

LA DIALECTOMETRITZACIÓ DELS QUATRE PRIMERS VOLUMS DE L'ALDC: UNA BREU PRESENTACIÓ

Hans GOEBL
Universitat de Salzburg

1. INTRODUCCIÓ

L'alta qualitat geolingüística de les dades de l'ALDC era, per a un dialectòmetre com jo, des de l'inici, un repte excepcional. La raó n'és que des de l'aparició del primer volum de l'ALDC tenia la convicció que els diversos volums de l'ALDC contindrien les famoses «estructures (dialectomètriques) de profunditat» amb una precisió i claredat comparables a les de l'ALF, de l' AIS o de l'ALPI, dels quals ja existeixen dialectometrizations amb resultats molt satisfactoris.¹

En efecte, ja havia expressat aquesta convicció en ocasió de la presentació del primer volum de l'ALDC el 21 març de 2002 al mateix Institut d'Estudis Catalans:

Les dades del primer volum podrien proporcionar excel·lents resultats: aquesta és la meva ferma convicció després d'haver consultat les dades brutes dels cent setanta-nou mapes integrals i dels dotze mapes d'escassa variació interna. Es podria avançar en la dialectometrization al ritme de l'aparició dels altres volums de l'ALDC i també considerar una comparació dialectomètrica paral·lela de les estructures de profunditat de la part catalana de l'ALPI, d'una banda, i de l'ALDC, de l'altra (2002b: 22).

Per això estic molt content que pugui presentar, en aquest article, la confirmació de la meua convicció i de les meves expectatives de 2002.

Un procés qualsevol de dialectometrization (segons els costums i les tradicions de la «Escola Dialectomètrica de Salzburg» [EDMS])² pressuposa, normalment, la disponibilitat de tres condicions:

- a) l'existència d'un equip de col·laboradors experimentats;
- b) d'un exemplar de l'atles en qüestió;

1. Per a l'ALF, cf. Goebel (1976, 1984, 2002a i 2003); per a l' AIS, cf. Goebel (1981, 1984, 2008a); per a l'ALPI, cf. Goebel (2013, en curs de publicació).

2. La primera menció i explicació de la sigla EDMS és del 2008; cf. Goebel (2008b: 35-37).

- c) de subsidis suplementaris com, per exemple, de llistes amb el contingut dels mapes de l'atles corresponent.

Pel que fa a la primera de les tres condicions esmentades, vaig tenir la bona fortuna, entre els anys 2006 i 2009, de comptar amb l'assistència de dues cooperadores molt valuoses —Emese Lörincz i Julia Hattinger—, que tenien experiència dialectomètrica prèvia amb el tractament de dades italianes, angleses, americanes i ibèriques.

Les dues altres condicions ens han estat subministrades gràcies a la generositat de Joan Veny i de l'Institut d'Estudis Catalans. Amb aquesta última organització vaig poder firmar, l'any 2007, un conveni que m'ha facilitat l'accés a llistes (al format Word) equivalents al contingut dels mapes dels volums I-IV de l'*ALDC*. Cal afegir que per als treballs de dialectometrització ha estat molt útil la disponibilitat paral·lela de mapes i de llistes corresponents.

En un mot, hi va haver, doncs, entre 2007 i 2009, les condicions ideals per a emprendre la dialectometrització regular dels volums llavors existents de l'*ALDC*.

2. BREU RECAPITULACIÓ DELS OBJECTIUS I MÈTODES DE LA DIALECTOMETRIA SEGONS ELS PRINCIPIS DE L'EDMS

L'objectiu principal de la DM³ és el descobriment, per via d'*inducció* (o de *síntesi*) de moltes dades *particulars*, d'estructures amagades en la profunditat de les dades de qualsevol atlas lingüístic de tipus tradicional.⁴ Les estructures de profunditat tenen un caràcter quantitatiu (i també visual) molt estable i ocupen, des del punt de vista epistemològic, el nivell *general*. La regularitat —sovint tan sorprenent com estètica— dels mapes dialectomètrics revela, a més a més, l'existència d'un factor lingüístic molt important, que és la capacitat dels locutors de gestionar l'espai que habiten mitjançant l'establiment d'un teixit molt puntualitzat de relacions de *similituds* i de *diferències* basilectals. La «gestió basilectal de l'espai per l'home» representa, doncs, el teorema de base de l'EDMS,⁵ mentre que la dialectometrització d'un nombre tan gran com sigui possible de diversos atlas lingüístics constitueix la seva empresa empírica principal.

3. Recordem que la gènesi de la DM (com a cosa i com a mot) es remunta a l'inici dels anys 70 de l'últim segle: cf. Ségy (1971 i 1973) (amb la primera menció del terme «dialectométrie»), com també Goebel/Winterleitner (1971).

4. Entenc per atlas (lingüístic) «tradicional» un atlas de tipus «monodimensional», és a dir, que conté només dades exclusivament basilectals. Aquesta selecció garanteix la bona intercomparabilitat de les dades i constitueix la clau de l'èxit *heurístic* de qualsevol anàlisi quantitativa de les mateixes dades.

5. Per als fonaments teòrics i pràctics de la DM cf. Goebel (1984) (en alemany); per aplicacions concretes dels mètodes de l'EDMS, em remeto als meus articles de presentació en diverses llengües romàniques apareguts des del 1976: cf. la Bibliografia a l'annex i el lloc web següent: https://www.sbg.ac.at/rom/people/prof/goebl/dm_publi.htm.

Un altre pilar molt important de l'EDMS consisteix en l'ús generalitzat de visualitzacions de diverses maneres (des de *mapes de similitud* fins a *dendrogrames* proveïts de representacions de l'espai apropiades). És que estem convençuts que el recurs a procediments de visualització (de tipus *quantitatiu*) representa una continuació directa de l'ús de mapes tipificats (de tipus *qualitatiu*), com era habitual en els temps de Jules Gilliéron i dels altres pioners de la geografia lingüística tradicional.

Cal afegir encara que la geografia lingüística de tipus gilliéronià ha rebut la seva raó de ser més important a partir de la presentació de les dades recollides sobre el terreny, no en taules o en llistes, sinó sobre mapes de gran format, i per tant en l'ús programàtic de la visualització com a mitjà heurístic.

Aquest fet explica per què l'EDMS es considera un element integral de la tradició de la geografia lingüística clàssica, amb l'única diferència que fa servir exclusivament mètodes *quantitatius*, de tipus *numèric*, d'una banda, i de tipus *icònic*, d'altra. En aquest sentit em remeto a la figura 1 que mostra l'arranjament dels mètodes-DM de l'EDMS.

Cal afegir encara que la realització de totes les etapes representades en la figura 1 s'ha confiat, a Salzburg, al programa informàtic VDM («Visual Dialectometry»), que va ser creat, els anys 1996-1999, pel meu amic Edgar Haimerl (ara domiciliat a Seattle, EUA).

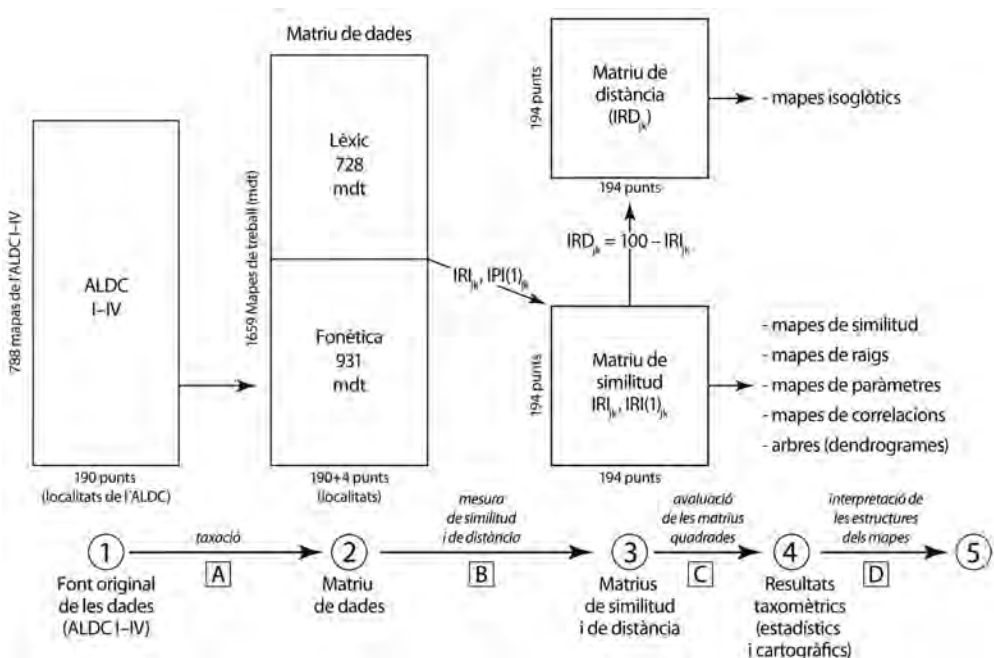


Figura 1: Cadena dels mètodes utilitzats per a l'«Escola Dialectomètrica de Salzburg».

3. DES DE L'ALDC A LA MATRIU DE DADES

Com es desprèn de la figura 1, la primera etapa (A) que condueix de la font original (ALDC) a la matriu de dades s'anomena «taxació». Es tracta de no res més que de la ben coneguda tipificació del contingut dels mapes de qualsevol atlas lingüístic, com havia estat realitzada moltes vegades pels grans mestres de la geografia lingüística tradicional.⁶ Cito, a propòsit, dues obres mestres de data recent, relatives, d'una banda, als mapes de l'atles francès *ALF*: Brun-Trigaud / Le Berre / Le Dù (2005), i, d'una altra, als de l'atles català *ALDC*: *PALDC* de Joan Veny (2008 i ss).

En ambdós casos, es tracta de col·leccions de mapes tipificats segons criteris lingüístics diversos, des de la *fonètica* fins al *lèxic* i a la *morfosintaxi*, els quals permeten de demostrar amb claredat el «suc» dels mapes originals en qüestió.

És clar que aquest treball de «taxació» només el pot efectuar un lingüista competent. Atesa la disponibilitat actual dels tres volums del *PALDC*, és fàcil oferir una idea precisa del treball de taxació dut a terme a Salzburg.

Els nostres mapes 1 i 2 corresponen, respectivament, als mapes 4 («timbre de la vocal del morfema dels femenins plurals») i 49 («el queixal del seny») del *PALDC* I. L'única diferència consisteix en el detall més gran de les nostres anàlisis, on es tractava abans que res d'aplegar tota la riquesa variacional dels mapes originals de l'*ALDC* sense pensar en simplificacions de matís didàctic.

El mapa 1, de caràcter fonètic, mostra *tres* tipus lingüístics o «taxats», que corresponen a altres tants succedanis catalans del nexa llatí -AS (en *AD+RUGAS). Es tracta, doncs, d'un mapa *trí*-nim. A causa de l'absència de l'ètim llatí *AD+RUGAS al Rosselló i a l'Alguer,⁷ aquesta zona figura al mapa 1 en blanc.

La configuració espacial del mapa 2, de caràcter *lèxic*, és molt més complexa. Es tracta d'una taxació *hendecà*-nim; és a dir, amb *onze* taxats. Joan Veny, en confeir el mapa 49 del *PALDC* I, s'ha acontentat, al contrari, amb una taxació només *pentà*-nima.

