

## IDIOMAS DE MESOAMERICA

### 1. NOTA INTRODUCTORIA

Mesoamérica es una de las 35 áreas culturales —aproximadamente— en que los antropólogos dividen el mundo aborigen americano con base en la distribución de rasgos y complejos culturales. Mesoamérica incluye la parte central, la parte sur y partes del noroeste de México, además Guatemala, Belice, El Salvador y partes de Honduras y Nicaragua. Mesoamérica es una área de especial interés porque fue allí, hacia el año 1000 a.C. donde primero surgió la civilización en el hemisferio occidental.

Desde esa época han florecido varios centros de civilización —algunas veces de modo concurrente— hasta la conquista española de México en 1519; sin embargo, Mesoamérica, como un todo, tiene una historia cultural más o menos común de dos mil quinientos años.

Es correcto tratar de los idiomas de esta área como de un solo grupo, puesto que la unidad cultural del área hace suponer que dichos idiomas comparten un alto grado de semejanza. Sin embargo, no todas las lenguas aborígenes actuales de Mesoamérica han existido en el área desde el año 1000 a.C. Por otra parte, es usual entre lingüistas ordenar grandes grupos de idiomas de acuerdo primordialmente con sus relaciones genéticas, y sólo de modo secundario según la distribución geográfica de dichos idiomas. De esta manera, en este trabajo

se consideran algunas lenguas que no se hablan en Mesoamérica propiamente, pero de las cuales se sabe que forman familias lingüísticas con lenguas que en realidad se hablan en el área.

En Mesoamérica existen en la actualidad unos setenta idiomas indígenas que se hablan por cerca de seis millones de personas.

Cuando los españoles conquistaron México en 1519, había unos veinte millones de personas en Mesoamérica. En el curso de los 100 años que siguieron a la conquista, la población indígena disminuyó en un ochenta por ciento como resultado de guerras, enfermedades, trabajos forzados y hambre. En el presente la población indígena ha recobrado sus altos índices, pero algunos de los idiomas se han extinguido; de ellos han sido documentados alrededor de diez. En la actualidad los idiomas con mayor número de hablantes son los siguientes:

1. Azteco	1.000.000	hablantes	Familia Yutonahua
2. Yucateco	350.000	}	Familia Mayance
3. Quiché-Tzutujil-Cakchiquel	900.000		
4. Mam	350.000		
5. Kekchí	300.000	}	Familia Otomangue
6. Mixteco	250.000		
7. Zapoteco	300.000		
8. Otomí	300.000		

### 2. ALGUNOS CONCEPTOS DE LINGÜISTICA HISTÓRICA

Antes de entrar a tratar del resto de los idiomas de Mesoamérica, se hace necesario introducir algunos conceptos básicos.

Se dice de algunos idiomas que están *genéticamente relacionados* (o simplemente *relacionados*), si se puede demostrar

que ellos se han desarrollado de una forma única de lenguaje que se hablaba en una época anterior. En algunos casos la forma anterior pudo haber sido documentada, como en el caso del latín —predecesor del francés, español, portugués, italiano, etcétera, a los cuales se les llama lenguas romances—; o en el caso del sánscrito —predecesor del indostaní, bengalí, etcétera, a los que se llama idiomas índicos o indo-arios. En la mayoría de casos, sin embargo, la forma anterior de lenguaje no fue registrada nunca. Registrada o no, el procedimiento para demostrar la relación consiste en comparar el vocabulario, los sonidos y los patrones gramaticales de los idiomas de que se trata.

Los idiomas relacionados muestran *equivalencias recurrentes* en los sonidos (fonemas) de palabras de significado idéntico o similar. Con base en tales equivalencias podemos postular un sistema de sonidos original y ancestral (reconstruido) e items de vocabulario o *elementos lexicológicos* que cumplen dos desideratas:

- (a) Las palabras pueden hacerse derivar de las formas postuladas mediante la aplicación a estas últimas de reglas por lo general sin excepción;
- (b) El sistema de sonidos y las formas de palabras que han sido postulados no violan constreñimiento alguno en los sonidos y en las combinaciones de éstos, que se usan en la actualidad entre idiomas realmente conocidos.

Los idiomas relacionados también muestran semejanzas sistemáticas en los patrones gramaticales, las cuales permiten la postulación de tipos de elementos gramaticales (*morfemas*) y de reglas para combinar dichos elementos en la formación de palabras y oraciones. Los sonidos, palabras y patrones gramaticales postulados se llaman *reconstrucciones*. La forma ante-

rior de lenguaje, la cual se postula como una forma 'ancestral' para dos o más idiomas 'registrados', se llama *protoidioma*. Las palabras 'registradas' en dos o más idiomas que tienen un origen común se llaman *cognadas*.

Los idiomas que están relacionados entre sí pueden estarlo en diferentes grados de proximidad. Por ejemplo, el francés y el italiano tienen más correspondencia, punto por punto, en todos los niveles de su estructura (vocabulario, fonología y gramática) que la que pueda tener cualquiera de ellos con el indostaní o el bengalí, aunque los cuatro se relacionan entre sí. Esta diferencia de grado en la relación se debe al hecho de que el francés y el italiano se separaron uno del otro en época relativamente más reciente —digamos hacia el año 500 d.C.— en comparación con el tiempo en que se separaron el predecesor del latín y el predecesor del sánscrito —posiblemente hacia el año 3,000 a.C.

Por consiguiente, se puede generalizar diciendo que cuando los idiomas se reconocen como mutuamente relacionados también se hace necesario determinar el grado de dicha relación.

Debemos mencionar aquí otras dos doctrinas lingüísticas.

