

Leopoldo Valiñas Coalla y Mario Cortina Borja  
*Contribución a la reconstrucción histórica a partir de métodos estadísticos en datos léxicos: El caso de las lenguas sonorenses*

XII Simposio de Historia y Antropología de Sonora. 387

CONTRIBUCION A LA RECONSTRUCCION HISTORICA A PARTIR DE METODOS ESTADISTICOS EN DATOS LEXICOS: EL CASO DE LAS LENGUAS SONORENSES

LEOPOLDO VALIÑAS COALLA  
MARIO CORTINA BORJA.

INTRODUCCION.

Hablar de la historia de los grupos del noroeste de México antes del siglo XVI es difícil. Las fuentes coloniales no nos permiten reconstruir las situaciones sociales y culturales que predominaban en el área para antes de ese siglo. Mientras más hacia atrás en el tiempo queramos ir, más problemático es el procedimiento de reconstrucción. La inferencia parece ser el unico recurso viable.

Se ha dejado a la arqueología la responsabilidad de dar cuenta de ese pasado que no fue narrado, del que no fue nada escrito. Nosotros presentamos en este trabajo, otro camino para la reconstrucción histórica previa al siglo XVI. Es a partir del uso de informaciones lingüísticas y mediante métodos estadísticos de análisis de datos que planteamos una proyección hacia el pasado.

Así pues, lo que proponemos es una posible evolución del noroeste de México anterior al siglo XVI. Trabajos posteriores podrán refutar o consolidar lo aquí dicho.

Una advertencia: este trabajo no es glotocronológico; no hablamos de fechas ni de siglos mínimos de divergencia.

EL PROBLEMA.

Los trabajos de lingüística histórica aplicados a las lenguas americanas han dividido a las lenguas yutoaztecas en cinco grandes grupos. Uno de ellos es el llamado sureño. Este incluye al náhuatl (con todas sus variantes), a los idiomas de Sonora (excluyendo al seri) y a los de la Sierra Madre Occidental, desde Chihuahua hasta Jalisco. Este grupo se divide en dos

Memoria del XII Simposio de Historia y Antropología, Volumen 1, pp.

387 - 412, 1987

Universidad de Sonora, Departamento de Historia y Antropología

© Derechos Reservados. Hermosillo, Sonora, México.

Leopoldo Valiñas Coalla y Mario Cortina Borja  
*Contribución a la reconstrucción histórica a partir de métodos estadísticos en datos léxicos: El caso de las lenguas sonorenses*

388

grandes troncos, el sonorense y el aztecaño (constituido éste por todos los dialectos nahuas, desde el mexicano de Durango y Nayarit hasta el pipil hablado en El Salvador).

Las lenguas sonorenses se dividen a su vez, en cuatro grupos principales: el tepimano (que incluye al pápago, névome, pima bajo, tepehuano del sur, tepehuano del norte y tepecano), el taracahita (que agrupa al tarahumara, guarijío, ópata, eu-deve, yaqui y mayo), el tubar (que es una lengua aislada) y el corachol (constituido por el cora y el huichol). En el mapa 1 se puede apreciar la distribución geográfica de las lenguas sonorenses para el siglo XVI.

Dentro de los supuestos de la teoría lingüística hay uno que, en especial, nos llama la atención: las lenguas de una misma familia o grupo deben ser vecinas. Es decir, se supone que un grupo o familia lingüística conforma una (llamémosle así) continuidad lingüística, y que ésta se manifiesta en una continuidad geográfica. Si observamos la distribución geográfica de las lenguas sonorenses, particularmente en el caso de las lenguas tepimanas, es claro que este supuesto no se cumple: el pápago y el névome están separados de los tepehuanos por lenguas tarahumaras.

#### HIPOTESIS.

Ante este problema, Sauer (1934) propone que esa discontinuidad se debe a una intromisión por parte de grupos tarahumaras y guarijíos en el área tepimana. Esta hipótesis parte del supuesto de un movimiento poblacional gradual de norte a sur por parte de todos los grupos sonorenses y de una penetración tarahumara y guarijía hacia el este.

Por su parte, Miller (1981) cree que en realidad tal separación se debe a un movimiento migratorio hacia el norte por parte de los pápagos y pimas. Él piensa que, en un momento histórico, las lenguas sonorenses (cuando aún no se habían diferenciado -cuando conformaban podemos decirlo, una sola lengua-), se hallaban concentradas en un área que podemos localizar al sur del estado de Sonora y norte de Sinaloa, en las faldas de la sierra y en extensión de los cañones bajos.

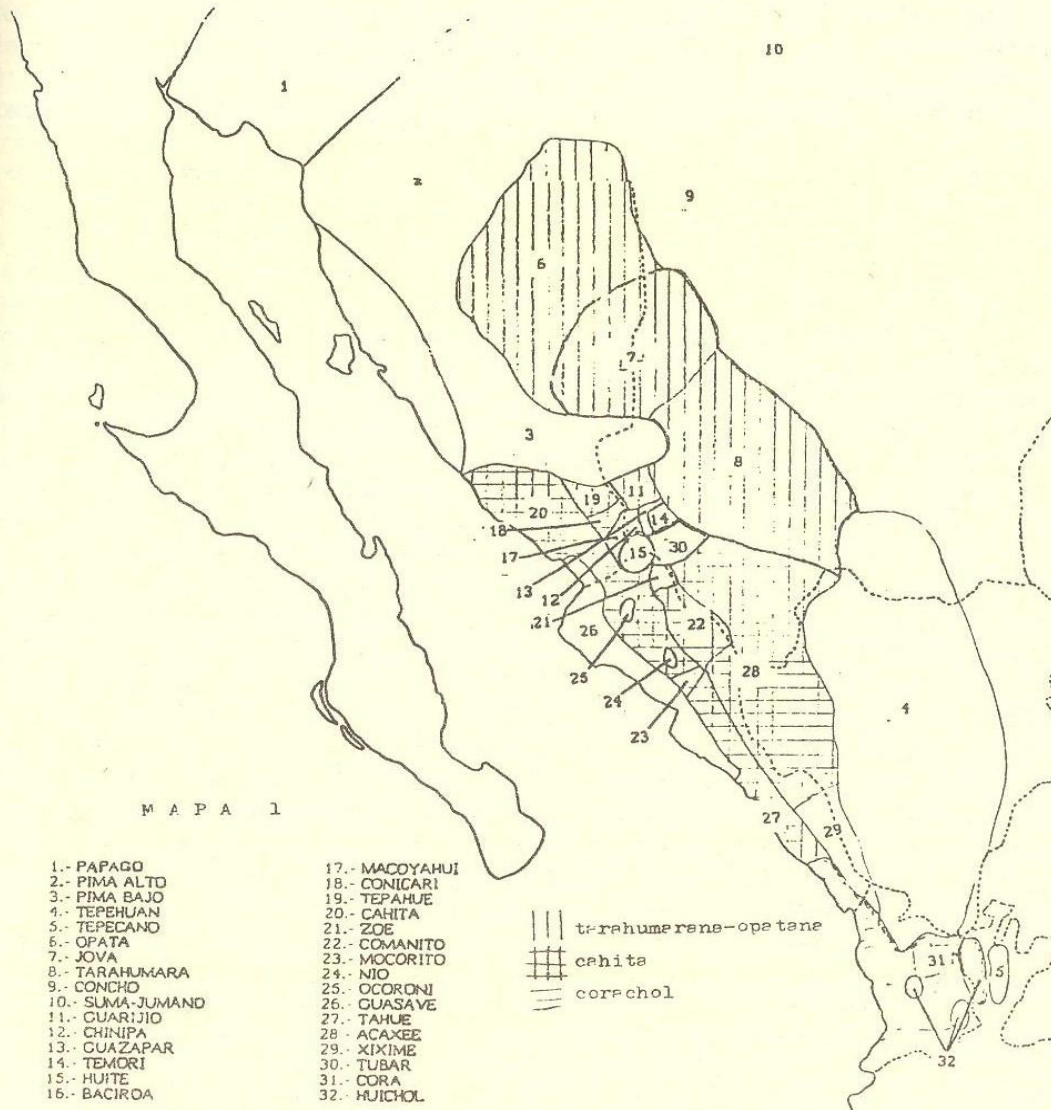
Memoria del XII Simposio de Historia y Antropología, Volumen 1, pp.

387 - 412, 1987

Universidad de Sonora, Departamento de Historia y Antropología

© Derechos Reservados. Hermosillo, Sonora, México.

Leopoldo Valiñas Coalla y Mario Cortina Borja  
*Contribución a la reconstrucción histórica a partir de métodos estadísticos en datos léxicos: El caso de las lenguas sonorenses*



Memoria del XII Simposio de Historia y Antropología, Volumen 1, pp.  
 387 - 412, 1987

Universidad de Sonora, Departamento de Historia y Antropología  
 © Derechos Reservados. Hermosillo, Sonora, México.

Leopoldo Valiñas Coalla y Mario Cortina Borja  
*Contribución a la reconstrucción histórica a partir de métodos estadísticos en datos léxicos: El caso de las lenguas sonorenses*

390

DATOS.

