

# **LAS HABLAS DE TERUEL DESDE UNA PERSPECTIVA DIALECTOMÉTRICA**



*José Luis Aliaga Jiménez*



# LAS HABLAS DE TERUEL DESDE UNA PERSPECTIVA DIALECTOMÉTRICA\*

José Luis Aliaga Jiménez\*\*

## RESUMEN

Abordamos aquí el estudio dialectométrico de las hablas de Teruel sobre la base de los datos proporcionados por el *Atlas Lingüístico Etnográfico de Aragón, Navarra y Rioja* (1979-1983). El trabajo presenta un interés doble. Por un lado, acomete el análisis de un ámbito lingüístico escasamente analizado hasta la fecha de forma autónoma. Por otro, utiliza con profusión los métodos cuantitativos al análisis dialectal.

**Palabras clave:** dialectometría, dialectología, geografía lingüística, estadística.

## ABSTRACT

*The linguistic varieties of Teruel from a dialectometric perspective.*

The present paper approaches the dialectometric study of the linguistic varieties of Teruel based on the data provided by the *Atlas Lingüístico Etnográfico de Aragón, Navarra y Rioja* (1979-1983). This work presents a double interest. On the one hand, it attacks the analysis of a linguistic environment rarely analyzed so far as an autonomous form, and on the other, it fully makes use of the quantitative methods as focused on the dialectal analysis.

**Key words:** dialectometric, dialectology, linguistic geography, statistics.

\* Este trabajo ha sido posible gracias a la ayuda recibida por el Instituto de Estudios Turolenses en el XVII Concurso de Ayudas a la Investigación de 1999.

\*\* jlaliaga@unizar.es

## INTRODUCCIÓN

La publicación del *Atlas Lingüístico Etnográfico de Aragón, Navarra y Rioja* (ALVAR y cols., 1979-1983; en adelante *ALEANR*) abrió unas expectativas desconocidas hasta entonces para la filología aragonesa, que se han ido colmando en los últimos años con numerosas investigaciones que han tomado la información plasmada en los mapas de esta obra como objeto de estudio de forma exclusiva o complementaria. Cuando se cumplieron diez años desde la aparición del primer volumen, la revisión bibliográfica de los frutos que había proporcionado su análisis arrojó un balance muy satisfactorio (CASTAÑER y ENGUITA, 1989). Desde entonces ha transcurrido otra década y los investigadores han continuado acudiendo a esa fuente de datos que no muestra por el momento síntomas de agotamiento<sup>1</sup>. Así pues, el *ALEANR* se sigue prestando a nuevas exploraciones desde diversos puntos de vista y la que aborda este trabajo resulta de interés, a nuestro juicio, desde dos perspectivas. En primer lugar, por aplicarse a las hablas turolenses, ya que hasta la fecha la provincia de Huesca ha sido objeto de los tres estudios dialectométricos sobre las hablas aragonesas que conocemos (GUTIER, 1983; MORENO, 1991 y GARCÍA MOUTON, 1991). En segundo lugar, porque el método dialectométrico (método cuantitativo aplicado al estudio de las diferencias y/o similitudes inter- o intralingüísticas) apenas ha sido explotado en la investigación sobre las modalidades lingüísticas españolas a diferencia de lo que ocurre para otras lenguas centroeuropeas. Por lo tanto, el trabajo que se desarrolla aquí desea contribuir, modestamente, a cubrir un hueco en la filología aragonesa desde el estudio básicamente cuantitativo de las modalidades lingüísticas de Teruel.

## BREVE ESTADO DE LA CUESTIÓN

### LOS ESTUDIOS DIALECTOLÓGICOS "TRADICIONALES" SOBRE LAS HABLAS TUROLENSES

Como decimos, el objetivo que se ha trazado el presente estudio consiste en dibujar un nuevo panorama de la composición interna de las distintas modalidades lingüísticas que conviven en la provincia de Teruel mediante la aplicación de técnicas cuantitativas de análisis a los datos propor-

1 No obstante, con el paso del tiempo, los materiales acopiados en el atlas tendrán que ser utilizados cada vez con mayor prevención si el investigador pretende extraer conclusiones fiables sobre el presente de la realidad lingüística aragonesa. No olvidemos que, de acuerdo con las ideas lingüísticas de la época, que primaban la búsqueda de pervivencias dialectales, se consideró «informante idóneo» para responder al cuestionario en cada localidad a quien, entre otros requisitos, tenía una edad de entre cincuenta y sesenta años. A ello debe añadirse que la mayoría de los mapas del *ALEANR* (1.394 sobre un total de 1.758) son de tipo léxico (el aspecto más cambiante de la lengua) y que las encuestas se llevaron a cabo entre 1963 y 1968. Cabe referirse también a las consecuencias que se hayan podido derivar en la obtención de una cierta imagen de la realidad lingüística aragonesa (y de otros ámbitos) por la exclusión casi total de la mujer como informante en la encuesta dialectal. Puede consultarse, a este respecto, el número monográfico de la revista *Orbis* (1952) dedicado a esta cuestión y, también, GARCÍA MOUTON (1988 y 1999: 33-49) y LOZANO (1995: 82-110).

cionados por el *ALEANR*. Las hablas de Teruel, como objeto de estudio, han sido incluidas generalmente en trabajos dialectológicos que engloban el conjunto del territorio aragonés o éste junto con otros ámbitos más amplios y han sido escasas las monografías centradas en esta provincia. De hecho, sólo contamos con el trabajo de J.M<sup>a</sup>. Enguita como obra de conjunto dedicada en exclusiva a las hablas vivas turolenses (ENGUITA, 1985). En ella, podemos encontrar un estado de la cuestión sobre la materia examinada en el que se pone de relieve que la zona nororiental de la provincia turolense –de habla catalana– y los territorios limítrofes con ella han centrado el interés de las investigaciones sobre la región<sup>2</sup>. Con posterioridad al trabajo de Enguita tampoco contamos con un número excesivo de estudios de carácter cualitativo (esto es, centrados en la descripción de rasgos lingüísticos) que nos puedan servir de piedra de toque con la que contrastar nuestras conclusiones<sup>3</sup>. Destacaremos, entre ellos, otro artículo del propio Enguita para el conjunto del territorio aragonés (ENGUITA, 1991a) estructurado de forma similar al dedicado a la provincia de Teruel<sup>4</sup>. Asimismo, nos servirá de instrumento contrastivo el apartado dedicado a la penetración de rasgos catalanes en las zonas de habla castellana incluido en un estudio de M.R. Fort sobre la Franja Oriental de Aragón (FORT, 1991). Y, finalmente, acudiremos a la comparación de nuestros mapas con las conclusiones a las que llega A. Llorente sobre la distribución de espacios lingüísticos en el interior de Aragón (LLORENTE, 1991).

Advertiremos, no obstante, que los resultados lexicométricos de carácter estadístico, como los obtenidos en nuestro análisis, no son directamente cotejables con las conclusiones alcanzadas mediante la aplicación de métodos dialectológicos que podemos denominar "tradicionales", es decir, orientados al estudio de características dialectales contrastadas con las de una norma estándar de referencia. A diferencia del sistema de trabajo habitual en dialectología, los datos que conforman el corpus de trabajo en la medición lexicométrica no están sometidos a una criba previa para seleccionar los que puedan resultar rentables desde el punto de vista de una descripción contrastiva. Por lo tanto, las representaciones cartográficas de signo dialectométrico no dan cuenta de áreas o espacios dialectales del mismo modo que las tradicionales porque se derivan del cómputo de diferencias y/o semejanzas lingüísticas, tipológicamente dialectales o no. Sin embargo, la aplica-

- 2 Por lo demás, el estudio de ENGUITA (1985), a partir de una selección de mapas del *ALEANR*, describe con precisión y por niveles de análisis (fonético, morfosintáctico y léxico) las huellas del antiguo romance aragonés en las hablas turolenses. Las conclusiones obtenidas quedan sintetizadas en un mapa-resumen (en la pág. 219) sobre el que volveremos una vez expuestos nuestros resultados.
- 3 Nos referimos a monografías para cuya confección se haya manejado un amplio conjunto de datos, entre los que se encuentren los de carácter léxico, y que se hayan orientado al establecimiento de áreas lingüísticas. A este respecto pueden consultarse los estados de la cuestión sobre los diferentes ámbitos de la lingüística aragonesa contenidos en ENGUITA (1999).
- 4 Viene a representar una síntesis de sus estudios parciales circunscritos a las provincias aragonesas por separado. A la de Teruel, ya mencionado (ENGUITA, 1985), a la de Huesca (ENGUITA, 1988) y a la de Zaragoza (ENGUITA, 1991b).

ción de este método también está encaminada a la detección de núcleos y zonas dialectales dado que la concentración de peculiaridades lingüísticas espaciales quedará reflejada con toda seguridad en el cálculo correspondiente. Nos hallamos ante un indudable ejemplo de aplicación de perspectivas lingüísticas complementarias sobre conjuntos de datos muy similares. En realidad, las mediciones estadísticas perderían a nuestro juicio buena parte de su interés si se prescindiera por completo del tipo concreto de fenómenos lingüísticos que constituye el trasfondo de los índices y porcentajes matemáticos. En definitiva, intentaremos, con las oportunas precauciones, sacar provecho de la confrontación entre los resultados estadísticos y los cualitativos contenidos en los trabajos comentados más arriba.

## LOS ESTUDIOS DIALECTOMÉTRICOS SOBRE LAS MODALIDADES LINGÜÍSTICAS ARAGONESAS

Tal como se ha dicho más arriba –y como se puede apreciar en la Bibliografía, donde tienen cabida algunos de los trabajos más representativos– la mayoría de los estudios dialectométricos ha sido desarrollada fuera de España. Pero la estadística ya se había aplicado de formas diversas a la medición de distancias lingüísticas antes de que se constituyera la disciplina denominada "dialectometría" (vid. GARCÍA MOUTON, 1991: 312). Y en lo que se refiere a nuestro país, en los años cincuenta M. Alvar hizo uso de procedimientos cuantitativos en un estudio sobre las interferencias lingüísticas en la frontera catalano-aragonesa a partir del *Atlas Lingüístico de Cataluña* (ALVAR, 1955). Después sólo encontramos tres estudios propiamente dialectométricos que tomen como objeto de estudio las hablas aragonesas. De ellos, dos versan sobre la provincia de Huesca (MORENO, 1991 y GARCÍA MOUTON, 1991) y otro, anterior, sobre la región cántabro-pirenaica donde queda incluida también la provincia más septentrional de Aragón (GUITER, 1983)<sup>5</sup>.

Por la proximidad de nuestro trabajo con ellos, nos interesan más concretamente los citados artículos de F. Moreno y de P. García Mouton, que incluyen una útil síntesis de las principales tendencias y principios dialectométricos<sup>6</sup> y toman como corpus de trabajo respectivamente la morfología y el léxico del ALEANR. Ambos siguen la misma metodología de cómputo consistente en medir

5 Además contamos con otro (SÉGUY, 1971) donde se analizan materiales españoles (en concreto del *Atlas Lingüístico de Cataluña* y del *Atlas Lingüístico Etnográfico de Andalucía*).

6 Allí se nos informa de que la dialectometría es "una variante más en la síntesis de mapas" (GARCÍA MOUTON, 1991: 312) que toma por objeto los datos geolingüísticos para elaborarlos estadísticamente. Esta metodología comenzó a desarrollarse en los años setenta y se afianzó una década después con un conjunto de trabajos elaborados, sobre todo, en Francia y en Bélgica y aplicados a distintos atlas lingüísticos del ámbito románico. Se puede decir que existen básicamente tres métodos, que no se excluyen mutuamente, de medir numéricamente las distancias dialectales desarrollados, respectivamente, por J. Séguy, E. Guiter y H. Goebl (GARCÍA MOUTON, 1991: 313-314; MORENO, 1991: 298-299).

diferencias lingüísticas comparando cada punto del atlas con todos los demás<sup>7</sup>. Efectuada la etapa de recuento se aplica una fórmula matemática para obtener un *índice de diferenciación dialectométrica* para cada localidad con un valor entre 0 y 1. Cuanto más cercano a 1 se encuentre dicho índice mayor será la distancia lingüística de esa población respecto del resto de poblaciones estudiadas. Aparte del *índice de diferenciación* los datos permiten alcanzar otras conclusiones estadísticas, como la media de rasgos diferenciadores o la desviación típica, esclarecedoras de ciertos aspectos de la diferenciación lingüística de las localidades examinadas (MORENO, 1991: 300-302). De los valores obtenidos con el manejo de los ochenta mapas de morfología verbal utilizados por Moreno se desprende una configuración dialectal de las hablas oscenses que se ajusta con sorprendente precisión a las alcanzada por medios cualitativos (1991: 302-305). No ocurre lo mismo, sin embargo, con los resultados derivados del cómputo de diferencias léxicas en cien mapas que acomete GARCÍA MOUTON (1991: 321-324). Por ejemplo, en este último caso, la representación gráfica deja entrever en la Franja Oriental de Aragón una diferenciación interna horizontal –además de la ya conocida orientada de norte a sur– que, en opinión de la citada investigadora, debe atribuirse entre otros factores al modo de difusión del léxico, divergente en buena medida del de la fonética o del de la morfología (1991: 324-325).

## METODOLOGÍA

Señala H. GOEBL (1987: 63) que la dialectometría es un procedimiento de clasificación numérica que comprende varias etapas en las que deben establecerse, sucesivamente, los datos que se van a clasificar; el modo de medición; el índice de similitud (o de diferenciación); los procedimientos estadísticos para explotar la tabla de similitud y el método de representar cartográficamente los resultados. Nuestro estudio se ajusta, en lo esencial, a la línea de trabajo y a los instrumentos de análisis empleados por Goebl. No obstante, dada la escasez de estudios dialectométricos publicados en España, nos inspiramos, en algunos aspectos, en las decisiones metodológicas de F. Moreno y P. García Mouton para facilitar la comparación de los resultados. Coincidimos, más concretamente, con el planteamiento de esta última quien, además, centra su estudio en el nivel léxico, al igual que en el presente estudio. A partir de ahí, en consonancia con el programa trazado por H. GOEBL (1981, 1983 y 1987), ampliamos sustancialmente la explotación estadística de los datos y la presentación gráfica de los resultados respecto de las llevadas a cabo para la provincia de Huesca.