Primera: Puesto que los idiomas cambian a lo largo del tiempo a una tasa más o menos estable, mientras más haya durado el proceso de variación o diversificación de un grupo dado de idiomas, menos semejantes se habrán desarrollado los idiomas que se separaron más temprano. Los métodos de que disponen en la actualidad los lingüistas para demostrar la relación, los capacitan para alcanzar probablemente no más de diez mil años. Dado que los seres humanos han estado hablando durante un tiempo mucho más largo que éste, deducimos que hay algunas relaciones genéticas reales que no

podemos demostrar, y otras que ni siquiera sospechamos. Por tal razón, la mayoría de los lingüistas acepta la doctrina de que aunque *es posible comprobar una relación, no es posible comprobar una no-relación*. Es posible, por supuesto, decir de una relación postulada que no ha sido probada.

Segunda: La mayoría de los lingüistas acepta la doctrina de que cada idioma tiene sólo un antecesor final, es decir, que no existen "idiomas mezclados". Un corolario de ello es que *a un solo idioma no se le puede asignar más de un grupo genético que no se traslapa*. Es necesario tener en mente que la noción "predecesor-descendiente" es tan sólo una metáfora en el campo de la lingüística genética, y que dicha metáfora no tiene precisamente connotación sexual alguna. Un idioma evoluciona de modo gradual, aun cuando puede experimentar el proceso de fisión análogo al de la reproducción amiboidea. El francés y el español, por ejemplo, son sencillamente formas modernas del latín hablado. Por conveniencia podemos dar otro nombre al latín en el punto en que se bifurca y empieza a seguir dos direcciones independientes, aunque cada variante es tan neolatina, como la otra.

De los idiomas relacionados se dice que forman *familias*; se usan otros términos tales como *grupo, tronco, rama (estirpe, linaje)*, etcétera, cuando se debe reconocer diferentes niveles u órdenes de relación dentro de una sola agrupación genética. El proceso de reconocer una relación genética se llama *clasificación y (sub)agrupación*.

### 3. GLOTOCRONOLOGIA

En 1951 los lingüistas norteamericanos Morris Swadesh y Robert Lees desarrollaron un procedimiento para medir el grado de diferencia entre idiomas relacionados, en términos de años de separación. (Se busca responder así a la pregunta siguiente: ¿en qué año se dividieron en dos comunidades sin comunicación los hablantes del común predecesor de dos idiomas determinados?).

Este método, llamado glotocronología o lexicostatística, se basa en la suposición de que todos los idiomas cambian a la larga en un grado más o menos igual durante un período determinado. No todos los lingüistas comparten este aserto, pero el mismo podría ser aceptado en principio por la mayoría de aquellos que han tenido experiencia con una gran variedad de idiomas de diferentes partes del mundo. La glotocronología asume además que existe una serie de nociones lexicográficas empíricamente definibles, para las que todos los idiomas tienen palabras y las cuales son por lo general resistentes a niveles fortuitos de cambio durante largos períodos; ellas por el contrario, están sujetas a índices regulares de cambio. Después de varios años de experimentación, Swadesh estableció una lista de cien ítems de vocabulario "básico" o "no cultural", respecto de la cual se pretendió lo siguiente:

- (a) 'Todos los ítems son conceptos que se pueden expresar en cualquier idioma.
- (b) Durante un período de mil años habrán sido seleccionadas nuevas palabras para expresar el 14% de los conceptos de la lista. (Por ejemplo, el concepto *comer* es *edere* en latín y *manger* en francés; asimismo, el concepto *perro* es *hund* en inglés antiguo y *dog* en inglés moderno).

La determinación de este índice de cambio se basó en el estudio de cerca de trece casos de idiomas que tienen documentación histórica escrita de por lo menos mil años de antigüedad; tal es el caso del inglés, árabe, latín-francés, sánscrito-hindú, egipcio-cóptico, etcétera. Si se aplica el método de la glotocronología a dos idiomas contemporáneos relacionados, la antigüedad de la relación entre ellos debe ser considerada según la fórmula siguiente:

$$t = \frac{\log c}{2 \log r}$$

c = porcentaje de items cognados entre los dos idiomas (*c* significa cognado)  
 t = milenios de separación (*t* significa tiempo)  
 r = .86 (*r* indica tasa de retención)  
 log = logaritmo.

La glotocronología pretende, respecto de parejas de idiomas, que se puede especificar la fecha aproximada en que los dos empezaron a desarrollarse independientemente el uno del otro.

Todos los idiomas cambian a lo largo del tiempo, pero una determinada comunidad parlante puede o no dividirse en algún punto de su historia. De los lingüistas interesados en este asunto alrededor de la mitad rechaza el método de la glotocronología como un método que se basa en falsas suposiciones o matemáticas deficientes. Cerca de la otra mitad acepta dicho método como un instrumento útil para obtener fechas aproximadas y una cronología relativa, si no absoluta, dentro de grupos de idiomas relacionados. Morris Swadesh, el lingüista-antropólogo norteamericano, ha efectuado cálculos léxico-estadísticos en relación con la mayor parte de los grupos idiomáticos de Mesoamérica. El autor de este estudio ha hecho

cálculos independientes en algunos casos. Para los propósitos de este ensayo, sin embargo, se citarán determinaciones disponibles de la glotocronología con el interés de contribuir en la tarea de la subagrupación de los idiomas del área.

#### 4. CLASIFICACION GENETICA

La clasificación de los idiomas que forman un grupo relacionado tiene por objeto demostrar cómo dichos idiomas en su variedad están relacionados entre sí. Cuando consideramos la distribución geográfica de una agrupación de idiomas relacionados, casi siempre reconocemos que el protoidioma debe haberse hablado en una área más pequeña, tal vez aun fuera del área en cuestión. La *separación* o *subdivisión* del idioma ocurre cuando las comunidades cesan de comunicarse y ello puede suceder ya sea por medio de fragmentación social o de migración (o ambas causas). Propiamente hablando, la clasificación refleja lo que se sabe sobre la historia de la diversificación del grupo de idiomas. En la práctica se puede reconocer las siguientes etapas en el estudio de tales grupos:

- (a) Reconocimiento de la relación genética debida a una semejanza preponderante en el vocabulario y en la gramática.
- (b) Establecimiento de "correspondencias" regulares entre sonidos de palabras cognadas en los diversos idiomas.
- (c) Subagrupación por medio de la glotocronología y otros medios para determinar la historia común.
- (d) Reconstrucción de palabras.
- (e) Reconstrucción de patrones gramaticales.