Tratando de resolver el problema, comenzamos a buscar la manera de correlacionar lo lingüístico con lo geográfico. Lo que pretendíamos era poder medir el grado de discontinuidad que se observaba.

Para hablar de continuidad lingüística, lo más adecuado era considerar las semejanzas existentes entre las diversas lenguas. Esta información la obtuvimos de una matriz léxica (constituida por una lista de 100 palabras en 32 lenguas yutoaztecas) y otra de similaridades (resultante de la anterior) publicadas por Miller en 1985. Esa lista léxica es una pequeña modificación hecha por Miller al llamado Vocabulario Básico de Swadesh. Esta clase de información podría considerarse como un buen indicador para medir la continuidad lingüística; esto, por tres razones. En primer lugar, porque ya se contaba con un trabajo de esta naturaleza (Miller, 1985). En segundo lugar, a esta matriz se le podían añadir fácilmente otras lenguas. Y, finalmente, tal léxico tiene la particularidad de poder mostrar los contactos con otros grupos y lenguas.

La información geográfica fue obtenida a partir de diversas obras que tratan sobre distribuciones actual e histórica de los diversos grupos étnicos de México (cf. Miller, 1981, Olivera et al., 1982, Mendizábal y Jiménez Moreno, 1937).

METODOLOGIA.

De la matriz léxica de Miller eliminamos (en un primer momento) las lenguas no sureñas, agregamos otras seis (tepehuano del sur, tepecano y cuatro variantes nahuas) y corregimos algunas entradas, reduciendo así el número de lenguas a 22, de las cuales 8 eran variantes nahuas. De esta matriz obtuvimos luego la de distancias o disimilaridades (en las que las parejas de lenguas más parecidas tendrían los menores índices: estarían menos separadas). Para cada par de lenguas, las distancias léxicas se calcularon con la siguiente fórmula:

$$d = 1 - (n - n(x) - n(0) / (100 - \& (0) ),$$

donde  $d$  es distancia léxica,  $n$  es el número de coincidencias,  $x$  corresponde a las no cognadas y  $0$  a las faltantes y  $\&$  es

Leopoldo Valiñas Coalla y Mario Cortina Borja  
*Contribución a la reconstrucción histórica a partir de métodos estadísticos en datos léxicos: El caso de las lenguas sonorenses*

XII Simposio de Historia y Antropología de Sonora. 391

la presencia en al menos una de las lenguas.

Para esto hay que aclarar que el vocabulario de una lengua se puede dividir en dos clases. Aquella que es reticente al reemplazo; que dado lo común, lo usado, tiende a mantenerse a lo largo del tiempo y del espacio, sufriendo principalmente cambios propios de los niveles lingüísticos (i.e. en su fonología y estructura) más que sustituciones. Dentro de este léxico podemos mencionar palabras como sol, luna, nube, uno, cabeza, blanco, dormir, etc. El vocabulario de la matriz léxica con la que trabajamos es de este tipo.

La otra parte del vocabulario es la más voluble, la más susceptible de ser sustituida o de mostrar interferencias léxicas de otras lenguas. Dentro de este tipo de léxico tenemos el referido, principalmente, a la flora y fauna, elementos que son cambiantes según el área. Este léxico aporta información más que lingüística, cultural: interrelaciones con otros grupos, visión del mundo, enfrentamiento con la naturaleza, etc.

Nosotros esperábamos que con el léxico del primer tipo las distancias lingüísticas estuvieran predominando sobre las distancias geográficas en nuestro análisis, lo que no hubiera sucedido si hubiéramos utilizado el vocabulario sobre flora y fauna.

Resumiendo, los datos con los que contábamos eran: a) la distribución geográfica, b) una matriz léxica de 100 palabras en 22 lenguas y c) una matriz de distancias léxicas entre esas lenguas.

#### METODOLOGIA ESTADISTICA.

Uno de los objetivos medios del trabajo era proponer una clasificación de las lenguas con el fin de utilizarla dinámicamente; esto es, para interpretarla en términos históricos. Con este fin, se utilizaron métodos de clasificación como análisis de cúmulos (Anderberg 1973, Espinosa *et. al.*, 1980) y escalamiento multidimensional en sus versiones clásica y ordinal (Shepard *et. al.*, 1972). En ambos se puede partir de una matriz de distancias.

En el análisis de cúmulos se tienen siete métodos diferentes para construir árboles de clasificación jerárquica (i.e.,

Leopoldo Valiñas Coalla y Mario Cortina Borja  
*Contribución a la reconstrucción histórica a partir de métodos estadísticos en datos léxicos: El caso de las lenguas sonorenses*

392

formando grupos anidados en grupos más generales). Esto se hizo utilizando el paquete ANALISIS/CUMULOS (Reyes *et. al.*, 1981).

El escalamiento multidimensional intenta construir una configuración que refleje las relaciones inducidas en los objetos por la matriz de datos en algún espacio de pocas dimensiones (dos o tres, como en este trabajo). Hay, fundamentalmente, dos maneras de generarla: la clásica y la ordinal. Fue posible efectuar los cálculos necesarios para ambos métodos gracias al conjunto de rutinas implementadas por Iraizos (1984).

En el escalamiento clásico se busca la configuración solución por medio de proyecciones obtenidas por la variación de la matriz de datos; en el ordinal se trata de conservar en la configuración solución el orden que tienen las distancias de cada objeto con los demás. En la primera forma, se evalúa la calidad del ajuste de la configuración propuesta con la matriz de distancias observada en términos de la proporción de varianza explicada por la solución. Como esta proporción es aditiva y decreciente respecto al número de dimensiones (es aditiva porque la proporción de varianza explicada aumenta al incrementarse el número de dimensiones y es decreciente debido a que la componente  $i + 1$  explica menos varianza que la componente  $i$ ), las primeras dimensiones pueden trabajarse como independientes de las superiores. En el escalamiento ordinal, la bondad de ajuste se mide en términos de un índice llamado stress, que puede considerarse como un porcentaje de discrepancias entre las distancias inducidas por la configuración solución y la matriz de distancias observada.

Casi todos los cálculos involucrados en este trabajo se realizaron en la Burroughs B7800 de la Dirección General de Cómputo Académico de la UNAM y en las microcomputadoras del IIA-UNAM.

#### PROCESO.

Lo primero que decidimos hacer fue obtener las distancias lingüísticas y su agrupación. Los resultados, además de corroborar la agrupación que Miller ha propuesto, mostraban ciertas situaciones interesantes, por ejemplo: a) la separación temprana

Memoria del XII Simposio de Historia y Antropología, Volumen 1, pp.  
387 - 412, 1987

Universidad de Sonora, Departamento de Historia y Antropología  
© Derechos Reservados. Hermosillo, Sonora, México.

Leopoldo Valiñas Coalla y Mario Cortina Borja  
*Contribución a la reconstrucción histórica a partir de métodos estadísticos en datos léxicos: El caso de las lenguas sonorenses*

na del tubar con respecto a las otras lenguas sonorenses, b) el hecho de que el cora y el huichol se diferencian mutuamente casi inmediatamente después de su separación como grupo, c) el prolongado mantenimiento de las lenguas tepimanas como un solo grupo, d) la separación del tarahumara y el guarijío del grupo taracahita antes que la del ópata, y, e) lo que buscábamos principalmente en este momento, la separación clara entre las lenguas nahuas y las sonorenses.

Con la información lingüística existente, decidimos que el dendrograma que mejor reflejaba la agrupación lingüística era el obtenido por el método de conexión simple y que aparece en la figura 1.

Cabe mencionar que el dendrograma fue interpretado en términos históricos; esto implica el supuesto de que mientras más pronto ocurra una separación en el dendrograma, ésta es más antigua.

Como se ve en la figura 1, las lenguas sonorenses se dividen, primero, entre la rama corachol y el resto de ellas. Dentro de éstas, el tubar se desprende casi inmediatamente después. Un hecho importante en la parte azteca es que la primera bifurcación se da entre el pochuteco y el resto de estas lenguas, lo que es consistente con las clasificaciones lingüísticas conocidas. Esta clase de consideraciones dieron elementos para validar a este dendrograma y así seleccionarlo de entre los siete dendrogramas obtenidos.

Posteriormente, utilizando escalamiento multidimensional en sus dos variantes, construimos configuraciones en 2 y 3 dimensiones con las que se obtuvo un reflejo satisfactorio (en términos de proporción de varianza explicada o de stress) de las interrelaciones de las lenguas según sus distancias léxicas. Lo que nos interesaba era comprobar (con un método independiente del de cúmulos) que las variantes nahuas quedaran efectivamente aisladas de las sonorenses; esto resultó ser cierto. Sin embargo, también se vió el aislamiento de la tepimanas y la proximidad del tubar a las lenguas taracahitas (fig. 2).

