527 4205

The New Encyclopædia Britannica

Volume 22

MACROPÆDIA

Knowledge in Depth

THE UNIVERSITY OF CHICAGO

"Let knowledge grow from more to more
and thus be human life enriched."

The *Encyclopædia Britannica* is published with the editorial advice
of the faculties of the University of Chicago.

Additional advice is given by committees of members drawn from
the faculties of the Australian National University, the universities
of Adelaide (Australia), British Columbia (Can.), Cambridge (Eng.),
Copenhagen (Den.), Edinburgh (Scot.), Florence (Italy), London
(Eng.), Marburg (Ger.), Melbourne (Australia), Oxford (Eng.),
Queensland (Australia), the Ruhr (Ger.), Sussex (Eng.), Sydney
(Australia), Toronto (Can.), Victoria (Can.), Waterloo (Can.), and
Western Australia; the Complutensian University of Madrid (Spain);
the Institute of Public Affairs (Australia); the Max Planck Institute
for Biophysical Chemistry (Ger.); the New University of Lisbon
(Port.); the School of Higher Studies in Social Sciences (Fr.); Simon
Fraser University (Can.); and the State University of Leiden (Neth.).

the other hand, thinks in terms of the cyclic recurrence of a single phenomenon. Whorf attempts to support this idea by reference to Hopi ceremonial behaviour, which involves repeated preparation for future events. If, in the Hopi view, each day is really a recurrence, rather than something new, then it is reasonable to believe that the daily repetition of ceremonial acts will have a cumulative effect on the future. As Whorf says, the Hopi belief is diametrically opposed to the English proverb that "To-morrow is another day."

More investigation is necessary to either prove or dis-prove the Whorfian hypothesis. In any case, the diversity of American Indian languages and cultures has contin-ued to provide a rich laboratory for investigation. A particularly interesting problem is found in the area of northwestern California, where several small tribes have very similar cultures, but use languages of very diverse types. These are Karok, genetically classified as Hokan; Yurok and Wiyot, which are Algonkian; and Hupa and Tolowa, Athabascan languages. By the Whorfian hypoth-esis, one might expect that the difference in languages would have produced a greater diversity in the cultures; or failing that, one might expect the languages to have grown more similar to each other. In fact, both linguis-tic diversity and cultural uniformity seem to have made modest accommodations to each other. As an example of Whorfian linguistic determinism, the systems of biologi-cal taxonomy of Yurok and Tolowa, referred to in the previous section, may be noted. The Yurok have a larger number of generic classifications, which means they have more choice in nomenclature, because either a generic or a specific term can be used. This is consistent with the high degree of choice afforded in Yurok grammar, in which word order is nearly free and many morphological categories are optional. The sparser taxonomy of Tolowa offers less choice, corresponding to a much more rigid grammatical structure.

A different kind of relationship between language and culture is of more interest to the student of North Amer-ican prehistory, namely, the fact that language retains traces of historical changes in culture and so aids in re-constructing the remote past. Here again the pioneering work was done by Sapir, who pointed out, for instance, that the original home from which a group of related languages or dialects has dispersed is more likely to be found in the area of great linguistic diversity; e.g., there are much greater differences in the English dialects of the British Isles than of the more recently settled areas such as North America or Australia. To take an Amer-ican Indian example, the Athabascan languages are now found in the Southwest (Navajo, Apache), on the Pacific Coast (Tolowa, Hupa), and in the Western Subarctic. The greater diversity of the Subarctic languages leads to the hypothesis that the original centre of Athabas-can migration was from that area. This northern origin of the Athabascans was further confirmed in a classic study by Sapir in which he reconstructed parts of pre-historic Athabascan vocabulary, showing, for example, how a word for "horn" had come to mean "spoon" as the ancestors of the Navajo migrated from the far north (where they made spoons of deerhorns) into the Southwest (where they made spoons out of gourds). The correlation of such linguistic findings with the data of archaeology holds great promise for the study of American Indian prehistory.

Writing and texts. Although a writing system was in use among the Mayas of Meso-America at the time of first European contact, none was known in North America. All writing systems that have been used for North Amer-ican Indian languages have resulted from the stimulus of European writing, or have actually been invented and introduced by whites. Perhaps the most famous system is that invented by Sequoyah, a Cherokee, for his native lan-guage. It is not an alphabet but a syllabary, in which each symbol typically stands for a consonant-vowel sequence. The forms of characters were derived in part from the English writing system, but without regard to their English pronunciation. Well suited to the language, the syllabary fostered widespread literacy among the Cherokee until

their society was disrupted by government action; its use, however, has never died out, and attempts are now being made to revive it.

Other writing systems, invented by missionaries, teachers, and linguists, have also included syllabaries; e.g., for Cree, Winnebago, and some northern Athabascan languages. Elsewhere, alphabetic scripts have been used, adapted from the Roman alphabet by the use of additional letters and diacritics. White educational policy, however, has gener-ally not encouraged literacy in Indian languages. A rich oral literature of American Indian myths, tales, and song texts has been in part published by linguists and anthro-pologists, and there is now increasing encouragement for the training of Indians to transcribe their own traditions— e.g., among the Navajo. It is possible that there may yet be a flowering of American Indian literature, not only in spoken but also in written form.   (W.O.B.)

Meso-American Indian languages

Meso-American, or Middle American, Indian languages are spoken in an area of the aboriginal New World that includes central and southern Mexico, Guatemala, Belize (British Honduras), El Salvador, parts of Honduras and Nicaragua, and part of northwest Mexico. Though various centres of civilization have flourished in the area, some-times concurrently, from 1000 BC down to the time of the Spanish conquest of Mexico in 1519, Meso-America as a whole has had a more or less common cultural history for 2,500 years.

Treatments of the languages of Meso-America are cus-tomarily organized on the basis of their genetic relation-ships, and only secondarily on that of geographical distri-bution. Thus, some languages treated as Meso-American are not in fact spoken in Meso-America proper but form linguistic families with languages that are spoken there. For information about languages of northeast, north cen-tral, and northwest Mexico that are not dealt with in this section, see above *North American Indian languages.* For languages of Central America not treated here, see below *South American Indian languages.*

Some 70 Indian languages are spoken today in Meso-America by perhaps 7,500,000 people. When the Spanish conquered Mexico in 1519, there may have been 20,-000,000 people in Meso-America. Within 100 years of the conquest, the Indian population had decreased by 80 percent as a result of war, disease, forced labour, and star-vation. Since then the Indian population has gone back to a higher level, but several languages—have become ex-tinct. Meso-American languages with the greatest number of speakers in the mid-20th century are:

language	number of speakers	family
Aztec	1,200,000	Uto-Aztecan
Yucatec	600,000	
Quiché-Tzutujil-Cakchiquei	1,200,000	Mayan
Mam	450,000	
Kekchi	375,000	
Mixtec	350,000	
Zapotec	400,000	Oto-Manguean
Otomi	450,000	