- (f) Reconstrucción de la diversificación y movimientos de población.

Hasta la fecha sólo la primera etapa ha sido alcanzada en muchas agrupaciones de idiomas. En familias bien estudiadas se ha llegado hasta la cuarta etapa (d). En familias excepcionalmente bien trabajadas se ha llegado hasta las etapas (e) y (f).

##### 5. EL ESTUDIO DE LOS IDIOMAS MESOAMERICANOS

Durante los siglos XVI y XVII algunos misioneros dominicos y franciscanos radicados en el hemisferio occidental se dedicaron al estudio de idiomas nativos para capacitar a los sacerdotes encargados de tratar asuntos religiosos con indígenas monolingües. Escribieron gramáticas siguiendo un modelo tomado del latín; idearon ortografías aplicando valores usados en español o en latín (ocasionalmente inventaron nuevas letras); tradujeron textos cristianos (confesionarios, sacramentos y sermones) a idiomas indígenas. Con excepción de una figura heroica, el sacerdote misionero español Bernardino De Sahagún, no coleccionaron ni fomentaron la recolección del folklore. Durante ese período se escribieron gramáticas y diccionarios en idiomas como el azteco, zapoteco, mixteco, tzeltal, yucateco, quiché-tzutujil-cakchiquel, chortí y otomí. Estas fuentes de información sirvieron a los sucesores de los primeros misioneros. Durante el siglo XVIII el ímpetu de tal trabajo disminuyó. Después de la independencia, en la primera parte del siglo XIX, los clérigos españoles fueron expulsados y el trabajo sobre idiomas indígenas cayó en manos de viajeros y de diletantes, en su mayor parte gente pobremente calificada para tales menesteres. Dichas personas, por supuesto, continuaron involucradas en esa actividad y muchas veces son todavía nuestra única fuente de información para algunos idiomas.

Durante el siglo XIX nació la ciencia de la lingüística, surgiendo de dos focos principales de estudio. El primero fue el reconocimiento de la relación genética entre idiomas (en oposición a la semejanza que resulta de la influencia mutua entre idiomas una vez adyacentes), explícitamente formulado por el filólogo inglés William Jones —en el último cuarto del siglo XVIII— con respecto al sánscrito, griego y latín —todos idiomas indoeuropeos. El segundo fue el interés en la fonética como campo de estudio. La mayoría de los estudiosos fueron alemanes y británicos. La lingüística norteamericana (de los Estados Unidos de América) se ha orientado desde la década de 1910 hacia el campo del estudio de los idiomas ágrafos, aunque tal orientación se ha desviado seriamente desde los primeros años de la década de 1960, cuando los lingüistas norteamericanos se empezaron a preocupar por la lingüística teórica y los universales del lenguaje. Debido a que ésta no se halla bien desarrollada en México y Guatemala, sus técnicas no se aplicaron en Mesoamérica sino hasta que los norteamericanos enfocaron su atención en el área, esto es en los años de la década de 1930 (misioneros protestantes) y en la década de 1940 (lingüistas académicos).

Desde entonces se ha hecho mucho trabajo lingüístico profesional en los idiomas mesoamericanos, en especial los de México. Un buen número de mexicanos ha recibido entrenamiento lingüístico, así como varios guatemaltecos. Casi cada uno de los idiomas de Mesoamérica ha sido estudiado cuando menos por un lingüista, aunque el tiempo utilizado y la competencia lingüística de los investigadores han variado considerablemente. Se ha reunido información gramatical y lexicográfica de la mayor parte de dichos idiomas. En su mayoría esta información permanece inédita, principalmente por los gastos que involucra el prepararla para su publicación, así como por la dificultad de encontrar canales editoriales para estudios lin-

güísticos que no están de moda y/o son teóricos. Sin embargo, se ha publicado un buen número de gramáticas y diccionarios calificados —aunque ninguno exhaustivo o definitivo. Se han coleccionado, y en algunos casos publicado, cuentos folklóricos, en un número menor de idiomas. En su mayoría los idiomas de Mesoamérica han tenido ortografías basadas en el español, ideadas para ellos en este siglo, pero la mayor parte de lo que se encuentra disponible para leer son traducciones de la Biblia, descripciones de cómo construir letrinas o instrucciones para evitar el paludismo. En resumen, queda mucho trabajo por hacer, a pesar de las apariencias en contrario. Por ejemplo, los siguientes idiomas han sido en extremo descuidados: jicaque, paya, chichimeco, ocuilteco, pame sur, tlapaneco, subtiaba (¿extinguido?), oluta, texistepeco, xinca oriental (¿extinguido?), lenca (¿extinguido?), chilanga, misquito, matagalpa, cacaopera, sumo, azteco occidental, guarijío, tepehuán sur, tepicano (¿extinguido?), pima bajo, chicomucelteco (¿extinguido?), itzá, teco, uspanteco.

Las siguientes lenguas muertas están relativamente bien documentadas: cuitlateco, subtiaba, (si es que ha muerto), chiapaneco, chilanga (si es que ha muerto).

Las siguientes lenguas muertas se encuentran pobremente documentadas: eudeve, ópata, jova, pochuteco, papabuco manguc, tapachulteco, lenca (si es que ha muerto), xinca oriental (si es que ha muerto), chicomucelteco (si es que ha muerto).

