

Atti
del 10° Congresso
dell'Associazione Italiana di Linguistica Applicata

Lingue e culture in contatto

In memoria di Roberto GUSMANI

Bolzano, 18-19 febbraio 2010

a cura di
Raffaella BOMBI, Mari D'AGOSTINO, Silvia DAL NEGRO,
Rita FRANCESCHINI

Perugia: Guerra Edizioni

2011

[Pubblicazioni
della
Associazione Italiana di Linguistica Applicata
vol. 9]

ISBN 978-88-557-0382-6

Hans GOEBL

Brevissima presentazione della dialettometria (con esempi tratti dall' AIS),
pp. 71-103.

Nota: In questo estratto le tavole 1–12 non vengono presentate in bianco e nero come nell'originale, bensì a colori. Viene così considerevolmente aumentata la loro utilità euristica. L'impaginazione originale rimane inalterata.

Hans Goebel *

BREVISSIMA PRESENTAZIONE DELLA DIALETTOMETRIA
(Con esempi tratti dall' AIS)

Ziel der seit bald 40 Jahren bestehenden Dialektometrie ist es, die in einer möglichst großen Anzahl von Karten eines Sprachatlasses auffindbare räumliche Variation mit quantitativen Mitteln zusammenfassend – also induktiv-synthetisch – darzustellen. Sie bedient sich dabei in statistischer Hinsicht eher einfacherer Methoden der “Numerischen Klassifikation” und in visualistischer Hinsicht speziell ausgewählter bzw. neu entwickelter Verfahren der quantitativen Kartographie. Zur möglichst benutzerfreundlichen Bemeisterung aller dabei anfallenden Probleme wurde im Jahr 2000 in Salzburg das Dialektometrie-Programm VDM (“Visual DialectoMetry”) entwickelt. Im vorliegenden Beitrag werden in aller Kürze die theoretischen Fundamente der Dialektometrie sowie die wichtigsten Etappen der dialektometrischen Verfahrenskette vorgeführt und einige der dabei anfallenden Ergebnisse anhand von acht – leider nur – schwarzweißen Graphiken exemplarisch dargestellt. Es handelt sich dabei um zwei Arbeits-Karten (von denen insgesamt 3911 Stück erstellt und danach synthetisiert wurden), drei Ähnlichkeits-Karten (mit Prüfbezugspunkten in Mailand, Neapel und der Heimat des Hochitalienischen), eine Parameter-Karte (Synopsis der Schiefen), zwei Zwischenpunkt-Karten (Isoglossen-Synthese und Strahlen-Karte), eine Baum-Analyse (umfassend sowohl den Baum als auch dessen Umlegung in den Raum) und zwei Korrelations-Karten, die sich zum einen auf die räumliche Parallel-Entwicklung von Sprache und Raum sowie zum anderen auf jene von Phonetik und Lexikon beziehen. Die analysierten Daten entstammen den acht Bänden des “Sprach- und Sachatlasses Italiens und der Südschweiz” (AIS) von K. Jaberg und J. Jud (Zofingen 1928-1940).

* Universität Salzburg

1. Osservazione preliminare

Questo contributo rappresenta una sintesi molto sommaria ed euristicamente “scarna” di una conferenza plenaria tenuta il 19 febbraio 2010 a Bolzano in occasione del X Congresso Internazionale di studi dell’Associazione Italiana di Linguistica Applicata. L’aggettivo *scarno* si spiega per il fatto che la suddetta conferenza è stata accompagnata da una ponderosa presentazione con 36 diapositive a colori. Considerata l’importanza euristica dell’uso di colori per una buona comprensione della dialettometria e soprattutto dei suoi risultati iconici tale qualifica risulta del tutto conveniente.

2. Che cos’è la dialettometria (DM)?

La DM rappresenta una disciplina di stampo induttivo che tramite la sintesi quantitativa dei dati di atlanti linguistici tradizionali cerca di studiare e individuare regolarità – anzi leggi – spaziali nascoste nella massa dei dati dei rispettivi atlanti linguistici. Nel quadro della “Scuola dialettometrica di Salisburgo”¹ (DM-S) si dà per assunto che le suddette leggi siano l’emanazione diretta di un atteggiamento speciale dei locutori dia- o basilettali rispetto allo spazio da loro abitato, chiamato da un paio di anni dalla DM-S “gestione basilettale dello spazio da parte dell’Homo loquens”.