The study of the Meso-American languages. During the 16th and 17th centuries, some Dominican and Franciscan missionaries devoted themselves to the study of native lan-guages so that priests could deal in religious matters with monolingual Indians. They wrote grammars following a Latin model, devised orthographies applying values used in Spanish or Latin (occasionally inventing new letters), made dictionaries (usually vocabularies or glossaries), and translated Christian texts (confessionals, sacraments, and sermons) into Indian languages. Except for one heroic figure, the Spanish missionary priest Bernardino de Sa-hagún, they neither collected nor fostered the collection of folklore. During this period grammars and dictionaries were written for such languages as Aztec, Zapotec, Mix-tec, Tzeltal, Yucatec, Quiché-Tzutujil-Cakchiquel, Chortí, and Northwestern Otomí. These collections of data served the successors of the first missionaries. During the 18th century, the momentum of such work decreased, and,

guage,
ure,

istory

Numbers
of lan-
guages and
speakers

Table 61: Meso-American Indian Languages

family, branch (or group), language	location	number of speakers	family, branch (or group), language	location	number of speakers
1. Uto-Aztecan (Uto-Nahuan) family 48c*			**5. Jicaque isolate**	NW Honduras	300
Shoshonean (Yutan, Oregonian) division† *34c*			(several dialects or languages)		
A. *Plateau group 18c*			**6. Tlapanec (Subtiaban, Tlapanecan) complex 8c**		
1. Mono, N Paiute—Bannock (complex?)			A. Tlapanec (Yope)	Guerrero	44,300
2. Shoshoni—Gosiute (Goshiute), Comanche (complex?)			B. Subtiaba (Nagrandan) ‖	Nicaragua	extinct?
3. Ute-Chemehuevi, S Paiute (complex?)			Maribio ‖	El Salvador	extinct?
B. Tubatulabal			**7. Oto-Pamean stock 55c**		
C. *Southern California branch 24c*			A. Chichimec (Meco, Jonaz)	Guanajuato	1,000
1. Serrano			B. *Pamean group 18c*		
2. Luiseño, Juaneño			N Pame	San Luis Potosí	3,600
3. *Gabrieleño complex 10c*			S Pame	Hidalgo	
Gabrieleño			C. *Matlatzinca complex 10c*	State of Mexico	
Fernandeño			Matlatzinca (Pirinda)		2,800
4. *Cahuilla complex‡*			Ocuiltec (Atzingo)		a few
Cahuilla			D. *Otomian group 16c*		
Cupeño			1. *Otomi complex 9c*	Hidalgo, Guanajuato,	432,000
D. Hopi			NW Otomi	State of Mexico,	
Sonoran (Mexican) division *39c*			NE Otomi	Querétaro	
E. *Piman group*			SW Otomi		
1. *Piman complex 8c*			Ixtenco Otomi		
Papago (Pima)	Arizona; Sonora	500	2. Mazahua	Michoacán,	221,000
Lower Pima (Nevome)	Sonora	900		State of Mexico	
Tepecano§	Jalisco	a few	**8. Popolocan (Mazatecan) family 25c**		
2. *Tepehuán complex‡*			A. *Chochoan group 13c*		
N Tepehuán§	Sonora	6,300	1. Ixcatec	NW Oaxaca (Santa	200
S Tepehuán§	Jalisco	17,700		María Ixcatlan)	
F. *Yaquian (Taracahitian) branch 23c*			2. *Chocho complex 8c*		
1. *Tarahumara complex 7c*			Popoloc	SE Puebla, NW	34,000
Tarahumara (Rarámuri)	Chihuahua	36,600		Oaxaca	
Guarillo		7,200	Chocho	NW Oaxaca	2,500
2. Tubar		extinct	B. *Mazatec complex 10c*		145,500
3. *Cáhita complex 15c*			Mazatec (1)	SE Puebla, N Oaxaca	
Eudeve (Heve)		extinct	Mazatec (2)		
Ópata, Jova		extinct	**9. Mixtecan family 42c**		
Yaqui, Mayo (Cáhita)§	Arizona; Sonora, Sinaloa	26,500	A. Amuzgo	E Guerrero, W Oaxaca	20,100
G. *Coran group 15c*			B. *Greater Mixtecan branch 25c*		
Cora§		?	1. *Mixtec group 15c*		335,100
Huichol§	Nayarit	10,900	Mixtec (1)	E Guerrero,	
H. *Nahuan group 15c*			Mixtec (2)	S Puebla,	
1. *Aztec complex 11c*			Mixtec (3)	W Oaxaca	
C. N Aztec (Nahuatl)§	State of Mexico, Puebla, Hidalgo	} 1,200,000	2. Cuicatec	NE Oaxaca	20,200
W Aztec (Nahual)§	Michoacán		C. Trique	W Oaxaca	18,700
E Aztec (Nahuat)§	Veracruz		**10. Zapotecan family 24c**	Oaxaca	
Pipil§	C America	2,000	A. *Zapotec group 14c*		407,600
2. Pochutec§	Oaxaca coast	extinct	Juárez Zapotec	Ixtlán	
2. Cuitlatec (Teco) isolate	Guerrero	extinct	Villalta Zapotec	Yatzachi	
3. Seri isolate	Sonora coast	400	S Mountain Zapotec	Cuixtla	
4. Tequistlatec complex or group	SE Oaxaca		Valley Zapotec	Mitla, Tehuantepec	
Huamelultec	coastal region	5,000	B. Soltec	Sola de Vega	extinct
Tequistlatec	mountain region	5,000	C. Papabuco	Elotepec	extinct
			D. Chatino	southwest	27,500

*Indicates centuries of separation. †Not spoken in Meso-America. ‡There is some doubt whether these groups should be given the status of complexes. §Sonoran languages spoken in Meso-America. ‖ Varieties of the same language spoken in different countries (and having different names).

after Mexico became independent in the first part of the 19th century, Spanish clerics were ousted, leaving further work on indigenous languages to travellers and gentlemen scholars—mostly people poorly qualified for such a task.

Extent of the studies of Meso-American languages
Modern linguistic techniques for the description of languages were not applied to Meso-American languages until North Americans turned their attention to the area in the 1930s and 1940s. Since then, much professional linguistic work has been done on these languages, especially those of Mexico. Almost every language of Meso-America has been worked on by at least one linguist, but the time spent and level of linguistic competence of the investigators have varied greatly. For most of the languages, grammatical and lexical data have been collected, much of which remains unpublished. A number of competent grammars and dictionaries have appeared; none of them however, is exhaustive or definitive. Folktales have been collected for a smaller number of languages. Spanish-based orthographies have been devised for most of the Meso-American languages in the 20th century, but not much reading matter is available in them. In short much work remains to be done.

CLASSIFICATION

Modern genetic groupings. The classification of Meso-American Indian languages presented here reflects generally accepted genetic groupings (as of the early 1970s), based on similarities in vocabulary and grammar and on the establishment of regular correspondences between sounds in cognate (related) words among the several languages. The languages grouped together are presumed to have developed from a common ancestor, called a protolanguage. Not all of the languages of Meso-America have been convincingly assigned to a specific group. A few of these languages are currently thought to be unrelated to any of the established genetic groupings and are listed individually in the table; these solitary languages are called isolates.