## 6. CLASIFICACION DE LOS IDIOMAS DE MESOAMERICA

A continuación se presenta una descripción de los idiomas de Mesoamérica según sus agrupaciones genéticas. Las agrupaciones genéticas generalmente aceptadas y apoyadas con

documentación adecuada, son admitidas como tales aquí. Los pocos idiomas que no han sido asignados de modo convincente a agrupaciones genéticas se hallan registrados individualmente y se les llama aislados. Dentro de una determinada agrupación genética pueden encontrarse varios niveles de relación y resultan útiles las categorías que se presentan a continuación, debido a que las mismas están correlacionadas según la glotocronología.

Término	Siglos de Separación	Porcentaje cognado
dialectos	0-5	86-100%
complejo de idiomas (término ideado por Swadesh)	7-11	71-81%
grupo de idiomas	13-17	60-68%
fama (o familia si no hay una categoría superordinal)	19-26	46-56%
familia de idiomas	35-45	26-35%
tronco (estirpe) o fílum	55-65	14-19%

Las agrupaciones dadas aquí por lo general han sido aceptadas hasta la fecha. Clasificaciones más antiguas se medirán en función del consenso actual. El orden en que se presentan las agrupaciones está diseñado para permitir referencias a clasificaciones más amplias e inclusivas que han sido propuestas en la actualidad o en épocas anteriores.

Cada familia, grupo (si se trata de menos de 17 siglos de separación), o idioma aislado, recibe un número separado de serie, del 1 al 21. Nuestra clasificación aparece en el Cuadro I.

Las siguientes anotaciones explican el contenido y convenciones del cuadro. Cada uno de los 21 rubros especifica el nombre de la agrupación, con nombres alternos entre paréntesis y la ubicación entre corchetes. En la primera columna

se halla el nombre del idioma, con nombres alternos entre paréntesis. Los nombres de los idiomas extinguidos han sido subrayados. Los números subscritos indican que hay varios idiomas estrechamente relacionados, todos llamados por el mismo nombre. Cada nombre con que empieza una línea es un idioma separado, excepto cuando el nombre previo termina en guiones, lo cual indica al lector que el próximo nombre es un dialecto del mismo idioma. La segunda columna indica dónde se habla el idioma (M=México, G=Guatemala, B=Belice). En la tercera columna se da el número aproximado de hablantes (m=1000). El resto del cuadro relativo a cada agrupación de idiomas especifica los diversos niveles de relación, incluyendo las cifras glotocronológicas (s=siglos); estas últimas cifras son de Swadesh, excepto las que se refieren al mixcoque, maya y xinca, las cuales son del autor del presente estudio.

Los nombres de familias y troncos se forman de tres maneras:

- (a) Un idioma típico; por lo general el más extensamente hablado, cuyo nombre se usa para designar a la familia, rama o grupo.
- (b) Dos nombres típicos son seleccionados y se unen para formar un nombre compuesto; ejemplo: mixcoque.
- (c) Se juntan partes de dos o más nombres de idiomas para formar un nuevo nombre; ejemplo: otomangué, otopame, misúlúa o misumalpa.

Se usan los nombres para las familias de idiomas cuando los grupos pueden ser subdivididos, pero no se hace tal cosa cuando los grupos han sido divididos previamente en lenguajes separados.

El mapa da la distribución geográfica aproximada de las agrupaciones y de los idiomas aislados de Mesoamérica. Ninguno de los extinguidos sin documentación está indicado. Los idiomas separados dentro de una agrupación no están localizados. Una *extensión* es un idioma que ha sido llevado por migración a un contexto lingüístico y cultural extraño; ejemplo, el mangue es una extensión chiapaneca en territorio misumalpa. El subtiaba es una extensión tlapaneca en territorio misumalpa. El pipil es una extensión nahua en territorio xinca, lenca y misumalpa.

#### 41 YUTONAHUA (1)

La familia yutonahua (YN) consiste de unos 27 idiomas los cuales se clasifican —como es reconocido universalmente— en 8 grupos o ramas. Dos de estos grupos, el tubatulabal y el jopi contienen sólo un idioma cada uno. Los primeros cuatro grupos se reconocen por lo general, aunque no universalmente, como integrantes de una división *shoshoni* o norteamericana, dentro de la familia. Ninguno de los idiomas shoshoni es hablado en Mesoamérica y no daremos información alguna sobre distribución o movimientos de población respecto de estos idiomas. Hay dos formas comunes para agrupar los idiomas restantes, lo cual se hace depender de la posición asignada al grupo nahua. Una forma es que se considera el nahua como separado y el resto como una división del *sonora*, lo cual produciría tres divisiones: shoshoni, sonora y nahua; otra forma es que se incluye el nahua dentro del sonora, produciendo así una dicotomía shoshoni-sonora, que es la que se sigue aquí. Varios investigadores creen que el concepto de "división" es deficiente aquí y que el yutonahua contiene ocho grupos y ramas que no se deben subclasificar en ninguna forma especial. Como se dijo, ningún idioma shoshoni es hablado en Me-

soamérica y sólo algunos idiomas sonora (indicados éstos por asteriscos en el cuadro). El extinguido tubar pertenece a la rama yaqui, pero no está claro si al complejo tarahumara, al complejo cáhita o a ninguno de éstos. El grupo nahua incluye al extinto pochuteco, anteriormente hablado en la costa de Oaxaca, México, el cual se halla pobremente documentado. El pochuteco es claramente muy divergente del resto del grupo. El complejo azteca es considerado por algunos como un solo idioma con varios dialectos. Tres idiomas aztecos se hablaban en el imperio azteca tal como éste estaba constituido en 1519. Hablantes pipiles, que también se refieren a su idioma como *nawat*, no formaron parte de la cultura azteca, y probablemente representan una expansión tolteca de varios siglos atrás (véase la sección 12 más adelante).