Tres dels nostres taxats excedents es refereixen als tipus francès (*dent de sagesse*), italià (*dente de giudizio*) i castellà (*muela del juicio*), que no figuren a la llegenda del mapa 49 del *PALDC* I. Els tres tipus restants (*queixal de abellanero*, *queixal del senderi* i *queixal extremer*) han estat incorporats per Joan Veny respectivament a les àrees dels tipus *queixal del judici*, *queixal del seny* i *extremer*, mentre que representen taxats particulars en la taxació feta a Salzburg.

Afegim encara el detall més aviat tècnic que durant el procés de taxació vam considerar sempre una sola resposta per punt d'enquesta. Quan, al mapa de l'*ALDC*, hi havia dues o tres respostes equivalents, ha estat presa en consideració només la primera de les respectives respostes múltiples.

Hi ha, doncs, una diferència capital entre els mapes *originals* i els mapes que se'n derivaren mitjançant el procés de *taxació*. Aquests mapes secundaris s'anomenen «mapes de treball». En general, un mapa original de caràcter *lèxic* genera un sol mapa de

6. Cf. p. e. Jaberg (1908) i Gilliéron (1918).

7. El Rosselló té els tipus *rides* i *rufes*, mentre que a l'Alguer es troba el tipus *frunza*.

treball, mentre que d'un mapa original de caràcter fonètic es poden derivar uns quants mapes de treball.

La taula presentada a sota⁸ mostra el balanç de les nostres taxacions per als corpus següents: corpus fonètic, corpus lèxic i corpus sencer. És clar que vam considerar encara unes altres categories lingüístiques, a la presentació i a la discussió de les quals, però, hem de renunciar aquí per manca d'espai.

ALDC	Fonètica			Lèxic			Corpus Sencer		
	Mapes originals considerats	Mapes de treball	Taxats/ Àrees taxadores	Mapes originals considerats	Mapes de treball	Taxats/ Àrees taxadores	Mapes originals considerats	Mapes de treball	Taxats/ Àrees taxadores
I	61	250	1.144	164	164	2.278	182	414	3.422
II	44	179	845	219	219	3.163	228	398	4.008
III	104	430	1.979	234	235	3.588	265	665	5.567
IV	18	72	373	109	110	1.974	114	182	2.347
Σ	227	931	4.341	726	728	11.003	789	1659	15.344

Cal precisar que el nombre dels taxats derivats per via de la taxació es correspon perfectament amb el nombre de les àrees taxadores respectives d'acord amb el principi segons el qual cada taxat disposa de la seva àrea geogràfica. La conseqüència és que ens trobem sempre davant d'una imbricació molt complexa d'un caramull d'àrees taxadores d'extensió i de configuració molt variables. És precisament de l'anàlisi d'aquesta situació complexa que depèn la bona comprensió de la formació de «dialectes», entesos com a entitats tipològicament superiors⁹.

Observeu, a propòsit, que cada àrea taxadora pot variar no tan sols segons la seva extensió geogràfica (entre 1 i N punts), sinó també segons la seva estructura espacial. Aquesta última pot oscil·lar entre una coherència espacial perfecta dels x polígons de l'àrea corresponent i llur disseminació (quasi-)aleatòria a l'espai.

En la perspectiva del vector d'atributs d'una sola localitat d'atles hi ha, doncs, una coexistència molt complexa d'àrees taxadores de *petita (micro-coremàtica)*, *mitjana*

8. Vam prendre en consideració dels mapes originals de l'ALDC i de les «l·listes des respostes amb escassa variació formal» només els mapes i l·listes amb un nombre molt reduït de dades absents (llindar màxim per l'acceptació de dades mancants: ca. 5 % de 190 (punts d'enquesta). Vet aquí el balanç per als quatre volums de l'ALDC (mapes + l·listes: nombre existent/nombre pres en consideració: vol. I: 191/182; vol. II: 276/228, vol. III: 334/265, vol. IV: 202/114). Cal precisar que a l'àmbit de càlculs taxomètrics cal sempre controlar amb rigorositat el nombre de dades absents perquè provoquen distorsions perilloses en els resultats dialectomètrics finals.

9. Cf. el llibre magistral, però malauradament poc conegut, de Karl von Etmayer (1924), on es troba el primer intent de comprensió del procés de formació tipològica dels dialectes.

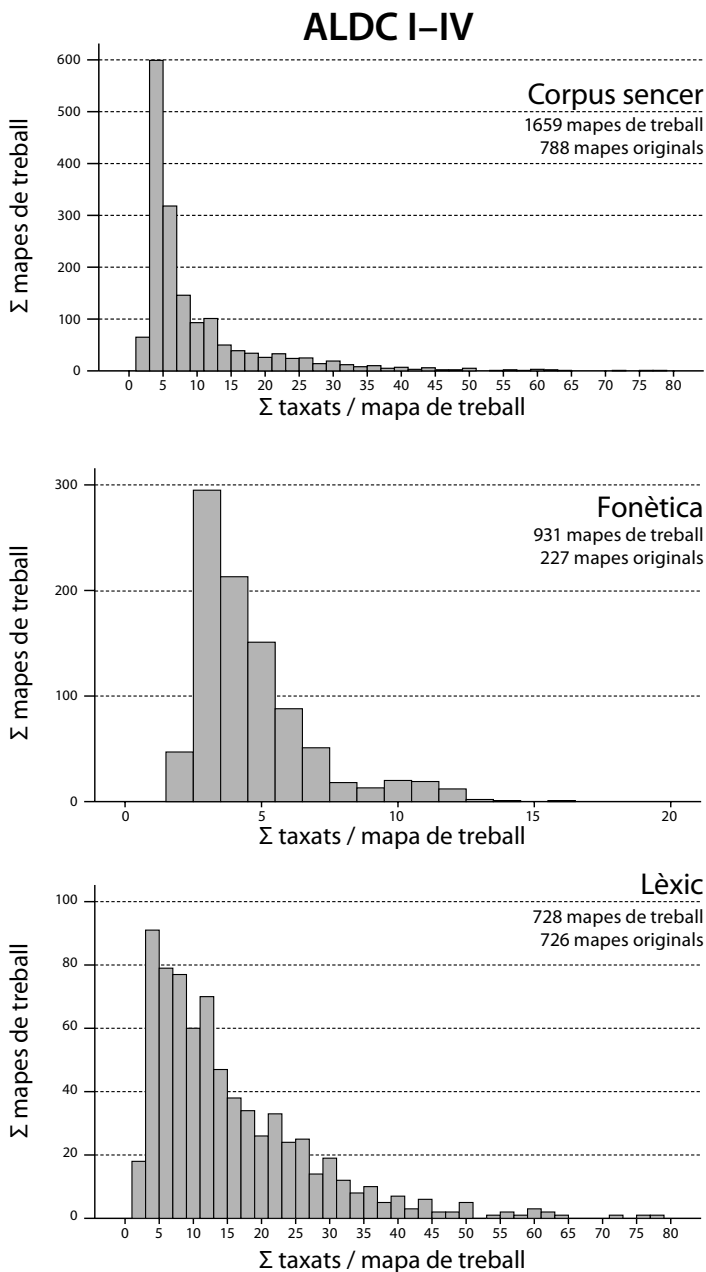


Figura 2: Estructures quantitatives dels mapes de treball analitzats durant la dialectometrització dels volums I-IV de l'ALDC: relacions entre el nombre i la polinòmia intrínseca dels mapes de treball per a tres corpus: corpus sencer, corpus fonètic i corpus lèxic.

(*meso-coremàtica*) i *gran (mega-coremàtica)* extensió. Atès que les diverses extensions de les àrees depenen sempre de múltiples factors extralingüístics i intralingüístics, és clar que la consideració sistemàtica dels percentatges de les àrees micro-, meso- i mega-coremàtiques pot ser molt rellevant a causa de la posició interaccional de la localitat en qüestió.

Una de les troballes completament inesperades de l'EDMS era la descoberta de l'ordre inherent als resultats dels treballs de taxació: vegeu per a aquesta finalitat la figura 2, de la qual es desprenen tres corbes (de tipus exponencial) molt regulars. L'ur missatge quantitatiu és doble: a) que la variació taxadora s'enceta amb el valor 2, i b) que s'acaba amb valors sovint molt alts: per a la fonètica: 16; per al lèxic: 78 taxats/mapa de treball.

Aquesta situació correspon perfectament a les nostres experiències prèvies amb la taxació de les dades de l'ALF, de l' AIS i de l'ALPI (i d'alguns altres atlas no-romànics).¹⁰

Cal afegir encara algunes observacions relatives a la confecció de la xarxa de la nostra recerca. Renuncio a fer un esment particular de la poligonització, que ha esdevingut una cosa absolutament normal en les recerques geolingüístiques. Recordo que per a la tipificació dels mapes de l'ALDC ja existeix un mapa poligonitzat de la xarxa d'aquest atlas.¹¹

El nombre dels punts originals de l'ALDC s'ha incrementat, però, amb quatre punts artificials, que es corresponen amb les llengües estàndard del català (P. 999), del francès (P. 997), de l'italià (P. 996) i del castellà (P. 998). La utilitat de llur inserció és molt gran perquè es pot visualitzar així el grau de penetració tipològica d'aquestes llengües estàndard en la totalitat dels dialectes de la xarxa de l'ALDC; vegeu a propòsit els mapes 7 i 8.

4. DE LA MATRIU DE DADES A LES MATRIUS DE SIMILITUD I DE DISTÀNCIA

Es tracta de l'etapa B en la figura 1. La mesura de les similituds o distàncies quantitatives entre els vectors diversos (de localitat) de la matriu de dades condueix a una transformació del caràcter fenomenològic de les dades de partida, des del *qualitatiu* al *quantitatiu*, i representa, a més d'això, la clau metòdica per a obtenir la síntesi de la variació interna d'un gran aplec de dades qualitatives; vegeu la figura 3.

L'índex de similitud estàndard de l'EDMS és molt simple: es tracta de l'IRI (*Índex Relatiu d'Identitat*, alem. *Relativer Identitätswert*, RIW), que mesura la relació, per a un parell de localitats d'atles, entre les sumes dels taxats «co-idèntics» (COI) i «co-diferents» (COD):

$$IRI_{jk} = \sum COI / \sum COI + \sum COD.$$

10. Cf. la nota 1.

11. Aquest mapa poligonitzat representa un present fet per la Romanística de Salzburg a l'ALDC l'any 2002.

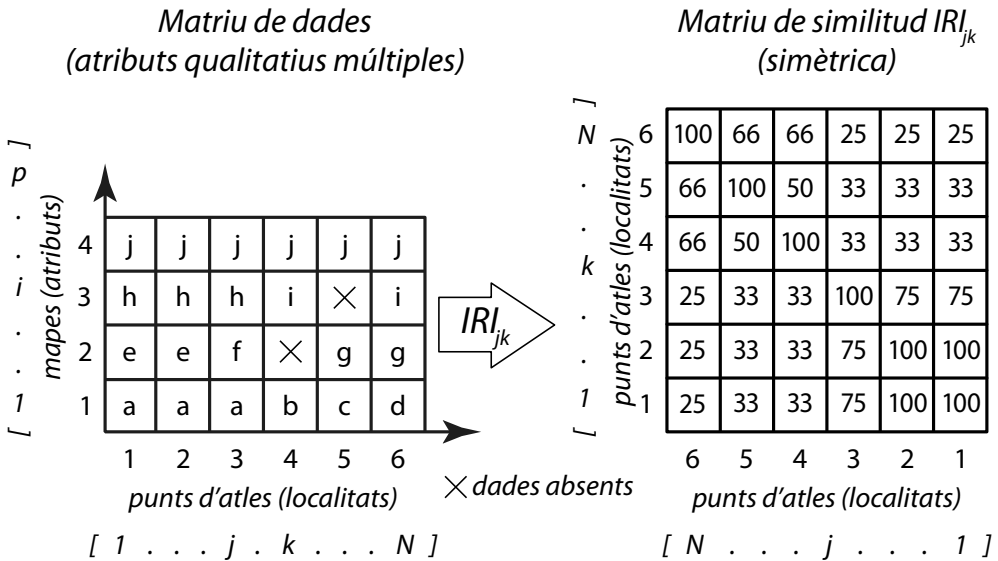


Figura 3: Esquema de la mesura de similitud entre dues localitats (j i k) segons l' IRI_{jk} (Índex Relatiu d'Identitat).