Within a given genetic grouping, there may be several levels of relatedness. Glottochronology (or lexicostatistics), developed by two United States linguists—Morris Swadesh and Robert Lees—is a controversial and not universally accepted procedure for measuring degrees of difference between related languages in terms of years of separation. Based on the assumption that all languages change more or less to the same degree in a given period of time the method employs a list of 100 items of "basic" or "non-cultural" concepts, which are assumed to be expressible by vocabulary items in any language. Over a period of 1,000 years, different words will have been substituted to express 14 percent of the 100 concepts every 1,000 years, two languages that separated 1,000 years ago will share 74

Table 61: Meso-American Indian Languages (continued)

family, branch (or group), language	location	number of speakers	family, branch (or group), language	location	number of speakers
11. Chinantecan group 15c*	N Oaxaca	80,000	2. *Greater Kanjobalan branch 21c*		
Chinantec (1)			a. *Chujean group 16c*		
Chinantec (2)			Tojolabal (Chaneabal)	Chiapas	19,000
Chinantec (3)			Chuj	NW Guatemala	30,000
Chinantec (4)			b. *Kanjobal proper group 15c*		
12. Manguean (Chorotegan, Chiapanec-Mangue) group 13c			i. *Kanjobal complex 7c*	NW Guatemala	
A. Chiapanec	Chiapas	extinct	Kanjobal (Conob. Solomec)		62,000
B. Mangue (Dirian, Nagrandan) ‖	Nicaragua	extinct	Acatec		13,000
Chorotega ‖	Honduras	extinct	Jacaltec		21,000
Nicoya (Orotiña) ‖	Costa Rica	extinct	ii. *Mochó complex* (Cotoque)	SE Chiapas	
13. Huave isolate	SE Oaxaca	25,300	Motozintlec		500
14. Mixe-Zoque (Zoquean, Mixean, Zoque-Mixe) family 36c			Tuzantec		100
			D. *Eastern division 34c*		
A. *Zoquean group 14c*			1. *Greater Mamean branch 26c*		
Zoque	Tabasco, Chiapas, Oaxaca	37,600	a. *Mamean proper group 15c*		
			Teco	SE Chiapas, W Guatemala	5,000
Sierra Popoluca	Veracruz	25,300	Mam	W Guatemala	434,000
Texistepec	Veracruz	3,000	b. *Ixilan group 14c*	NW Guatemala	
B. *Mixean group 13c*			Aguacatec		15,000
1. Sayula	Veracruz	1,000	Ixil		60,000
Oluta	Veracruz	1,000	2. *Greater Quichéan branch 26c*		
E. W Mixe	E Oaxaca	77,500	a. Uspanteco	NW Guatemala	15,000
2. Tapachultec	SE Chiapas coast	extinct	b. *Quiché complex 10c*	C Guatemala	
15. Totonacan family 2bc			Quiché (Achi)		680,000
Totonac	Veracruz, Puebla	239,000	Sacapultec		3,000
Tepehua	Veracruz, Hidalgo	18,800	Sipacapa		3,000
16. Mayan family 41c			Cakchiquel		434,000
A. *Huastec complex 9c*			Tzutujil		50,000
Huastec	San Luis Potosi, N Veracruz	101,000	c. *Pocom complex 10c*	EC Guatemala	
			Pocomam		30,000
Chicomuceltec (Coxoh)	Chiapas	a few?	Pocomchi		75,000
B. *Yucatec (Maya) Complex 10c*			d. Kekchi		374,000
Yucatec	Yucatán, Campeche, Quintana Roo, N Guatemala, Belize	605,000	**17. Tarasco isolate**	SW Michoacán	72,000
			18. Xinca complex 10c	SE Guatemala	
			Eastern Xinca	Yupiltepeque, Jutiapa	extinct
Lacandón	Chiapas	200	Northern Xinca	Jumaytepeque	50
Itzá	N Guatemala	500	Southern Xinca	Chiquimulilla	100
Mopán	N Guatemala, Belize	6,000	Western Xinca	Guazacapan	100
			19. Lencan family 20c		
			Lenca	SW Honduras	25
C. *Western division 30c*			Chilanga	E El Salvador	a few
1. *Greater Tzeltalan branch 19c*			**20. Paya complex 10c**	N Honduras	300
a. *Cholan proper group 14c*					
Chontal (Yocotán)	Tabasco	51,000	**21. Misumalpan (Misuluan) family 43c**		
Chol	Tabasco, Chiapas	109,000	A. Mosquito (Miskito)	Nicaragua, Honduras	115,000
Chorti	Honduras, E Guatemala	64,000	B. *Matagalpa complex 10c*		
b. *Tzeltalan group 14c*	Chiapas		Matagalpa	Nicaragua, Honduras	100
Tzotzil (Quelén)		123,000	Cacaopera	El Salvador	?
Tzeltal		123,000	C. *Sumo complex 11c*		
			Sumo, Ulua, Tahuajca	Nicaragua	200

*Indicates centuries of separation.   †Not spoken in Meso-America.   ‡There is some doubt whether these groups should be given the status of complexes.
§Sonoran languages spoken in Meso-America.   ‖Varieties of the same language spoken in different countries (and having different names).

percent cognates (86 percent of 86 is 74 percent). The following are terms and categories for degree of relatedness, correlated with glottochronological time depths, that will be used to describe the various Meso-American language groups. The figures given are minimal bounds.

term	centuries of separation	percentage of cognates
dialects	0–5	86–100
language complex	7–11	71–81
language group	13–17	60–68
branch (or family if there is no superordinate category)	19–26	45–56
language family	35–45	26–35
stock or phylum	55–65	14–19

In Table 61 every family (group) and isolate has a separate number from 1 to 21. Each of the 21 headings specifies the name of a grouping, with alternative names. Numbers in parentheses following language names indicate that there are several closely related languages all referred to by the same name. For each language grouping the various levels of relatedness are specified, including glottochronological figures (c = centuries), which are Swadesh's, except for Mixe-Zoque, Mayan, and Xincan, which are those of the U.S. linguist Terrence Kaufman. Family and stock names are formed in the following ways: (1) A typical language, usually the most widely spoken, is

suffixed with -an (e.g., Mixtecan). (2) Two typical names are chosen and compounded (e.g., Mixe-Zoque). (3) Parts of two or more language names are joined, and -an is suffixed (e.g., Oto-Manguean, Oto-Pamean, Mis-Uluan/Misumalpan).

Group names end in -an if the groups are further subgrouped but do not end in -an if they are immediately divided into discrete languages.

The map gives the approximate geographical distribution of the 21 language groupings and isolates of Meso-America. None of the extinct undocumented languages is indicated. Except for some outliers, separate languages within a grouping are not localized. An outlier is a language that has been carried into a foreign cultural and linguistic context by migration; e.g., Mangue is a Chiapanec outlier in Misumalpan territory, Subtiaba is a Tlapanec outlier in Misumalpan territory, Pipil is a Nahua (Aztec) outlier in Quichéan, Xinca, Lencan, and Misumalpan territories.

Outliers

In the following paragraphs the numbers in parentheses refer to groupings in Table 61.