En 1859, el filólogo alemán Johann Carl Buschmann identificó correctamente todos los idiomas yutonahua hasta entonces conocidos como parte de una sola familia. En 1883, el filólogo francés Hyacinthe de Charencey dividió el yutonahua en oregoniano (=shoshoni) y mexicano (=sonora). En 1891, el antropólogo norteamericano Daniel Brinton reconoció el shoshoni y dividió nuestro sonora en náhuatl (=nahua) y lo que él llamó sonora (=sonora menos nahua); el biólogo norteamericano John Wesley Powell siguió a Brinton en su clasificación de los idiomas norteamericanos.

Muchos investigadores creen que varios grupos étnicos indígenas del norte de México ahora aculturados y racialmente absorbidos, cuyos nombres se conocieron hasta en época reciente (a partir del siglo XVI y XVII) hablaron idiomas YN, como el suma, jumano, lagunero, cazcán, tecuexe, guachichil y zacateco.

Los especialistas aceptan por lo general que el YN está relacionado con la familia kayowetaño y que con ésta forman el tronco azteco-taño.

El mapa da la distribución geográfica aproximada de las agrupaciones y de los idiomas aislados de Mesoamérica. Ninguno de los extinguidos sin documentación está indicado. Los idiomas separados dentro de una agrupación no están localizados. Una *extensión* es un idioma que ha sido llevado por migración a un contexto lingüístico y cultural extraño; ejemplo, el mangue es una extensión chiapaneca en territorio misumalpa. El subtiaba es una extensión tlapaneca en territorio misumalpa. El pipil es una extensión nahua en territorio sinca, lenca y misumalpa.

#### 4.1 YUTONAHUA (1)

La familia yutonahua (YN) consiste de unos 27 idiomas los cuales se clasifican —como es reconocido universalmente— en 8 grupos o ramas. Dos de estos grupos, el tubatulabal y el jopi contienen sólo un idioma cada uno. Los primeros cuatro grupos se reconocen por lo general, aunque no universalmente, como integrantes de una división *shoshoni* o norteamericana, dentro de la familia. Ninguno de los idiomas shoshoni es hablado en Mesoamérica y no daremos información alguna sobre distribución o movimientos de población respecto de estos idiomas. Hay dos formas comunes para agrupar los idiomas restantes, lo cual se hace depender de la posición asignada al grupo nahua. Una forma es que se considera el nahua como separado y el resto como una división del *sonora*, lo cual produciría tres divisiones: shoshoni, sonora y nahua; otra forma es que se incluye el nahua dentro del sonora, produciendo así una dicotomía shoshoni-sonora, que es la que se sigue aquí. Varios investigadores creen que el concepto de "división" es deficiente aquí y que el yutonahua contiene ocho grupos y ramas que no se deben subclasificar en ninguna forma especial. Como se dijo, ningún idioma shoshoni es hablado en Me-

soamérica y sólo algunos idiomas sonora (indicados éstos por asteriscos en el cuadro). El extinguido tubar pertenece a la rama yaqui, pero no está claro si al complejo tarahumara, al complejo cáhita o a ninguno de éstos. El grupo nahua incluye al extinto pochuteco, anteriormente hablado en la costa de Oaxaca, México, el cual se halla pobremente documentado. El pochuteco es claramente muy divergente del resto del grupo. El complejo azteca es considerado por algunos como un solo idioma con varios dialectos. Tres idiomas aztecos se hablaban en el imperio azteca tal como éste estaba constituido en 1519. Hablantes pipiles, que también se refieren a su idioma como *nawat*, no formaron parte de la cultura azteca, y probablemente representan una expansión tolteca de varios siglos atrás (véase la sección 12 más adelante).

En 1859, el filólogo alemán Johann Carl Buschmann identificó correctamente todos los idiomas yutonahua hasta entonces conocidos como parte de una sola familia. En 1883, el filólogo francés Hyacinthe de Charencey dividió el yutonahua en oregoniano (=shoshoni) y mexicano (=sonora). En 1891, el antropólogo norteamericano Daniel Brinton reconoció el shoshoni y dividió nuestro sonora en náhuatl (=nahua) y lo que él llamó sonora (=sonora menos nahua); el biólogo norteamericano John Wesley Powell siguió a Brinton en su clasificación de los idiomas norteamericanos.

Muchos investigadores creen que varios grupos étnicos indígenas del norte de México ahora aculturados y racialmente absorbidos, cuyos nombres se conocieron hasta en época reciente (a partir del siglo XVI y XVII) hablaron idiomas YN, como el suma, jumano, lagunero, cazcán, tecuexe, guachichil y zacateco.

Los especialistas aceptan por lo general que el YN está relacionado con la familia kayowetaño y que con ésta forman el tronco aztecaño.

Buschmann (1859) y el lingüista-antropólogo norteamericano Edward Sapir (1915) contribuyeron al estudio comparativo del YN, reuniendo un número considerable de juegos o conjuntos de palabras cognadas. El lingüista Benjamín Lee Whorf (1935) codificó las correspondencias de sonido y propuso algunas reconstrucciones del protoyutonahua. Los lingüistas-antropólogos norteamericanos Carl Voegelin, Florence Voegelin y Kenneth Hale (1962) mejoraron las investigaciones de Whorf. Wick Miller (1967) —también lingüista-antropólogo de los Estados Unidos—, reunió una colección de unas 500 etimologías, usando fuentes fonológicas modernas y confiables.

### 4.3 CUITLATECO (2)

El idioma cuitlateco, ahora extinguido, no ha sido relacionado de modo convincente con ningún otro idioma o familia, aunque tanto Sapir como Swadesh mantenían la idea de que pudiera estar relacionado con el yutonahua.