En general, sólo utilizamos los resultados de escalamientos clásicos; en todos los casos, con los primeros dos componentes

Leopoldo Valiñas Coalla y Mario Cortina Borja  
*Contribución a la reconstrucción histórica a partir de métodos estadísticos en datos léxicos: El caso de las lenguas sonorenses*

394

se tuvo arriba del 80% de varianza explicada, porcentaje que alcanzaba hasta el 95% al añadir una dimensión más. Igualmente, las medidas de stress estuvieron siempre por debajo del 12%, lo que indica un buen ajuste.

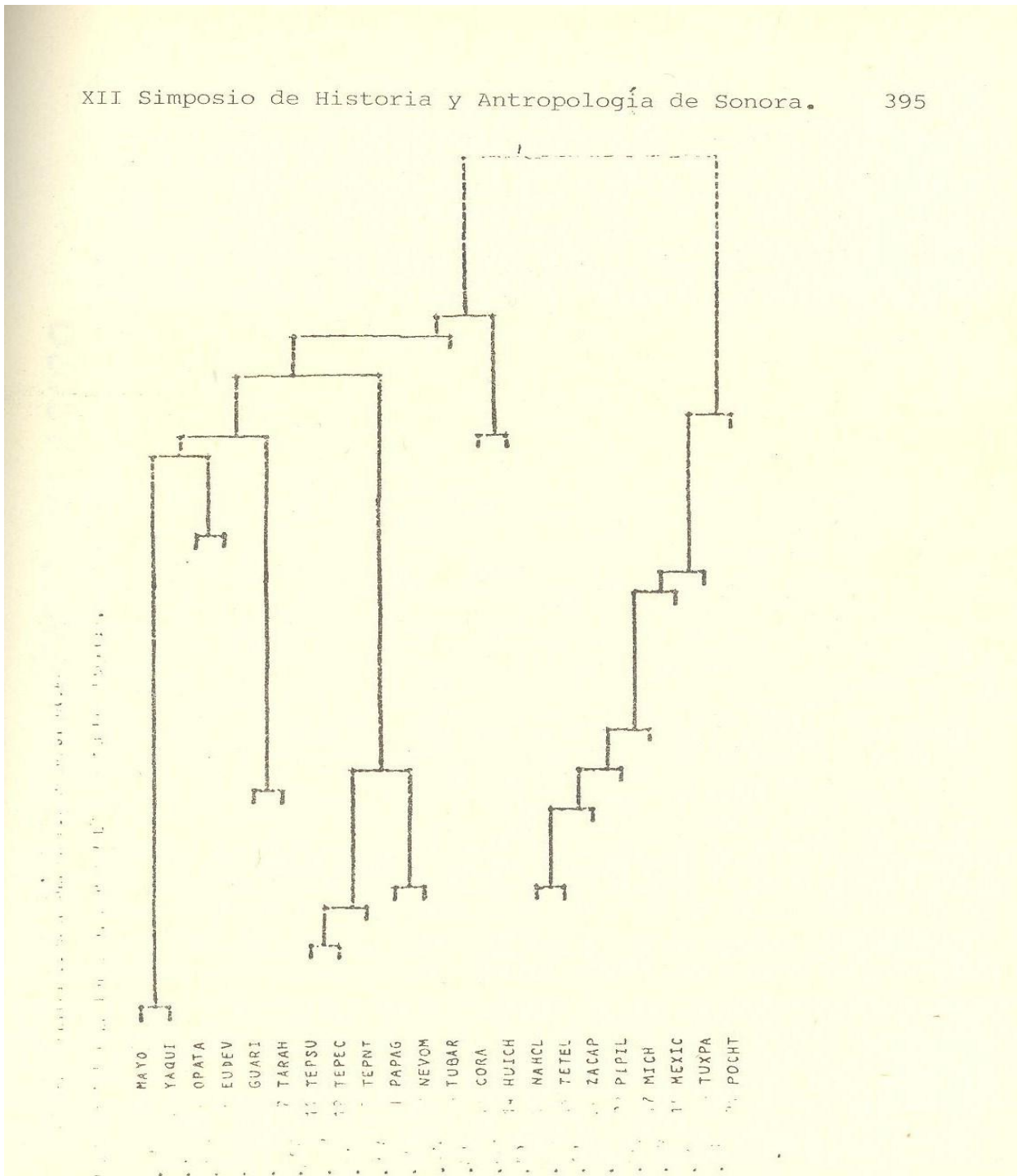
La figura 3 muestra las primeras tres coordenadas obtenidas con las 22 lenguas. Como se ve, las primeras dos coordenadas caracterizan bastante bien a los tres principales grupos: tepimano, taracahita-tubar-corachol y azteca. La curva de la tercera coordenada alcanza a diferenciar claramente al corachol (valores muy altos) y al taracahita (valores negativos).

Una de las posibilidades analíticas del escalamiento multidimensional consiste en que puede permitir dar una interpretación a las dimensiones obtenidas asociándolas con características no observables a partir de la matriz de distancias porque, se sabe, inciden en ella. Como resultado de una búsqueda en esta dirección obtuvimos estadísticas descriptivas para las distancias de cada lengua (tabla 1). Al graficar la desviación estándar contra la media (fig.4), se ve que algunas lenguas nahuas aparecen mezcladas con los tepimanos, si bien otras de aquel grupo (el pochuteco, el mexicanero y el náhuatl de Tuxpan) se encuentran totalmente aisladas. Algo similar ocurre con el cora. Estos hechos pueden considerarse como debido a que los tepimanos y algunos nahuas, al tener las mayores desviaciones estándares y las menores medias, son grupos de alta cohesión y con grandes diferencias con el resto de las lenguas yutoaztecas. Lo contrario sucede con el cora, el tubar, el ópata y el huichol, que al tener las menores desviaciones estándares junto con las mayores medias, representan a lenguas que están consistentemente muy separadas del resto. La aparición del ópata en este grupo de lenguas puede explicarse más bien por la gran cantidad de valores "0" que tuvo esta lengua que por un aislamiento real.

En vista de que la presencia de lenguas nahuas en el análisis reducía las distancias existentes dentro del grupo de lenguas sonorenses (pues las distancias entre éstas y aquellas eran relativamente mayores), decidimos eliminarlas y de esta manera, únicamente trabajar con 14 lenguas. El mismo proceso seguido para las 22 lenguas se efectuó con estas 14.



Leopoldo Valiñas Coalla y Mario Cortina Borja  
*Contribución a la reconstrucción histórica a partir de métodos estadísticos en datos léxicos: El caso de las lenguas sonorenses*



Memoria del XII Simposio de Historia y Antropología, Volumen 1, pp.  
387 - 412, 1987

Universidad de Sonora, Departamento de Historia y Antropología  
© Derechos Reservados. Hermosillo, Sonora, México.

Leopoldo Valiñas Coalla y Mario Cortina Borja  
*Contribución a la reconstrucción histórica a partir de métodos estadísticos en datos léxicos: El caso de las lenguas sonorenses*