Uto-Aztecan (1). The Uto-Aztecan family consists of some 27 languages that are universally recognized to fall into eight groups or branches—the Plateau group, Tubatulabal, the Southern California branch, Hopi, the Piman group, the Yaquian branch, the Coran group, and the Nahuan group. Tubatulabal and Hopi contain just one language each. The first four groups are commonly, but

not universally, recognized as forming a Shoshonean division within the family. None of the Shoshonean languages is spoken in Meso-America, and no distribution or population data is cited for them in Table 61 (see above *North American Indian languages*). There are two common ways of grouping the remaining languages, depending on the position assigned the Nahuan group. Either Nahuan is considered as separate and the rest as forming a Sonoran division, thereby producing three divisions—Shoshonean, Sonoran, and Nahuan—or else Nahuan is included within Sonoran, thereby producing a Shoshonean versus Sonoran dichotomy, which is the arrangement used in this article. Several scholars believe that the "division" concept is faulty here and that Uto-Aztecan contains eight groups and branches that are not to be further grouped in any special way.

Only some Sonoran languages are spoken in Meso-America (indicated by signs [§] in Table 61). The extinct Tubar belongs to the Yaquian branch, but whether to the Tarahumara complex, the Cáhita complex, or neither, is not clear. The Nahuan group includes the extinct Pochutec, formerly spoken on the coast of Oaxaca, Mexico, and poorly documented; Pochutec is clearly very divergent from the rest of the group. The Aztec complex is considered by some to be a single language with several dialects. The three Aztec languages were spoken within the Aztec Empire as it was constituted in 1519. Pipil speakers, who also refer to their language as *nawat*, were not a part of the Aztec culture and probably represent a Toltec expansion from several centuries earlier.

In 1859, Johann Karl Buschmann, a German philologist, correctly identified all the then-known Uto-Aztecan languages as forming a family. In 1883 a French philologist, Hyacinthe de Charencey, divided Uto-Aztecan into Oregonian (=Shoshonean) and Mexican (=Sonoran), and, in 1891, in the United States, anthropologist Daniel Brinton recognized Shoshonean and divided the Sonoran division (of this article) into Nahuatlan (=Nahuan) and Sonoran (=the Sonoran of this article minus Nahuan). Brinton's division was followed by the United States biologist John Wesley Powell in his classification of North American languages.

Buschmann in 1859 and United States anthropological linguist Edward Sapir in 1915 contributed to the comparative study of Uto-Aztecan by assembling sizable numbers of cognate sets.

A number of now-acculturated and racially absorbed Indian ethnic groups of northern Mexico are believed by many to have spoken Uto-Aztecan languages, although only the language names are known, and not the languages themselves. These are: Suma, Jumano, Lagunero, Cazcán, Tecuexe, Guachichil, and Zacatec.

Uto-Aztecan is generally accepted by specialists as related to the Kiowa-Tanoan family of North America and with it to form the Aztec-Tanoan stock (or phylum).

Cuitlatec (2). The now extinct Cuitlatec language has not been linked convincingly with any other language or family, though the idea that it might be related to Uto-Aztecan has been entertained.

The Hokan hypothesis (3–5). In 1919 two United States anthropologists, Roland Dixon and Alfred Kroeber, tried to improve on an older North American classification by reducing the multiplicity of language groupings in California (about 50) to a manageable number of families and stocks. Working over a period of several years, they developed the hypothesis that most California languages belong to one of two great groupings (called phyla or superstocks), Hokan and Penutian. The formulation was accepted and extended by others. Hokan included Shasta, Achumawi, Atsugewi, Chimariko, Karok, Yanan, Pomoan, Washo, Esselen, Yuman, Salinan, and Chumashan. By 1891/92 it had been suggested that Yuman, Seri (3), and Tequistlatec (4) were related. In 1915 the matter was re-examined in the light of the Hokan hypothesis, and it was concluded that all of the languages named above are related. Since then most scholars familiar with Yuman languages have believed that Seri and Yuman are related, and many who accept the Hokan hypothesis believe that Seri and Yuman form a special group within Hokan.

Jicaque (5), which is very poorly documented, though still spoken, has plain, aspirated, and glottalized stops (different varieties of consonant sounds), as do many Hokan languages. In 1953 it was suggested that Jicaque is a Hokan language. The general acceptance of the proposition may have been uncritical, because the available data on Jicaque is hardly reliable.

Extinct languages of northeast Mexico. All of the several languages once spoken in northeast Mexico and South Texas have become extinct. Documented languages of Mexico are: Coahuilteco, Comecrudo, Cotoname, Naolan, and Tamaulipec (or Maratino). Those of Texas are Karankawa (and Klamkosh), Atakapa, and Tonkawa. John Wesley Powell classified the first three as forming a Coahuiltecan family. The other Mexican languages were unknown until recently. Each of the three Texan languages was considered by Powell to be an isolate. In 1920 Coahuiltecan was redefined to include Karankawa and Tonkawa and to be coordinate with Hokan in a Hokan-Coahuiltecan (=Hokaltecan) superphylum.

Tlapanec (6). The Tlapanec complex was first correctly identified by Walter Lehmann, a German physician, in 1920. In 1925 Edward Sapir tried to establish Subtiaba as a Hokan language, proposing some Proto-Hokan reconstructions that could account for the Subtiaba forms. This classification is generally accepted. More recently, however, Calvin Rensch, a U.S. missionary linguist, tried to validate the Oto-Manguean hypothesis (see below) by means of full-scale phonological reconstruction. He believed Tlapanec to be Oto-Manguean; others considered it to be intermediate between Oto-Manguean and Hokan. It must be kept in mind that most of the specialists who have immersed themselves in the study of large numbers of American Indian languages believe that almost all of them are genetically related to one another. This relationship derives from a period, perhaps 20,000 to 30,000 years ago, when some of the languages were still spoken in Asia. With such a point of view, correct grouping (or degree of relationship) is a more interesting question than genetic relatedness.

Oto-Pamean (7). The Oto-Pamean stock contains four groups and complexes, Chichimec, Pamean, Matlatzinca, and Otomian, of which only the last two are spoken within Meso-America. The exact number of languages within the Otomí complex is not yet determined, though there seem to be four. Oto-Pamean was first correctly identified in 1892.

Popolocan (8). The Popolocan family (which might more appropriately be called Mazatecan) was correctly identified in 1926. The exact number of languages within the Mazatec complex has not yet been determined, though there are at least two.

Mixtecan (9). There is some difference of opinion as to how the various languages here included within Mixtecan are to be grouped. The main problem is whether Amuzgo is Mixtecan or a separate branch within Oto-Manguean. It has been included within Mixtecan in some systems and excluded from it in others. There seem to be three languages within the Mixtec group, a subdivision of Mixtecan.

Zapotecan (10). The Zapotecan family was correctly identified by William Mechling in 1912, but only Francisco Belmar, a Mexican philologist, correctly recognized that Papabuco is a separate language, neither Zapotec nor Chatino (in 1905). Belmar, however, incorrectly included Chinantec within Zapotecan. The Chatino language has several dialects. Within the Zapotec complex there are at least four languages, and perhaps more.

Chinantecan (11). The Chinantecan group contains approximately four languages, the exact number as yet undetermined. The separateness of Chinantecan within Oto-Manguean was recognized in 1912.