### 4.3 LA HIPOTESIS JOKA: GRUPOS 3-5

En 1919 los antropólogos norteamericanos Roland Dixon y Alfred Kroeber trataron de mejorar la clasificación que para los idiomas norteamericanos había ofrecido Powell, reduciendo para ello la multiplicidad de agrupaciones de los idiomas de California (cerca de 50) a un número manejable de familias y troncos. En un período de varios años desarrollaron la hipótesis de que la mayoría de idiomas californianos pertenece a una o dos grandes agrupaciones (llamadas *filums* o *super-troncos*): *joka* y *penuti*. Esta proposición fue aceptada y ampliada por Sapir, así como por muchos otros. El *joka* incluía el *shasta*, *achumawi*, *atsugewi*, *chimáriko*, *karok*, familia *yana*, familia *pomo*, *washo*, *esselen*, familia *yuma*, *salina* y *chumash*.

En 1891/2 Brinton había ya sugerido que el yuma, el seri (3) y el tequistlateco (4) estaban relacionados. En 1915 Kroeber reexaminó la cuestión a la luz de la hipótesis joka y concluyó que todos los idiomas mencionados antes están relacionados. Desde entonces, la mayoría de los estudiosos que se hallan familiarizados con los idiomas yuma creen que el seri y el yuma están relacionados y muchos de ellos, que aceptan la hipótesis joka, opinan que el seri y el yuma forman un grupo especial, dentro del joka.

El jicaque (5), pobremente documentado aunque todavía se habla, tiene oclusivas plenas, aspiradas y glotalizadas, como muchos idiomas joka. En 1953 Swadesh y Joseph Greenberg, sugirieron que el jicaque es un idioma joka. Esto ha sido generalmente aceptado, pero quizás sin espíritu crítico debido a que la información disponible acerca de dicho idioma no es muy confiable.

#### 6.4 IDIOMAS EXTINGUIDOS DEL NORDESTE DE MEXICO

Todos los idiomas, menos uno (el tónkawe), hablados otrora en el noreste de México y el sur de Texas se han extinguido. Los idiomas documentados de México son: el coahuilteco, comecrudito, cotoname, naolán y maratino (o tamaulipeco). Los de Texas son: el karánkawe (y klamkosh) y atákapa, y el tónkawe.

Powell clasifica los tres primeros como una familia coahuilteca. Los otros idiomas mexicanos no fueron conocidos sino recientemente. Cada uno de los tres idiomas texanos fue considerado por Powell como un idioma aislado. En 1920 Sapir redefinió la familia coahuilteca incluyendo en la misma el karánkawe y el tónkawe, y como una familia coordinada con el joka, constituyendo así un superfílum joka-coahuilteco (=jokalteco).

#### 4.5 TLAPANECO (6)

El complejo tlapaneco fue correctamente identificado en 1920 por el médico alemán Walter Lehmann y verificado en 1933 por el lingüista-antropólogo norteamericano Paul Radin. En 1925 Sapir trató de establecer el subtiaba como un idioma joka. Propuso algunas reconstrucciones protojoka que podrían explicar las formas del subtiaba. La clasificación es por lo general aceptada.

En un artículo publicado en la *Enciclopedia Británica* en 1929, Sapir hizo del joka, el tlapaneco y el coahuilteco, ramas coordinadas del jokalteco. Sin embargo, recientemente el lingüista misionero Calvin Rensch ha tratado de dar validez a la hipótesis otomangue (OM) por medio de la reconstrucción fonológica de toda la escala (véase adelante). Rensch cree que el tlapaneco es otomangue mientras que Swadesh lo considera como un idioma intermedio entre el otomangue y el joka. La hipótesis de Sapir puede salvarse únicamente suponiendo (a) que el tlapaneco (6) sea otomangue con elementos prestados del joka; (b) que el tlapaneco (6) sea joka con elementos prestados del otomangue; (c) que este último (OM) y el joka están relacionados.

Debe tenerse en cuenta que la mayor parte de los especialistas que se han dedicado al estudio de gran número de idiomas indígenas americanos creen que casi todos éstos están genéticamente relacionados entre sí. Esta relación se hace derivar de un período lejano que se remonta quizás a unos 20,000 a 30,000 años, cuando algunos de los idiomas todavía se hablaban en Asia. Desde tal punto de vista, la correcta agrupación es una cuestión más interesante que la mera relación genética.

## 6.6 OTOPAME (7)

La estirpe otopame contiene cuatro grupos y complejos, de los cuales sólo los dos últimos (con asteriscos en el cuadro IV) se hablan en Mesoamérica. El número exacto de idiomas dentro del complejo otomí no se ha determinado todavía, aunque parecen ser cuatro.

En 1892, Brinton definió el otopame correctamente, aunque trató de compararlo con el atabasco. El lingüista-antropólogo Stanley Newman y el lingüista germano-mexicano Robert Weitlaner, efectuaron un estudio de fonología comparada con el otomí (1950); el último de dichos investigadores (1953) lo hizo con el matlatzinca y la lingüista misionera Doris Bartholomew (1960) hizo lo mismo con el tronco completo.

## 6.7 MAZATECO (8)

La familia mazateca (generalmente llamada popoloca) fue correctamente identificada en 1962 por Pablo González Casanova. El número exacto de idiomas dentro del complejo mazateco no se ha determinado todavía, aunque hay cuando menos dos. La lingüista mexicana María Teresa Fernández de Miranda efectuó un estudio de fonología comparada para el grupo chocho (1951). La lingüista misionera Sara Gudschinsky hizo lo propio respecto del complejo mazateco y respecto de toda la familia (1953).

## 6.8 MIXTECO (9)

Existe cierta disparidad de opiniones en cuanto a cómo deben agruparse los diversos idiomas que aquí se incluyen en el mixteco. El principal problema es si el amuzgo pertenece al grupo mixteco o forma una rama separada del otomangue.