## SELECCIÓN DEL CORPUS Y ETAPA DE RECUENTO

En la selección del corpus de trabajo no sólo nos hemos ceñido a las consideraciones de GARCÍA MOUTON (1991: 315) –quien coincide con GUITER (1973) en la cifra de cien mapas como la cantidad

<sup>7</sup> En este aspecto siguen el procedimiento de H. GOEBL (1976, 1981, 1983 y 1987), que difiere de los empleados por Séguy y Guter, aunque Goebl orienta sus recuentos al cálculo de la similitud lingüística y no de la diferencia.

mínima que asegura un rendimiento estadístico fiable- sino que nos ha parecido oportuno, además, analizar exactamente los mismos mapas léxicos del *ALEANR* que utiliza la citada autora porque resultan totalmente pertinentes los criterios de elección aleatoria manejados (1991: 315-316)<sup>8</sup>. Al mismo tiempo, ello permitirá perfilar un estudio dialectométrico sustantivamente equivalente para cada una de las tres provincias aragonesas<sup>9</sup>. Así pues, el corpus de datos originales que componen nuestro trabajo responde a las siguientes coordenadas:

- 1) Datos dialectales brutos: 100 mapas del *ALEANR*.
- 2) Objetos (puntos del atlas): 36 puntos de la provincia de Teruel.
- 3) Criterio discriminatorio para la identificación de *tipos léxicos diferenciados* o "taxats": diferencias léxicas.
- 4) Número total de *tipos léxicos diferenciados* ("taxats"): 685.
- 5) Número de "taxats" por mapa: 6,85.

Una vez transcrita la información de cada mapa en la *matriz de datos* (cf. GOEBL, 1981: 352) la etapa de recuento, como se ha señalado más arriba, se basa en la comparación de cada punto del atlas con todos los demás para, en la posterior etapa de cálculo, lograr índices estadísticos para cada localidad comparables con los de cualquiera de las otras. Así pues, en nuestro caso, Te 100 se compara con Te 101, Te 102, etc., hasta llegar a Te 601. A continuación se coteja Te 101 con Te 102, Te 103, etc. El último escalón viene representado por la confrontación de Te 600 con Te 601. Por lo tanto, el modo de proceder en el cómputo es sencillo pero extremadamente laborioso porque se realiza de forma manual de principio a fin. El material necesario para proceder a la etapa de cálculo exige, en primer lugar, la transcripción de 3.600 formas léxicas (36 puntos de Teruel en el *ALEANR* de 100 mapas) que quedan registradas en la *matriz de datos* (véase un detalle en la tabla 1). Posteriormente se procede a contabilizar las 630 distancias de diferenciación en 100 mapas y, en nuestro caso, a calcular otros 630 *índices relativos de identidad*, también sobre 100 mapas (vid. tabla 2 y tabla 4, respectivamente)<sup>10</sup>.

- 
- 8 Los cien mapas que integran nuestro corpus llevan en el *ALEANR* la siguiente numeración: 16, 20, 24, 58, 81, 92, 128, 149, 156, 161, 166, 174, 222, 236, 239, 252, 254, 270, 286, 323, 335, 336, 361, 418, 434, 439, 468, 483, 495, 515, 523, 541, 546, 560, 567, 577, 586, 592, 602, 614, 621, 627, 645, 657, 661, 689, 702, 723, 726, 739, 741, 765, 797, 814, 854, 861, 870, 872, 936, 937, 942, 944, 946, 960, 968, 981, 982, 992, 1008, 1035, 1047, 1096, 1116, 1118, 1124, 1129, 1133, 1179, 1184, 1211, 1228, 1230, 1238, 1240, 1246, 1260, 1261, 1268, 1270, 1272, 1286, 1291, 1310, 1331, 1342, 1367, 1385, 1388, 1393, 1396. Puede consultarse el nombre de cada mapa en GARCÍA MOUTON (1991: 317).
- 9 Nos hallamos trabajando en la actualidad, con los mismos criterios y datos, sobre las modalidades lingüísticas de la provincia de Zaragoza. Conviene advertir, en todo caso, que la suma de los estudios dedicados a cada provincia no proporcionará un panorama dialectométrico de conjunto sobre Aragón. Para ello, lógicamente, debería acometerse una exploración que tomara como objeto de análisis todos los puntos aragoneses del *ALEANR* al mismo tiempo. No ocurre lo mismo, sin embargo, con los estudios de carácter cualitativo (cf. ENGUITA, 1991a).
- 10 Desde un punto de vista técnico se calcula, en realidad, el doble de índices de similitud y de distancias de diferenciación (aquí supondrían un total de 1.260) ya que se comparan las similitudes (o diferencias) tanto entre,

TABLA 1  
Matriz de datos (detalle)

MAPA/LOCALIDAD	335	336	361	418	434	439	468	483	495	515
100 NOGUERAS —— escabezar	esgarrar	albericoque	margarita	tocinico	sanguijuela	cuchareta	bacalau	esperadera	serrano	
101 FERRERUELA —— escabezar	esgarrar	alberge/ abaricoque	XXX	ciempés	sangrijuela	cuchareta	abadejo/ bacalao	espera	serrano	
102 MUNIESA —— escabezar	rasgar	albericoque	grillo	gusano de ciengarras	sangrijuela	ranueco/ cuchareta	abadejo/ bacalao	esperadero	serrano	
103 BELLO —— escabezar	rasgar	albericoque	rezadora	milpiés	sanguijuela	renacuajo	abadejo/ bacalao	esperadera	serrano	
104 BARRACHINA —— escabezar	esgarrar	albericoque	señorita	ciempés	sangrijuela	cuchareta	abadejo/ bacalao	esperadero	serrano	
200 HÍJAR —— rebajar	esgarrar	alberge	maría	ciengarras	sangrijuela	gucharón	abadejo/ bacalau	XXX	montañés/ serrano	

TABLA 2  
Matriz de diferenciación (diferencias léxicas)

	100	101	102	103	104	200	201	202	203	204	205	206
100												
101	31											
102	33	32										
103	43	32	33									
104	28	24	27	27								
200	47	44	48	44	41							
201	50	50	52	53	48	45						
202	86	86	85	87	86	85	85					
203	46	42	41	42	40	39	49	82				
204	68	62	64	65	63	65	57	69	58			
205	79	76	77	78	76	77	78	51	74	58		
206	44	40	39	42	32	41	47	81	38	59	72	
207	81	79	78	78	76	78	78	51	71	60	43	73
300	34	39	40	44	32	42	46	82	47	61	75	32
301	50	40	45	37	38	55	55	84	49	63	74	47
302	41	39	44	39	29	49	53	83	45	66	78	42
303	45	40	46	43	37	54	50	89	50	63	81	53
304	44	37	41	43	37	55	57	85	50	63	76	49
305	35	35	37	41	30	53	49	82	42	64	77	41
306	45	40	38	41	36	52	54	86	49	64	76	47
307	54	52	54	59	49	62	59	81	49	64	68	52
308	43	38	38	42	36	55	56	82	50	66	77	47
400	50	45	46	51	44	56	45	83	46	59	76	45
401	58	52	52	57	48	53	59	80	54	69	73	42
402	38	32	33	36	28	41	47	84	37	57	70	28
403	49	47	52	54	45	56	55	83	50	64	76	50
404	44	43	45	47	38	53	52	79	49	60	65	43
405	52	53	50	55	43	56	59	76	51	64	69	46
406	48	48	44	47	40	53	56	84	44	61	74	48
500	55	46	52	53	45	58	61	86	57	72	79	57
501	46	47	45	52	44	54	60	85	55	63	75	46
502	52	42	45	45	43	52	49	84	48	62	78	43
503	53	53	48	52	47	60	61	82	55	62	73	50
504	50	46	50	55	48	58	57	85	54	62	74	54
600	53	53	53	53	48	58	61	78	51	60	69	48
601	53	52	53	57	50	59	64	78	54	67	73	53

	207	300	301	302	303	304	305	306	307	308	400	401
100												
101												
102												
103												
104												
200												
201												
202												
203												
204												
205												
206												
207												
300	72											
301	77	45										
302	74	37	39									
303	83	46	43	35								
304	76	43	31	40	41							
305	74	39	38	32	42	33						
306	79	40	38	41	38	41	41					
307	73	53	46	45	56	51	41	44				
308	80	46	43	44	43	35	34	32	45			
400	78	50	50	48	50	46	44	50	51	45		
401	69	51	61	53	59	61	52	59	55	60	50	
402	74	34	41	37	45	41	34	35	44	36	38	49
403	74	49	55	41	49	51	41	45	50	47	46	50
404	69	43	50	45	49	45	38	42	47	40	44	50
405	71	49	51	43	56	54	47	51	43	50	51	45
406	79	46	41	40	54	47	32	43	39	42	46	55
500	81	48	53	46	53	51	50	47	53	53	56	56
501	74	45	56	47	54	53	47	49	47	50	55	53
502	75	44	48	44	46	41	38	42	54	42	52	57
503	72	49	54	45	56	50	43	52	45	52	55	54
504	50	46	50	55	48	58	57	85	54	62	74	54
600	69	53	53	49	61	57	50	47	46	54	53	51
601	73	53	56	42	57	56	45	51	47	55	60	54

	402	403	404	405	406	500	501	502	503	504	600	601
100												
101												
102												
103												
104												
200												
201												
202												
203												
204												
205												
206												
207												
300												
301												
302												
303												
304												
305												
306												
307												
308												
400												
401												
402												
403	38											
404	29	33										
405	41	45	34									
406	33	43	39	43								
500	52	54	52	53	53							
501	43	50	44	49	48	53						
502	40	50	45	51	49	49	49					
503	46	51	45	46	44	51	37	46				
504	46	47	42	52	42	52	41	52	44			
600	36	41	37	42	32	61	51	51	48	53		
601	49	46	49	45	39	58	40	56	46	46	43	

No existe, que sepamos, un programa informático capaz de resolver los problemas metodológicos que surgen en el momento de decidir qué se contabiliza exactamente (con independencia de que se tomen en cuenta las coincidencias o las diferencias). Interviene necesariamente en esta fase el criterio subjetivo (que no arbitrario) del lingüista, que debe identificar los "taxats", los tipos léxicos diferenciados (o morfosintácticos, en su caso) entre el polimorfismo que ofrecen habitualmente los mapas del atlas. Claro que un hipotético programa informático no contendría sino las instrucciones de las que le dotara un lingüista, aunque las ejecutara con mayor precisión. Resulta significativo comprobar cómo la bibliografía especializada apenas ofrece orientaciones prácticas en este punto crucial del análisis dialectométrico. Contamos con un escueto ejemplo proporcionado por GOEBL (1987: 64, n. 5) mientras que García Mouton, por su parte, facilita unas indicaciones útiles pero inevitablemente genéricas. A esta autora no le resulta totalmente convincente ni el criterio de agrupar todas las formas léxicas que comparten un mismo étimo ni el de establecer un corte allí donde el lingüista cree ver dificultada la comprensión de la palabra. Así, se decanta por otra solución: "hemos contado diferencias allí donde la fonética distorsionaba la identificación o un sufijo de diminutivo parecía lexicalizado" (1991: 318). Todavía podría apuntarse alguna otra pauta de actuación que no hemos visto señalada por ningún autor. Consistiría en articular un proceso de lematización que emulara el camino seguido para establecer la macroestructura de un diccionario a partir de un corpus textual sintetizando en una forma-tipo las unidades lingüísticas flexionadas y conjugadas del mismo paradigma. Los lematizadores automáticos empleados en lexicografía se desenvuelven ya con asombrosa precisión pero cabe preguntarse si admitirían, sin dejar de ser rentables, el grado de minuciosidad que exige el proceso dialectométrico en el que se manipulan, en buena medida, unidades léxicas alejadas del estándar. En cualquier caso, por muy acertada que pueda estarse cualquiera de las medidas señaladas, hasta la fecha ninguna exime al dialectólogo de adoptar soluciones dictadas por su dominio de la materia en los casos problemáticos que, según nuestra experiencia, asoman con mayor frecuencia de lo señalado por GARCÍA MOUTON (1991: 318)<sup>11</sup>.

Tampoco existe unanimidad a la hora de tratar los casos de ausencia de respuesta en un punto del mapa, como Te 101 en el mapa 418 «Santateresa» o Te 200 en el 495 «Puesto» (vid. tabla 1). En estas situaciones el cálculo estadístico se verá afectado tanto si se opta por no contar una diferencia, posibilidad por la que se decanta MORENO (1991: 301), como si se adopta la solución contraria (GARCÍA MOUTON, 1991: 318-320) por la que nosotros nos hemos inclinado. Finalmente, no contamos diferencia en los casos en los que un punto de encuesta presenta varias respuestas y sólo una

por ejemplo, Te 100 y Te 101, como entre Te 101 y Te 100. Dado que la distancia de ambos recorridos es idéntica sólo es necesario anotar una de ellas. De ahí la forma asimétrica de la tabla 2 y de la tabla 4.

11 Veamos un sencillo ejemplo. En el mapa 236 «Acemite» del ALEANR encontramos en la provincia de Teruel la serie de respuestas *cabisuela*, *cabesuela*, *cabezuela*. Parece claro que no debe contabilizarse ninguna diferencia entre ellas. Pero en el mismo mapa hallamos también la serie *moyuelo*, *remoyuelo*, *remoelo*, *remoñuelo*, *remuñuelo*, *rebuñuelo*, *remogüelo*. Reparando en este último caso, ¿mostrarían unanimidad todos los dialectólogos en que la fonética distorsiona la identificación de la palabra?

de ellas coincide con la del punto con el que se compara (GARCÍA MOUTON, 1991: 320). Así pues, no se anotaría una diferencia en la comparación de los puntos Te 100 y Te 200 con Te 101 en el mapa 361 «Albaricoque».

## ETAPA DE CÁLCULO

Las operaciones estadísticas a las que dan lugar el cómputo de diferencias y el de similitudes, respectivamente, están claramente diferenciadas y orientadas a la consecución de resultados complementarios. Por ello tratamos por separado cada una de las dos perspectivas. Nos limitamos en este tercer apartado a reseñar brevemente los aspectos técnicos reservando para el siguiente la interpretación de los datos más relevantes una vez comentado, asimismo, el método de representar cartográficamente los datos.