396

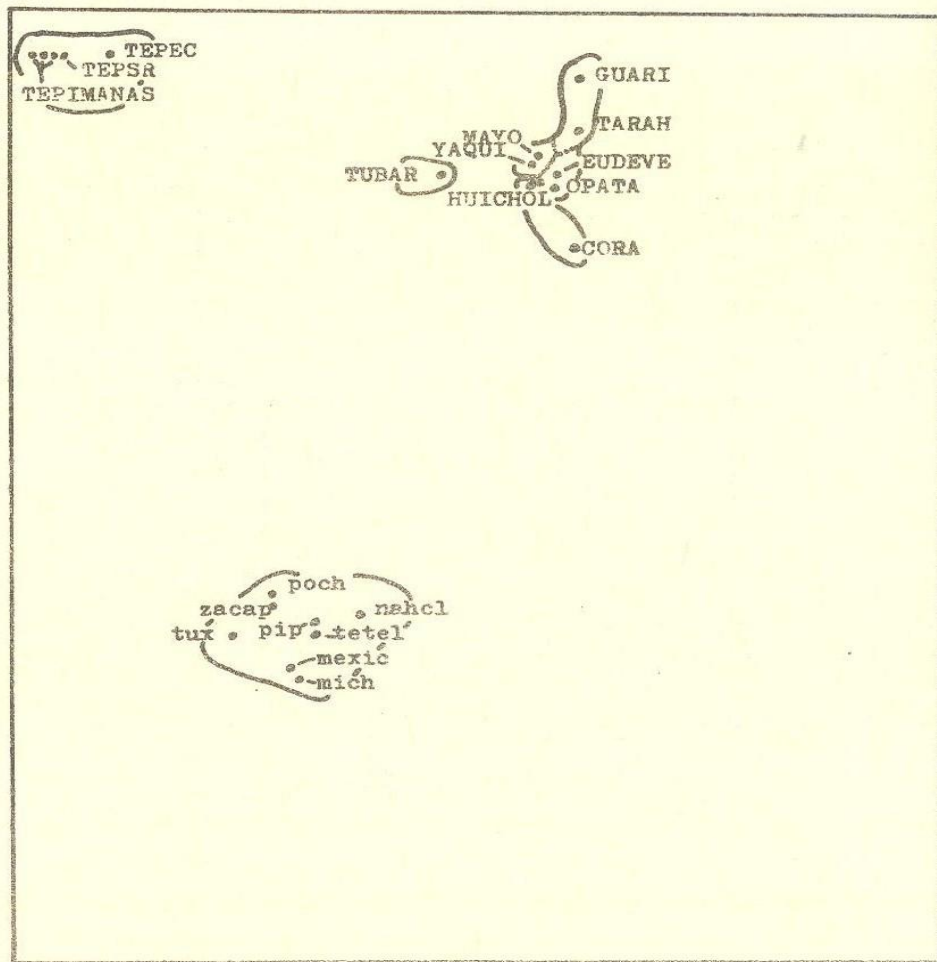
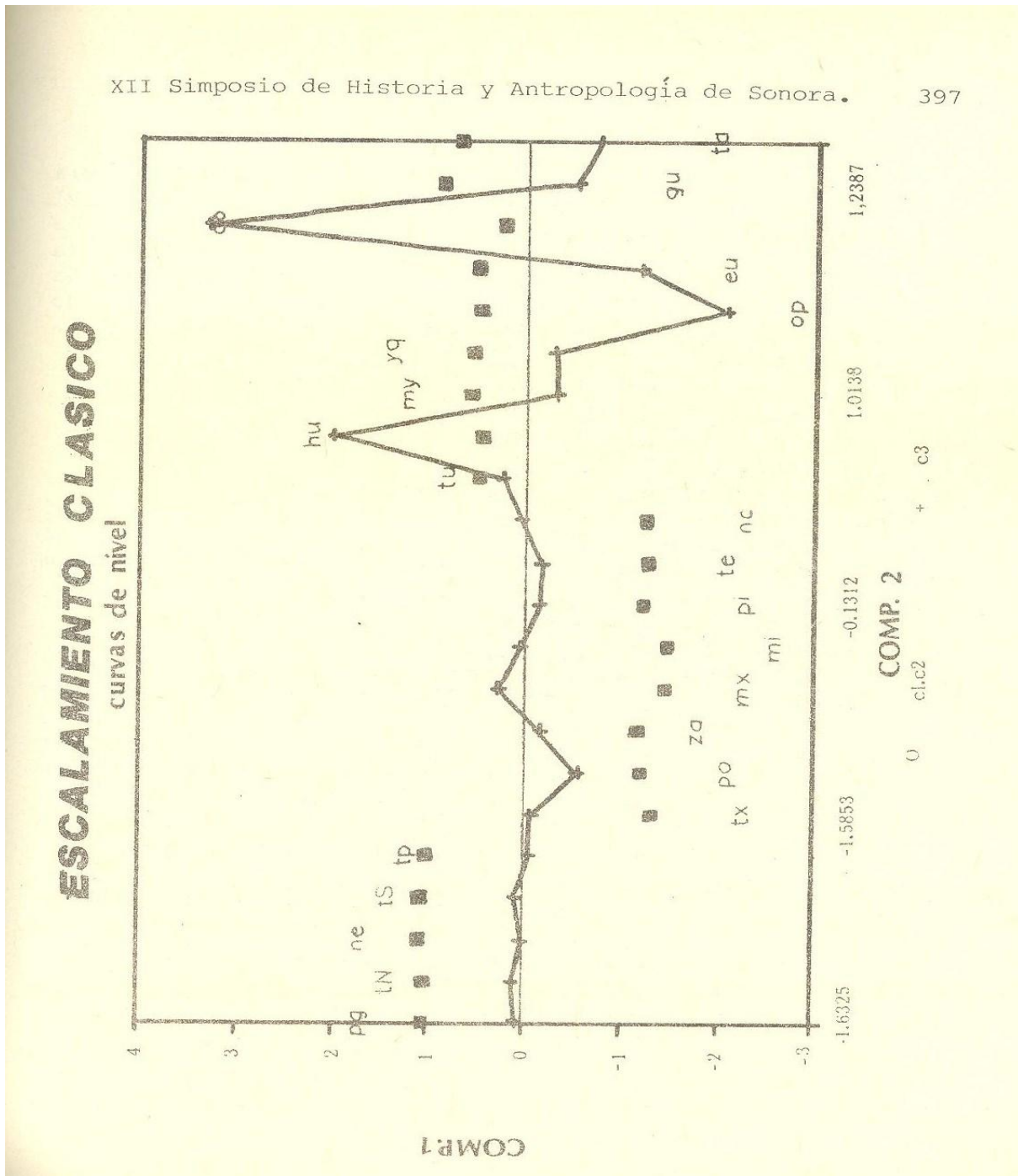


FIGURA 2: Escalamiento clásico 2-Dim. (22 lenguas)



Memoria del XII Simposio de Historia y Antropología, Volumen 1, pp.  
 387 - 412, 1987

Universidad de Sonora, Departamento de Historia y Antropología  
 © Derechos Reservados. Hermosillo, Sonora, México.

Leopoldo Valiñas Coalla y Mario Cortina Borja  
*Contribución a la reconstrucción histórica a partir de métodos estadísticos en datos léxicos: El caso de las lenguas sonorenses*

398

El dendrograma que aportaba más claramente elementos para la clasificación de las lenguas fue también el obtenido con el método de conexión simple (fig.5). La figura 6 muestra la configuración en dos dimensiones y la figura 7 añade la curva de la tercera dimensión para el escalamiento clásico. Llama la atención la gran cohesión dentro del grupo tepimano así como lo marcado de su separación con el resto de las lenguas y la diferenciación que hay en la tercera dimensión para el tubar-mayo-yaqui y el guarijío-tarahumara.

Se obtuvieron las estadísticas descriptivas para cada lengua y se graficó la desviación estándar contra la media. Esto aparece en la fig. 8. Como era de esperarse, aquí aparecen las relaciones principales entre las distancias de las lenguas sonorenses mucho más claramente que en la figura 4: los grupos están muy bien definidos y hay una obvia tendencia a disminuir la media de las distancias conforme aumenta la desviación estándar.

La parte geográfica del análisis se hizo con base en reducir la localización de las lenguas a una sola dimensión utilizando escalamiento multidimensional. Era necesaria una sola dimensión debido a que ésta sería incorporada al análisis como variable respuesta en función de la parte léxica. La matriz de distancias necesaria se obtuvo caracterizando a las lenguas por un punto sobre el mapa, buscando que correspondiera a las locaciones dadas en los trabajos consultados. Esto generó una configuración en dos dimensiones y sobre ella se calcularon distancias euclidianas entre cada punto. La proporción de varianza explicada por este índice fue de 96%.

El siguiente paso fue correlacionar lo léxico con lo geográfico. Esto se hizo expresando al índice geográfico en función de las primeras tres componentes dadas por el escalamiento multidimensional clásico con un modelo de regresión lineal. Esperábamos que no hubiera correlación, dada la separación geográfica que existía en el grupo tepimano. El resultado fue que las correlaciones, si bien no eran despreciables, tampoco eran significativas en vista de contar sólo con 14 datos. Esto nos llevó a considerar que el supuesto de no correspondencia entre lo léxico y lo lingüístico era plausible.

Leopoldo Valiñas Coalla y Mario Cortina Borja  
*Contribución a la reconstrucción histórica a partir de métodos estadísticos en datos léxicos: El caso de las lenguas sonorenses*