Aplicant aquesta fórmula a les localitats 3 i 4 de la matriu de dades de la figura 3, obtenim el resultat següent:

$$IRI_{3,4} = 1 / 1 + 2 = 33\%.$$

El valor de 33 % es troba a dins de la matriu de similitud als dos costats de la diagonal, que conté només valors (simètrics) de 100 %.

Com es pot desprendre de la figura 1, la conversió matemàtica dels valors de *similitud* (s) en valors de *distància* (d) és fàcil: cal aplicar només la fórmula següent:

$$s + d = 1 \text{ (o } 100\%).$$

Fora de l' IRI_{jk} , el bagatge metodològic de l'EDMS disposa encara d'altres índexs de similitud la utilitat classificadora dels quals és molt variable. Com que tots els mapes d'aquest article es basen exclusivament en l' IRI_{jk} , renunciem aquí a la presentació d'altres índexs de similitud o de distància.

Fem ara una altre cop d'ull a la figura 1: es veu amb claredat que l'etapa C es correspon amb l'avaluació del contingut quantitatiu de les matrius (quadrades) de similitud i de distància, les quals, segons els principis de l'EDMC, s'han de fer visualment, és a dir, mitjançant visualitzacions apropiades. Aquesta pràctica visual és normal en moltes ciències naturals, socials i humanes (angl. «imaging»; alem. «bildgebende Verfahren»).

És, però, molt important la bona selecció de les tècniques de visualització, per a les quals les necessitats i expectatives de la geografia lingüística tradicional ens han servit de fil conductor.

Novament, la figura 1 ofereix les informacions necessàries: el programa VDM permet de produir els esquemes visuals següents: mapes de similitud, mapes interpuntuals (= mapes isoglòtics i mapes de raigs), mapes de paràmetres, mapes de correlacions i dendrogrames. Com que cadascun dels esquemes esmentats es pot visualitzar de diverses maneres, la riquesa de les visualitzacions possibles és efectivament considerable.¹² El programa VDM permet de navegar-hi amb facilitat.

El rendiment final de tot el procediment-DM consisteix en la interpretació de les estructures dels mapes: vegeu, en la figura 1, l'espai entre els punts 4 i 5, que es correspon amb l'etapa D.

5. BREU COP D'ULL A ALGUNS MAPES DE SIMILITUD

Vegeu els mapes 3-8.

El mapa de similitud representa el pilar heurístic central de l'EDMS.

Des del punt de vista *estadístic*, es tracta de la visualització —en la forma d'un mapa coroplètic— d'un dels N vectors de la matriu de similitud.

Des del punt de vista *geolingüístic*, cadascun dels N mapes de similitud calculats assenyala, d'una banda, la posició relacional del potencial lingüístic d'un punt d'enquesta respecte a tots els altres, i prova, d'una altra, la gran regularitat, amb la qual les similituds geolingüístiques es despleguen en l'espai.

Des de la perspectiva *visual*, cal afegir encara dos detalls de caràcter cartogràfic:

- a) Cada mapa de similitud té un punt de referència el polígon del qual roman sempre en blanc.
- b) La visualització dels N-1 valors de similitud es fa mitjançant un algorisme ben definit, que el dialectòmetre ha d'escollir segons les seves finalitats científiques i segons la seva disposició òptica i cognitiva.

Un compromís sovint provat entre aquestes dues necessitats és l'algorisme MINMW-MAX, que compta amb sis graons cromàtics (MINMWMAX 6-tuple). Aquest algorisme agrupa els N-1 valors per sota i per sobre de la respectiva mitjana aritmètica. Els llinars dels tres intervals inferiors (intervals 1-3) es calculen a través d'una tripartició de la dife-

12. Afegeixo només un exemple per a un sol mapa de tipus coroplètic: la configuració de qualsevol mapa depèn —en prescindir del corpus pròpiament dit— dels tres factors següents (entre parèntesis es troben les possibilitats ofertes per VDM): de l'índex de similitud (ca. 20), de l'algorisme d'intervalització (3) i del nombre de graons cromàtics (10): 20 per 3 per 10 = 600 estructures (sovint molt variables).

rència entre el valor mínim (al mapa 3: 28,02) i la mitjana aritmètica (al mapa 3: 69,17). La mateixa operació per als valors superiors (interval 4-6) es basa en la diferència entre els valors màxims (al mapa 3: 89,83) i —novament— la mitjana aritmètica. L'algorisme MINMWMAX —l'expressivitat visual del qual es manifesta de la millor manera possible mitjançant sis graus— permet, a més a més, la comparació visual d'un gran nombre de mapes de similitud.

Esmento entre parèntesis que el nostre programa VDM permet també de fer altres visualitzacions, sigui aplicant altres algorismes d'intervalització, sigui aplicant un altre nombre de graus cromàtics.¹³ Amb VDM, els respectius canvis es poden efectuar en fraccions de segons.

Afegim encara una observació de caràcter cartogràfic: cada mapa es compon de tres parts: a) del cartograma policrom pròpiament dit, b) de la llegenda numèrica corresponent i c) de l'histograma de la respectiva distribució de freqüència. El cartograma i l'histograma són, doncs, dues formes diverses de visualització del mateix potencial quantitatiu.

5.1. PRESENTACIÓ I INTERPRETACIÓ DEL MAPA 3, RELATIU AL P.-ALDC 62 (BARCELONA)

Vegeu el mapa 3.

L'estructura del perfil coroplètic del mapa revela la posició geo-relacional del dialecte de Barcelona respecte als altres de la xarxa de l'ALDC i als quatre punts artificials. Les similituds més grans (en vermell) cobreixen els territoris de la Catalunya Vella; la zona vermella és envoltada, a l'oest, primer per alguns polígons en taronja, i, després, per una faixa compacta en groc.

Recordem que els polígons que tenen colors *calents* visualitzen els valors-IRI situats per *sobre* de la mitjana aritmètica. No sorprèn que el valor màxim (marcat per un ratllat blanc) es trobi al costat del punt de referència. La zona verda representa un llarg cinturó que va des del Rosselló al nord fins al punt més meridional de la xarxa. No meravella ni tan sols que, entre les llengües estàndard afegides com a punts artificials, el català sigui molt similar al potencial lingüístic del barceloní.

És, al contrari, digne de menció que les altres tres llengües estàndard ocupin posicions molt allunyades respecte al punt de referència 62.

L'estructura general del perfil coroplètic del mapa revela, a més, que existeix una relació sistemàtica entre l'augment de la distància quilomètrica i la disminució de la similitud lingüística respecte al punt de referència. Farem servir aquesta relació més endavant per al càlcul i la confecció dels mapes de correlacions.

13. Cf. la nota precedent.

5.2. PRESENTACIÓ I INTERPRETACIÓ DEL MAPA 4, RELATIU AL P.-ALDC 163 (VALÈNCIA)

Vegeu el mapa 4.

De l'observació de la silueta espacial del mapa 3 ja es desprenen algunes previsions relatives al mapa 4: a) que la zona vermella es concentrarà a l'entorn de València, b) que els polígons amb colors calents cobriran el territori del català occidental, c) que el valor màxim de la distribució de similitud (metafòricament parlant: «el millor amic») es trobarà a prop del punt de referència (aquí: del P.-ALDC 164, Alfafar).

No era previsible, però, que el valor mínim (metafòricament parlant: «el pitjor enemic») romanguí en la mateixa posició (= P. 997, francès estàndard). En resum, es pot constatar que els perfils coroplètics dels mapes 3 i 4 són complementaris l'un de l'altre. Cal observar encara la diferència entre les siluetes dels dos histogrames, que reflecteix divergències estadístiques considerables existents entre les dues distribucions de similitud. Utilitzarem més endavant aquest fet per al càlcul i la confecció de mapes de paràmetres.

5.3. PRESENTACIÓ I INTERPRETACIÓ DEL MAPA 5, RELATIU AL P.-ALDC 86 (ESTERRI D'ÀNEU)

Vegeu el mapa 5.

El mapa 5 mostra un perfil típicament nord-occidental: els polígons vermells i de color de taronja cobreixen la majoria del domini del català nord-occidental, mentre que els polígons grocs ocupen, d'una banda, la Catalunya Vella, i efectuen, d'una altra, la conjunció amb el valencià. Cal observar també que el valor mínim de similitud es troba sempre en el lloc del francès estàndard i que el castellà i l'italià estàndard continuen tenint una posició relacional molt modesta respecte al potencial lingüístic del punt de referència.

5.4. PRESENTACIÓ I INTERPRETACIÓ DEL MAPA 6, RELATIU AL P.-ALDC 79 (PALMA)

Vegeu el mapa 6.

El mapa 6 reflecteix, d'una banda, els lligams lingüístics i històrics ben coneguts que existeixen entre les Balears i la Catalunya Vella, i, d'una altra, la sòlida cohesió lingüística interna de les tres illes balears representades a la xarxa de l'ALDC. Cal observar, a més, que la configuració de la zona verda s'assembla molt a la del mapa 3 i que la posició relacional del dialecte de l'Alguer continua sent molt modesta (interval/classe 2). Aquest fet és degut a la pregona assimilació lingüística de l'alguerès a l'italià i al sard.

5.5. PRESENTACIÓ I INTERPRETACIÓ DEL MAPA 7, RELATIU AL P.-ALDC ARTIFICIAL 999 (CATALÀ ESTÀNDAR)

Vegeu el mapa 7.

La visualització dels mapes 7 i 8 ha estat realitzada mitjançant vuit graons cromàtics, en lloc de sis, i aplicant un altre algorisme d'intervalització (MEDMW), que fa ressaltar millor els desnivells del perfil coroplètic. El propòsit central d'aquest mapa és de visualitzar el grau de penetració tipològica del català estàndard a la xarxa de l'ALDC. Es desprèn del mapa amb tota claredat que aquest fenomen es manifesta sobretot a la Catalunya Vella i a la majoria de les terres del català nord-occidental.

5.6. PRESENTACIÓ I INTERPRETACIÓ DEL MAPA 8, RELATIU AL P.-ALDC ARTIFICIAL 998 (CASTELLÀ ESTÀNDAR)

Les normes cartogràfiques i la finalitat heurística del mapa 8 concorden perfectament amb les del mapa 7. De la repartició dels polígons en vermell i en color taronja es desprenen amb molta claredat les zones de castellanització tipològica més gran.

Es dedueix igualment molt bé dels mapes 7 i 8 que l'efecte de correlació entre l'augment de la distància geogràfica i la disminució de la similitud lingüística val també per a la difusió espacial del català i del castellà, malgrat llur naturalesa de punts d'atles artificials.

6. PRESENTACIÓ DE LA DIALECTOMETRIA ISOGLÒTICA

Vegeu els mapes 9-10.

En l'àmbit de la DM isoglòtica, es tracta, d'una banda, de mesurar i de visualitzar no *similituds* sinó *distàncies* lingüístiques, i, d'una altra, d'efectuar les mesures de distància només en situacions de veïnatge immediat dels punts d'enquesta, és a dir, *entre dos punts veïns* o *contigus*. Aquesta condició explica l'altre qualificatiu d'aquest segment de la DM: *inter-puntual*.