### Índice de diferenciación y desviación típica

La tabla 3 presenta para cada localidad las operaciones efectuadas a partir de la *matriz de diferenciación* (tabla 2). La primera columna indica el número de la localidad en el ALEANR; la segunda, la suma total de diferencias entre esa localidad y todas las demás; la tercera, la media de esas diferencias (el total de diferencias de cada punto dividido por 35, es decir, por el número de localidades menos uno); la cuarta, la *desviación típica* y la quinta, el *índice de diferenciación léxica*. La *desviación típica* representa el alejamiento de una serie de números de su valor medio y se calcula promediando la suma de los cuadrados de todas las desviaciones individuales respecto de la media y tomando después la raíz cuadrada. La *desviación típica* nos informa, en un estudio dialectométrico, de si las diferencias de una localidad respecto de las demás es constante o si presenta variaciones acusadas según el punto con la que se compare en cada caso. El *índice de diferenciación* es igual a la suma total de diferencias de una localidad respecto de las demás dividida por el número de localidades estudiadas menos uno multiplicado por el número de mapas (cf. MORENO, 1991: 301). Cuando se analizan cien mapas, como en este estudio, el índice de diferenciación se obtiene de forma más simple dividiendo por cien la media de diferencias de cada localidad (vid. tabla 3, columnas tercera y quinta). Como ya adelantamos más arriba, la modalidad lingüística de una localidad se halla más alejada de las demás cuanto más cercano a 1 se encuentre el valor del *índice de diferenciación* (vid. fig. 1).

### Índice relativo de identidad

Señala GARCÍA MOUTON (1991: 318) que la *matriz de diferenciación* puede leerse como una *matriz de similitud* dándole la vuelta, lo cual resulta particularmente sencillo cuando se manejan cien mapas porque no es necesario operar para el cálculo del porcentaje. Por ejemplo, si como refleja la tabla 2, las diferencias entre Te 202 y Te 203 suman 82 sobre 100 (82%) la similitud entre esos dos puntos sería del 18%. Sin embargo, esto es así siempre que no existan casillas vacías (puntos

TABLA 3  
Elaboración estadística de la matriz de diferenciación

	SUMATORIO	MEDIA	DESVIACIÓN TÍPICA	ÍNDICE DE DIFERENCIACIÓN	LOCALIDAD
100	1728	49,37	12,75	0,493	Nogueras
101	1617	46,2	13,06	0,462	Ferreruela
102	1660	47,42	12,56	0,474	Muniesa
103	1727	49,34	12,73	0,493	Bello
104	1503	42,94	13,94	0,429	Barrachina
200	1898	54,22	10,21	0,542	Híjar
201	1957	55,91	9,05	0,559	Alcañiz
202	2835	81	8,22	0,81	Calaceite
203	1758	50,22	9,51	0,502	Alloza
204	2204	62,97	3,45	0,629	La Codoñera
205	2545	72,71	7,75	0,727	Valderrobres
206	1671	47,74	10,75	0,477	Esteruel
207	2568	73,37	7,87	0,733	Peñarroya de Tastavins
300	1656	47,31	10,91	0,473	Montalbán
301	1757	50,2	11,28	0,502	Torrijo del Campo
302	1627	46,48	11,76	0,464	Visiedo
303	1827	52,2	12,1	0,522	Villar del Salz
304	1726	49,31	11,91	0,493	Santa Eulalia del Campo
305	1565	44,71	12,28	0,447	Alfambra
306	1684	48,11	11,93	0,481	Noguera
307	1823	52,08	8,91	0,52	Cedrillas
308	1705	48,71	12,09	0,487	Teruel
400	1818	51,94	9,54	0,519	Mas de las Matas
401	1959	55,97	7,52	0,559	Bordón
402	1492	42,62	12,26	0,426	Aliaga
403	1777	50,77	10	0,507	Tronchón
404	1629	46,54	9,73	0,465	Fortanete
405	1786	51,02	8,58	0,51	La Iglesuela del Cid
406	1676	47,88	11,54	0,478	Alcalá de la Selva
500	1956	55,88	9,42	0,558	Masegoso
501	1807	51,62	9,78	0,516	La Puebla de Valverde
502	1764	50,4	10,18	0,504	Riodeva
503	1827	52,2	8,99	0,522	Manzanera
504	1838	52,51	9,54	0,525	Arcos de las Salinas
600	1823	52,08	9,16	0,52	Puertomingalvo
601	1879	53,68	9,01	0,536	Olba
Total 66072		Total 1887,63			
Media 1835	Media 52,43	Media 10,28		Media 0,524	

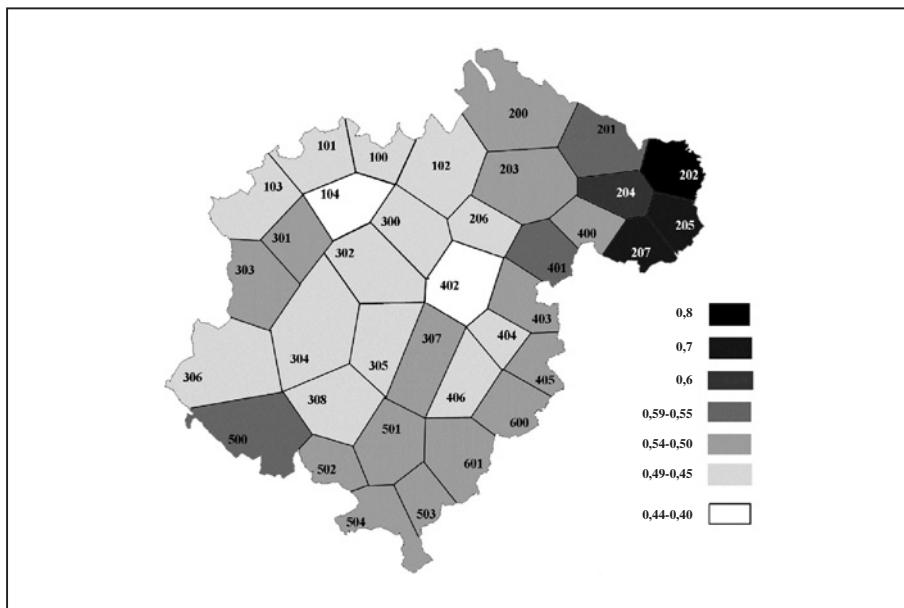


Fig. 1. Mapa poligonal de la distribución del índice de diferenciación léxica.

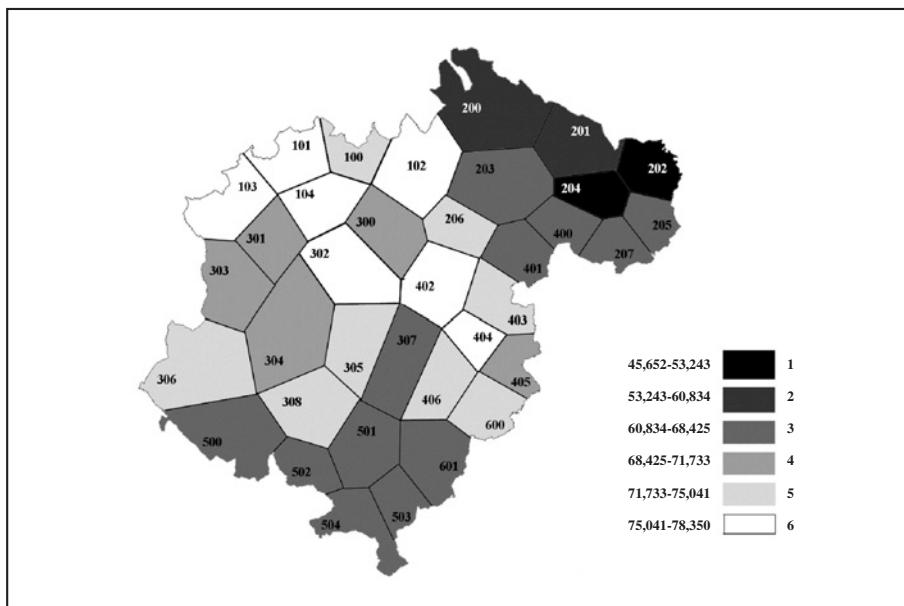


Fig. 2. Mapa poligonal de la distribución de los valores máximos del IRI (MINMWMAX).

sin respuesta) en las localidades que se confrontan, al menos en la metodología estadística utilizada por GOEBL (1981: 357-358 y 1987: 68-69) al que seguimos en este aspecto. Así, para cada par de puntos comparados, el *índice relativo de identidad (IRI)* propuesto por Goebl se obtiene dividiendo el número de co-identidades por el de co-ocurrencias y multiplicando por cien el resultado. El número de co-ocurrencias equivale al número de mapas del corpus descontados los casos en que no existe respuesta en una localidad o en ninguna de las dos comparadas. Además, de producirse esta última circunstancia no se computa ni una semejanza ni una diferencia de modo que, en la práctica, se utiliza un mapa menos. Siguiendo con la muestra anterior, el *IRI* entre Te 202 y Te 203 es el 18,681% (vid. tabla 4), ligeramente superior, por lo tanto, al 18% que resultaría de una estimación especular de la *Matriz de diferenciación*. Este porcentaje responde al hecho de que entre una localidad y otra suman 91 co-ocurrencias (ocho casillas vacías y una novena, la correspondiente al mapa 1228 «Maestro», que no tiene respuesta en ninguno de los dos puntos). Por lo tanto, el número de diferencias es de 82 y el número de semejanzas es de 17 (99 en total). El *IRI* (18,681%) resulta, así, de dividir 17 por 91 y multiplicar por 100 la cifra resultante<sup>12</sup>.

### Otras operaciones estadísticas significativas

Conviene precisar, antes de continuar, que el conjunto de valores de similitud respecto de un punto de referencia constituye una *distribución de similitud*. Así, existen tantas distribuciones de similitud como puntos del atlas estudiados y cada distribución está compuesta por tantos valores como localidades menos uno (precisamente, el punto de referencia). En nuestro estudio contamos con 36 distribuciones de similitud de 35 valores cada una (vid. tabla 4). Cada distribución de similitud puede dar lugar a la elaboración de, al menos, un mapa sinóptico, pero a partir de ella son posibles tantas presentaciones cartográficas como algoritmos de visualización sean utilizados (vid. el apartado siguiente). En la práctica, se suelen reproducir gráficamente en mapas las distribuciones de similitud cuyo punto de referencia pertenece a alguna localidad en la que se presentan indicios estadísticos que pueden resultar reveladores para la comprensión de la configuración lingüística de

12 El *IRI* no es la única posibilidad para medir la similitud entre dos puntos del atlas. H. Goebl se ha servido en otros estudios de otro tipo de medición, denominada *índice ponderado de identidad (IPI)*. Mientras que en el cálculo del *IRI* se les adjudica el valor 1 a todas las co-identidades entre dos puntos del atlas, con el *IPI* se persigue otorgar un mayor peso a las co-identidades cuanto menor sea la frecuencia absoluta en el resto de los puntos del atlas de los "taxats" que generan esas co-identidades. Ello responde al principio estadístico de que la coincidencia de aspectos extraños o singulares cuenta con menos probabilidades de suceder que la de aspectos usuales o comunes y, por ello, al primer tipo de coexistencia se le debe otorgar un mayor peso cuantitativo (cf. GOEBL, 1987). Así, si nos ciñéramos a la matriz de datos presentada en la tabla 1, la co-identidad en el mapa 336 «Desgajar» entre los puntos Te 102 y Te 103 alcanzaría un mayor valor numérico, en un índice ponderado, que la co-identidad de esas mismas localidades en el mapa 515 «Serrano», porque la respuesta *serrano* (515) se repite en todos los demás puntos de encuesta mientras que *rasgar* (336) sólo aparece en Te 102 y Te 103.

TABLA 4

## Matriz de similitud-Índice relativo de identidad (similitudes léxicas)

	100	101	102	103	104	200	201	202	203	204	205	206
100												
101	71.134											
102	69.072	69.387										
103	58.762	69.387	68.367									
104	75	78.35	75.257	75.257								
200	54.639	57.152	53.061	57.142	60.824							
201	51.612	52.688	52.173	51.086	57.142	59.782						
202	15.217	15.053	16.129	13.978	15.217	16.129	15.909					
203	57.446	61.052	62.105	59.375	63.829	64.21	57.303	18.681				
204	34.042	40	37.894	36.842	39.361	36.842	45.054	32.967	44.086			
205	21.276	25.263	24.21	23.157	25.531	24.21	23.333	52.747	28.26	45.652		
206	58.333	61.855	62.886	59.793	70.833	60.824	58.241	19.354	65.957	43.617	29.787	
207	20.652	22.58	23.655	23.655	26.086	23.655	23.863	52.222	29.347	44.444	61.538	29.347
300	67.346	61.616	60.606	56.565	69.387	58.585	58.064	19.148	55.208	40.625	26.041	69.387
301	52.083	60.204	56.701	64.948	64.583	46.391	46.236	16.129	54.255	37.894	27.659	55.208
302	62.765	62.5	58.947	64.21	75.531	53.684	51.111	18.888	58.064	36.956	23.913	61.702
303	61.111	64.13	59.34	62.637	70	50.549	55.172	11.494	53.846	38.888	21.59	52.222
304	57.731	64.285	60.204	58.163	64.948	45.918	46.739	16.129	52.631	38.947	25.263	52.577
305	67.01	66.326	64.285	60.204	72.164	47.959	53.763	18.085	61.052	36.458	24.21	60.824
306	57.894	62.105	65.957	62.765	68.817	51.063	50.561	14.444	56.043	39.56	26.373	55.319
307	48.936	48.958	47.916	43.157	55.255	40	42.857	21.111	55.434	39.13	34.782	51.063
308	61.29	65.957	65.957	61.702	68.817	47.872	50	17.582	54.945	37.362	25.274	56.989
400	54.945	59.782	58.695	53.26	60.869	47.826	60.227	17.977	58.888	44.444	26.966	60.439
401	46.153	50.537	52.173	46.739	55.434	51.086	44.318	21.59	50	34.831	30.337	61.956
402	63.265	68.686	67.676	64.646	73.469	59.595	56.989	17.021	65.625	44.791	31.25	73.469
403	57.954	59.55	54.545	50.561	63.218	50	50.588	16.279	55.172	39.08	28.235	53.932
404	59.139	61.29	59.139	55.319	67.391	50.537	54.022	23.863	54.945	44.444	38.888	62.637
405	51.063	49.473	52.631	47.368	60.638	46.315	46.067	25.274	51.612	39.13	33.956	58.064
406	54.736	54.166	58.333	53.608	63.157	48.453	47.252	16.304	58.51	40.425	27.956	54.736
500	51.724	59.55	54.545	51.685	61.363	48.275	44.047	15.294	45.454	31.395	24.705	49.425
501	60	58.241	60.439	52.747	62.222	50.549	45.348	16.091	51.136	42.045	26.666	60
502	51.612	61.702	58.51	56.842	61.29	51.063	56.179	16.666	55.434	40.217	24.175	61.29
503	50	49.473	54.736	50.526	56.382	42.105	43.82	18.681	48.913	41.304	27.956	53.191
504	53.191	56.842	52.631	47.368	55.319	44.21	46.666	16.666	50	41.304	28.26	48.936
600	50.537	50	50	48.421	55.913	44.68	42.696	23.333	52.173	43.956	34.065	52.631
601	52.808	53.333	51.648	47.777	56.179	45.054	41.176	21.59	51.136	37.931	29.545	49.45