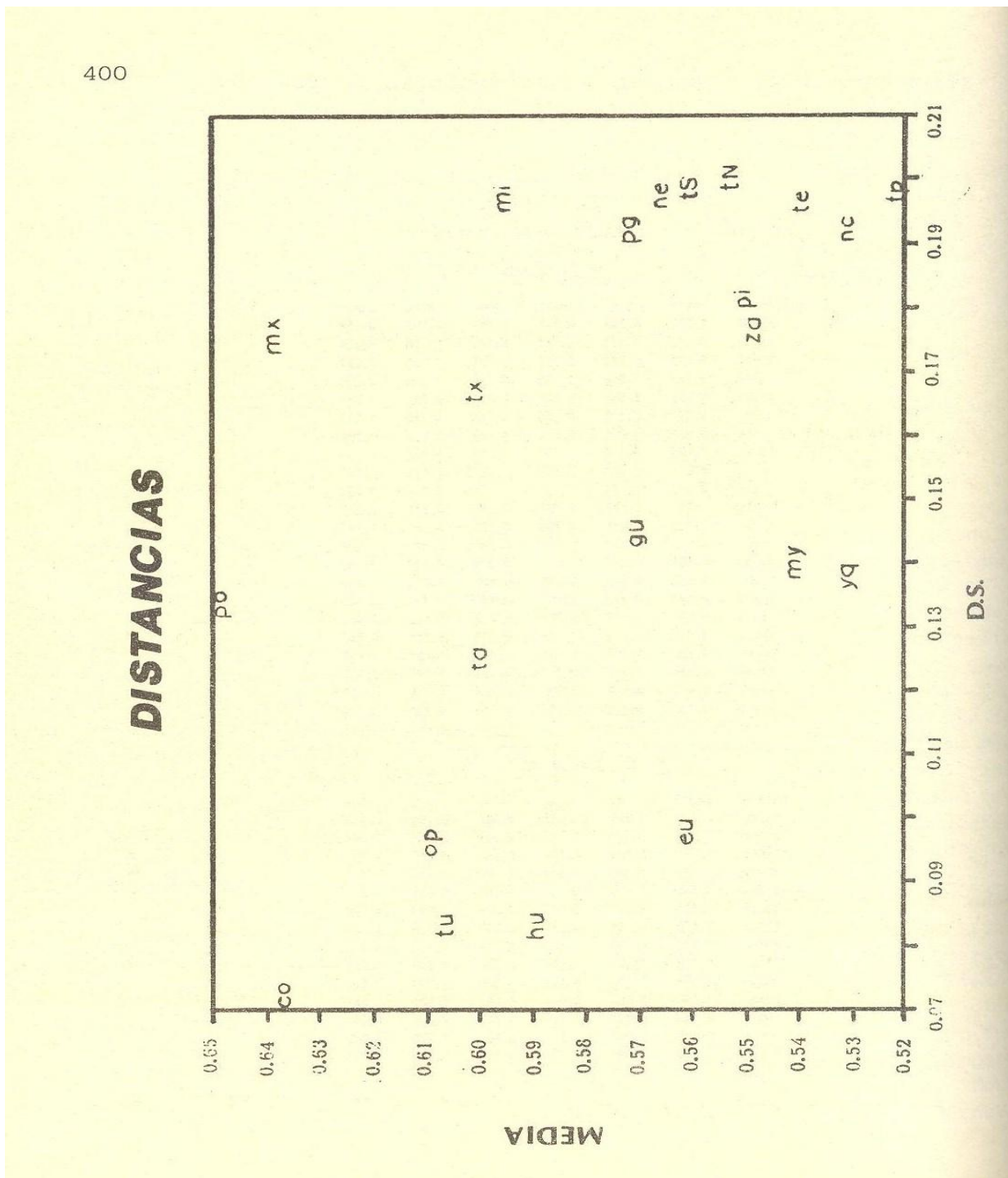
T A B L A 1  
 ESTADÍSTICAS DE DISTANCIAS

\*\* 22 LENGUAS \*\*

nombr	media	d.s.	c.v.	max.	min.	rango
papag	0.572	0.192	0.336	0.789	0.150	0.639
nevom	0.566	0.197	0.348	0.800	0.150	0.650
tepNT	0.553	0.200	0.362	0.777	0.130	0.647
opata	0.609	0.096	0.158	0.750	0.359	0.391
eudev	0.561	0.098	0.174	0.723	0.359	0.364
guari	0.571	0.145	0.254	0.768	0.210	0.558
tarah	0.600	0.125	0.208	0.768	0.210	0.558
tubar	0.607	0.083	0.137	0.769	0.474	0.295
mayo	0.541	0.140	0.259	0.757	0.070	0.687
yaqui	0.531	0.138	0.259	0.726	0.070	0.656
tepSR	0.561	0.199	0.354	0.779	0.109	0.670
tepec	0.522	0.198	0.380	0.756	0.109	0.647
cora	0.637	0.072	0.113	0.747	0.423	0.324
huich	0.589	0.083	0.141	0.729	0.423	0.306
mexic	0.639	0.176	0.275	0.800	0.326	0.474
tuxpa	0.601	0.167	0.277	0.745	0.340	0.405
mich	0.596	0.197	0.331	0.768	0.242	0.526
nahcl	0.531	0.192	0.361	0.710	0.150	0.560
tetel	0.540	0.197	0.364	0.710	0.150	0.560
zacap	0.549	0.176	0.321	0.710	0.190	0.520
pocht	0.649	0.134	0.206	0.768	0.434	0.334
pipil	0.550	0.181	0.329	0.710	0.220	0.490

\*\* 14 LENGUAS \*\*

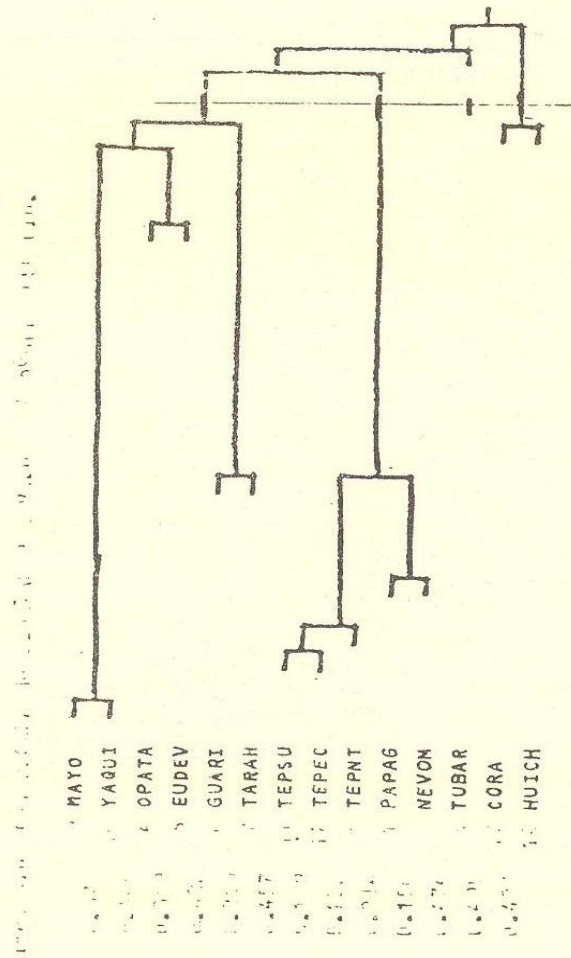
nombr	media	d.s.	c.v.	max.	min.	rango
papag	0.473	0.181	0.383	0.660	0.150	0.510
nevom	0.181	0.179	0.389	0.660	0.150	0.510
tepnt	0.383	0.389	0.417	0.653	0.130	0.523
opata	0.660	0.660	0.653	0.688	0.359	0.329
eudev	0.150	0.150	0.130	0.359	0.359	0.233
guari	0.510	0.510	0.523	0.329	0.233	0.370
tarah	0.529	0.104	0.197	0.630	0.210	0.420
tubar	0.557	0.045	0.081	0.618	0.474	0.144
mayo	0.468	0.128	0.274	0.570	0.070	0.500
yaqui	0.459	0.124	0.270	0.550	0.070	0.480
tepsu	0.458	0.183	0.400	0.650	0.109	0.541
tepec	0.413	0.174	0.421	0.630	0.109	0.521
cora	0.605	0.071	0.117	0.688	0.423	0.265
huich	0.536	0.049	0.091	0.588	0.423	0.165



Memoria del XII Simposio de Historia y Antropología, Volumen 1, pp. 387 - 412, 1987

Universidad de Sonora, Departamento de Historia y Antropología  
© Derechos Reservados. Hermosillo, Sonora, México.

Leopoldo Valiñas Coalla y Mario Cortina Borja  
*Contribución a la reconstrucción histórica a partir de métodos estadísticos en datos léxicos: El caso de las lenguas sonorenses*

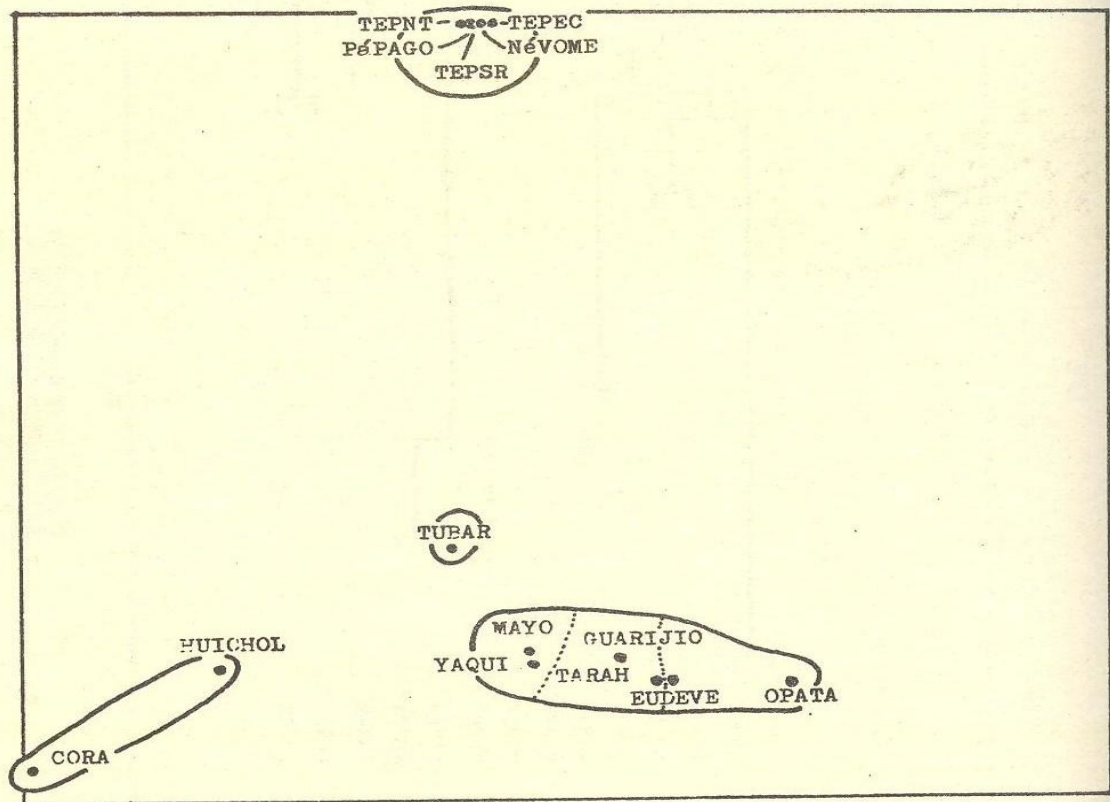


Memoria del XII Simposio de Historia y Antropología, Volumen 1, pp.  
387 - 412, 1987

Universidad de Sonora, Departamento de Historia y Antropología  
© Derechos Reservados. Hermosillo, Sonora, México.