El missatge icònic dels dos mapes es basa en la interacció òptica de 453 costats de polígon i de llur gruixària i coloració. Entre la gruixària i la coloració existeix una correlació molt simple: més gruixut = més fred (com a coloració) i més prim = més calent. És clar que els principis heurístics d'aquesta anàlisi-DM es corresponen perfectament amb els de la tècnica tradicional del dibuix de feixos d'isoglosses combinades.

El primer intent de produir amb mitjans-DM un mapa isoglòtic del domini català va ser publicat per Lluís Polanco Roig l'any 1984. El seu mapa es basava en una xarxa de 131 punts d'enquesta extrets dels 529 punts de l'ALPI, i de 108 mapes de caràcter sobretot lèxic, la majoria dels quals es trobaven als arxius personals de M. Sanchis Guarner (cf. Polanco Roig [1984: 22]).

Els nostres dos mapes tenen estructures molt clares: el mapa 9 és de caràcter *fonètic* i mostra l'estructuració isoglòtica del domini català sobre la base de 931 mapes de treball i de 4.341 àrees taxadores de tipus exclusivament *fonètic*.

El mapa 10, al contrari, es basa en 728 mapes de treball de caràcter *lèxic* i es refereix, per tant, exclusivament a l'estructuració isoglòtica del domini català a la llum de les fronteres exteriors de 11.003 àrees taxadores *lèxiques*.

Cal observar, al mapa 10, la desaparició total del feix isoglòtic que separa el català nord-occidental i el català oriental, la més gran compartimentació isoglòtica de la província de València i finalment l'aplanament isoglòtic interior de Mallorca.

De la consideració comparada dels mapes 9 i 10 es desprèn amb claredat que l'estructuració isoglòtica del domini català varia considerablement segons les categories de la *fonètica* i del *lèxic*.

7. PRESENTACIÓ D'UN MAPA DE PARÀMETRE: LA SINOPSI DELS «COEFICIENTS D'ASIMETRIA DE FISHER» (CAF)

Vegeu els mapes 11 i 12.

Com ja havia mencionat prèviament, és possible utilitzar els valors característics (o paràmetres) de les diverses distribucions de similitud per a unes quantes anàlisis-DM més avançades, que permeten d'explorar noves dimensions del funcionament del llenguatge a l'espai. Una de les millors solucions en aquest sentit és la sinopsi dels N valors del «coeficient d'asimetria de Fisher» (CAF) —anomenat també en català *biaix* (cf. alem. *Schiefe*, angl. *skewness*) que agafa l'asimetria més o menys gran d'una distribució de similitud.

Observem per a aquesta situació els histogrames dels mapes 3 (relatiu al dialecte de Barcelona) i 5 (relatiu al d'Esterrí d'Àneu). A primer cop d'ull, es veu que les siluetes dels histogrames són molt diferents i que tenen un altre grau de simetria. Mentre que l'histograma del mapa 3 té dos cims (i és, doncs, «bimodal»), la silueta de l'histograma del mapa 5 és molt regular i té només una sola culminació. En general, la simetria dels histogrames s'origina per una interacció complexa entre els diversos valors de la distribució de similitud i la mitjana aritmètica, per a la mesura de la qual es fa servir el paràmetre del CAF, que agafa, des del punt de vista lingüístic, el nivell de l'*abast comunicatiu* d'un dialecte.

Al mapa 3, relatiu a Barcelona, dels 193 valors-IRI visualitzats, 104 (= 53,8 %) es troben per sota, i 89 valors (= 46,2 %) per sobre de la mitjana aritmètica. Al mapa 5, relatiu a la localitat d'Esterrí d'Àneu, al contrari, aquesta relació és invertida: només 87 dels 193 valors-IRI es troben per sota la mitjana aritmètica, mentre que 106 (= 53,8 %) se situen per sobre d'aquest llindar. Això significa que l'*abast comunicatiu* del dialecte d'Esterrí d'Àneu (P. 86) és millor que el del dialecte de Barcelona (P. 62).

Una sinopsi cartografiada de N biaixos esdevé, doncs, una temptativa molt prometedora per a l'anàlisi de la repartició espacial dels abasts comunicatius de tots els dialectes de la xarxa de l'ALDC.

El mapa 11 mostra el resultat basat sobre els 931 mapes de treball de caràcter *fonètic*, mentre que el mapa 12 manifesta l'efecte de la mateixa operació-DM mitjançant els 728 mapes de treball de caràcter *lèxic*. La repartició espacial dels sis colors és molt clara i, a més, no està desproveïda d'una certa harmonia.

Cal, però, definir amb precisió el significat *lingüístic* dels colors que es corresponen amb les diverses mesures del biaix. El color vermell simbolitza dialectes tipològicament més aviat autònoms i dotats d'una certa estabilitat diacrònica, mentre que el color blau remet a un alt grau d'intercanvi respecte a la totalitat de la xarxa examinada. Vaig anomenar a aquesta capacitat, des de fa molts anys, *compromís lingüístic* (fr. *compromis linguistique*, alem. *Sprachausgleich*).

Les zones blaves dels mapes 11 i 12 representen, doncs, llocs on hi havia, ara i en els temps passats, intercanvis lingüístics molt intensos.¹⁴ És significatiu, entre altres qüestions, que l'illa lingüística de l'Alguer —exposada des d'alguns segles a influències sardes i italianes— tingui dues vegades el color blau fosc.

Cal notar, a més a més, que la posició geogràfica de les tires en blau fosc sobre els mapes 11 i 12, es correspon en una gran part a l'emplaçament dels feixos d'isoglosses més gruixuts dels mapes 9 i 10. Aquest fet sembla paradoxal només a primer cop d'ull; en una darrera anàlisi prova que els fenòmens de la *demarcació* i de la *interacció* lingüístiques són, de fet, dues cares de la mateixa moneda.

8. DOS COPS D'ULL A LA DIALECTOMETRIA CORRELATIVA

Vegeu els mapes 13-14 i la figura 4.

La DM correlativa representa una de les eines més eficaces de l'EDMS. Des del punt de vista estadístic, es basa en la comparació quantitativa —mitjançant l'aplicació del «coeficient de correlació de [Auguste] Bravais i [Karl] Pearson» ([r(BP)])— dels N vectors de dues matrius de similituds de dimensions estretament idèntiques: vegeu la part central de la figura 4.

El resultat d'aquesta comparació estadística de N parells de vectors és sempre N valors de correlació, que seran cartografiats després de la manera habitual: vegeu els mapes 13 i 14 que, novament, forneixen dos perfils coroplètics clarament configurats.

14. Davant dels resultats de les meves mesures dialectomètriques, tinc la impressió que el domini del català nord-occidental sencer ha estat una zona molt més proveïda de contactes i de conflictes lingüístics de tota mena que totes les altres subagrupacions de la resta del domini català. Em sembla que algunes de les reflexions diacròniques d'A. M. Badia i Margarit exposades al seu llibre del 1981 sobre la formació de la llengua catalana (p. 140 s. i p. 174 s.) van en la mateixa direcció. Cf., a més, Veny / Massanell i Messalles (2002, passim); Veny (1989: 120 s.); Veny (1991, passim), així com Lüdtke (1991, passim).

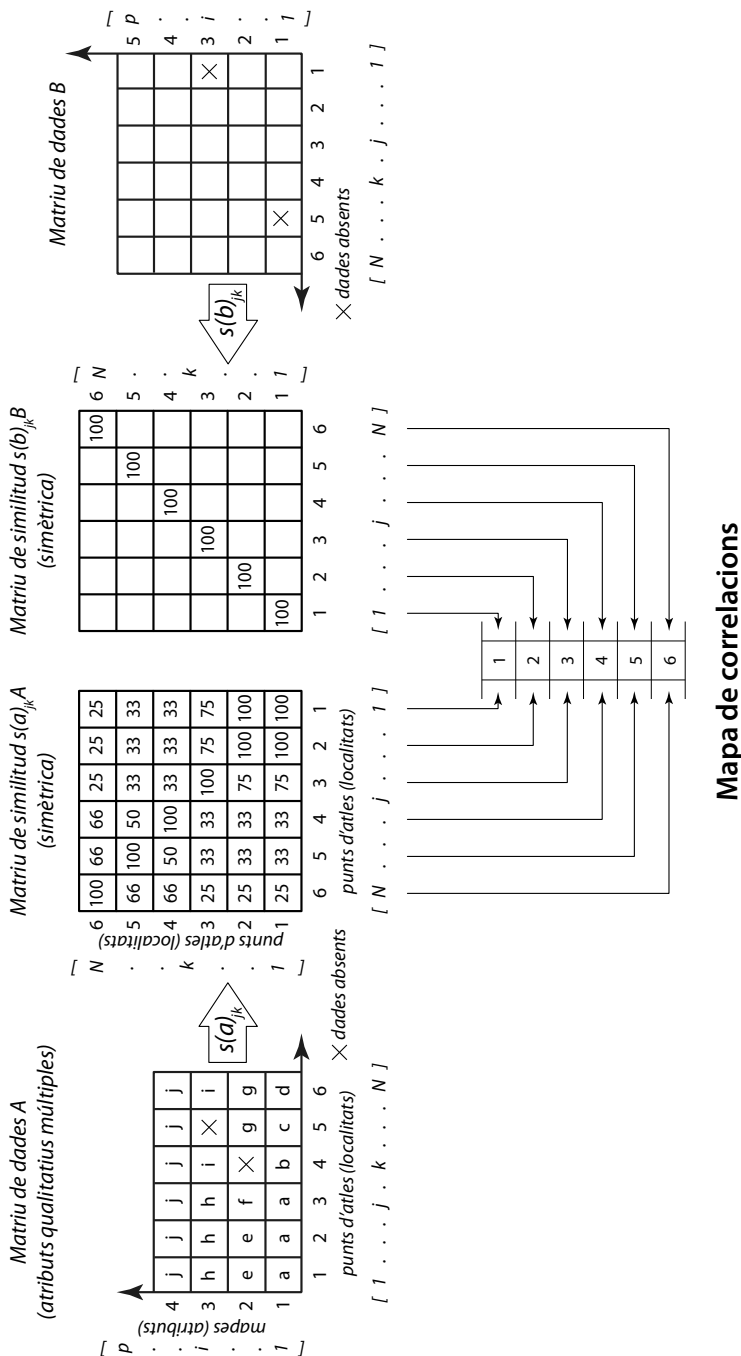


Figura 4: Esquema del càlcul dels mapes de correlacions.

8.1. LA CORRELACIÓ ENTRE LES GESTIONS LINGÜÍSTICA I EUCLIDIANA DE L'ESPAI

Vegeu el mapa 13.

Les variables relacionades són, d'una banda, les similituds *geolingüístiques* (segons l'IRI_{jk}) i, d'una altra, les proximitats *geogràfiques* (calculades segons el teorema ben conegut de Pitàgores). Al mapa 13, els valors rellevats del r(BP) oscil·len entre -0,25 i +0,89, mentre que la desviació teòrica del r(BP) se situa entre -1 i +1.

Quina és, doncs, la significació *lingüística* d'aquests *valors de correlació*, o, millor, dels *colors corresponents*?