	207	300	301	302	303	304	305	306	307	308	400	401
100												
101												
102												
103												
104												
200												
201												
202												
203												
204												
205												
206												
207												
300	29,787											
301	25	56,122										
302	27,472	65,625	63,157									
303	19,54	58,695	59,782	65,591								
304	25,806	57,575	71,134	63,157	64,835							
305	27,956	61,616	62,244	71,578	61,956	68,367						
306	23,333	63,157	64,893	63,043	67,415	62,765	62,765					
307	28,571	48,958	55,789	58,064	48,314	51,578	62,105	58,064				
308	22,471	56,842	61,29	61,538	65,517	67,368	70,212	73,626	58,695			
400	25,287	53,763	53,26	58,426	55,172	58,695	59,139	56,818	53,333	60,674		
401	29,67	52,688	41,304	51,111	46,511	42,391	52,173	44,943	47,252	45,454	58,139	
402	27,659	66	60,204	65,625	59,782	59,595	66,666	68,421	58,333	67,368	66,666	54,838
403	29,761	57,303	50	67,441	58,333	53,932	67,045	61,627	58,139	59,302	61,176	57,142
404	35,227	60,638	54,347	59,34	57,471	56,842	66,666	65,168	58,888	65,555	62,5	57,471
405	29,347	53,125	52,127	60,215	48,314	48,421	55,789	53,846	61,956	54,945	55,056	58,241
406	23,076	55,67	60,416	62,765	48,351	55,208	69,072	61,956	65,591	63,043	58,241	50
500	20	58,426	52,272	59,77	52,941	55,681	56,818	61,176	49,438	55,952	49,411	48,235
501	27,272	59,782	48,888	58,426	52,941	51,648	58,241	56,666	55,681	52,941	53,488	
502	28,089	58,947	54,255	61,538	60,227	62,765	64,21	62,637	50,549	64,444	52,808	48,863
503	29,347	53,125	48,936	59,782	50	52,631	60	52,747	59,782	52,747	50,561	51,685
504	25,274	55,208	52,127	52,173	45,454	56,25	58,947	56,043	58,064	56,521	51,685	45,555
600	33,333	49,473	50,537	56,043	44,827	45,744	53,191	55,434	57,608	51,111	53,409	52,222
601	27,272	51,648	49,438	62,921	46,511	48,888	61,111	52,808	59,09	52,325	47,619	49,425

	402	403	404	405	406	500	501	502	503	504	600	601
100												
101												
102												
103												
104												
200												
201												
202												
203												
204												
205												
206												
207												
300												
301												
302												
303												
304												
305												
306												
307												
308												
400												
401												
402												
403	69,662											
404	75,531	74,418										
405	61,458	62,79	71,428									
406	69,072	62,5	63,44	59,574								
500	53,932	55	54,117	51,724	51,136							
501	61,956	61,728	63,218	54,444	56,666	52,38						
502	63,157	54,022	60	53,846	52,127	58,823	56,818					
503	56,25	57,647	61,111	56,989	60,215	57,647	64,13	57,608				
504	56,25	60,465	62,637	52,173	62,365	54,651	64,444	52,747	60,869			
600	67,368	62,5	68,888	63,736	72,043	43,023	54,545	51,086	57,142	49,462		
601	56,043	59,523	58,139	59,55	66,666	46,987	66,279	48,863	59,55	59,55	56,382	

la zona estudiada<sup>13</sup>. En nuestro trabajo nos hemos decantado por las distribuciones correspondientes a los puntos de referencia Te 104 (Barrachina), Te 202 (Calaceite) y Te 204 (La Codoñera), localidades en las que concurrían características cuantitativas de acusada personalidad, como se explica más abajo. Cada distribución de similitud ha sido proyectada en tres mapas aplicando tres algoritmos distintos de visualización (vid. *infra*). Todo ello ha dado lugar, por lo tanto, a la confección de nueve presentaciones gráficas diferenciadas (figs. 5 a 13).

Al margen de lo anterior, entre los elementos significativos, estadísticamente hablando, con que contamos en cada distribución de similitud se encuentra el *valor máximo*, esto es, el porcentaje de coincidencias lingüísticas más elevado de una localidad con otra (vid. tabla 5). El valor máximo de una distribución suele aparecer, por lo general, en la comparación con una localidad próxima al punto de referencia. El tratamiento cartográfico de los valores máximos de cada distribución (vid. fig. 2) permite identificar la localización y extensión de los diversos núcleos dialectales del

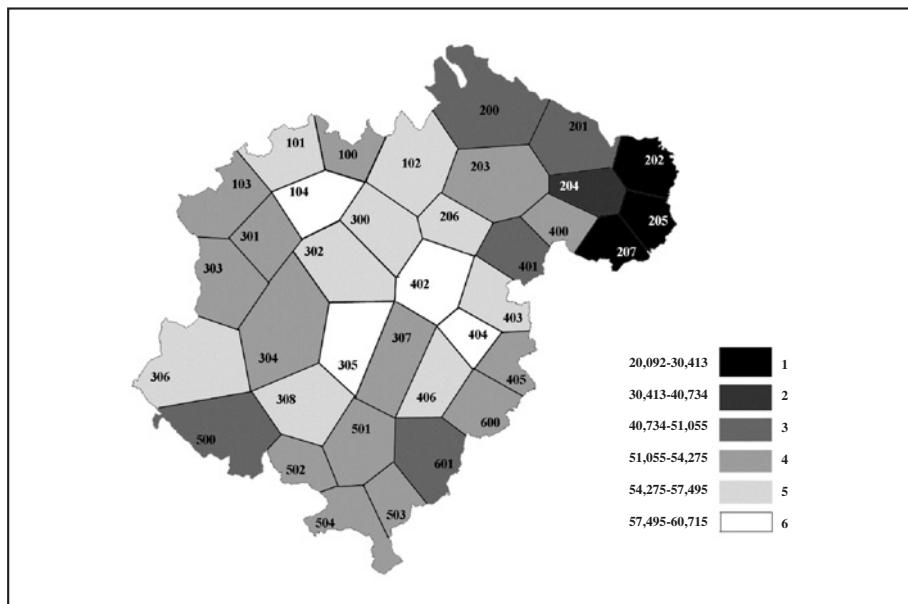


Fig. 3. Mapa poligonal de la distribución de la media del IRI (MINMWMAX).

13 Se pueden ensayar también presentaciones y montajes más ambiciosos y elaborados, con un efecto cinematográfico, como los que Goebel ha puesto en práctica en alguna ocasión proyectando consecutivamente el resultado cartográfico de diversas distribuciones de similitud pertenecientes a puntos de referencia insertos en un recorrido lingüístico representativo (GOEBL, 1981: 374).

TABLA 5

## Elaboración estadística de la matriz de similitud-Índice relativo de identidad

	MEDIA	MÍNIMA	MÁXIMA	ÍNDICE DE SIMETRÍA	LOCALIDAD
100	53,727	15,217	75	1,395	Nogueras
101	56,36	15,053	78,35	1,39	Ferreruela
102	55,423	16,129	75,257	1,357	Muniesa
103	53,086	13,978	75,257	1,417	Bello
104	60,715	15,217	78,35	1,29	Barrachina
200	48,292	16,129	60,824	1,259	Híjar
201	48,059	15,909	60,227	1,253	Alcañiz
202	20,092	11,494	52,747	2,625	Calaceite
203	53,203	18,681	65,957	1,239	Alloza
204	39,769	31,395	45,652	1,147	La Codoñera
205	29,515	21,276	61,538	2,084	Valderrobres
206	55,322	19,354	73,469	1,328	Estercuel
207	28,616	19,54	61,538	2,15	Peñarroya de Tastavins
300	55,05	19,148	69,387	1,26	Montalbán
301	52,28	16,129	71,134	1,36	Torrijo del Campo
302	57,23	18,888	75,531	1,319	Visiedo
303	52,556	11,494	70	1,331	Villar del Salz
304	53,28	16,129	71,134	1,335	Santa Eulalia del Campo
305	58,006	18,085	72,164	1,244	Alfambra
306	56,011	14,444	73,626	1,314	Noguera
307	51,24	21,111	65,591	1,28	Cedrillas
308	55,898	17,582	73,626	1,317	Teruel
400	53,117	17,977	66,666	1,255	Mas de las Matas
401	48,113	21,59	61,956	1,287	Bordón
402	59,952	17,021	75,531	1,259	Aliaga
403	55,444	16,279	74,418	1,342	Tronchón
404	58,418	23,863	75,531	1,292	Fortanete
405	52,591	25,274	71,428	1,358	La Iglesuela del Cid
406	55,052	16,304	72,043	1,308	Alcalá de la Selva
500	49,628	15,294	61,363	1,236	Masegoso
501	53,568	16,091	66,279	1,237	La Puebla de Valverde
502	53,526	16,666	64,444	1,203	Riodeva
503	51,645	18,681	64,13	1,241	Manzanera
504	51,151	16,666	64,444	1,259	Arcos de las Salinas
600	51,357	23,333	72,043	1,402	Puertomingalvo
601	50,977	21,59	66,666	1,307	Olba
	Media 51,055	Media 18,027	Media 68,425	Media 1,38	

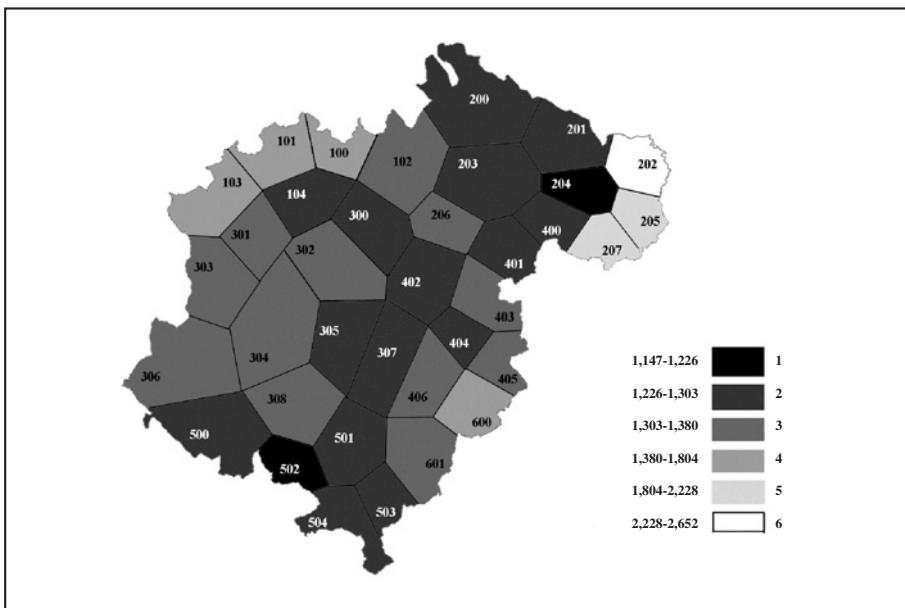


Fig. 4. Mapa poligonal de la distribución del índice de simetría (MINMWMAX).

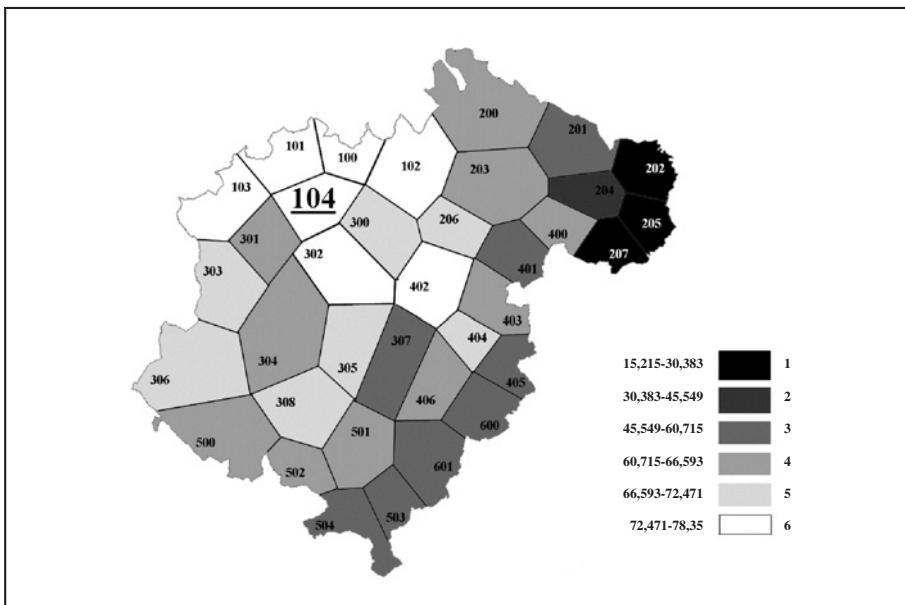


Fig. 5. Mapa poligonal de la distribución de similitud relativa al punto 104 (Barrachina) (MINMWMAX).

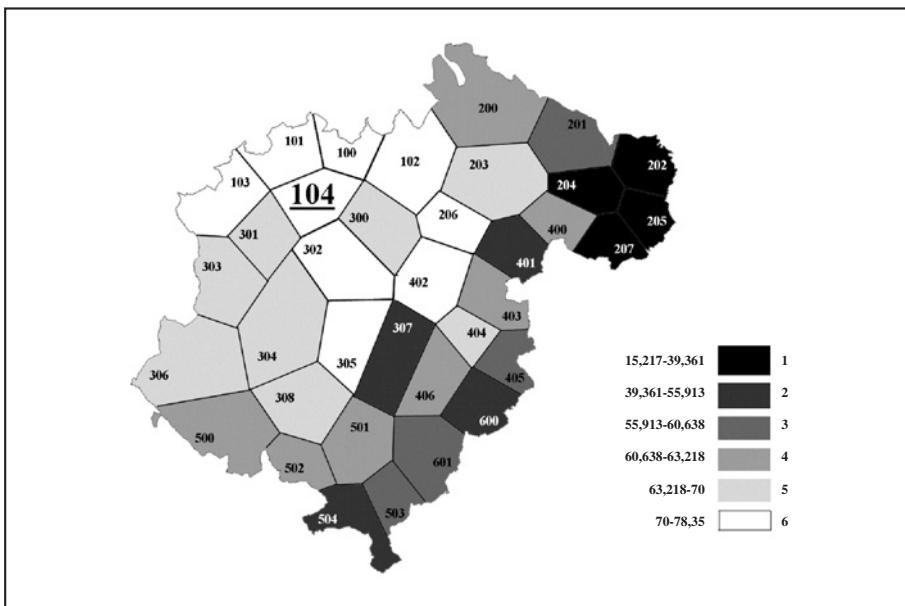


Fig. 6. Mapa poligonal de la distribución de similitud relativa al punto 104 (Barrachina) (MEDMW).