Swadesh incluye el amuzgo dentro del mixteco; el lingüista misionero Robert Longacre se muestra vacilante pero finalmente lo excluye. Swadesh, por otra parte, titubea entre incluir o no el trique dentro del mixteco. Parece haber tres idiomas dentro del grupo mixteco. Longacre (1957) y la lingüista mexicana Evangelina Arana de Swadesh (1960) han hecho estudios de fonología comparada en la familia mixteca. Longacre y la lingüista misionera Cornelia Mak (1960) han trabajado en la fonología comparativa del grupo mixteco.

## 6.9 ZAPOTECO (10)

La familia zapoteca fue identificada correctamente en 1912 por William Mechling, pero sólo el filólogo mexicano Francisco Belmar (1905a) reconoció correctamente que el papabuco es un idioma separado, ni zapoteco, ni chatino. Sin embargo, Belmar incluyó de manera incorrecta el chinanteco (11) dentro del zapoteco. El chatino tiene varios dialectos. Dentro del complejo zapoteco hay cuando menos cuatro idiomas y quizás más. La fonología histórica del complejo zapoteco ha sido investigada por Swadesh (1947) y Fernández (Ms. sin fecha); la del chatino por Longacre y el lingüista misionero B. W. Upson (1965). Los dos últimos lingüistas clasifican incorrectamente el papabuco como un dialecto del chatino y, sin buen éxito, tratan de derivar formas del papabuco de las formas reconstruidas del protochatino.

## 6.10 CHINANTECO (11)

El grupo chinanteco contiene aproximadamente cuatro idiomas; el número exacto aún no se ha determinado. En 1912 Mechling reconoció la separación del chinanteco dentro del otomangue. La fonología comparada del chinanteco ha sido

investigada por los lingüistas-antropólogos norteamericanos Angulo y Freeland (1933-34) y Sapir (1925).

#### 6.11 MANGUE (12)

El grupo mangue fue correctamente identificado por Belmar (1905b). Fernández y Weitlaner han estudiado la fonología comparada del mangue (1961).

#### 6.12 LA HIPOTESIS OTOMANGUE.

##### GRUPOS 7-12 (6-13)

Empezando con la de Brinton (1891) se han hecho reiteradas propuestas a efecto de unir dos o más de las familias anteriores (7-12). En 1920, Lehmann unificó los grupos 7, 8, 9, 12, y en 1925 Jaime de Angulo hizo lo propio con los grupos 8, 9, 10 y 11; desde el punto de vista del último autor, sin embargo, ello no constituía necesariamente una agrupación genética.

Más o menos desde 1925 ha sido aceptado generalmente por los especialistas que los grupos 7-12 forman una mayor agrupación genética (filum) comúnmente identificada como otomangue (OM). (Sapir [1929] se refiere al otomangue como mixteco-zapoteco-otomí, lo cual indica que la hipótesis fue formulada en efecto desde 1927). Esta puede ser llamada la "formulación clásica del otomangue". Desde 1950 se ha estado trabajando en la reconstrucción de idiomas emparentados con cada uno de las familias y grupos constituyentes. En 1959 Gudschinsky comparó el popoloca y el mixteco. En 1961 Fernández y Weitlaner compararon el mangue con estos últimos. Tanto Calvin Rensch, en su tesis doctoral de la Universidad de Pensilvania (1966), como Morris Swadesh (1964a 1964b), han trabajado en la fonología comparada del otomangue como un todo.

Desde 1961 se han propuesto dos revisiones en la formulación de lo que constituye el otomangue:

- (1) Rensch, en su reconstrucción de la fonología otomangue, incluye el tlapaneco (6). Swadesh, sobre bases glotocronológicas, concluye de modo independiente que el tlapaneco está estrechamente relacionado con el otomangue.
- (2) Swadesh (1964a) ha sostenido la opinión de que el huave (13) es un idioma otomangue y Longacre, el decano de los estudiosos del otomangue, se halla inclinado a aceptar esa idea. Actualmente, por tanto, la mayoría de los especialistas en OM, considera que éste consiste de los grupos 6-13.

El estudio comparativo que se ha dedicado al filo otomangue ha dado como resultado el primer caso, en el hemisferio occidental, en que el remoto predecesor común de varias familias de idiomas ha sido fonológicamente reconstruido. La lingüística comparada a nivel de filo ha sido en extremo infructuosa con otros supertruncos postulados. Con excepción del grupo mangue, todos los idiomas otomangue se hablan en México central.

#### 6.13 HUAVE (13)

Antes de Swadesh (1964a), las únicas propuestas serias acerca de las conexiones genéticas del huave fueron formuladas por Radin (1916, 1924), y en ellas se indicaba la unión del huave con el mixezoque (14) y el maya (16). Ello no fue aceptado por los especialistas en los idiomas 14 ó 16, pero ha sido repetido mecánicamente en la mayor parte de las compilaciones hechas desde entonces. Swadesh ha presentado una

proposición razonablemente bien documentada, a efecto de que se tenga al huave como un idioma OM. Rensch (tesis doctoral) no incluye al huave como idioma OM, pero su opinión actual, si fuere distinta, no se conoce.

6.14 MIXEZOQUE (14)

La familia mixezoque consiste de ocho idiomas, los cuales como lo sugiere la fonología comparada y la gramática, forman dos ramas —un grupo zoque y un grupo mixe, incluyendo el tapachulteco. Las cifras de la glotocronología, sin embargo, sugieren una división tripartita, como se presenta en el Cuadro I. La familia mixezoque fue correctamente identificada por Charencey en 1883. Estudios fonológicos comparativos han sido realizados por el misionero-lingüista norteamericano William Wonderly (1949) y el lingüista-antropólogo norteamericano Terrence Kaufman (1963 Ms.). Se han hecho estudios gramaticales por Wonderly y el misionero-lingüista norteamericano Benjamín Elson (1953; marcadores de pronombres personales), y por Kaufman (Ms. morfología en general). El antropólogo norteamericano George Foster ha sostenido (1942) —y Kaufman ha presentado la evidencia (1964)— que el tapachulteco es propiamente un miembro del grupo mixe.