Leopoldo Valiñas Coalla y Mario Cortina Borja  
*Contribución a la reconstrucción histórica a partir de métodos estadísticos en datos léxicos: El caso de las lenguas sonorenses*

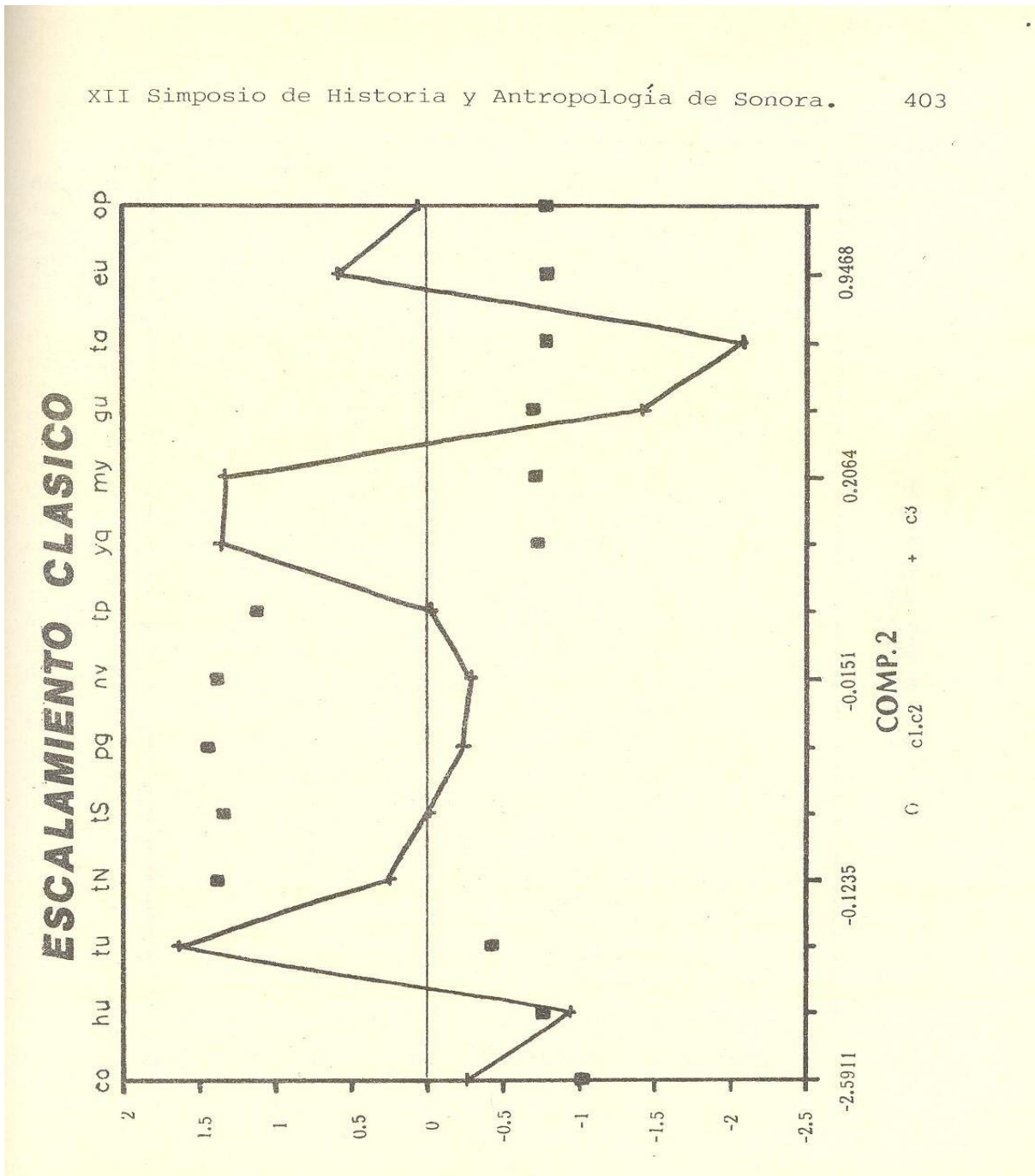
402

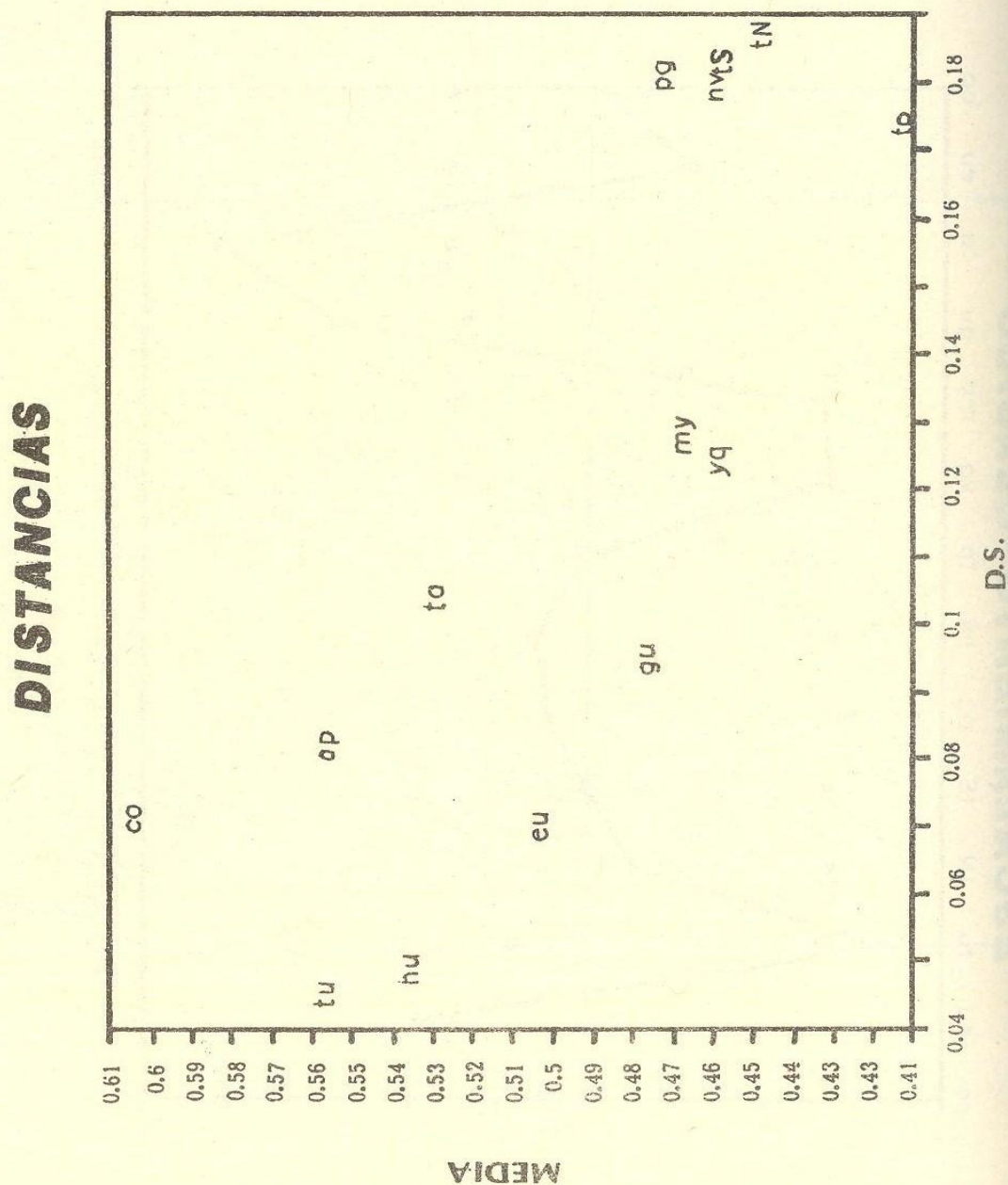


Memoria del XII Simposio de Historia y Antropología, Volumen 1, pp.  
387 - 412, 1987

Universidad de Sonora, Departamento de Historia y Antropología  
© Derechos Reservados. Hermosillo, Sonora, México.