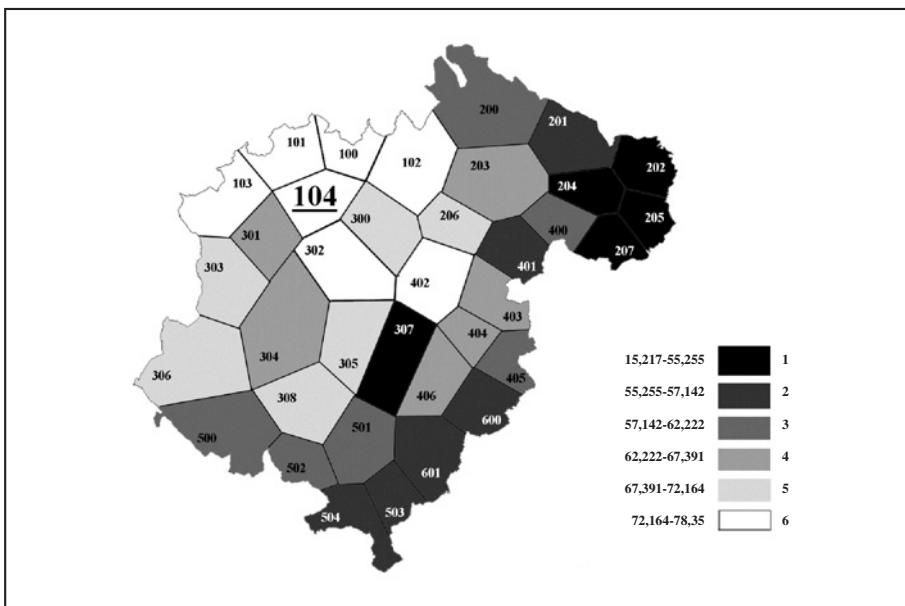


Fig. 7. Mapa poligonal de la distribución de similitud relativa al punto 104 (Barrachina) (MED).

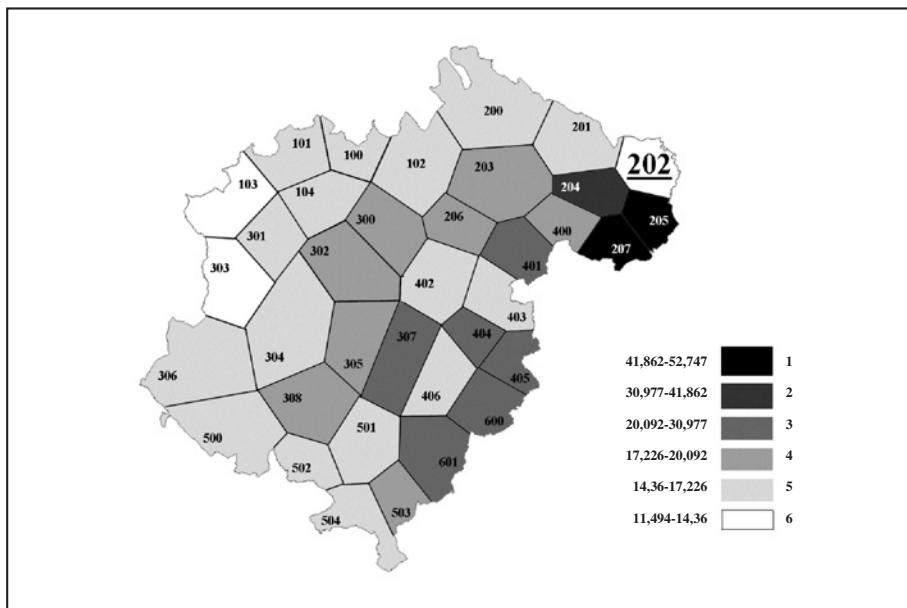


Fig. 8. Mapa poligonal de la distribución de similitud relativa al punto 202 (Calaceite) (MINMWMAX).

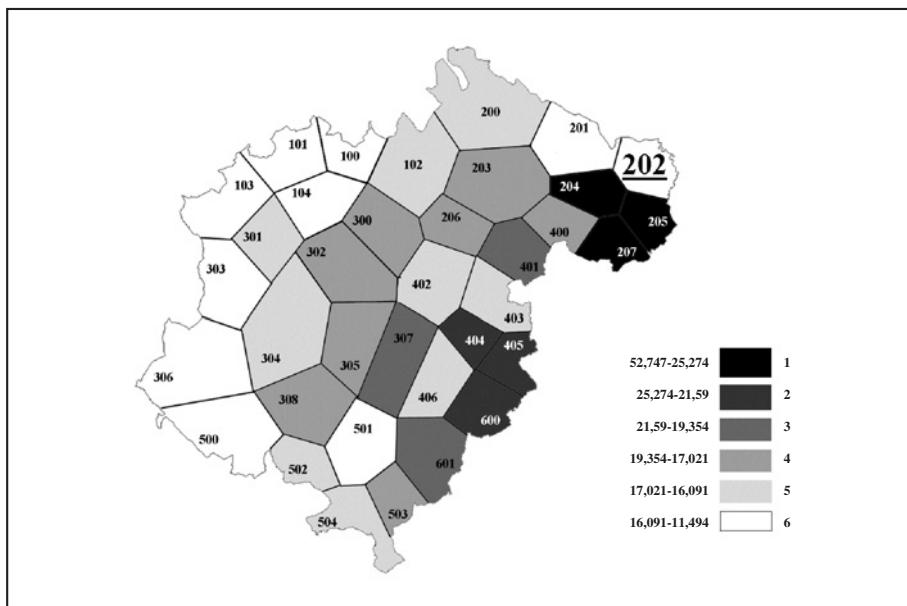


Fig. 9. Mapa poligonal de la distribución de similitud relativa al punto 202 (Calaceite) (MEDMW).

territorio estudiado (GOEBL, 1981: 383). No menos interés ofrece el análisis de la *media aritmética* de cada distribución (vid. tabla 5), cuya sinopsis gráfica nos informa sobre las zonas cuya modalidad lingüística ocupa una posición central en la red de puntos examinada y desempeña una función de mediación que facilita la intercomprensión comunicativa en la red de puntos examinada (GOEBL, 1981: 388; vid. fig. 3). Finalmente, es necesario referirse al cálculo y valor explicativo del *índice de simetría*. Aunque contamos con procedimientos sumamente complejos para determinar un coeficiente de simetría riguroso –como el propuesto por R.A. Fisher– es posible obtener un índice de simetría aceptable desde el punto de vista estadístico recurriendo al sencillo cálculo del cociente entre el valor máximo y la media aritmética de una distribución de similitud (GOEBL, 1981: 394 y 406), opción de la que nos hemos servido en este estudio (vid. tabla 5 y fig. 4). La medición de la simetría pone de relieve los vínculos, la conexión de un punto en el seno de una red determinada, conexión que será mayor cuanto más cercano a cero se encuentre el resultado de la operación descrita.

## LA REPRESENTACIÓN GRÁFICA

Es preciso, en este punto, referirse a dos aspectos de distinto orden. Por un lado, el puramente gráfico, la disposición que adopta el fondo del mapa sobre el que se vierten los resultados de las operaciones matemáticas. De entre las opciones manejadas habitualmente por las diferentes escuelas dialectométricas, la triangulación y el trazado poligonal de Thiessen, hemos elegido la confección geométrica en polígonos que transforma la representación discreta, por puntos, de un atlas lingüístico en otra de carácter continuo<sup>14</sup>. Conviene puntualizar, no obstante, que la triangulación y el trazado poligonal no son disposiciones gráficas completamente equivalentes o con las mismas capacidades explicativas (GOEBL, 1983).

Por otro lado, hay que referirse al procedimiento empleado para organizar de forma discreta en el mapa poligonal una variación numérica continua, en nuestro caso la variación numérica de cada distribución de similitud. Para cada una de las distribuciones de similitud estudiadas aquí –la de Te 104, Te 202 y Te 204– hemos aplicado los tres *algoritmos de visualización* (o *intervalización*) descritos por Goebel, a los que este autor identifica, respectivamente, con las siglas MINMWMAX, MEDMW y MED (1981: 361-368 y 1987: 79-90)<sup>15</sup>. Como en el caso de los tipos de articulación gráfica del fondo del mapa, cada algoritmo sirve a una finalidad taxométrica particular. Las diferencias básicas consisten en el hecho de que el algoritmo MINMWMAX está concebido para generar una visualización poco accidentada de la distribución dialectométrica de un territorio. El MEDMW y el MED, por su parte, tienden a poner de relieve las microzonas dialectales del territorio estudiado

14 Puede consultarse el modo de construir los polígonos de Thiessen en GOEBL (1981: 363-364 y 1983).

15 Para el cartografiado de los valores máximos, de la media aritmética y del índice de simetría nos hemos ceñido al algoritmo MINMWMAX (vid. figs. 2, 3 y 4).

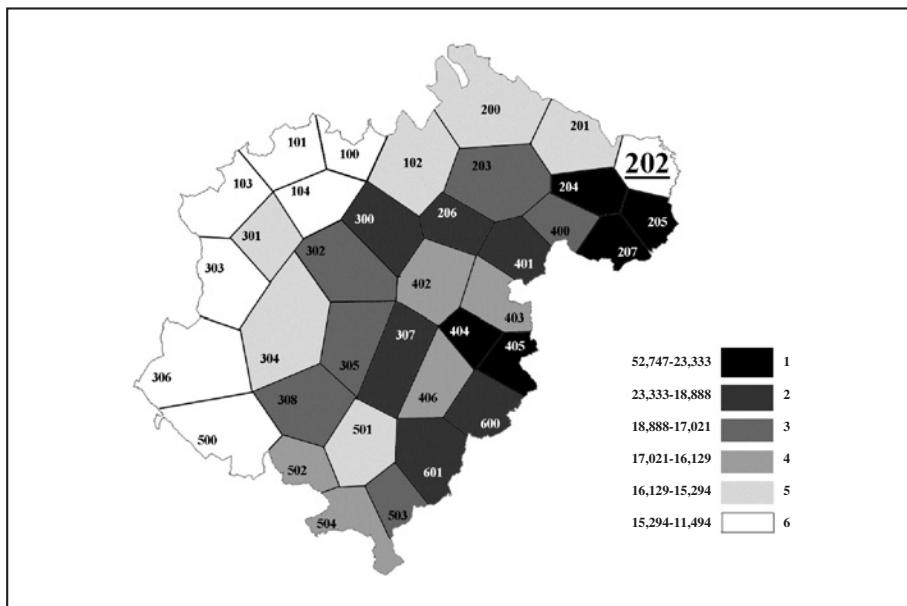


Fig. 10. Mapa poligonal de la distribución de similitud relativa al punto 202 (Calaceite) (MED).

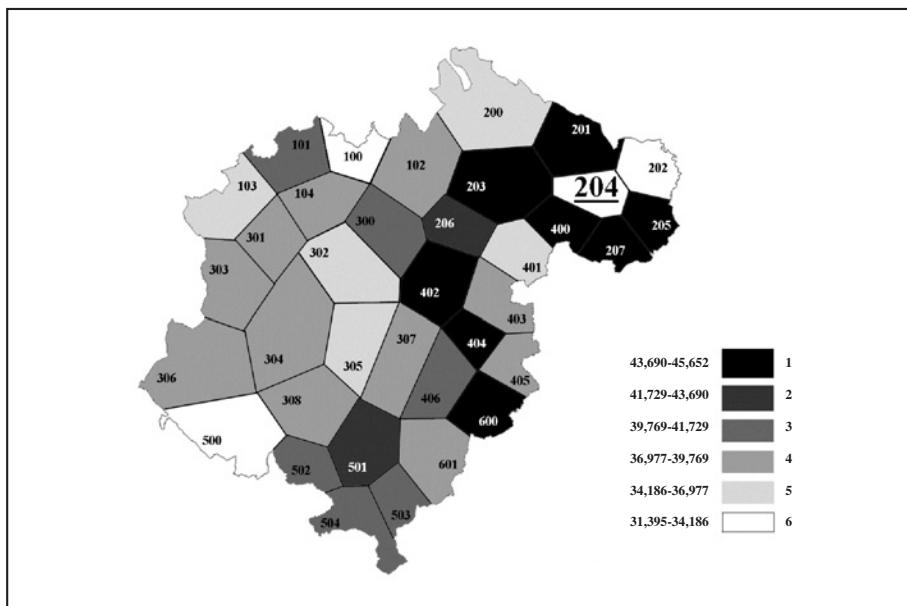


Fig. 11. Mapa poligonal de la distribución de similitud relativa al punto 204 (La Codoñera) (MINMWMAX).

pero el primero de forma progresiva y menos abrupta que el segundo. Los tres algoritmos comparten, eso sí, la propiedad de convertir en clases discretas la variación numérica continua de toda distribución de similitud. No está prefijado el número de clases más adecuado para segmentar una variación numérica cualquiera, aunque Goebl, con el que coincidimos aquí, ha utilizado por lo general una parcelación en seis intervalos asignando a cada uno de ellos un tipo diferenciado de trama tintada que nosotros hemos reemplazado por una escala de grises (vid. figs. 2 a 13).