El color vermell —equivalent a graus alts del r(BP)— simbolitza una harmonia intrínseca entre el desplegament de les «forces» *lingüístiques* i *euclidianes* en l'espai, mentre que el color blau fosc representa el contrari: fortes friccions, conflictes i divergències entre aquestes dues dimensions dinàmiques. Una tal situació es deu remuntar a contactes i conflictes lingüístics intensos escaiguts en un passat més o menys allunyat. És molt probable que la dinàmica expansiva ocorreguda durant la Reconquesta en l'àmbit del regne d'Aragó hagi estat el factor històric dotat de la més alta repercussió que hagi generat conflicte.¹⁵

Les zones vermelles, al contrari, simbolitzen un desenvolupament molt menys agitat, més harmònic i, sobretot, més conforme a la lògica euclidiana del canvi lingüístic ocorregut en el passat.

A la llum d'aquestes definicions, la zona del català nord-occidental —amb inclusió del francès, de l'italià i dels dialectes de l'Alguer i d'Eivissa— sembla haver estat el teatre de desenvolupaments trasbalsadors, mentre que les zones del català oriental, de les illes Balears (sense Eivissa) i del valencià semblen haver gaudit d'una història (geo)lingüística més tranquil·la.

8.2. LA CORRELACIÓ ENTRE DUES CATEGORIES LINGÜÍSTIQUES DIFERENTS: FONÈTICA I LÈXIC

Vegeu el mapa 14.

Una altra aplicació fascinant de la DM correlativa és la correlació espacial entre dues categories lingüístiques diverses. Aquest procediment dona una resposta geoquantitativa precisa a moltes preguntes corresponents que s'han fet més d'una vegada en l'àmbit de la geografia lingüística tradicional.¹⁶ El mapa 14 mostra la consonància i la dissonància espacial de la *fonètica* (representada per 931 mapes de treball i 4.341 taxats o àrees taxadores) amb el *lèxic* (representat per 728 mapes de treball i 11.003 taxats o àrees taxadores).

Novament, la zona del català nord-occidental es perfila com a teatre de les disharmònies més grans entre les gestions *fonètica* i *lèxica* de l'espai per part dels locutors autòc-

15. Cf. la nota precedent.

16. Em remeto de nou a l'obra cabdal de Karl von Etmayer de 1924, on el paper jugat per les diverses categories lingüístiques en la formació dels dialectes és tractat de manera pionera.

tons o, en altres termes, com a una veritable zona pivot del domini català sencer. En mesura menor, l'italià estàndard i els dialectes de l'Alguer, de Menorca i d'Eivissa participen d'aquesta conflictivitat.

Per raons de manca d'espai, no podem presentar aquí altres correlacions categorials¹⁷ que són igualment molt interessants i que demostren abans que res l'alta complexitat, proveïda, no gensmenys, d'una magnífica lògica espacial, de la gestió humana de l'espai segons les diverses categories (intra)lingüístiques.

9. PRESENTACIÓ DE LA DIALECTOMETRIA DENDROGRÀFICA

Vegeu els mapes 15 i 16.

La DM dendrogràfica constitueix una eina analítica amb un altíssim valor heurístic. Del punt de vista estadístic es tracta d'aplicar a les dades de la matriu de similitud un algorisme d'agrupació que crea estructures ramificades amb una jerarquització sempre binària, les quals es corresponen molt de prop amb el nostre vell concepte metafòric de l'«arbre genealògic».

En l'àmbit de la taxonomia numèrica, aquests algorismes pertanyen a la «classificació jeràrquica ascendent» (CJA). Un dels algorismes més coneguts de la CJA és el mètode de [Joe] Ward [Jr.], que permet de crear agrupacions (*clusters* o *dendremes*) ramificades, les quals es corresponen, després de la transposició de llur contingut a l'espai (és a dir, a la respectiva xarxa d'investigació), amb agrupacions espacials (o *coremes*) molt coherents¹⁸.

Els dos arbres dels mapes 15 i 16 han estat calculats mitjançant l'algorisme de [Joe] Ward [Jr.] Els càlculs respectius es corresponen amb aglomeracions (o fusions) binàries que enceten al parell de valors-IRI més semblant de la matriu de similitud i que avancen, mitjançant N-1 etapes aglomeradores, fins a l'«arrel» (o fins al «tronc») de l'arbre. El criteri de semblança que s'ha aplicat depèn directament de l'algorisme en qüestió. L'aglomeració s'efectua per tant des de les «fulles» fins a l'«arrel» de l'arbre: vegeu les fletxes a l'esquerra dels arbres dels mapes 15 i 16, que indiquen el sentit de l'aglomeració.

És clar que la informació classificadora de l'arbre disposa de més d'un nivell analític. La bifurcació més baixa —és a dir, més veïna a l'arrel— produeix a l'espai en ambdós casos una bipartició molt clara del domini català: al mapa 15, basat sobre el nostre corpus *fonètic* (amb 931 mapes de treball), el conjunt dels dendremes 1-4 ocupa la zona del català occidental (amb inclusió del castellà, de l'italià, del francès i de l'alguerès), mentre que els dendremes 5-8 es concentren a l'oest de la xarxa de l'ALDC.

El programa VDM permet de canviar lliurement la coloració dels dendremes i dels coremes a qualsevol nivell de l'arbre tot descomponent-lo, sigui en *superfície*, sigui en *profunditat*.

17. Cf. Goebel (2005), on aquesta branca de la DM es va presentar per primer cop.

18. Hom troba una descripció detallada de la DM dendrogràfica a Goebel (2003: 84 s.).

Hi ha també una altra possibilitat d'interpretació dels arbres-CJA calculats que és de caràcter diacrònic. Aquesta possibilitat parteix de la hipòtesi —sostinguda abans que enlloc en l'àmbit de la lexicoestadística i de la glotocronologia¹⁹— segons la qual la ramificació creixent de l'arbre reflecteix el desenvolupament fragmentador del geotipus lingüístic analitzat, del qual s'admet que posseeix d'antuvi una coherència perfecta, que, en transcórrer el temps, aniria fragmentant-se progressivament: vegeu, per a aquest propòsit, les fletxes a la dreta dels arbres dels mapes 15 i 16 amb la menció «fragmentació diacrònica».

La representació espacial del dendrograma del mapa 16 mostra una visió bifurcadora una mica diferent del territori català, que es correspon àmpliament amb el missatge geolingüístic del mapa 10 (síntesi isoglòtica). És que el conjunt dels dendremes 5-8 del mapa 16 ocupa una superfície més àmplia que el bloc dels dendremes 5-8 del mapa 15.

Cal observar també la diversa afiliació del francès, de l'italià i de l'alguerès, que, sobre el mapa 16, pertanyen tots al dendrema-corema 8 i, per tant, a la meitat oriental del domini català.

Cal subratllar, però, que les solucions aquí presentades no són ni exclusives ni definitives. Com que hi ha encara altres algorismes-CJA molt útils per a aplicacions-DM, és imprescindible que el dialectòmetre treballi amb una pluralitat de solucions i amb la consciència que diverses finalitats classificadores sempre es corresponen amb altres mitjans d'agrupació. En aquest context, el programa VDM permet, d'altra banda, l'aplicació ràpida i alternativa de sis algorismes-CJA diferents.

10. CONCLUSIÓ

Els resultats de la dialectometrització de les dades dels quatre primers volums de l'*ALDC* s'arreglaren perfectament amb els dels nostres treballs anteriors relatius als atles romànics *ALF*, *AIS* i *ALPI*, i demostren amb claredat l'existència —darrere les estructures sovint «caòtiques» de mapes de treball aïllats— d'estructures de profunditat molt regulars, espacialment ben ordenades i proveïdes, a més a més, d'una estabilitat considerable. Aquesta última particularitat permet la predicció que l'anàlisi-DM dels volums restants de l'*ALDC* encara no publicats produirà resultats molt semblants.

11. AGRAÏMENTS

Sense la cooperació i l'ajut de moltes persones i institucions, ni la dialectometrització de l'*ALDC* (vols. I-IV) ni la redacció del present article no haurien pogut ser dutes a terme.

19. Per a més informacions, cf. Goebel (2003: 86).

Taxació dels volums I-IV de l'ALDC: Emese Lörincz i Julia Hattinger (Salzburg).

Finançament: projecte 13 365 del «Fonds zur Förderung der wissenschaftlichen Forschung in Österreich» (FWF), Wien (*Fundació austríaca de la recerca científica*, Viena).

Conveni sobre la utilització de llistes relatives a l'ALDC: Institut d'Estudis Catalans (IEC) (signat el 7 de març de 2007).

Manutenció i aplicació del programa VDM: Edgar Haimerl (Seattle, EUA) i Pavel Smečka (Salzburg).

Producció dels gràfics del present article: Werner Goebel (Viena).

Correcció estilística: Maria Pilar Perea (Barcelona).

Que totes les persones citades i els organismes FWF i IEC rebin aquí l'expressió de la meua sincera gratitud.

BIBLIOGRAFIA I ABREVIATURES

AIS: JABERG, Karl / JUD, Jakob (1928-1940): *Sprach- und Sachatlas Italiens und der Südschweiz*. 8 vols. Zofingen: Ringier (reimpresió: Nendeln: Kraus 1971).

ALDC: VENY, Joan / PONS I GRIERA, Lúdia (2001-2010): *Atlas Lingüístic del Domini Català*. 5 vols. Barcelona: Institut d'Estudis Catalans.

ALF: GILLIÉRON, Jules / EDMONT, Edmond (1902-1910): *Atlas linguistique de la France*. 10 vols. Paris: Champion (reimpresió: Bologna: Forni 1968).

ALPI: (1962): *Atlas lingüístico de la Península Ibérica*. Vol. I: Fonética. Madrid: CSIC.

BADIA I MARGARIT, Antoni M. (1981): *La formació de la llengua catalana. Assaig d'interpretació històrica*. Barcelona: Publicacions de l'Abadia de Montserrat.

BRUN-TRIGAUD, Guylaine / LE BERRE, Yves / LE DÛ, Jean (2005): *Lectures de l'Atlas linguistique de la France de Gilliéron et Edmont. Du temps dans l'espace. Essai d'interprétation des cartes de l'Atlas linguistique de la France de Jules Gilliéron et Edmond Edmont augmenté de quelques cartes de l'Atlas linguistique de la Basse Bretagne de Pierre Le Roux*. Paris: CTHS.

CAF: coeficient d'asimetria de [Ronald A.] Fisher.

CJA: classificació jeràrquica ascendent (classe particular d'algorismes dendrogràfics).

COD: co-diferència (diferència de dos taxats comparats en la mateixa línia de la matriu de dades).

COI: co-identitat (identitat de dos taxats comparats en la mateixa línia de la matriu de dades).

DM: dialectometria, relatiu a la dialectometria, dialectomètric.