Memoria del XII Simposio de Historia y Antropología, Volumen 1, pp.  
387 - 412, 1987

Universidad de Sonora, Departamento de Historia y Antropología  
© Derechos Reservados. Hermosillo, Sonora, México.

Leopoldo Valiñas Coalla y Mario Cortina Borja  
*Contribución a la reconstrucción histórica a partir de métodos estadísticos en datos léxicos: El caso de las lenguas sonorenses*

XII Simposio de Historia y Antropología de Sonora. 405

Es interesante hacer notar que las correlaciones parciales de la primera y la tercera componente con el índice geográfico fueron prácticamente cero. Fue la segunda componente la que mejor correspondió con lo geográfico, sin embargo, no fue suficiente para dejar de considerar a la separación locacional de grupo tepimano como significativa.

Esta falta de correlación nos permitió plantearnos el origen de tal discontinuidad. A partir de las hipótesis de Miller y Sauer nos propusimos simular los desplazamientos poblacionales que la causaron. Cortamos el dendrograma en un momento en el que las diferencias lingüísticas entre las lenguas eran mínimas y obtuvimos la matriz de distancias correspondiente y las configuraciones en dos y tres dimensiones para las lenguas. Considerando las hipótesis, ubicamos geográficamente a las lenguas construyendo un índice geográfico análogo al mencionado antes.

Al efectuar las regresiones vimos que los modelos propuestos no explicaban satisfactoriamente la presencia del cora y del pápago; esto a pesar de mostrar una altísima correlación. Es decir, o lo geográfico había sido mal planteado, o lo léxico no correspondía con las hipótesis.

Buscando solución a este problema, resolvimos proponer una tercera hipótesis: Para esto, hubo tres hechos que determinaron nuestra propuesta: a) la separación del tepimano debió haber sido muy rápida y cercana al siglo XVI, b) las distancias del cora con respecto a las otras lenguas eran demasiado grandes, no así las del huichol y c) el tubar estaba más cercano a sus lenguas vecinas (yaqui, mayo, eudeve y guarijío) que al resto.

Los argumentos que nos permitieron no considerar a plenitud la hipótesis de Sauer fueron la presencia de los ópatas al este de los pápagos (hecho que no es apoyado por el dendrograma). Por otro lado, la hipótesis de Miller muestra, a nuestro ver, un inconveniente: mientras que había constantes movimientos poblacionales (reflejados por la diversidad lingüística en la región), los tepimanos se hallaban, aparentemente, muy homogéneos y casi inmóviles, para luego empezar a moverse aceleradamente, unos hacia el norte y otros hacia el sur, lo que resulta difícil de explicar.

Memoria del XII Simposio de Historia y Antropología, Volumen 1, pp.  
387 - 412, 1987

Universidad de Sonora, Departamento de Historia y Antropología  
© Derechos Reservados. Hermosillo, Sonora, México.

Leopoldo Valiñas Coalla y Mario Cortina Borja  
*Contribución a la reconstrucción histórica a partir de métodos estadísticos en datos léxicos: El caso de las lenguas sonorenses*

406

En vista de todo lo anterior proponemos la siguiente secuencia evolutiva para las lenguas sonorenses hasta el siglo XVI.

Nuestra propuesta se puede ver en las figuras 9a-9i. Esta está hecha a partir del dendrograma que jerarquiza a las 14 lenguas sonorenses (fig. 5) y del escalamiento multidimensional (fig. 6 y 7).

En estas figuras se propone un momento inicial en el que todas las lenguas sonorenses forman una sola lengua apenas separada de las demás yutoaztecas; podemos ubicarlo geográficamente en el norte de Sonora y el sur de Estados Unidos.

En el movimiento del grupo hacia el sur, se realiza una primera separación: un pequeño grupo sureño se desprende. Más adelante, otra separación sureña. El primer grupo logra avanzar por las planicies y se separa rápidamente. Es el corachol. El segundo grupo es el tubar.

Al paso del tiempo, el grupo mayor sufre otra división: las lenguas taracahitas logran diferenciarse de las tepimanas.

Esta situación se mantiene por algún tiempo. En su rápido movimiento por la parte no serrana del norte de Sinaloa, el cora ya se diferencia del huichol.

Por ese tiempo, debido a dificultades de movimiento de algunos grupos taracahitas, se escinden las tarahumaras del resto del grupo. Probablemente a causa del tubar.

Las cahitas se separan, a su vez, de las opatanas. Tal vez las lenguas seris influyen en esta separación y en un movimiento hacia el este por parte de las opatanas. Esto hace que se dividan éstas en ópata y eudeve. Las tarahumaras se desplazan hacia el este y el sur.

Esta situación se mantiene por largo tiempo, durante el cual las lenguas tepimanas comienzan a moverse hacia la costa y hacia el sur. Esto hace que, después de un largo tiempo (y ya estando en el actual estado de Sonora) se dividan los tepehuanos, quienes van moviéndose rápidamente hacia el sur. Este movimiento divide a los tarahumaras de los guarijíos.

En el sur, los coras se ven detenidos por otros grupos que poseen una cultura más compleja, lo que hace que los huicholes se muevan hacia el este, hacia la sierra de Jalisco.

Memoria del XII Simposio de Historia y Antropología, Volumen 1, pp.  
387 - 412, 1987

Universidad de Sonora, Departamento de Historia y Antropología  
© Derechos Reservados. Hermosillo, Sonora, México.

Leopoldo Valiñas Coalla y Mario Cortina Borja  
*Contribución a la reconstrucción histórica a partir de métodos estadísticos en datos léxicos: El caso de las lenguas sonorenses*

XII Simposio de Historia y Antropología de Sonora. 407

Esta situación perduró por algún tiempo. Durante éste, los tepehuanos lograron penetrar en el actual estado de Durango. Los tepehuanos ya han rebasado a los grupos tarahumaras.

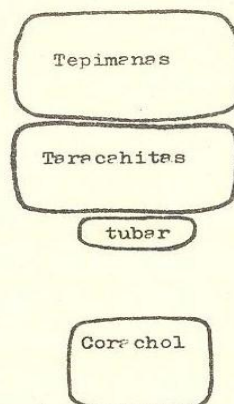
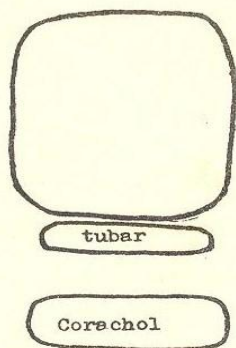
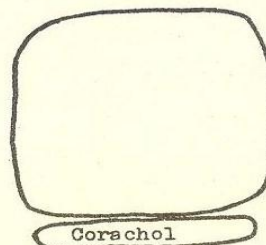
Mientras tanto, en el norte, los pimas se escinden en altos y bajos. Estos últimos ya comienzan un movimiento hacia el sureste.

Sabemos que estamos pasando por alto informaciones etnológicas, históricas y arqueológicas, pero lo que nos interesa, principalmente, es la construcción de hipótesis a partir de datos lingüísticos. En este sentido, la independencia tanto de los datos como del método permiten que este trabajo pueda utilizarse para efectos de contrastación en estudios de otras especialidades. A la larga, el conjunto de todas esas investigaciones servirán para hacer la historia.