Manguean (12). The Manguean group was correctly identified by Belmar in 1905. Its members, formerly spoken in Chiapas (Mexico), and in Nicaragua, Honduras, and Costa Rica, are now extinct.

The Oto-Manguean hypothesis (7–12 or 6–13). Ever since 1891, it has been proposed that two or more of the above families (7–12) should be linked. Since about 1925,
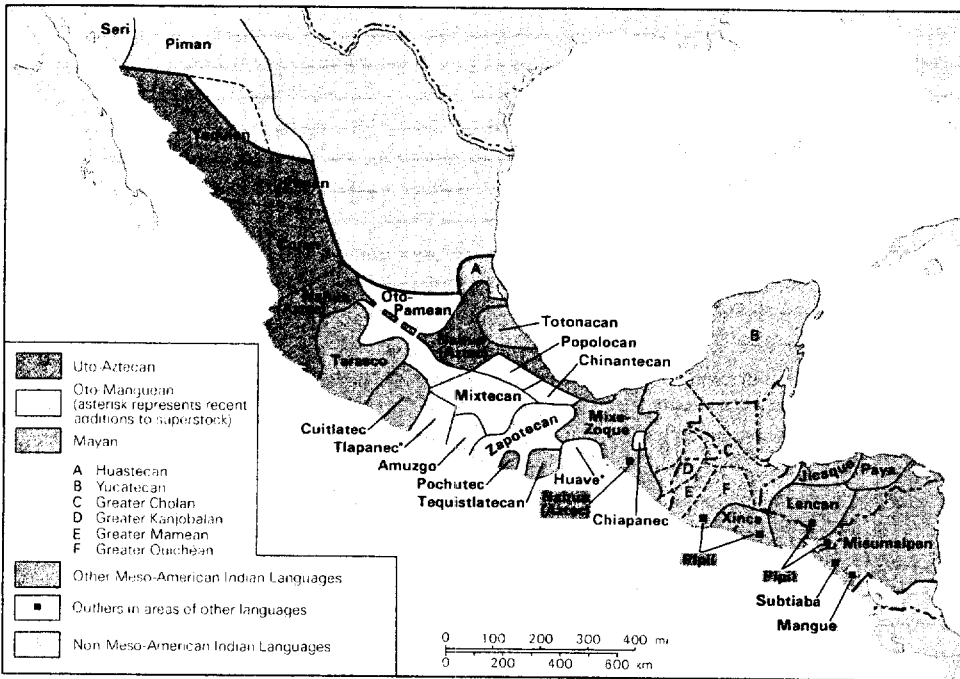
Figure 35: Distribution of Meso-American Indian
languages c. AD 1500. Boundaries are schematic.

it has been generally accepted by specialists that the Oto-Pamean, Popolocan, Mixtecan, Zapotecan, Chinantecan, and Manguean groups form a larger genetic grouping (phylum), commonly labelled Oto-Manguean. This may be called the "classical Oto-Manguean formulation." Since 1950, work has been going on in the reconstruction of parent languages for each of the constituent families and groups. Since 1961, two revisions have been proposed in the formulation of what constitutes Oto-Manguean: the Tlapanec language complex has been recognized as included in or closely related to Oto-Manguean, and Huave has been proposed as an Oto-Manguean language. In the early 1970s, therefore, most Oto-Manguean specialists considered the grouping to consist of groups 6–13.

The comparative study of the Oto-Manguean phylum has resulted in the first case in the Western Hemisphere in which the remote common ancestor of several language families has been phonologically reconstructed. Comparative linguistics at the phylum level has been largely unsuccessful with other postulated superstocks because of the relatively small number of cognates that can be identified. Except for Manguean, all Oto-Manguean languages are spoken in central Mexico.

Huave (13). Early proposals linked Huave to Mixe-Zoque and Mayan. Although this has not been generally accepted by many specialists, it has been uncritically repeated in most compilations. Recently, Morris Swadesh presented a reasonably well documented proposal for Huave as an Oto-Manguean language.

Mixe-Zoque (14). The Mixe-Zoque family consists of eight languages, which, comparative phonology and grammar suggest, form two branches—a Zoquean group, and a Mixean group including Tapachultec. Glottochronological figures, however, suggest a three-way division, as shown in the Table. The Mixe-Zoque family was correctly identified by Hyacinthe de Charencey in 1883. The Texistepec, Sayula, and Oluta languages of this family are all locally called Populuca.

Totonacan (15). The Totonacan family contains just two languages, of which one (Totonac) has at least three dialects. Possibly, Totonac is a complex.

Mayan (16). The Mayan family was correctly identified by a German ethnographer, Otto Stoll, in 1884. This family, with 24 languages and nearly 3,500,000 speakers, is the most diversified and populous language family of Meso-America. The Huastec language is separated by more than 1,000 miles from the nearest other Mayan language. Taken with the fact that the Huastecs did not

share in the Classic Maya civilization, this requires a historical explanation involving the separation of Huastec from the rest of the family more than 2,500 years ago. Though the geographical extent of the Mayan languages is considerable, the Mayan peoples, languages, and cultures (as contrasted with those of the Aztecs), have never been particularly expansionist.

A number of attempts have been made to classify the Mayan languages, each one availing itself of more data than the last. The classification given here as of 1971 recognizes, at the lowest level, ten groupings. Specialists have disagreed on the precise positions of Tojolabal and Chuj, Motozintlec, Aguacatec, Uspantec, and Kekchí and have held no firm opinions about the Yucatec or Huastec complexes. Not much comparative work on the Mayan family has seen print, but much data has recently been collected. The main contributors to Mayan comparative studies have been the U.S. linguists Norman McQuown (1950s and 1960s) and Terrence Kaufman (1960s).

The Macro-Mayan and Macro-Penutian hypotheses. In 1931 L.S. Freeland, a U.S. anthropological linguist, tried to show that Mixe (Zoque) is related to the "Penutian" languages, a superstock that up until then had been limited to California, Oregon, Washington, and British Columbia. In 1935 it was suggested that the similarities between Uto-Aztecan, Tanoan, Kiowa, Penutian, Mixe-Zoque, and Mayan were such as to indicate the existence of a superstock, which it was proposed to call Macro-Penutian. This hypothesis had favour for a period but was never demonstrated nor taken very seriously by specialists. Since then the first three have been generally joined in Aztec-Tanoan. In 1942 it was suggested that Mixe-Zoque and Totonacan might be related genetically to each other and the two in turn might be related to Mayan, the resultant superstock to be called Macro-Mayan. Recently it has been claimed that Tarasco (17) probably belongs in Macro-Mayan as well, though the attempt to prove this has not been convincing to most Mayanists, to whom, minus Tarasco, the Macro-Mayan hypothesis seems as reasonable as the Hokan hypothesis.

Tarasco (17). Tarasco has been linked genetically by some not only to Marco-Mayan but also to both Zuni (in North America) and Quechua (in South America), but without general scholarly acceptance.

Xinca and Lencan (18–19). It has been suggested that Xinca and Lencan are related and that one or both of them is related to Mayan (16), Chibchan (in South Amer-

(margin left:) Reconstruction of Oto-Manguean phonology

(margin right:) Hypotheses concerning superstocks

ica), or Uto-Aztecan (1). None of these hypotheses has been demonstrated as probable.