El algoritmo MINMWMAX se obtiene del siguiente modo: 1) Se calcula la media aritmética de la distribución de similitud. 2) Se divide por tres la diferencia entre la media aritmética y el valor mínimo de la distribución para obtener la cifra que corresponde a la extensión de los intervalos situados por debajo de la media. Con la dimensión del intervalo resultante se establecen los límites de los tres intervalos por debajo de la media aritmética y, en consecuencia, los valores de la distribución (puntos del atlas) que aglutina cada intervalo. 3) Se realiza la misma operación para la diferencia entre la media aritmética y el máximo de la distribución de similitud. Presentado en una tabla, el resultado de la aplicación del algoritmo MINMWMAX a la distribución que tiene como punto de referencia a Te 202 es el siguiente (vid. tabla 4 y fig. 8):

INTERVALO	DE	HASTA	TAMAÑO DEL INTERVALO	NÚMERO DE PUNTOS DEL ATLAS POR INTERVALO
1	52.747	41.862	10,885	2
2	41.862	30.977	10,885	1
3	30.977	20.092*	10,885	6
4	20.092	17.226	2,866	8
5	17.226	14.36	2,866	16
6	14.36	11.494	2,866	2

\*Media aritmética

Total: 35

Por su parte, el algoritmo MEDMW resulta de las siguientes operaciones: 1) Se ordenan de menor a mayor los 35 valores de la distribución de frecuencia elegida. 2) Se busca el valor más próximo a la media aritmética. 3) El conjunto de valores que se encuentran entre el mínimo y el valor más próximo, y entre éste y el valor máximo forman, respectivamente, un grupo cada uno. 4) Cada uno de estos dos grupos se divide en tres clases iguales, en la medida de lo posible, tomando como criterio el número de puntos del atlas por clasificar. La tabla resultante para el punto Te 202 presenta las características que se señalan a continuación (vid. tabla 4 y fig. 9):

INTERVALO	DE	HASTA	TAMAÑO DEL INTERVALO	NÚMERO DE PUNTOS DEL ATLAS POR INTERVALO
1	52,747	25,274	27,473	3
2	25,274	21,59	3,684	3
3	21,59	19,354*	2,236	3
4	19,354	17,021	2,333	8
5	17,021	16,091	0,93	9
6	16,091	11,494	4,597	9
*Valor próximo a la media aritmética				Total: 35

Finalmente, el algoritmo MED requiere los cálculos que ahora se detallan: 1) Se ordenan de menor a mayor los 35 valores de la distribución de frecuencia elegida. 2) Los 35 valores se reparten en seis grupos iguales, en la medida de lo posible, ya que pueden ser necesarios algunos ajustes, como para Te 202, para que valores iguales no queden encuadrados en intervalos diferentes. El manejo del algoritmo MED para clasificar la distribución de similitud del punto Te 202 da lugar a la siguiente tabla (vid. tabla 4 y fig. 10):

INTERVALO	DE	HASTA	TAMAÑO DEL INTERVALO	NÚMERO DE PUNTOS DEL ATLAS POR INTERVALO
1	52,747	23,333	29,414	5
2	23,333	18,888	4,445	6
3	18,888	17,021	1,867	6
4	17,021	16,129	0,892	5
5	16,129	15,294	0,835	6
6	15,294	11,494	3,8	7
Total: 35				

## LA INTERPRETACIÓN DE LOS RESULTADOS

### LA DISTANCIA LINGÜÍSTICA ENTRE LAS HABLAS TUROLESES

Para cartografiar el *índice de diferenciación léxica* hemos distinguido siete intervalos, como puede apreciarse en la figura 1, con el objeto de lograr una percepción de las distancias lingüísticas

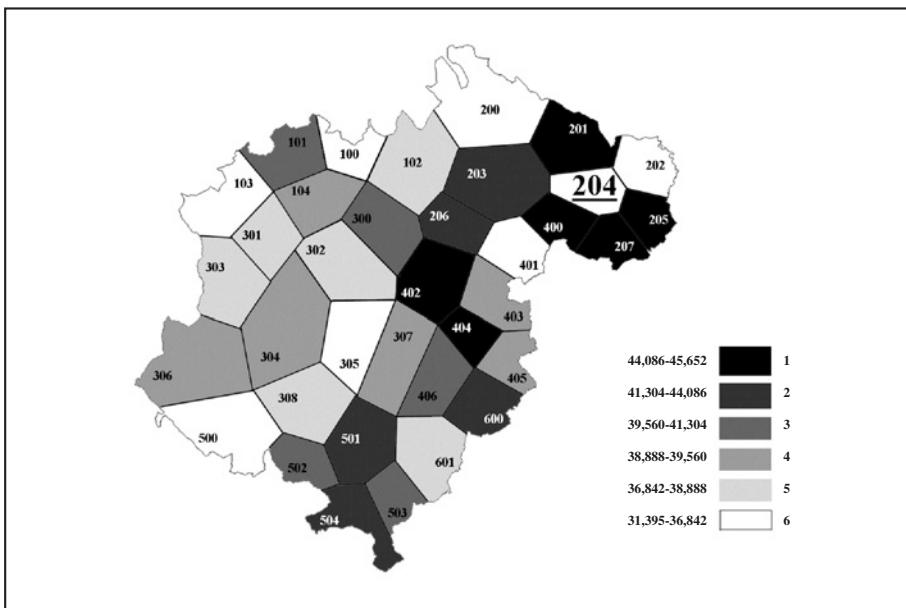


Fig. 12. Mapa poligonal de la distribución de similitud relativa al punto 204 (La Codoñera) (MEDMW).

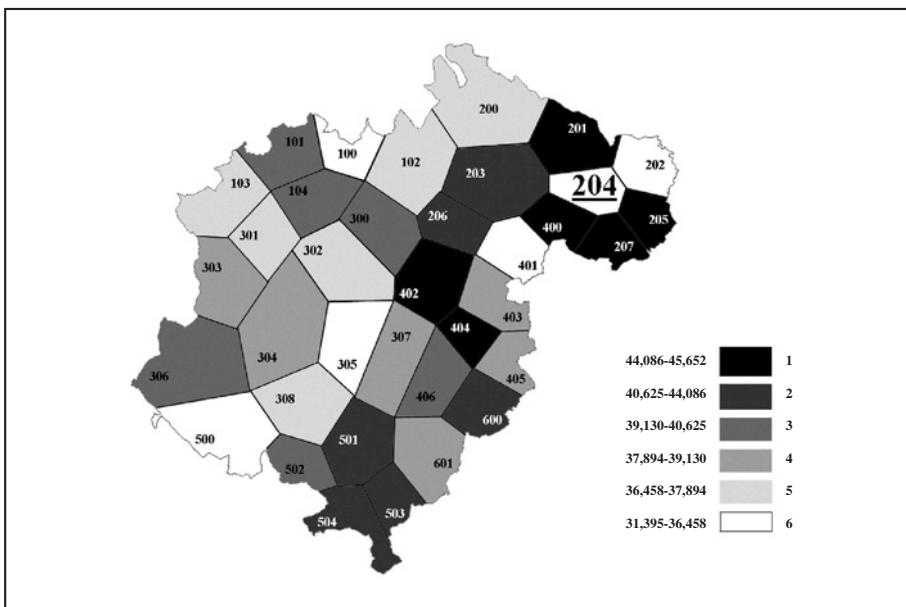


Fig. 13. Mapa poligonal de la distribución de similitud relativa al punto 204 (La Codoñera) (MED).

algo más matizada de la que hubiera surgido de discriminar sólo dos grupos, en lugar de cuatro, entre los valores más bajos. No procedía, sin embargo, efectuar la misma operación para los puntos que ofrecían los valores más altos de diferenciación ya que no había localidades con índices entre, por ejemplo, 0,79 y 0,75 y, de hecho, podían distinguirse tres intervalos para sólo cuatro localidades, extremo que da fe por sí solo de las peculiaridades locales del habla de las poblaciones de la Franja Oriental de Aragón. Así, mientras Calaceite (Te 202) alcanza el mayor índice de diferenciación de toda la provincia (0,81), Valderrobres (Te 205) y Peñarroya (Te 207) presentan índices elevados, homogéneos entre sí (0,727 y 0,733, respectivamente) pero sensiblemente inferiores al de Calaceite (vid. tabla 3)<sup>16</sup>. Por su parte, La Codoñera (Te 204), con un índice muy alejado del de las otras poblaciones de habla catalana (0,629), comienza a mostrar un perfil estadístico altamente peculiar que se verá confirmado por todos los demás indicadores manejados en este estudio.

Si atendemos a la comparación entre las cuatro poblaciones de habla catalana deteniéndonos en los datos de la tabla 2, vemos que Calaceite se distingue en 51 casos tanto de Valderrobres como de Peñarroya las cuales, de nuevo, se comportan de un modo muy similar y sólo se separan entre sí por 43 diferencias. Pero nuestro recuento arroja un total de 69, 58 y 60 diferencias entre La Codoñera y Calaceite, Valderrobres y Peñarroya, respectivamente. Esto quiere decir que La Codoñera se distancia de Calaceite más que cualquier localidad de habla castellana de otra con la misma modalidad, aunque en ciertas ocasiones –escasas– se superan entre éstas las 60 diferencias (61 diferencias entre Bordón –Te 401– y Santa Eulalia –Te 304– o entre Puertomingalvo –Te 600– y Villar del Salz –Te 303–). Se puede argüir, con razón, que ninguna población castellanohablante se acerca tanto como La Codoñera al habla de Calaceite, ya que respecto de esta última las diferencias oscilan entre las 76 de La Iglesuela del Cid (Te 405) y las 89 de Villar del Saz (Te 303). Pero adoptando otro punto de vista se puede argumentar, con igual razón, que la modalidad lingüística de La Codoñera, estadísticamente y de acuerdo con los datos del ALEANR, tiene menos relación con Calaceite que con cualquier otra localidad de habla castellana si exceptuamos Masegoso (Te 500) con la que mantiene 72 diferencias.

A propósito de los valores indicados en la tabla 3 es obligado, antes de cualquier otra consideración, subrayar la considerable magnitud de los porcentajes medios (columna tercera) y, por lo tanto, de los índices de diferenciación (columna quinta) de muchas localidades castellanohablantes de Teruel. De hecho, si comparamos nuestros resultados con los obtenidos por GARCÍA MOUTON en su análisis de Huesca (1991: 321) –recuérdese que se han examinado los mismos mapas del ALEANR en ambos trabajos– llama poderosamente la atención que la media del porcentaje de diferencias léxicas y del índice de diferenciación en el conjunto de Teruel alcanzan cifras más elevadas (52,43%.

16 Recordaremos aquí la coincidencia con M. ALVAR (1955) quien, con distinto método, cifró en un 79,5% el componente catalán del habla de Calaceite a partir del *Atlas Lingüístic de Catalunya* (Barcelona, Institut d'Estudis Catalans, 1923-1936). Véanse también los comentarios de F. MORENO sobre este aspecto en la provincia de Huesca (1991: 303).

y 0,524) que las medias obtenidas para la provincia norteña de Aragón (50,26% y 0,502). A primera vista, no deja de constituir un dato cuando menos llamativo si consideramos que la considerable fragmentación lingüística de las modalidades altoaragonesas es uno de los rasgos distintivos de las hablas norteñas frente a las meridionales de Zaragoza y Teruel. Podría pensarse, en primer lugar, en una diferencia metodológica consistente en una discriminación de tipos léxicos ("taxats") por nuestra parte más fraccionada que la realizada por García Mouton<sup>17</sup>. Pero un intento de valorar este aspecto presentaría no pocos inconvenientes. Por un lado, el catálogo de tipos léxicos es una información que, como señalamos más arriba, no se facilita en los estudios consultados. Y en lo que se refiere concretamente al *ALEANR*, la diversidad de respuestas en el mismo mapa no tiene por qué coincidir –y, de hecho, rara vez coincide– entre las localidades de una y otra provincia<sup>18</sup>.

No obstante, aun en el caso de que la agrupación de ciertas variantes léxicas lograra disminuir en algunas centésimas la media de diferenciación de las hablas turolenses, el porcentaje citado continuaría siendo notablemente elevado y merecedor de otras reflexiones. En este sentido, no puede obviarse el hecho de que, cualitativamente, el *ALEANR* certifica la existencia de un extraordinario grado de diversidad léxica en las hablas turolenses, diversidad que quizás pueda explicarse por la pervivencia de vocabulario tipológicamente dialectal o por influjo catalán en ciertas zonas, pero que responde asimismo a una notoria variación interna del español regional. De los 100 mapas estudiados, 73 acumulan cinco o más tipos léxicos y ninguno de los restantes resultó ser mononímico. La media de "taxats" por mapa, como se apuntó más arriba, alcanzó la cifra de 6,85. Reparemos ahora en otro aspecto. Según las estimaciones de MORENO (1991: 304), una media de diferenciación por encima del 70% apuntaría a la presencia de un sistema lingüístico diferente. Si el porcentaje no alcanza el 50% podríamos considerar que se trata de un habla de tipo castellano y los valores intermedios darían cuenta de un espacio claramente dialectal. En una línea similar, aunque de modo menos preciso, se pronuncia GARCÍA MOUTON (1991: 324). De acuerdo con estas apreciaciones –y al margen del comentario que requieren de nuevo los datos de La Codoñera (Te 204)– debemos llamar la atención sobre las diecisiete localidades castellanohablantes de Teruel cuya media de diferenciación rebasa el 50% de diferencias léxicas<sup>19</sup>. Todas ellas, a excepción de Torrijo (Te 301) y

17 Si un investigador aísla, por ejemplo, tres tipos léxicos donde otro distingue cinco, el número final de diferencias léxicas resultará algo mayor en el segundo caso. Si esta circunstancia se repite siempre en el mismo sentido la imagen final de las diferentes distancias lingüísticas mostrará quizás variaciones de cierta importancia sin que, necesariamente, una versión pueda calificarse como más adecuada o fiel que la otra si están justificados los criterios utilizados por ambos dialectólogos.

18 Por ello, aunque hemos adoptado los criterios diferenciadores expuestos por la citada autora (1991: 315-320) nos hemos enfrentado a materiales de trabajo parcialmente divergentes. Sólo una investigación paralela sobre idénticos materiales –los mismos mapas de una sola zona– podría poner de manifiesto la entidad de las hipotéticas discrepancias de método a las que nos referimos.

19 Son las siguientes (vid. tabla 3): Te 200 Hijar, Te 201 Alcañiz, Te 203 Alloza, Te 301 Torrijo, Te 303 Villar del Salz, Te 307 Cedrillas, Te 400 Mas de las Matas, Te 401 Bordón, Te 403 Tronchón, Te 405 La Iglesuela del Cid, Te 500 Masegoso, Te 501 La Puebla de Valverde, Te 502 Riodeva, Te 503 Manzanera, Te 504 Arcos de las

Villar del Salz (Te 303) se sitúan significativamente en la mitad oriental de la provincia, precisamente en la zona en la que los estudios cualitativos –que sopesan, no olvidemos, la fonética, la morfosintaxis y el léxico– han detectado mayor índice de rasgos dialectales aragoneses (ENGUITA, 1985) o de influjo catalán (FORT, 1991).