- EDMS: Escola Dialectomètrica de Salzburg.
- ETTMAYER, Karl von (1924): «Das Wesen der Dialektbildung, erläutert an den Dialekten Frankreichs». *Denkschriften [Memories] der Akademie der Wissenschaften in Wien*, philosophisch-historische Klasse, 66/3, p. 1-56, 7 taules.
- GILLIÉRON, Jules (1918): *Généalogie des mots qui désignent l'abeille d'après l'Atlas linguistique de la France*. Paris: Champion.
- GOEBL, Hans (1976): «La dialectométrie appliquée à l'ALF (Normandie)». Vàrvaro, Alberto (ed.): *Atti del XVI Congresso Internazionale di Linguistica e Filologia Romanza* (Nàpols 1974). Nàpols: Marchiaroli, Amsterdam: Benjamins, vol. 2, p. 165-195.
- GOEBL, Hans (1981): «Eléments d'analyse dialectométrique (avec application à l'ALS)». *RLiR*. Vol. 45, p. 349-420.
- GOEBL, Hans (1983): «Parquet polygonal et treillis triangulaire. Les deux versants de la dialectométrie interponctuelle». *RLiR*. Vol. 47, p. 353-412.
- GOEBL, Hans (1984): *Dialektometrische Studien. Anhand italomannischer, rätoromanischer und galloromanischer Sprachmaterialien aus ALS und ALF*. 3 vols. Tübingen: Niemeyer.
- GOEBL, Hans (2002a): «Analyse dialectométrique des structures de profondeur de l'ALF». *RLiR*. Vol. 66, p. 5-63.
- GOEBL, Hans (2002b): «Presentació del primer volum de l'Atles lingüístic del domini català». *Publicacions de la Presidència de l'Institut d'Estudis Catalans*. Vol. 15, p. 11-23.
- GOEBL, Hans (2003): «Regards dialectométriques sur les données de l'Atlas linguistique de la France (ALF): relations quantitatives et structures de profondeur». *Estudis Romànics*. Vol. 25, p. 59-96.
- GOEBL, Hans (2005): «La dialectométrie corrélative: un nouvel outil pour l'étude de l'aménagement dialectal de l'espace par l'homme». *RLiR*. Vol. 69, p. 321-367.
- GOEBL, Hans (2008a): «La dialettometrizzazione integrale dell'ALS. Presentazione dei primi risultati». *RLiR*. Vol. 72, p. 25-113.
- GOEBL, Hans (2008b): «Le laboratoire de dialectométrie de l'Université de Salzburg. Un bref rapport de recherche». *ZFSL*. Vol. 118, p. 35-55.
- GOEBL, Hans (2013): «La dialectometrization del ALPI: rápida presentació de los resultados», dins CASANOVA HERRERO, Emili / CALVO RIGUAL, Cesáreo (ed.) *Actes del XXVI^e Congrés Internacional de lingüística i Filologia Romàniques* (València 2010). Berlin; Boston: Walter de Gruyter, VI p. 143-154.
- GOEBL, Hans / WINTERLEITNER, Norbert (1971): «Projekt einer sprachstatistischen Auswertung von in Sprachatlanten gespeicherter linguistischer Information mit Hilfe elektronischer Rechenanlagen». *Linguistische Berichte*. Vol. 14, p. 60-61.
- IRD_{jk}: Índex Relatiu de Distància (entre les localitats *j* i *k*).
- IRI_{jk}: Índex Relatiu d'Identitat (entre les localitats *j* i *k*).
- JABERG, Karl (1908): *Sprachegeographie. Beitrag zum Verständnis des Atlas linguistique de la France*. Aarau: Sauerländer (versió castellana: *Geografía lingüística. Ensayo de interpretación del «Atlas lingüístico de Francia»*). Traducció d'A. Llorente i M. Alvar. Granada: Universidad de Granada 1959).

- LRL (1988-2005): HOLTUS, Günter / METZELTIN, Michael / SCHMITT, Christian (eds.): *Lexikon der Romanistischen Linguistik*. 8 vols. Tübingen: Niemeyer.
- LÜDTKE, Jens (1991): «Katalanisch: Externe Sprachgeschichte. Histoire externe de la langue». *LRL*. Vol. V/2, p. 232-242.
- MEDMW: algorisme d'intervalització (basat en la mitjana aritmètica, amb la formació de n intervals als dos costats de la mitjana aritmètica, la dimensió numèrica [en nombres de punts d'enquesta] dels quals és igual).
- MINMWMAX: algorisme d'intervalització (basat en la mitjana aritmètica, amb la formació de n intervals als dos costats de la mitjana aritmètica, la desviació numèrica [relativa als llindars inferiors i superiors de cada interval] dels quals és igual).
- PALDC: VENY, Joan (2008-2011): *Petit Atlas Lingüístic del Domini Català*. 3 vols. Barcelona: Institut d'Estudis Catalans.
- POLANCO ROIG, Lluís B. (1984): «Llengua o dialecte: solucions teòriques i aplicació al cas català». *Sociolinguistique des langues romanes. Actes du XVII^{ème} Congrès International de Linguistique et Philologie Romanes* (Aix-en-Provence 1983)». Aix-en-Provence: Université de Provence, Marseille: Laffitte, vol. 5, p. 13-31.
- r(BP) : coeficient de correlació de [Auguste] Bravais i [Karl] Pearson.
- SÉGUY, Jean (1971): «La relation entre la distance spatiale et la distance lexicale». *RLiR*. Vol. 35, p. 335-357.
- SÉGUY, Jean (1973): «La dialectométrie dans l'Atlas linguistique de la Gascogne». *RLiR*. Vol. 37, p. 1-24.
- VDM: Visual DialectoMetry (programa dialectomètric utilitzat a Salzburg; lliurement disponible sota petició).
- VENY, Joan (1989⁸): *Els parlars catalans*. Palma de Mallorca: Editorial Moll.
- VENY, Joan (1991): «Katalanisch: Areallinguistik. Áreas lingüísticas». *LRL*. Vol. V/2, p. 243-261
- VENY, Joan / MASSANELL I MESSALLES, Mar (2002): *Dialectes occidentals. Català nord-occidental i valencià*. Barcelona: Universitat Oberta de Catalunya.

RESUM

El present article representa un primer balanç de la dialectometrització del volums I-IV de l'*Atlas Lingüístic del Domini Català* (ALDC) segons els principis teòrics i pràctics de l'«Escola Dialectomètrica de Salzburg» (EDMS). Després d'una descripció de la «taxació» (és a dir, de la mesura taxadora) de les dades originals de l'ALDC, de l'establiment de la matriu de dades i de la matriu de similituds, segueixen la presentació i la interpretació detallades de 16 gràfics policroms generats mitjançant el programa dialectomètric «Visual DialectoMetry» (VDM). Els 16 gràfics esmentats contenen, majoritàriament sota la forma cartogràfica de mapes coroplètics, dos mapes de treball, sis mapes de similitud, dos mapes isoglòtics, dos mapes de paràmetres, dos mapes de correlació i dues anàlisis dendrogràfiques amb representació espacial. La naturalesa empírica i la impor-

tància científica dels resultats s'arregleren perfectament a les de totes les altres dialectometritzacions fetes d'atles lingüístics romànics i no romànics sota els auspicis de l'EDMS i confirma l'alta qualitat geolingüística de les dades brutes de l'*ALDC*.

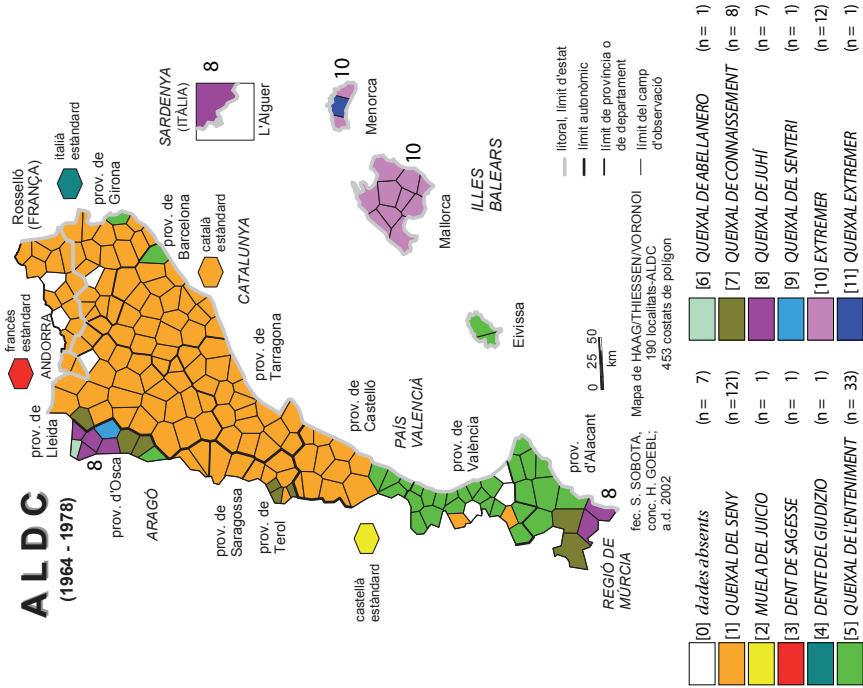
MOTS CLAU: dialectologia, dialectometria, geografia lingüística, gestió basilectal de l'espai per l'home, classificació numèrica.

ABSTRACT

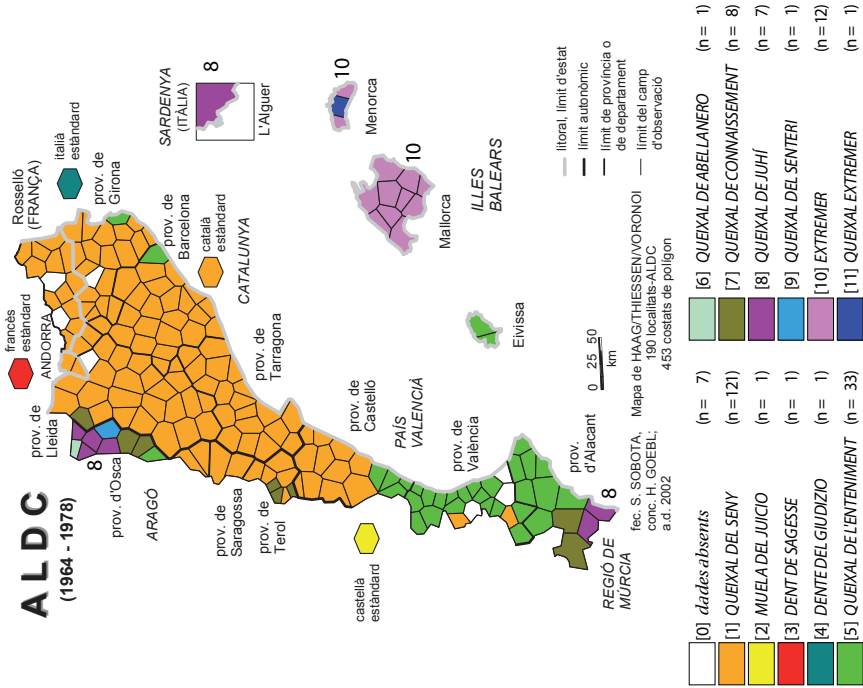
The dialectometrization of the first four volumes of the *ALDC*: a brief introduction

The present article constitutes a preliminary report on the dialectometrization of the first four volumes of the Catalan linguistic atlas *ALDC* («Atlas Lingüístic del Domini Català») according to the theoretical and practical principles of the «Salzburg School of Dialectometry». After a brief description of the techniques used to evaluate the original data in the *ALDC* and to create the tables for the data and the measurements showing similarity between paired character-vectors in the data, there is an introduction followed by 16 colour maps produced by the dialectometric programme «Visual DialectoMetry» (VDM). These 16 graphics, mostly in the form of choropleth maps, contain two working maps, six similarity maps, two isoglossic maps, two parameter maps, two correlation maps and two dendrograms with their spatial equivalents («spatialization maps»). The empirical nature and scientific importance of the results are perfectly in line with all the other dialectometrizations of Romance and non-Romance linguistic atlases carried out under the auspices of the EDMS, and confirm the high geolinguistic quality of the raw data from the *ALDC*.