Languages outside Meso-America proper. The Paya language (20) and the Misumalpan family (21) are Central American languages spoken outside of the cultural area of Meso-America proper, though they have Meso-American outliers in their territory. Paya (20) has been linked in hypotheses to Chibchan and Cariban (both in South America), and perhaps to others, but not convincingly. The Misumalpan family (21) has been recognized since 1895. Since that date some scholars have believed that the three languages and complexes listed are coordinate, and others have believed that the first two constitute one group and the other consitutes a second group. Although the family relationship can be verified on inspection, no supporting comparative work has been published. Previous comprehensive classifications of the Meso-American Indian languages were presented by the U.S. anthropologists Cyrus Thomas and John R. Swanton in 1911 in *Indian Languages of Mexico and Central America and Their Geographical Distribution,* by Edward Sapir in the 14th edition of *Encyclopædia Britannica* (1929), and by Morris Swadesh in 1967 in *Handbook of Middle American Indians.*

Newly discovered languages and reconstructions. Although there are probably no uncharted areas in Meso-America, it is not necessarily the case that all the Indian languages of Meso-America have been correctly identified, and there are probably some multilingual Indian communities as well that are not known to be such. In 1967 Terrence Kaufman discovered a hitherto undocumented Mayan language spoken by several hundred Indians in four or five towns in southeast Chiapas and west central Guatemala. Although it appears to be closely related to Mam, Kaufman considered it a separate language and christened it Teco. Kaufman identified two more new Mayan languages in the course of a linguistic survey of Guatemala. These two new languages—Sacapultec (formerly considered Quiché) and Sipacapa (formerly assumed to be Mam)—are not documented in print and both belong to the Quiché complex.

Reconstruction of earlier forms of the Meso-American Indian languages has focussed primarily on phonology and vocabulary. Phonological and lexical comparative studies as well as reconstruction have been done for the following groups: Uto-Aztecan; Oto-Manguean—Oto-Pamean, Popolocan, Mixtecan, Zapotecan, Chinantecan, Manguean; Mixe-Zoque; and Mayan (in part). A small amount of grammatical comparison has been done within Oto-Manguean and Mixe-Zoque. In addition, some studies have been done of reconstructed vocabulary for the purpose of hypothesizing about the culture of the speakers of the protolanguages.

RELATION OF LANGUAGES TO HISTORICAL
AND CULTURAL INFLUENCES

Pre-Columbian diffusion. The following are some of the important civilizations that have flourished in Meso-America:

civilization	period	location
Olmec	1200 BC–400 BC	Gulf Coast, Mexico
Monte Albán	400 BC–AD 700	Oaxaca, Mexico
Teotihuacán	AD 100–600	Central Mexico
Classic Maya	AD 300–900	Chiapas, Mexico; Petén, Guatemala
Toltec	900–1200	Central Mexico
Aztec	1300–1500	Central Mexico

The Aztecs spoke Nahuatl, as did the Toltecs. The Classic Maya probably spoke two or three Mayan languages, and the people of Monte Albán probably spoke one or more Zapotecan languages. No one knows what either the Teotihuacán people or the Olmecs spoke, but it has been surmised that at least some Olmecs spoke Mixe-Zoque languages and that the Teotihuacán people may have spoken Otomían languages (though an Aztec tradition says Totonac).

In the pre-Columbian period, there was naturally contact among Meso-American languages and occasional borrowing of vocabulary and other linguistic features. Partly be-

cause of the unavailability of grammars and dictionaries, actual cases of such diffusion have not been much studied.

Some of the known contacts resulting in borrowing are the following: (1) Mixe-Zoque languages (Olmecs?) have given words to Mayan, Mixtecan, Zapotecan, Otomían, Aztec, Lencan, Xinca, and Jicaque; (2) Zapotecan languages (Monte Albán) have given words to Huastec and Yucatec; (3) Mayan languages (Mayas) have given words to Xinca, Lencan, and Jicaque; and (4) Nahuatl (Toltecs and Aztecs) has given words to Mayan, Lencan, other Uto-Aztecan languages, as well as to other Meso-American languages. Words diffused from these sources provide evidence that contact took place. Scholars know that contact must have taken place at particular times and places, and therefore can form hypotheses about where certain languages may have been spoken in the more remote past.

External relationships and contacts. Various scholars have suggested that some Meso-American language or family is related to a language or family (other than Uto-Aztecan) outside of Meso-America. These suggestions are mostly parts of larger attempts to synthesize the language classification of the New World, or of the whole world, and are usually based on the sometimes unexpressed view that all the languages of the Western Hemisphere or even of the whole world are ultimately genetically related. Although the assumption may be true, the proposed connections have been unconvincing to specialists in Meso-American languages. The only generally accepted larger groupings are Hokan and Penutian. Most scholars do not have the breadth of knowledge to be able to evaluate these vast proposals.

One proposal of external relationship probably has some merit. In 1961 it was suggested that Chipaya—a language spoken on the shores of Lake Titicaca in Bolivia—is genetically related to the Mayan languages. The hypothesis, proposed by Ronald Olson, a U.S. missionary linguist, was based on 120 sets of lexical comparisons between Chipaya and Proto-Mayan. The data cited are subject to more than one interpretation, because many of the comparisons involve semantic notions and word forms that are widespread in the Western Hemisphere; also, Chipaya has been so influenced grammatically by Aymara (which all Chipayas can speak) that any grammatical peculiarities it may once have shared with Mayan have disappeared. Because a core of data showing regular sound correspondences remains, it is probably necessary to assume that there is a historical connection between Chipaya and Mayan, possibly, but not demonstrably, a genetic relationship. The connection may have been direct—presumably from Meso-America to Bolivia via land—or there may be other languages in western South America that show prehistoric contacts with Mayan. The acceptance of a prehistoric linguistic connection, neither extremely remote nor extremely recent, between Meso-America and the Andes is quite provocative, inasmuch as other evidence exists for early culture contact between Meso-America and the Andes, Meso-America generally being the donor and the Andes generally being the beneficiary; e.g., in the case of corn. Later diffusion from South America to Meso-America also occurred; e.g., witness the transference of peanuts, metallurgy, hammocks.

Interaction between Spanish and Indian languages. In modern Meso-America, the dominant European language is Spanish. The speakers of all Meso-American Indian languages include some who are bilingual; and a few languages are spoken by almost totally bilingual populations. Most Indian languages spoken by sizable populations have at least 50 percent monolingual speakers. All Meso-American languages with a significant number of bilingual speakers have been influenced by Spanish, primarily in the areas of vocabulary, particles, and word order. Since the Spanish conquest, Meso-American languages have been borrowing words from Spanish, and, because the kind of Spanish spoken has changed somewhat over the years, both in vocabulary and pronunciation, different historical periods are usually distinguishable in lexical borrowings. For a variety of reasons, certain function words, primarily conjunctions and adverbs, are frequently borrowed from Spanish; e.g., *ya* "already," *pero* "but," *hasta* "until," *y*

"and," *o* "or," *ni* "not even," *hasta* "even," *si* "if," *cuando* "when," *porque* "because," *por eso* "therefore, so," *entonces* "then." Some languages have assimilated the Spanish word order of subject–verb–object.