Cabe inducir de ello que, en términos globales, la medición cuantitativa viene a complementar la aproximación tradicional para confirmarla: los valores situados entre el 50 y el 55 por ciento de diferenciación darían cuenta, así, de unas hablas dialectalmente peculiares por uno u otro condicionante. Sin embargo, no resultaría acertada una simplificación excesiva porque algunas situaciones no encajan en una explicación general como la anterior. Por ejemplo, no se distancian llamativamente del resto ni Munesa (Te 102) ni Estercuel (Te 206) –con índices de 0,474 y 0,477, respectivamente– pese a ser dos de los puntos más conservadores de caracteres aragoneses junto con Hijar, Alloza y Mas de las Matas, según la aproximación de ENGUITA (1985: 204). En sentido contrario, Alcañiz (Te 201), cuya modalidad cataloga este mismo investigador entre las más castellanizadas de Teruel y pertenece, como las anteriores, al rincón noreste de la provincia, arroja un índice de 0,559, la distancia lingüística más elevada de toda la zona castellanohablante de Teruel junto con la de Bordón (Te 401), asimismo con 0,559 (vid. nota 20). En cuanto a la zona de influjo catalán, cabe decir que las hablas de Fortanete (Te 404) y Alcalá (Te 406) poseen un índice de diferenciación poco relevante (0,465 y 0,478) y, sin embargo, ambas localidades pertenecen al grupo de las que exteriorizan la mayor impronta del catalán de toda la provincia (FORT, 1991: 197-198), aspecto este último que también se pone de relieve en nuestro corpus, en especial, en el caso de Fortanete como se verá luego.

Contamos, por otro lado, con el caso contrario. El elevado índice (0,558) de Masegoso (Te 500) no responde, a juzgar por nuestros datos, a la destacada presencia en su habla de rasgos catalanes detectada por FORT (1991: 197-198) sino a la de usos léxicos distintivos en el marco del español regional –extremo ratificado por su relativamente baja *desviación típica*– y, del mismo modo, al hecho de que en once mapas el informante no respondió. Estas casillas vacías representan una diferencia fija en la comparación con el resto de los puntos salvo que la respuesta falte en los dos. Algo muy similar sucede con La Puebla de Valverde (Te 501), cuyo índice de 0,516 es atribuible en parte a los ocho mapas sin respuesta y, en menor medida, con Tronchón (Te 403) aunque en once ocasiones el encuestador no consiguiera la información requerida. Son todas ellas, localidades poco afines al catalán según el corpus manejado aquí, lo que no se corresponde con las conclusiones de una aproximación cualitativa (FORT, 1991: 197-198). Sin embargo, a pesar de las once casillas vacías de

Salinas, Te 600 Puertomingalvo y Te 601 Olba. Ciertamente, algunas superan por escaso margen esa cantidad (Alloza, Torrijo, Tronchón, La Iglesuela del Cid y Riodeva) pero otras alcanzan el 55% o se acercan a esa cifra (Hijar, Alcañiz, Bordón, Masegoso y Olba). Interesa, por ello, subrayar que sólo seis de esas diecisiete localidades sobrepasan el porcentaje y el índice medio de diferenciación (52,43% y 0,524): Te 200 Hijar, Te 201 Alcañiz, Te 401 Bordón, Te 500 Masegoso, Te 504 Arcos de las Salinas y Te 601 Olba. A éstas se suman las localidades catalanohablantes, por encima asimismo de los valores medios: Te 202 Calaceite, Te 204 La Codoñera, Te 205 Valderrobres y Te 207 Peñarroya.

Olba (Te 601) y de las siete de Bordón (Te 401) ambas localidades se siguen mostrando próximas al catalán<sup>20</sup>. En cualquier caso, no deseamos magnificar las discrepancias, que asoman sobre un dibujo de fondo claramente común a las dos perspectivas de análisis, según venimos diciendo, y no constituyen necesariamente contradicciones. Más bien debe pensarse, para estos casos, en un estrato de discrepancias (y afinidades) entre las distintas hablas que capta una aproximación cuantitativa pero no una cualitativa y viceversa. También cabe anticipar una consecuencia crítica relativa a la metodología adoptada: quizás deba compensarse el corpus de trabajo, perteneciente en exclusiva al nivel léxico, con el tratamiento conjunto de mapas morfológicos y fonéticos, posibilidad por la que se decantó H. GUITER en algún trabajo (1973) a diferencia de la práctica observada por J. SÉGUY (1971).

Sin embargo, en el extremo opuesto de los indicadores de la tabla 3 se perfila un cuadro altamente esclarecedor y mucho más acorde con las descripciones tradicionales. Las nueve localidades con menor media e índice de diferenciación se encuentran en torno al eje que podemos trazar entre Barrachina (Te 104) y Aliaga (Te 402), los dos puntos menos diferenciados de Teruel o, en otras palabras, con una habla menos específica (índices de 0,429 y 0,426). Predominan en este grupo las poblaciones del cuadrante noroeste de la provincia (Te 101 Ferreruela, Te 102 Muniesa, Te 300 Montalbán, Te 302 Visiedo) con una ligera penetración hacia el sur (Te 305 Alfambra) y hacia el este (Te 206 Esteruel y Te 404 Fortanete). Ocupan el ámbito en el que convergen todas las tendencias lingüísticas de la provincia sin que ninguna ostente una posición de dominio. Por un lado, seis de esas nueve poblaciones (Esteruel, Visiedo, Aliaga, Alfambra y Barrachina) forman parte de una zona intermedia de penetración occidental de rasgos catalanes (FORT, 1991: 197-198). Por otro, se sitúan en el límite entre la zona más conservadora de rasgos aragoneses y la más castellanizada (ENGUITA, 1985: 204-205 y 1991a: 125-151). Esta equidistancia explica la escasa magnitud de sus valores de diferenciación y también las elevadas *desviaciones típicas*, que nos advierten de la pobre peculiaridad de las hablas de esta franja. Este último dato no puede ser más revelador. A excepción de la de Fortanete –a cuya particularidad ya nos hemos referido más arriba– todas las hablas de esta área, que podemos denominar *zona de convergencia*, alcanzan unas desviaciones típicas por encima de la media, situada en 10,28, empezando por Barrachina (Te 104) con la más alta de toda la provincia (13,94)<sup>21</sup>.

- 20 Los ocho mapas sin respuesta explican, asimismo, el índice (0,522) de Villar del Salz (Te 303), localidad del extremo occidental de Teruel. Por su parte, Torrijo (Te 301), también en la parte occidental, sólo excede en dos milésimas del índice del 0,5 (0,502) por lo que, dentro de su peculiaridad, se mueve en los límites de las localidades de su entorno. Finalmente, el índice (0,559) de Alcañiz (Te 201) se ve afectado al alza por las siete casillas vacías del corpus, pero no podemos perder de vista que, como Masegoso, Alcañiz posee una desviación típica (9,05) muy inferior a la media (10,28) indicativa de un habla marcadamente singular.
- 21 El trazado de áreas lingüísticas propuesto para Aragón por R. CASTAÑER (1990: 371) viene a confirmar plenamente nuestra interpretación. El ámbito que hemos denominado *zona de convergencia* se corresponde casi con total precisión con un área en torno a Te 206 Esteruel, Te 300 Montalbán y Te 402 Aliaga, donde se encuentran las isoglosas que delimitan los cuatro espacios dialectales del español que Castañer descubre en la provincia meridional de Aragón.

Casi todas las demás desviaciones típicas que superan la media conciernen a un grupo homogéneo de localidades emplazadas en el cuadrante centro-occidental, en contacto estrecho con las hablas castellanas de Guadalajara (Te 103 Bello, Te 301 Torrijo, Te 303 Villar del Salz, Te 304 Santa Eulalia, Te 306 Noguera y Te 308 Teruel). Son hablas muy cercanas entre sí (menos de cuarenta diferencias de media) y muy alejadas del resto, en especial, como podía preverse, de la mitad oriental. En sentido contrario, entre los puntos con una menor desviación típica ocupan un lugar destacado los de habla catalana, al igual que en la provincia de Huesca (GARCÍA MOUTON, 1991: 321). Son, lógicamente, hablas que, al margen del influjo que ejercen de modo más o menos acusado en el entorno castellanohablante, se oponen de forma muy parecida a casi todas las demás, incluyendo en ocasiones a las propias de la Franja Oriental. Al margen de esto, el cuadrante suroriental reúne a un grupo nutrido de hablas singularizadas por su reducida desviación típica: Cedrillas (Te 307), La Iglesuela del Cid (Te 405), Manzanera (Te 503), Puertomingalvo (Te 600) y Olba (Te 601). Pero son particularmente llamativas las desviaciones típicas más reducidas del conjunto. Bordón (Te 401), con 7,52, posee la segunda más baja de Teruel, menor incluso que la resultante en las poblaciones de la Franja si exceptuamos la atingente a La Codoñera (Te 204), con tan sólo un 3,45, punto que merece un comentario aparte que recapitule las características de su habla a la vista de su personalidad estadística. Bordón, por su parte, comparte con Alcañiz el mayor índice de diferenciación del español regional de Teruel (0,559) e igual cantidad de mapas –siete– sin respuesta. Pero, además, el informante de Bordón facilitó un número apreciable de formas léxicas muy peculiares que aclaran la reducida desviación típica.

### EL VALOR MÁXIMO, LA MEDIA ARITMÉTICA Y EL ÍNDICE DE SIMETRÍA

Decíamos más arriba que el cartografiado de los *valores máximos* de una distribución de similitud permitía detectar la ubicación de focos dialectales relevantes. Así lo confirman las características de la figura 2, que permiten reconocer con bastante nitidez las siguientes zonas:

1) Área de la Franja Oriental, de habla catalana, a cuya delimitación poco se puede añadir salvo quizás el hecho de que el elevado nivel de similitud (*IRI* superior al 61%) entre Valderrobres (Te 205) y Peñarroya (Te 207) sitúa a estos puntos en el tercer intervalo de visualización y no en el primero, al que se adscriben Calaceite y La Codoñera.

2) Área conservadora de rasgos autóctonos aragoneses (ENGUITA, 1985: 204-205 y CASTAÑER, 1990: 371), menos destacada en el mapa que en los datos (vid. tabla 4) y aglutinada en torno a Esteruel (Te 206), localidad frente a la cual alcanzan el valor máximo de similitud Híjar (Te 200), Alloza (203) y Bordón (Te 401). A su vez, el *IRI* más elevado de Esteruel se emplaza en Aliaga (Te 402), al igual que el de Mas de las Matas (Te 400).

3) Área del rincón noroeste de Teruel, perfectamente caracterizada en los estudios tradicionales por su acentuada castellanización, su vinculación lingüística con el extremo suroeste de Zaragoza y por quedar al margen, casi por completo, de la influencia del catalán. Estas propiedades, conjunta o individualmente, se ponen de relieve en las investigaciones de CASTAÑER (1990: 371), FORT (1991:

197-198), ENGUITA (1991a: 150 y 1991b: 220) y LLORENTE (1991: 181). El epicentro de este ámbito parece ser Barrachina (Te 104) que congrega el *IRI* máximo en las distribuciones de similitud de Nogueras (Te 100), Ferreruela (Te 101), Muniesa (Te 102), Bello (Te 103), Montalbán (Te 300) y Visiedo (Te 302). A su vez el valor máximo de Barrachina se origina en la comparación con Ferreruela.

4) Área del sector sureste de Teruel de influencia valenciana (CASTAÑER, 1990: 317; FORT, 1991: 197-198 y LLORENTE, 1991: 180-181). De acuerdo con nuestros datos sus centros serían Alcalá de la Selva (Te 406) y La Puebla de Valverde (Te 501). En la primera se agrupan las máximas de Puertomingalvo (Te 600), Olba (Te 601) y Cedrillas (Te 307) y en la segunda las de Manzanera (Te 503) y Arcos de las Salinas (Te 504). La máxima de Alcalá queda emplazada, a su vez, en Puertomingalvo y la de La Puebla de Valverde en Olba.

La validez del procedimiento taxométrico seguido se pone de manifiesto, de igual modo, en las figuras 3 y 4. En la primera se representa la distribución de la media aritmética del *índice relativo de identidad* de cada punto del *ALEANR* en Teruel. La disposición centrífuga resultante (de una magnitud mayor en el centro a una menor en la periferia) nos está indicando la posición relativa de las hablas turolenses de una perspectiva comunicativa mediadora entre las distintas corrientes dialectales. No por casualidad cobra de nuevo protagonismo la región distinguida más arriba como *zona de convergencia*, ahora con un peso mayor de las hablas-puente hacia el sur y el sureste concretadas en Alfambra (Te 305) y Fortanete (Te 404). Por su parte, la figura 4, síntesis gráfica de la distribución del *índice de simetría*, adopta una disposición francamente estética en forma de cruz o, mejor, de aspa. Quedan así trazados los ejes a través de los cuales discurren las zonas de contacto entre los espacios dialectales de la provincia, surcada, por un lado, por el eje noreste-suroeste, verdadero nexo de enlace entre el ámbito más vinculado lingüísticamente al oriente de la Península y el más castellanizado. Los núcleos de mayor conectividad aparecen en los extremos del eje, en La Codoñera (Te 204) y Riodeva (502). Por otro, esta línea se ve atravesada en dirección noroeste-sureste por el eje Barrachina-Montalbán-Aliaga-Fortanete (que reaparece para proporcionar nueva información) donde se concentran las hablas que sirven de engarce entre el área más conservadora de rasgos que entroncan con el antiguo dialecto, en el noreste, y la que parece más distanciada del romance autóctono, en el suroeste.