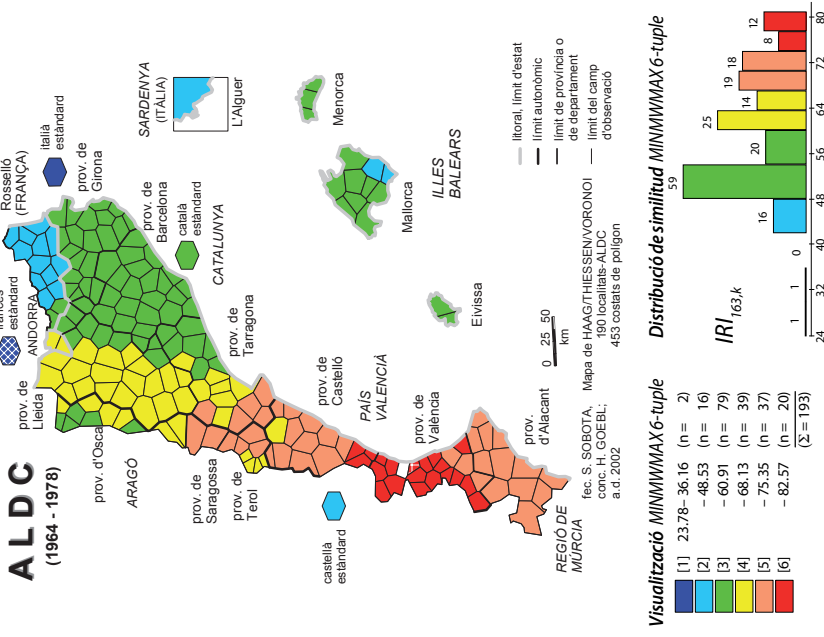
KEY WORDS: dialectology, dialectometry, linguistic geography, basilectal management of space by man, numerical classification.



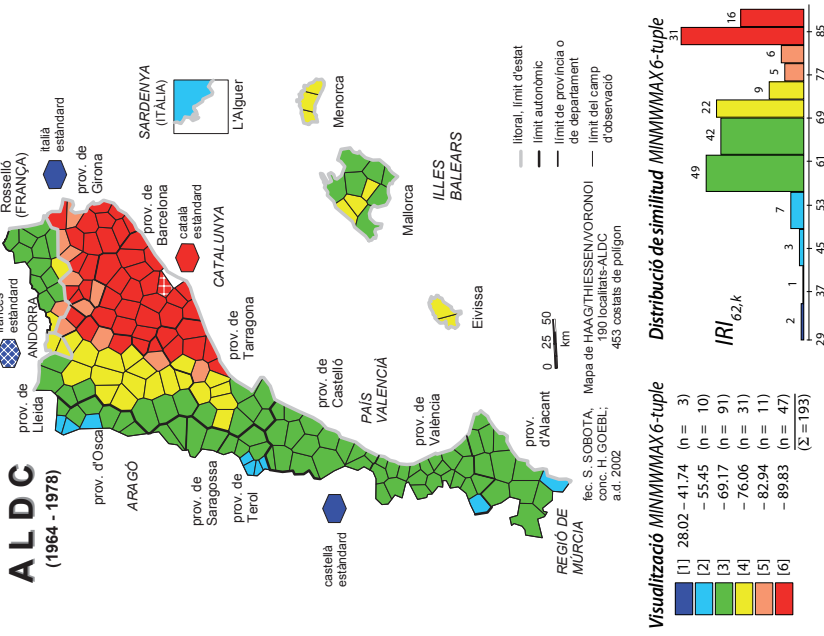
Mapa 1: Mapa de treball de caràcter fonètic: timbre de la vocal del morfema llatí –AS en *AD+RUGAS (cf. també ALDC I, 8 i PALDC I, 4).



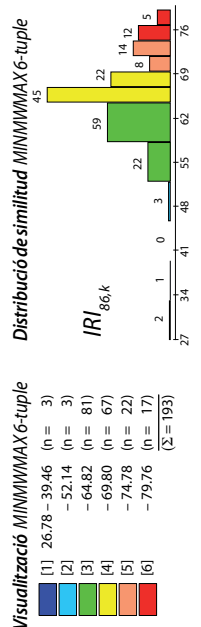
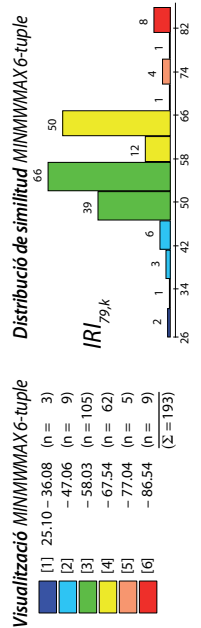
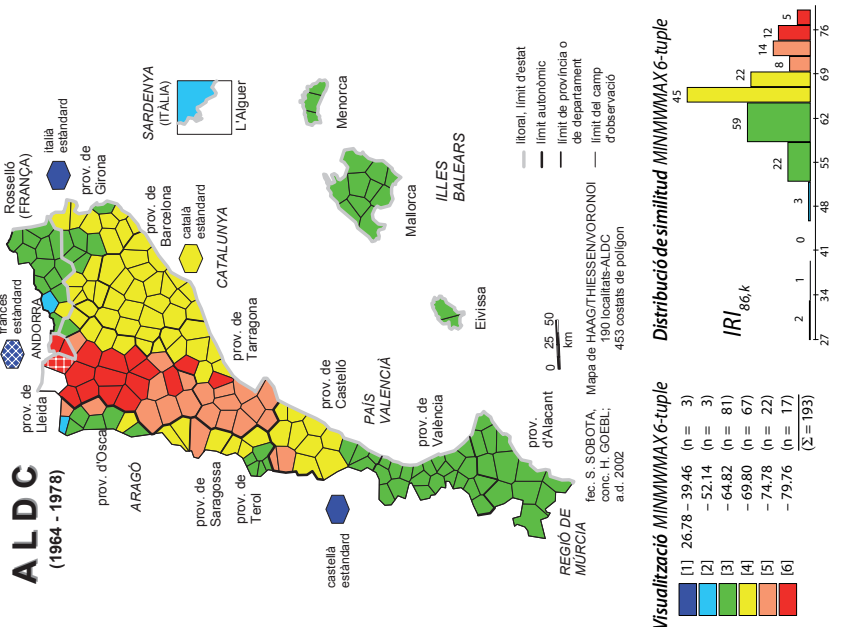
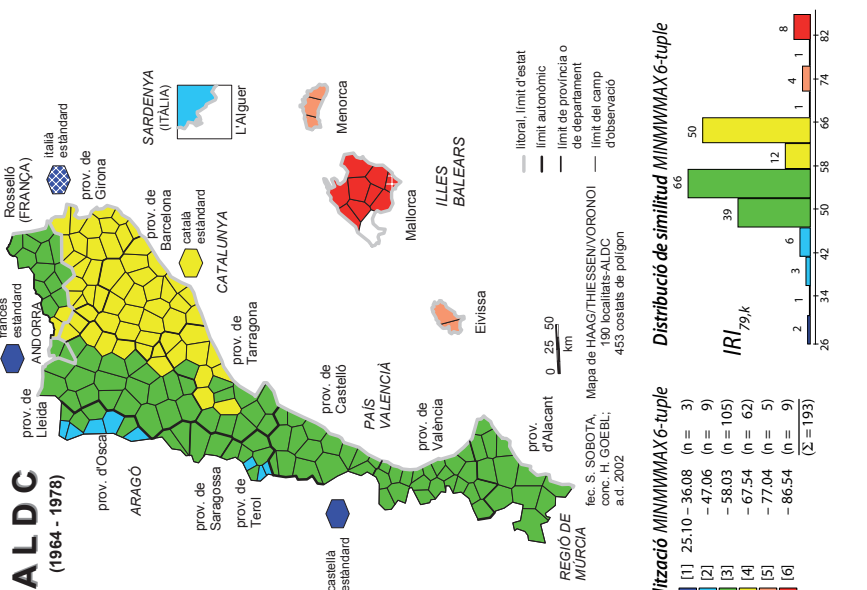
Mapa 2: Mapa de treball de caràcter lèxic: les denominacions catalanes del *queixal del seny*. (cf. també ALDC I, 52 i PALDC I, 49).



Mapa 4: Mapa de similitud relatiu al punt 163 de l'ALDC: València.
Corpus: 1.659 mapes de treball (corpus sencer).
Algorisme d'intervalització: MINMMWMAX 6-tuple.

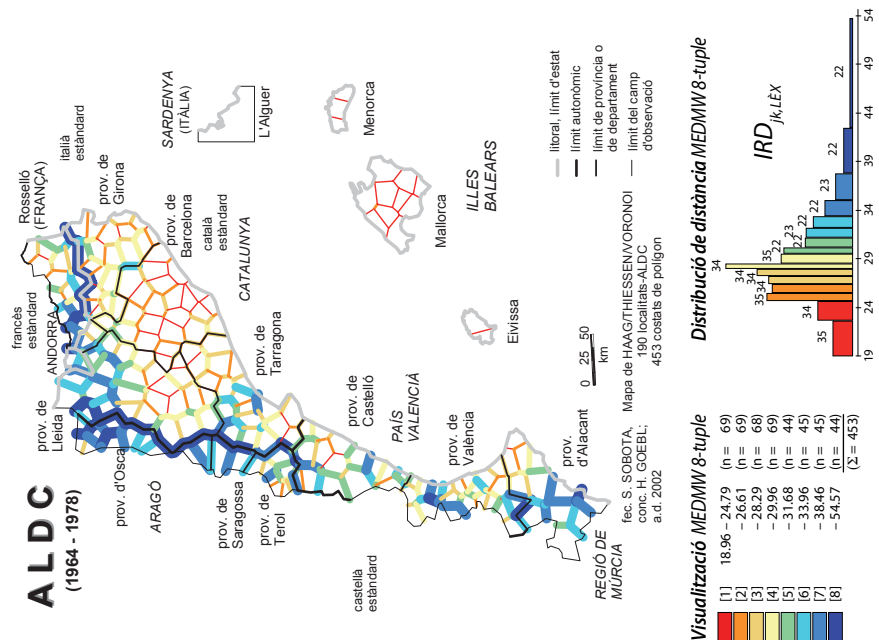


Mapa 3: Mapa de similitud relatiu al punt 62 de l'ALDC: Barcelona.
Corpus: 1.659 mapes de treball (corpus sencer).
Algorisme d'intervalització: MINMMWMAX 6-tuple.

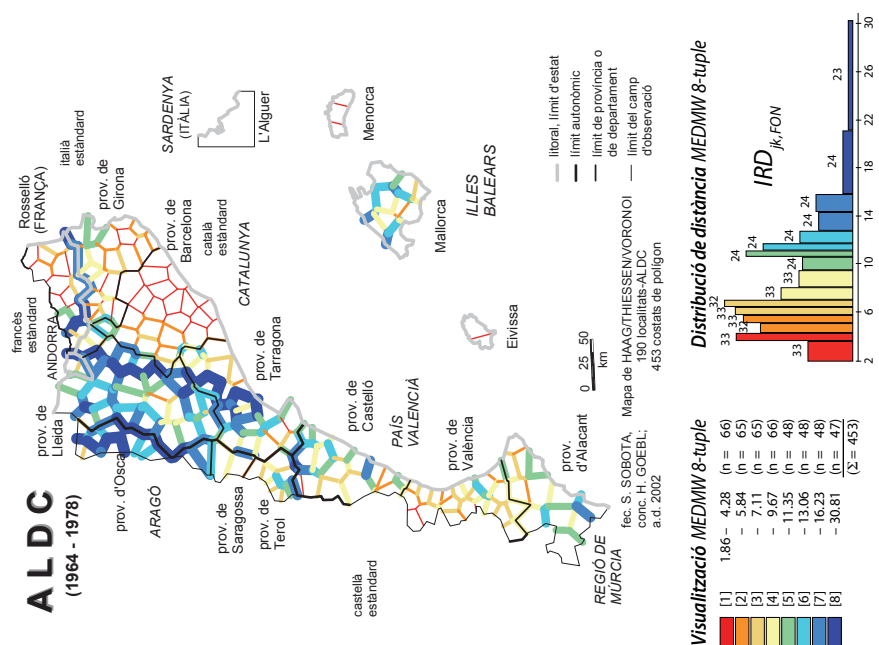


Mapa 6: Mapa de similitud relatiu al punt 79 de l'ALDC; Palma de Mallorca.
 Corpus: 1.659 mapes de treball (corpus sencer).
 Algorisme d'interval·lització: MINMMWMAX 6-tuple.