Conversely, the Spanish of Meso-America has been the recipient of vast amounts of lexical material from local languages, primarily Nahuatl. The borrowing has provided names of plants, animals, artifacts, and social forms indigenous to Meso-America and lacking names in Spanish. Among the reasons that Nahuatl has been the primary source is that the Aztecs were the first Meso-American people conquered by the Spaniards; the Aztecs had outposts in many parts of Meso-America; the Spaniards recruited Aztecs, particularly as guides, into their military force to assist their venture of subduing the rest of Meso-America; and, for several decades, Aztec, written in Roman orthography, was used in many parts of Meso-America to keep official records, such as deeds, wills, and censuses.

Many of the words borrowed into Spanish from Aztec have since passed in turn into English; *e.g.,* chili, chile, or chilli (Spanish *chile*), avocado (Spanish *aguacate*), chicle, chocolate, peyote, coyote, tomato (Spanish *tomate*), ocelot (Spanish *ocelote*), guacamole, mescal.

In some parts of Meso-America, because of economic and social conditions, an Indian may speak one or more Indian languages besides his own. This is common in Guatemala, where some areas have been recently colonized by speakers of more than one language, or some communities have received outside settlers in the more remote past.

The names used in this article for the Meso-American Indian languages are English versions of the Spanish terms for them. Only in a few cases are these names the ones actually used by the people who speak the languages in question. First, most of the names are of Aztec origin, because at the outset the Spanish learned of local phenomena primarily via Aztec. Secondly, some languages have no special name of their own, simply being called "our language."

Pre-Columbian writing. Most of the Meso-American cultures shared a mathematical notation and calendrical system that had been developed and diffused in the distant past, probably before 500 BC. At the time of European contact the Aztecs, Zapotecs, Mixtecs, Otomis, Mayans, and perhaps some others were all producing records on stone (inscriptions) and on a type of homegrown paper (produced from the amate tree, *Ficus glabrata*), these latter being commonly called codices. Except for the Mayan system, which probably originated before AD 1, the records cannot properly be called writing, in that it was not possible to represent all of speech, but only numbers, dates, and names (pictographically). The Mayan system, besides representing all these, was also used to represent morphemes (words and word elements) and phonemes (distinctive sounds). Presumably the symbols used in this system (called glyphs) represent individual phonemes, syllables, and morphemes; and they give semantic information as well to take the ambiguity out of homophonous readings. Several scholars have devoted much time to the study of Mayan writing, but, to date, the results have not been very impressive. A few scholars outside the Meso-American field believe the Mayan writing system is purely ideographic and hence inherently undecipherable without a bilingual inscription or text in a known language. All specialists within the Mayan field hold that the Mayan is a mixed ideographic and phonological system.

What may be delaying progress in the deciphering of Mayan writing is the absence of reconstructions for intermediate groupings within the Mayan family (*e.g.,* Proto-Yucatecan, Proto-Cholan, and others) and ignorance of Mayan languages other than colonial Yucatec on the part of the investigators. Efforts are being made to correct these deficiencies, particularly by Mexican specialists. It is not known whether Mayan writing was used to write more than one language and, if so, what the languages were. If only one, it was probably either Proto-Cholan or Proto-Yucatecan. The symbols used in all the pre-Columbian notation systems are obviously pictographic in origin, as was the case in the ancient Egyptian, Sumerian, ancient Chinese, and Indus Valley writing systems.

LINGUISTIC CHARACTERISTICS

In general, all the languages of a particular family are typologically similar to one another both in phonology and grammar. Among the 21 language groupings in Meso-America, there are several types of sound systems and grammatical systems. Because study in this area has hardly begun, nothing very secure can be asserted here, but some general characteristics can be outlined on the basis of data for the following reasonably well-documented languages: Tequistlatec, Otomí, Mazatec, Mixtec, Zapotec, Chinantec, Aztec, Zoque, Totonac, Quiché, and Tarasco.

Phonologically, there is a wide diversity among Meso-American languages. Voiced spirants—*i.e.,* sounds like English *v, z,* or *th* in "then"—are missing from all Meso-American languages. Other phonological features in these languages include a voiceless lateral spirant sound, *lh* (in Tequistlatec and Totonac); a lateral affricate, *tl* (in Aztec and Totonac); a postvelar stop, *q,* in contrast with a velar stop, *k* (in Quiché and Totonac); glottalized vowels (in Zapotec, Zoque, Aztec, and Totonac); glottalized consonants (in Tequistlatec, Quiché, Otomí, and Mazatec); aspirated stops (in Tarasco, Otomí, and Mazatec); voiced stops (in Tequistlatec, Otomí, Mazatec, and Chinantec); prenasalized stops (in Otomí, Mazatec, and Mixtec); nasalized vowels (in Otomí, Mazatec, Mixtec, and Chinantec); a labiovelar stop, *kw,* sometimes contrasting with a bilabial stop, *p* (in Otomí, Mazatec, Mixtec, Aztec); tone and stress accent (tone in Otomí, Mazatec, Mixtec, Chinantec, Zapotec; stress in Tarasco and Tequistlatec); and initial and final consonant clusters (in Tequistlatec).

Grammatically, Meso-American languages are rather diverse, but, according to available data, they fall into three main types: Type A, an Oto-Manguean type, is rightward expanding (*i.e.,* modifiers follow the elements they modify) and synthetic to a low degree (*i.e.,* characterized by relatively few morphemes per word). It employs prefixes and prepositions, and it seldom uses compounding to form words. Type B, an intermediate type, is prepositional, like A, and averagely synthetic, making some use of prefixes (subjects, objects, and possessors) and much use of suffixes. It is mildly leftward expanding (*i.e.,* modifiers precede the elements they modify) and is mainly represented by Mayan and Uto-Aztecan languages but partially by Mixe-Zoque and Totonacan. Type C, a leftward expanding type, is highly synthetic with great use of suffixes and postpositions and active ablaut (an interchange among consonants and vowels for the purpose of derivation or inflection). It is represented by Tarasco and, partially, by Totonacan and Mixe-Zoque.

There are a number of grammatical generalizations that can be made about all, or most, Meso-American Indian languages. (1) The genitive relationship between nouns or noun phrases is (except for Tarasco) expressed by means of a possessive pronoun with the possessed noun; *e.g.,* "the dog's fleas" is expressed as "his fleas the dog." (2) Locative notions, such as "above," "below," "in," "on," "beside," are not expressed by prepositions and adverbs, as in European languages, but by means of location nouns (meaning "aboveness," "belowness," "belly," "surface," "side," and so forth), which are always combined with a possessive pronoun, the function of which is to indicate the "object" of the prepositional–adverbial notion. Most languages, however, have at least one generic relational particle that is combined in a phrase with a location noun and its object and has "generic prepositional" function; thus "on the table" is expressed "at (generic particle) its-top the table," or "in the box" is expressed "at its-inside the box." Whereas in most languages the generic relational particles are prepositions, Zoque and Tarasco have postpositions, which are in part related to location nouns.