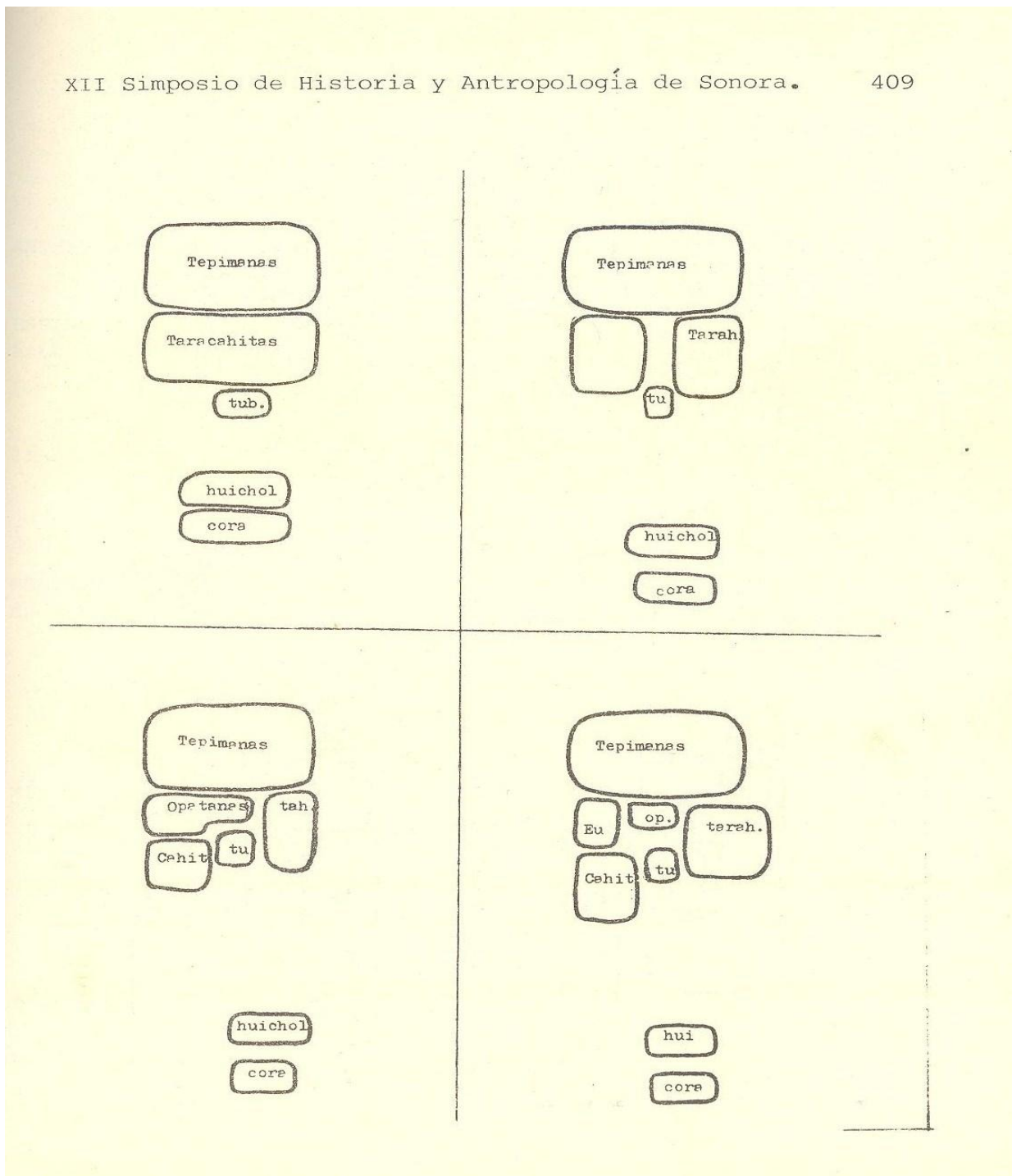
Nuestra hipótesis, pensamos, puede ser el punto de partida para investigaciones arqueológicas y etnohistóricas del área, las que darán, sin duda, mayor claridad al problema. Por otra parte, esperamos que el presente trabajo sea un aporte metodológico que pueda ser retomado por estudios lingüísticos similares.

Leopoldo Valiñas Coalla y Mario Cortina Borja  
*Contribución a la reconstrucción histórica a partir de métodos estadísticos en datos léxicos: El caso de las lenguas sonorenses*

408

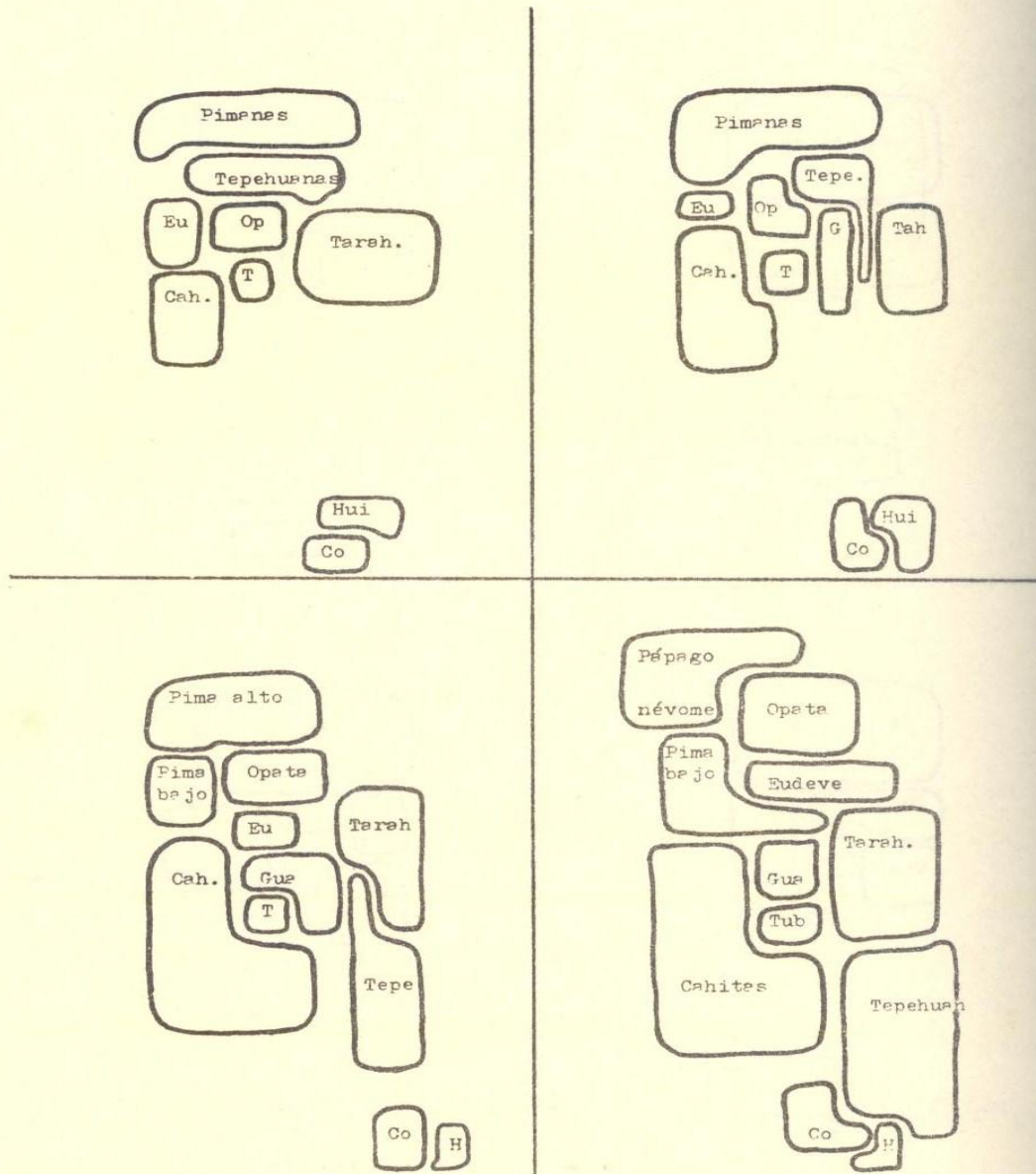


Leopoldo Valiñas Coalla y Mario Cortina Borja  
*Contribución a la reconstrucción histórica a partir de métodos estadísticos en datos léxicos: El caso de las lenguas sonorenses*



Leopoldo Valiñas Coalla y Mario Cortina Borja  
*Contribución a la reconstrucción histórica a partir de métodos estadísticos en datos léxicos: El caso de las lenguas sonorenses*

410





Leopoldo Valiñas Coalla y Mario Cortina Borja  
*Contribución a la reconstrucción histórica a partir de métodos estadísticos en datos léxicos: El caso de las lenguas sonorenses*

XII Simposio de Historia y Antropología de Sonora.

411

BIBLIOGRAFIA

- ANDERBERG, M. Cluster Analysis for applications. EUA, Academic Press 1973.
- ESPINOZA, G. y Arturo López Introducción a los métodos de clasificación jerárquica, México, IIMAS, UNAM, 1980.
- IRAIZOS, Manuel, Implementación de programas para escalamiento multidimensional, Tesis de actuario, UNAM, Facultad de Ciencias, 1984.
- MENDIZABAL, Miguel O. de y W. Jiménez Moreno, Distribución prehispánica de las lenguas indígenas de México, publicación del Instituto Panamericano de Geografía e Historia en colaboración con el Instituto Mexicano de Investigaciones Lingüísticas, 1937.
- MILLER, Wick, A note Extinct Languages of Northwest Mexico of Supposed Uto-Aztec Affiliation, 1981, copia mim.
- , "The Classification of the Uto-Aztec Languages Based on Lexical Evidence", IJAL, 1985.
- REYES, Lilia, Guillermo ESPINOZA y Arturo LOPEZ, ANALISIS/CUMULOS: un programa para clasificación jerárquica, México, IIMAS-UNAM, 1981.

Leopoldo Valiñas Coalla y Mario Cortina Borja  
*Contribución a la reconstrucción histórica a partir de métodos estadísticos en datos léxicos: El caso de las lenguas sonorenses*

412

SAUER, Carl,

The distribution of Aboriginal Tribes and Languages in Northwestern Mexico, Ibero-Americana # 5, EUA, Berkeley, 1934.

SHEPARD, R. N., A. K.  
ROMNEY y S.A. NERLOVE,

Multidimensional scaling: Theory and practice in the behavioral sciences, EUA, Seminar Press, 1972.