El textistepeco, el sayula y el oluta son localmente denominados popoluca.

6.15 TOTONACO (15)

La familia totonaca (15) contiene sólo dos idiomas, de los cuales uno (tononaco), tiene cuando menos tres dialectos. Quizás el totonaco sea un *complejo*. Sapir (1929) registra correctamente el totonaco como un grupo separado, aunque el precedente en que se basa no está claro. Arana (1935) ha re-

construido la fonología prototonaca, basando su estudio en tres dialectos del totonaco y uno del tepehua.

6.16 MAYANCE (16)

La familia mayance (16) fue correctamente identificada por el etnógrafo alemán Otto Stoll en 1884. Esta familia que tiene de 24 a 30 idiomas y cerca de 2.5 millones de hablantes, es la familia lingüística más diversificada y populosa de Mesoamérica. El idioma huasteco está separado por cerca de 1,000 kilómetros del idioma mayance más cercano. Este hecho, tomando en cuenta que los huastecos no compartieron la civilización maya clásica, requiere una explicación histórica que comprenda la separación del huasteco del resto de la familia, hace más de 2,500 años. Aunque la extensión geográfica de los idiomas mayances es considerable, los pueblos, idiomas y culturas mayances (en contraste con los aztecas), nunca han sido particularmente expansionistas.

Se han hecho numerosos intentos para clasificar los idiomas mayances, cada uno con más información que el anterior. Cabe mencionar los estudios de Stoll (1884); Kroeber (1939a); el lingüista norteamericano Abraham Halpern (1942); el lingüista-antropólogo norteamericano Norman McQuown (1955); Swadesh (1960d) y Kaufman (1970). La clasificación que se presenta aquí es la de Kaufman, 1971. Esta reconoce, en el nivel más bajo, diez agrupaciones. Los especialistas han estado en desacuerdo en cuanto a las posiciones precisas del tojolabal y el chuj, el motozintleco, aguacateco, uspanteco y kekchí; tampoco hay opiniones firmes en cuanto al yucateco y al huasteco. No se ha publicado mucho trabajo comparativo acerca de la familia mayance. Stoll (1884) y el geógrafo alemán Karl Sapper (1912) reunieron listas comparativas de palabras en varios idiomas mayances, pero quienes han contribuido más a

los estudios mayas comparativos han sido McQuown (en las décadas de 1950 y de 1960) y Kaufman (en la de 1960). Los próximos diez años deberán mostrar un incremento importante en la extensión del trabajo descriptivo y comparado en el campo de los idiomas mayances, pues se ha reunido y organizado mucha información, todavía inédita.

### 6.17 LAS HIPOTESIS MACROMAYA Y MACROPENUTI

En 1931, L. S. Freeland trató de demostrar que el mixe (zoque) está relacionado con los idiomas "penuti", un supertronco que hasta entonces había estado limitado a California, Oregón, Washington y Columbia Británica.

En 1935, Whorf sugirió que entre el yutonahua, el taño, kayowe, penuti, mixezoque y mayance, existían semejanzas tales que indicaban la existencia de un supertronco, el cual propuso que se llamara macropenuti. La hipótesis obtuvo una favorable acogida durante algún tiempo, pero nunca fue demostrada, ni tomada muy en serio por los especialistas. Desde entonces los primeros tres idiomas han sido incorporados al aztecotaño (Whorf y George Trager 1937). En 1942 McQuown sugirió que el mixezoque y el totonaco podrían estar relacionados genéticamente y que los dos, a su vez, lo estarían respecto del mayance —el supertronco resultante sería llamado *macromaya*. Esto fue aceptado por Swadesh (1954b, 1956), aunque había sido rebatido por Wonderly (1953). Con posterioridad Swadesh (1966) sostuvo que el tarasco (17) probablemente pertenece al macromaya, aunque sus intentos para probar dicha tesis no resultan convincentes para los mayistas. Con la exclusión del tarasco, la hipótesis macromaya parece tan razonable a los especialistas como lo es la hipótesis joka a sus proponentes.

### 6.18 TARASCO (17)

Además de la unión del tarasco al macromaya, Swadesh ha sugerido que tanto el zuni como el quechua son parientes genéticos del primero, pero ello no ha tenido aceptación general.

### 6.19 XINCA Y LENCA (18-19)

Se ha sugerido (Lehmann 1920) que el xinca (18) y el lenca (19) están relacionados mutuamente, y que uno o ambos se hallan relacionados con el maya (16), el chibcha o el yutonahua (1). Ninguna de estas conjeturas ha sido demostrada como probable.

### 6.20 IDIOMAS FUERA DE LO QUE ES MESOAMERICA

Los siguientes idiomas de Centroamérica están fuera de lo que es Mesoamérica propiamente dicha: el paya (20) y la familia misumalpa (21). Se ha formulado la hipótesis de que el paya (20) está unido al chibcha y al caribe y quizás a otros idiomas más, pero ello no se ha hecho de modo convincente. La familia misumalpa (21) ha sido reconocida desde 1895 (Brinton). Desde esa fecha algunos estudiosos han creído que los tres idiomas y los complejos están coordinados (Lehman 1910, Swadesh 1967); otros creen que los dos primeros se agrupan *vis a vis* respecto del tercero (Brinton 1895, Conze-mius 1929) aunque la relación de familia puede comprobarse por medio de la inspección, no se ha publicado trabajo comparativo alguno en que dicha relación pueda apoyarse.

El caribe (garífuna) de la costa caribe de Belice, Guatemala y Honduras es un idioma arahuaco, introducido a Centroamérica durante el siglo XVIII.