(3) Within the verbal system, aspect (type of action—*e.g.,* ongoing, habitual, finished, potential, and so forth) is well developed, and tense (time—*e.g.,* now, in the past, in the future) is generally weakly developed. (4) The copula, or equational verb "be," is not expressed in most Meso-American languages. (5) Case suffixes are generally absent,

being present in just three languages: Tarasco has a genitive case, an objective case, and various locational cases; Aztec and Zoque have only locational cases, and these are usually related to location nouns. (6) A relative clause that modifies a noun follows it in all the languages of the sample above; *e.g.*, "the man whom I saw (on the street yesterday)." (7) Some Oto-Manguean languages and some Mayan languages distinguish an inclusive pronoun "we" ("I and you") from an exclusive pronoun "we" ("I and he/they"). (8) Gender, or inflectional agreement of other word classes in the noun phrase with the noun itself, is rare in Meso-American languages and is limited to some Oto-Manguean languages. (9) Noun subclassification in the context of possession is not uncommon. In some languages, some nouns undergo form changes when possessed; these languages, therefore, have at least two classes of nouns. In other languages, the possessive pronouns differ in form according to how they are associated with different classes of nouns. In languages in which the semantic motivation for such a subdivision is clear, the main kind of distinction is between intimate possession (body parts, kinship terms, articles of clothing) and casual possession (domestic animals, tools). (10) Some languages (Mayan, Mixe-Zoque) distinguish between the subject (actor) of a transitive verb and that of an intransitive verb by the form of the associated affixed pronoun. (11) Most Meso-American languages average more than one morpheme per word, and Tarasco and Totonac average more than two morphemes per word. (12) Most Meso-American languages (except Aztec) have consonantal or vocalic ablaut, or else show in their vocabulary sets of words that seem to be related through a formerly functional ablaut system.

(13) The numeral systems are vigesimal–decimal; that is, counting is from 1 to 10, then from 11 to 20, then from 21 to 40 (adding 1–20 to 20), then from 41–60 (adding 1–20 to 40), and so on, with special terms for 400 (20 × 20), 8,000 (20 × 20 × 20), 160,000 (20 × 20 × 20 × 20), and so on. In most languages (except Mayan) the numeral expressions for 6 through 9 (sometimes 5 through 9) are compounds of 5 + 1, 5 + 2, 5 + 3, 5 + 4, or the like. (14) In all the languages referred to here, a numeral precedes the noun it quantifies. (Te.K.)

South American Indian languages

South American Indian languages once covered and to-day still partially cover all of South America, the Antilles, and Central America to the south of a line from the Gulf of Honduras to the Nicoya Peninsula in Costa Rica. Estimates of the number of speakers in that area in pre-Columbian times vary from 10,000,000 to 20,000,000. In the early 1980s there were approximately 15,900,000, more than three-fourths of them in the central Andean areas. Language lists include around 1,500 languages, and figures over 2,000 have been suggested. For the most part, the larger estimate refers to tribal units whose linguistic differentiation cannot be determined. Because of extinct tribes with unrecorded languages, the number of languages formerly spoken is impossible to assess. Only between 550 and 600 languages (about 120 now extinct) are attested by linguistic materials. Fragmentary knowledge hinders the distinction between language and dialect and thus renders the number of languages indeterminate.

Because the South American Indians originally came from North America, the problem of their linguistic origin involves tracing genetic affiliations with North American groups. To date only Uru-Chipaya, a language in Bolivia, is surely relatable to a Macro-Mayan phylum of North and Meso-America. Hypotheses about the probable centre of dispersion of language groups within South America have been advanced for stocks like Arawakan and Tupian, based on the principle (considered questionable by some) that the area in which there is the greatest variety of dialects and languages was probably the centre from which the language groups dispersed at one time; but the regions in question seem to be refuge regions, to which certain speakers fled, rather than dispersion centres.

South America is one of the most linguistically differentiated areas of the world. Various scholars hold the plausible view that all American Indian languages are ultimately related. The great diversification in South America, in comparison with the situation of North America, can be attributed to the greater period of time that has elapsed since the South American groups lost contact among themselves. The narrow bridge that allows access to South America (*i.e.*, the Isthmus of Panama) acted as a filter so that many intermediate links disappeared and many groups entered the southern part of the continent already linguistically differentiated.

Investigation and scholarship. The first grammar of a South American Indian language (Quechua) appeared in 1560. Missionaries displayed intense activity in writing grammars, dictionaries, and catechisms during the 17th century and the first half of the 18th. Data were also provided by chronicles and official reports. Information for this period was summarized in Lorenzo Hervás y Panduro's *Idea dell' universo* (1778–87) and in Johann Christoph Adelung and Johann Severin Vater's *Mithridates* (1806–17). Subsequently, most firsthand information was gathered by ethnographers in the first quarter of the 20th century. In spite of the magnitude and fundamental character of the numerous contributions of this period, their technical quality was below the level of work in other parts of the world. Since 1940 there has been a marked increase in the recording and historical study of languages, carried out chiefly by missionaries with linguistic training, but there are still many gaps in knowledge at the basic descriptive level, and few languages have been thoroughly described. Thus, classificatory as well as historical, areal, and typological research has been hindered. Descriptive study is made difficult by a shortage of linguists, the rapid extinction of languages, and the remote location of those tongues needing urgent study. Interest in these languages is justified in that their study yields basic cultural information on the area, in addition to linguistic data, and aids in obtaining historical and prehistorical knowledge. The South American Indian languages are also worth studying as a means of integrating the groups that speak them into national life.

Classification of the South American Indian languages. Although classifications based on geographical criteria or on common cultural areas or types have been made, these are not really linguistic methods. There is usually a congruence between a language, territorial continuity, and culture, but this correlation becomes more and more random at the level of the linguistic family and beyond. Certain language families are broadly coincident with large culture areas—*e.g.*, Cariban and Tupian with the tropical forest area—but the correlation becomes imperfect with more precise cultural divisions—*e.g.*, there are Tupian languages like Guayakí and Sirionó whose speakers belong to a very different culture type. Conversely, a single culture area like the eastern flank of the Andes (the Montaña region) includes several unrelated language families. There is also a correlation between isolated languages, or small families, and marginal regions, but Quechumaran (Kechumaran), for instance, not a big family by its internal composition, occupies the most prominent place culturally.

Most of the classification in South America has been based on inspection of vocabularies and on structural similarities. Although the determination of genetic relationship depends basically on coincidences that cannot be accounted for by chance or borrowing, no clear criteria have been applied in most cases. As for subgroupings within each genetic group, determined by dialect study, the comparative method, or glottochronology (also called lexicostatistics, a method for estimating the approximate date when two or more languages separated from a common parent language, using statistics to compare similarities and differences in vocabulary), very little work has been done. Consequently, the difference between a dialect and language on the one hand, and a family (composed of languages) and stock (composed of families or of very differentiated languages) on the other, can be determined only approximately at present. Even genetic groupings recognized long ago (Arawakan or Macro-Chibchan) are probably more differentiated internally than others that have been questioned or that have passed undetected.

Margin notes:
Research on the languages

Number of languages and speakers