache, Toboso, and Olive are eliminated from the map of Mexico on the supposition that they came into Mexico after the conquest (Mason, 1940, pp. 60, 62). (3) A few Mendizábal-Jiménez distributions are modified as follows: (a) The area of Tepehuan-Tepecan is

[34] As stated in note 1, I now regard Amuzgo as constituting a seventh branch of Otomanguean. Although Amuzgo is so represented on the map I have not tried to bring the body of the article in line with the map in this regard.

enlarged somewhat at the expense of Cozcan (following Johnson) on its southern boundary. (b) The Guachichil enclave in Tamaulipec, a Southern Tepehuan enclave in Cora, and a Chiapanec enclave in Zapotec are omitted following Johnson. (4) A few Johnson distributions are modified also: (a) Maribichicoa is put in Salvador (following McQuown) rather than in Nicaragua.

A complete key for the map is given in the caption to figure 15.

REFERENCES

Arana Osnaya, 1953, 1957
Arndt, 1959
Bartholomew, 1959, 1960
Bernal and Dávalos Hurtado, *see* Huastecos, Totonacos y sus Vecinos
Bright, 1956
Dahl, 1953
Ecker, 1939
Ellegard, 1959
Fernández de Miranda, 1951, 1960
—— and Weitlaner, 1961
Gamboa, 1942
Greenberg and Swadesh, 1953
Gudschinsky, 1955, 1958a, 1959c
Haas, 1959
Hale, 1958, 1959
Hall, 1958
Hamp, 1958, 1960
Hay, *see* The Maya and their Neighbors
Hoeningswald, 1960
Hymes, 1959, 1960
Johnson, F., 1940
Kirchhoff, 1943
Kroeber, 1960a
Lamb, 1958, 1959
Longacre, 1957, 1961a, 1961b
McQuown, 1942, 1955a, 1956
Mak and Longacre, 1960
Mason, 1940, 1952
Mayers, 1960
Millon and Longacre, 1961
Newman and Weitlaner, 1950a, 1950b
Pike, E. V., 1954
Radin, 1916
Ravicz and Romney, n.d.
Smith, P., and Weitlaner, 1957
Stewart, 1957
Swadesh, 1947, 1956, 1959b, 1959e, 1960a
Tax, 1960
Thieme, 1958
Vivó, 1941
Wallis, 1956
Weinreich, 1958
Weitlaner, 1942
Whorf, 1935
Wonderly, 1949, 1953