## LAS DISTRIBUCIONES DE SIMILITUD RELATIVAS A LOS PUNTOS TE 104, TE 202 Y TE 204

Hemos escogido para su cartografiado, de entre todas las posibles, las distribuciones de similitud relativas a tres puntos cuyos valores despuntan por alguna propiedad estadística. Ya nos hemos referido en varias ocasiones a los indicadores de La Codoñera (Te 204). Parecía de interés conocer las vinculaciones lingüísticas de una modalidad con un rendimiento estadístico tan distanciado del de todas las demás. En el caso de Barrachina (Te 104) y Calaceite (Te 202) porque se trataba de los puntos que arrojan los valores extremos de la media aritmética del *IRI* (60,715 y 20,092, respectivamente). Esto es, por representar respecto del resto de Teruel el habla más similar (Barrachina) y la menos parecida (Calaceite). De acuerdo con el corpus manejado, podemos interpretar el dato de

Calaceite como el indicador de que nos hallamos ante la modalidad *más catalana*, en el sentido de la más cercana al catalán estándar y menos impregnada de castellanismos (QUINTANA, 1976-80: 190-191). En consecuencia, su distribución de similitud nos orienta sobre qué variedades dialectales presentan más cercanía con el catalán. Sin embargo, no podemos deducir, a partir de nuestro corpus, que Barrachina posea la modalidad más castellanizada de la provincia sino sólo la menos peculiar o, dicho de otro modo, la más parecida a todas las demás. Sabemos, no obstante, que a esta localidad alcanza muy débilmente el influjo tanto catalán como aragonés y, al mismo tiempo, se encuentra en la órbita de las hablas castellanas limítrofes con Aragón<sup>22</sup>. Creemos, pues, que todo ello autoriza, con las debidas precauciones, a tomar Barrachina como término de referencia con el que valorar la mayor o menor castellanización de las hablas turolenses.

Y lo cierto es que, aunque el juego de grises se ha utilizado con criterio inverso, los mapas correspondientes a la distribución de similitud de Barrachina (figs. 5, 6 y 7) sugieren una imagen casi invertida de los correspondientes a Calaceite (figs. 8, 9 y 10). En términos generales, las concomitancias con Calaceite disminuyen conforme se avanza hacia el oeste mientras que con Barrachina se degradan en sentido opuesto. Cabe matizar, claro está, que las coincidencias lingüísticas con Barrachina quedan de manifiesto, sobre todo, en el cuadrante noroccidental y en la llamada *zona de convergencia* –y son algo inferiores con el suroeste de Teruel– y que, sin embargo, las discrepancias más acusadas respecto de Calaceite quedan patentes a lo largo de toda la frontera con Cuenca, Guadalajara y el suroeste de Zaragoza. Nótense, por otro lado, cómo los algoritmos de visualización acentúan las diferencias, más difuminadas con la aplicación del MINMWMAX, más abruptas con el MED. En la distribución de Barrachina, por ejemplo, el polígono de Cedrillas (Te 307) asciende de intervalo con cada algoritmo, alejándose del punto de referencia, hasta encuadrarse en el primero, junto con los puntos de habla catalana en el mapa de la figura 7. En la distribución de Calaceite, las hablas de Fortanete (Te 404) y La Iglesuela del Cid (Te 405) se muestran más *catalanizadas* con el algoritmo MEDMW y, en un grado mayor, con el MED. Puede apreciarse, también, que el algoritmo MED ayuda a perfilar la personalidad de la frontera con el valenciano en el sureste de Teruel (figs. 7 y 10).

Queda, por último, referirse a la distribución de similitud de La Codoñera (Te 204). Se ha comentado ya la singularidad de los valores estadísticos de este punto de la Franja, tanto en el índice de diferenciación, como en el índice de simetría o en la media aritmética del IRI. Con el algoritmo MINMWMAX, tampoco se comporta igual que las otras poblaciones de habla catalana turolenses del ALEANR, como se aprecia en los mapas poligonales correspondientes a Barrachina y Calaceite (figs. 5 y 8). Así, se puede afirmar que, estadísticamente y en el marco de las hablas turolenses, la modalidad lingüística de La Codoñera está menos diferenciada que la de Calaceite,

22 Para calcular numéricamente la distancia del habla de cada localidad respecto del estándar –en este caso, respecto del español estándar– sin recurrir a las descripciones cualitativas podría hacerse uso de la solución manejada habitualmente por H. Goebel. Consiste en la introducción, como una población más, de un punto artificial representante del estándar. Sus datos estarían integrados por los nombres de cada mapa, pertenecientes, por lo general, a la variedad de referencia.

Valderrobres y Peñarroya de Tastavins, ocupa una posición menos marginal que la de éstas y su nivel de conectividad con las hablas de su entorno es el máximo de la provincia mientras que el de Calaceite, Valderrobres y Peñarroya es mínimo. Se acerca más al habla castellana de Barrachina que las otras variedades catalanas y se parece menos a la modalidad catalana de Calaceite que las hablas de Valderrobres y Peñarroya. El cartografiado de la distribución de similitud de La Codoñera (figs. 11, 12 y 13) viene a ratificar los comentarios precedentes al mostrar una compleja red de afinidades y discordancias que no nos atrevemos a calificar de irregular. En una primera aproximación habría que subrayar, lógicamente, que su modalidad se encuentra más estrechamente vinculada a las del sector oriental de Teruel. Pero no debe perderse de vista el número de poblaciones occidentales que ocupan los intervalos tercero y cuarto, así como los puntos orientales emplazados en los tres últimos intervalos, empezando por Calaceite. El arco que describen los puntos taxométricamente más próximos a La Codoñera, dejando aislados algunos puntos limítrofes con el valenciano, hace emerger la conexión de esta población con la *segunda línea* de influjo catalán en las hablas castellanas de Teruel. No por casualidad, con los algoritmos MEDMW y MED particularmente (figs. 12 y 13), la distribución de similitud de La Codoñera presenta evidentes puntos de contacto con la del *índice de simetría* (fig. 4), donde se asocia con localidades de carácter lingüístico fronterizo antes que con otras de contornos más delimitados.

En definitiva, si tuviéramos que juzgar en exclusiva por las implicaciones cuantitativas, calificáramos el habla de La Codoñera como típicamente fronteriza, subdialectal y de pronunciada singularidad al mismo tiempo. Salvando las distancias, desempeñaría en Teruel el papel que ocupan en Huesca la modalidad de Benasque o, en cierta medida, la de Gistaín. Pero las investigaciones dialectológicas cualitativas no apuntan a una caracterización de esa naturaleza al describir el habla de La Codoñera. Es conocida la controversia sobre la adscripción lingüística del benasqués al aragonés o al catalán, es decir, su carácter híbrido, al menos desde el punto de vista taxonómico. La filiación aragonesa del chistavino no se pone en tela de juicio aunque algunos rasgos, como el perfecto perífrastico, lo vinculen en alguna medida al romance peninsular más oriental. Sin embargo, a propósito del habla de La Codoñera, A. QUINTANA (1976-80: 6) concluye que "no té justificació parlar ací de català de transició a l'aragonès o al castellà o bé al revés" sino de una "variant del català occidental". Admite, claro está, sus propiedades distintivas de carácter comarcal y local. En sus aspectos más significativos, éstas se concretan, en el plano fónico, en un sistema vocálico de cinco vocales (en lugar de siete), en un sistema consonántico en el que tiene cabida el fonema /θ/ interdental fricativo sordo, en la presencia de diptongación [já] en lugar de la [é] abierta del catalán occidental y, en el nivel léxico, en la considerable concurrencia de localismos, arcaísmos y castellanismos. La Codoñera está ubicada en el valle del río Guadalupe compartido con poblaciones castellanohablantes entre las que se encuentra el importante centro comarcal de Alcañiz. Sin embargo, Calaceite, Valderrobres y Peñarroya de Tastavins se encuentran enclavadas en el valle, más oriental, del río Matarraña, en contacto directo con las hablas catalanas de Tarragona, territorio con el que Calaceite y otros pueblos próximos mantuvieron vínculos históricos en el pasado al estar adscritos al obispado de Tortosa y no al de Zaragoza.

## CONCLUSIONES

No sabríamos decir si la mirada propuesta aquí sobre las hablas de Teruel encaja en la categoría de nueva perspectiva sobre datos dialectales actuales o en la de revisión de viejos materiales (GARCÍA MOUTON, 1991: 326). El coste temporal en la confección de los atlas lingüísticos y el tipo de informante, de edad avanzada, privilegiado por la geografía lingüística, obligan a que, transcurridos unos años, la interpretación de las informaciones tan laboriosamente recopiladas y transcritas deba incluir alguna referencia a la vigencia sincrónica de las informaciones. Todo ello no resta validez, claro está, a los resultados que se puedan alcanzar. Sólo relativiza su extrapolación cronológica.

En el marco del método dialectométrico las hablas de Teruel han sido examinadas a lo largo de estas páginas en sus diferencias y semejanzas. No hemos agotado, por supuesto, todos los cálculos estadísticos que permitían los materiales recopilados, que son susceptibles todavía de aprovechamientos ulteriores. Pero, aun ciñéndonos a las operaciones que hemos considerado más relevantes, hemos logrado que las modalidades lingüísticas turolenses dispongan de la descripción dialectométrica más detallada que, hasta donde conocemos, se ha ensayado para las hablas peninsulares. Con ella se han visto ratificados muchos de los aspectos anticipados por la dialectología tradicional. No obstante, se han sacado a la luz ciertas discrepancias que, en parte, podrían ser imputables a la selección exclusiva de un corpus de mapas léxicos, pero que se derivan también, no se olvide, de que dialectología y dialectometría *operan de distinta manera* y, en definitiva, cada una de ellas crea *un objeto propio* con su particular punto de vista.

## BIBLIOGRAFÍA

- ALVAR, M. (1955), «Catalán y aragonés en las regiones fronterizas», *Actas del VII Congreso Internacional de Lingüística Románica*, Barcelona, pp. 737-778. (También en M. ALVAR., *La frontera catalano-aragonesa*, Zaragoza, IFC, 1976).
- ALVAR, M. y cols. (1979-1983). *Atlas Lingüístico Etnográfico de Aragón, Navarra y Rioja*, 12 vols., IFC-CSIC.
- CASTAÑER, R. (1990). *Estudio del léxico de la casa en Aragón, Navarra y Rioja*, Zaragoza, DGA.
- CASTAÑER, R. y ENGUITA, J.M.<sup>a</sup> (1989). «Una década de estudios sobre el ALEANR», *Archivo de Filología Aragonesa*, 42-43, pp. 241-257.
- ENGUITA, J.M.<sup>a</sup> (1985). «Rasgos dialectales aragoneses en las hablas de Teruel», *Teruel*, 74, pp. 179-219.
- (1988). «Panorama lingüístico del Alto Aragón», *Archivo de Filología Aragonesa*, 41, pp. 175-191.
  - (1991a). «Modalidades lingüísticas del interior de Aragón», *Actas del I Congreso de Lingüistas Aragoneses*, Zaragoza, DGA, pp. 103-151.
  - (1991b). «Las hablas de Zaragoza», en J.M. ENGUITA (ed.), *I Curso de Geografía Lingüística de Aragón*, Zaragoza, IFC, pp. 201-239.
- ENGUITA, J.M.<sup>a</sup> (ed.) (1999). *Jornadas de Filología Aragonesa en el L Aniversario del AFA*, 2 vols., Zaragoza, IFC.
- FORT, M.<sup>a</sup> R. (1991). «Hablas orientales», en J.M. ENGUITA (ed.), *I Curso de Geografía Lingüística de Aragón*, Zaragoza, IFC, pp. 185-199.

- GARCÍA MOUTON, P. (1988), «Sobre la mujer en la encuesta dialectal», *Revista de Dialectología y Tradiciones Populares*, 43, pp. 291-299.
- (1991), «Dialectometría y léxico en Huesca», en J.M. ENGUITA (ed.), *I Curso de Geografía Lingüística de Aragón*, Zaragoza, IFC, pp. 311-326.
  - (1999), *Cómo hablan las mujeres*, Madrid, Arco/Libros.
- GOEBL, H. (1976), «La dialectométrie appliquée à l'ALF (Normandie)», en A. VARVARO (ed.), *Atti del XIV Congresso Internazionale di Lingüística e Filología Romanza*, II, Amsterdam, pp. 165-195.
- (1981), «Éléments d'analyse dialectométrique (avec application à l'AIS)», *Revue de Linguistique Romane*, 45, pp. 349-420.
  - (1983), «Parquet polygonal et treillis triangulaires: les deux versants de la dialectométrie interponctuelle», *Revue de Linguistique Romane*, 47, pp. 353-412.
  - (1987), «Points chauds de l'analyse dialectométrique: pondération et visualisation», *Revue de Linguistique Romane*, 51, pp. 63-118.
- GUITER, H. (1973), «Atlas et frontières linguistiques», en G. STRAKA y P. GARDETTE (eds.), *Les dialectes de France à la lumière des atlas régionaux (Colloque de Strasbourg, 1971)*, Paris, pp. 61-109.
- (1983), «Aproximació lingüística a la cadena cantabro-pirenenc», *Miscel.lania Aramon i Serra*, III, Barcelona, pp. 230-254.
  - (1985), «Les méthodes quantitatives en géolinguistique sont-elles équivalentes?», *Actes del XVI Congreso Internacional de Lingüística y Filología Románicas*, II, Barcelona, Ed. Moll, pp. 355-367.
  - (1987), «Étalonnage d'une méthode géolinguistique», *Revue de Linguistique Romane*, 51, pp. 55-62.
- LLORENTE, A. (1991), «Fronteras lingüísticas internas en territorio aragonés», en J.M. ENGUITA (ed.), *I Curso de Geografía Lingüística de Aragón*, Zaragoza, IFC, pp. 165-184.
- LOZANO, I. (1995), *Lenguaje femenino, lenguaje masculino. ¿Condiciona nuestro sexo la forma de hablar?*, Madrid, Minerva Ediciones.
- MORENO, F. (1991), «Morfología en el ALEANR: aproximación dialectométrica», en J.M. ENGUITA (ed.), *I Curso de Geografía Lingüística de Aragón*, Zaragoza, IFC, pp. 289-309.
- QUINTANA, A. (1976-80), «El parlar de La Codonyera. Resultats d'unes enquestes», *Estudis Romànics*, 17, pp. 1-253.
- SARAMAGO, J. (1986), «Differentiation lexique (un essai dialectométrique appliqué aux matériaux portugais de l'ALE)», *Géolinguistique*, 2, pp. 1-31.
- SÉGUY, J. (1971), «La relation entre la distance spatiale et la distance lexicale», *Revue de Linguistique Romane*, 35, pp. 335-357.
- (1973), «La dialectométrie dans l'ALG», *Revue de Linguistique Romane*, 37, pp. 1-24.
- VERLINDE, S. (1988), «La dialectométrie et la détection des zones dialectales: l'architecture dialectale de l'Est de la Belgique romane», *Revue de Linguistique Romane*, 51, pp. 151-172.
- VIERECK, W. (1988), «The Computerisation and Quantification of Linguistic Data: Dialectometrical Methods», en A.R. THOMAS (ed.), *Methods in Dialectology*, Clevedon, Phi., Multilingual Matters, pp. 406-427.

Recibido el 25 de abril de 2001

Aceptado el 19 de junio de 2001