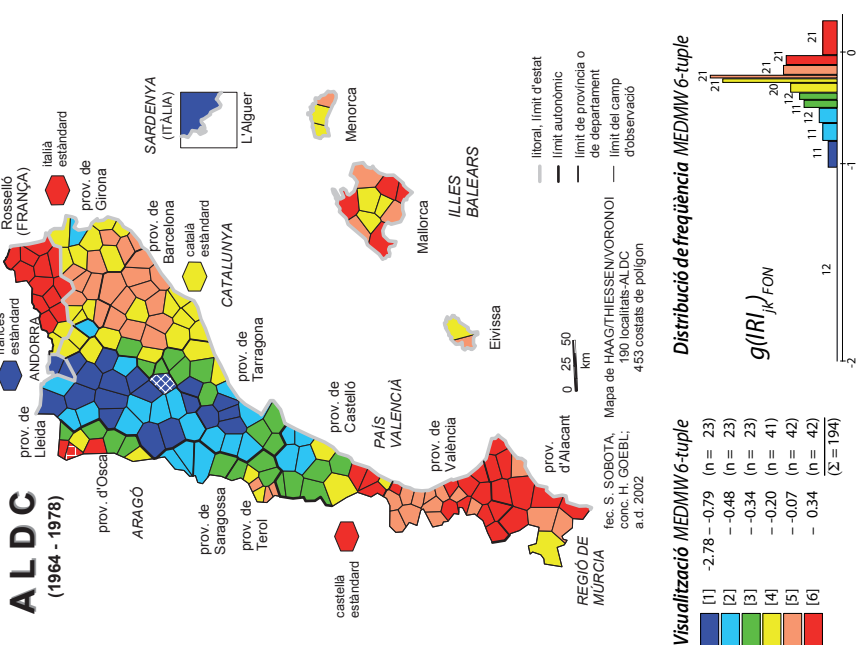
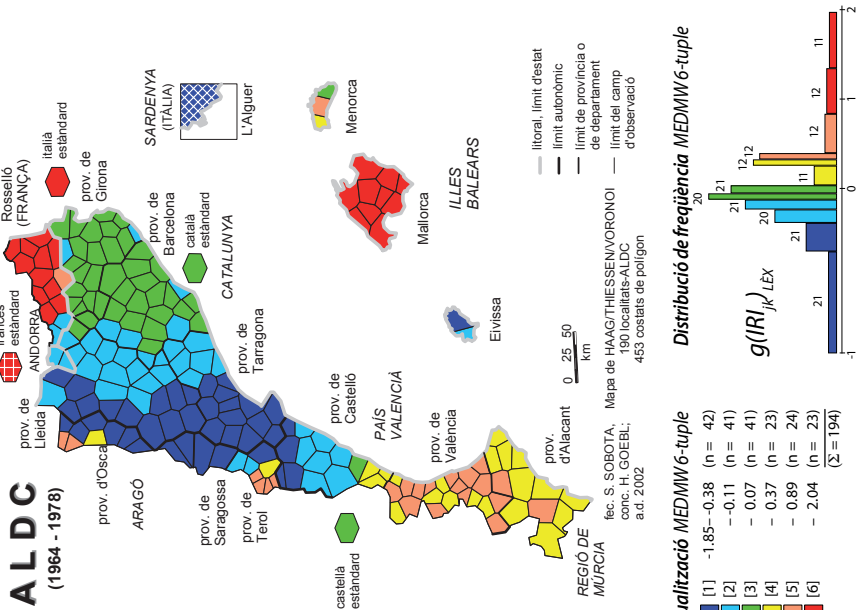
Mapa 5: Mapa de similitud relatiu al punt 86 de l'ALDC; Esterrí d'Àneu.
 Corpus: 1.659 mapes de treball (corpus sencer).
 Algorisme d'interval·lització: MINMMWMAX 6-tuple.



Mapa 10: Mapa isoglotic.
Corpus: 728 mapes de treball *lèxics*.
Algorisme d'intervalització: MEDMW 8-tuple.

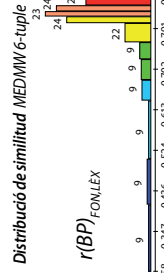
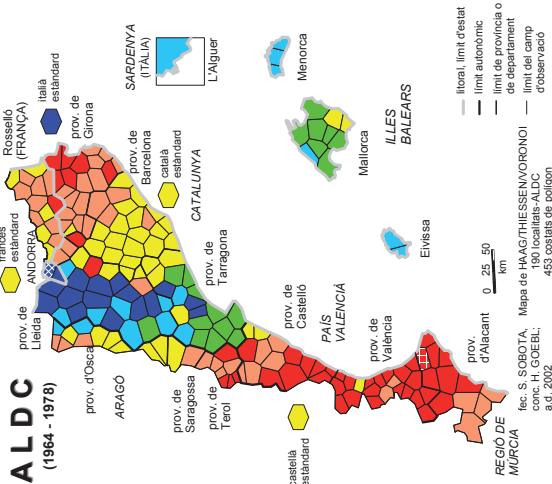


Mapa 9: Mapa isoglotic.
Corpus: 931 mapes de treball *fonètics*.
Algorisme d'intervalització: MEDMW 8-tuple.

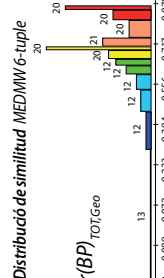
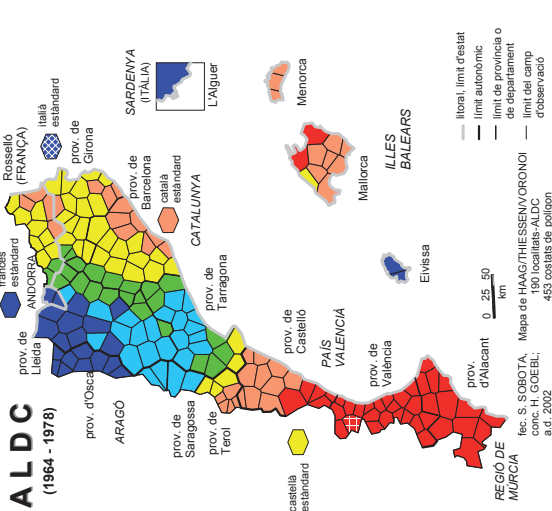


Mapa 12: Sinopsi de 194 coeficients d'asimetria (biaix).
 Corpus: 728 mapes de treball *lèxics*.
 Algorisme d'intervalització: MEMMW 6-tuple.

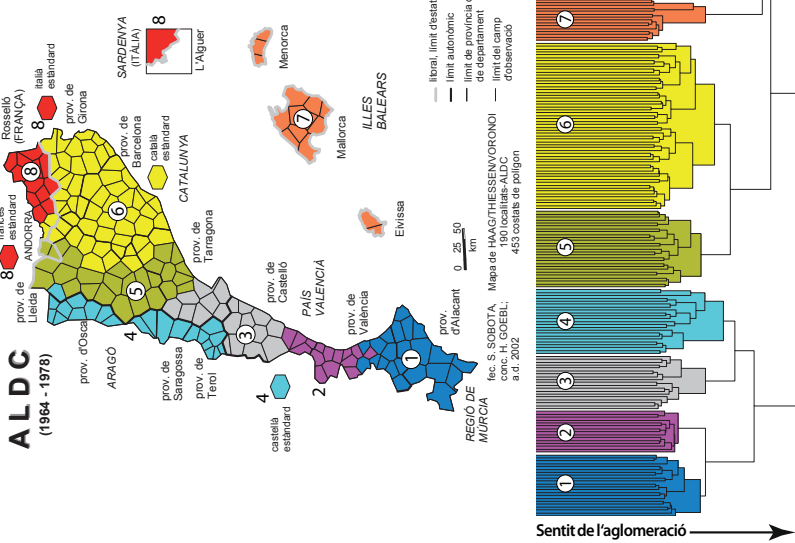
Mapa 11: Sinopsi de 194 coeficients d'asimetria (biaix).
 Corpus: 931 mapes de treball *fonètics*.
 Algorisme d'intervalització: MEMMW 6-tuple.



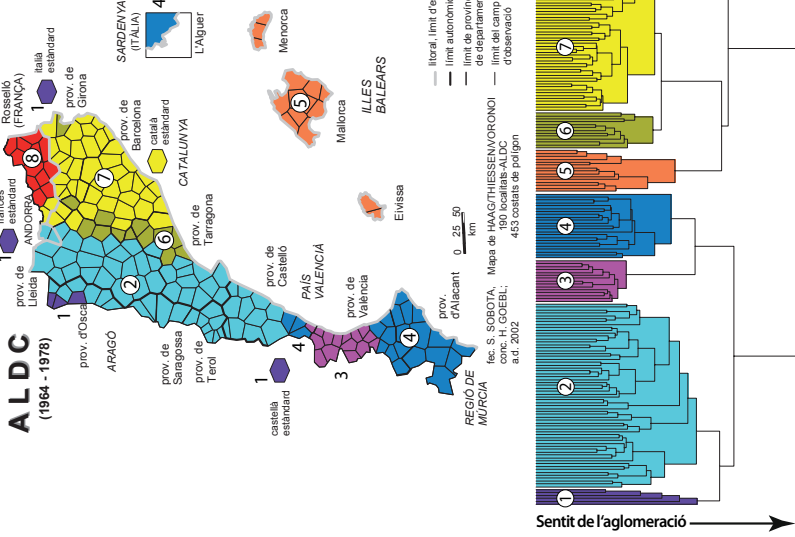
Mapa 14: Correlació entre les gestions fonètica i lèxica de l'espai de l'ALDC.
 Corpus A: 931 mapes de treball fonètics, similituds (lingüístiques) segons l'IRI_{jk}.
 Corpus B: 728 mapes de treball lèxics, similituds (lingüístiques) segons l'IRI_{jk}.
 Algorisme d'intervalització: MEDMW 6-tuple.



Mapa 13: Correlació entre les gestions lingüística i euclidi-ana de l'espai de l'ALDC.
 Corpus A: 1.659 mapes de treball, similituds (lingüístiques) segons l'IRI_{jk}.
 Corpus B: proximitats (geogràfiques) segons el teorema de Pitàgores.
 Algorisme d'intervalització: MEDMW 6-tuple.



Mapa 15: Anàlisi dendrogràfica (mitjançant l'algorisme de Joe A. Ward Jr.) i representació espacial corresponent. Subagrupació en 8 dendremes. Corpus: 931 mapes de treball (corpus *fonètic*), similituds (lingüístiques) segons l'IRI_{jk}.



Mapa 16: Anàlisi dendrogràfica (mitjançant l'algorisme de Joe A. Ward Jr.) i representació espacial corresponent. Subagrupació en 8 coremes. Corpus: 728 mapes de treball (corpus *lèxic*), similituds (lingüístiques) segons l'IRI_{jk}.