Dal punto di vista meramente metodico la DM si serve, sul versante qualitativo, di tutta la panoplia teorica, concettuale e metodica della geografia linguistica (a base d’atlanti) tradizionale e, su quello quantitativo, di metodi appositamente selezionati, della tassonomia numerica nonché della geo- e cartografia quantitativa moderna.

Dal 1999 tutte le analisi-DM effettuate a Salisburgo (ed in parte anche altrove) si avvalgono dell’eccellente software VDM (“Visual DialectoMetry”) concepito e realizzato dal nostro amico Edgar Haimerl. VDM consente l’esecuzione informatizzata di pressoché tutti gli anelli della catena metodica della DM-S: l’immissione dei dati in una banca dati-*Access*, il calcolo delle similarità e distanze, nonché degli schemi dendrografici e la realizzazione istantanea di tutte le visualizzazioni previste.

3. La dialettometrizzazione integrale dell’AIS

Dopo le nostre previe esperienze limitate alla parte settentrionale della rete dell’AIS e

¹ Quanto ad altri lavori realizzati in seno alla DM-S rinvio ai contributi dialettometrici di Bauer (2009) e Videsott (2009). Rimando inoltre ai lavori dialettometrici della “Scuola di Groningen”, capeggiata da John Nerbonne (cfr. <http://www.let.rug.nl/~nerbonne/>), seppure svolti secondo altri principi metodici e metodologici.

alla base dei dati (di stampo prevalentemente lessicale) dei volumi I, II e IV dell' AIS², grazie anche agli ottimi risultati ottenuti nel quadro di una dialettometrizzazione integrale dei dati dell'atlante linguistico francese ALF³, ci siamo accinti, negli anni 2007-2009, a realizzare un progetto analogo per la totalità dei dati dell' AIS⁴.

L' AIS dispone, come si sa, nei suoi otto volumi di formato *in-folio*, di 1705 carte e di una rete di 407 punti di rilevamento. Il nostro proposito era di analizzare la variabilità geolinguistica reperibile su tutte le carte dell' AIS secondo le categorie linguistiche della fonetica, della morfologia, della sintassi e del lessico, considerando non solo le carte propriamente dette dell' AIS, ma anche tutti i materiali intabellati presentati alla fine dell'ottavo volume ("carte" 1683-1701).

4. Dalle carte originali dell' AIS alle "carte di lavoro" (CL) ed alla matrice dei dati

Si veda la figura 1.

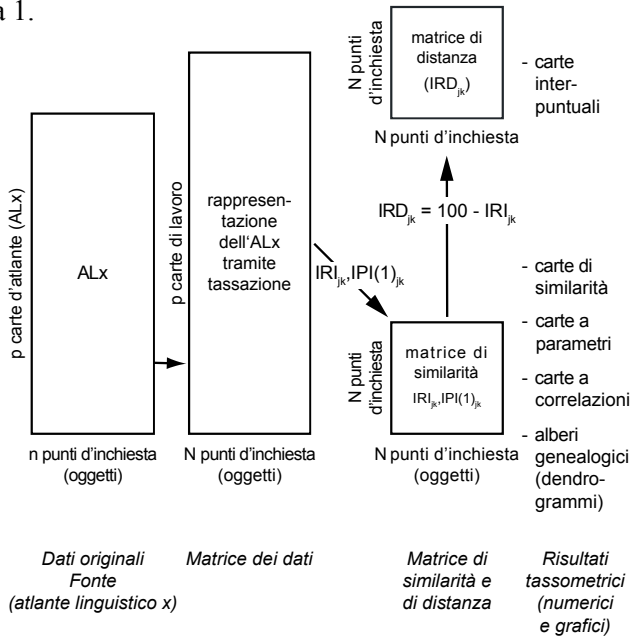


Fig. 1 - Diagramma dei metodi di misurazione e di visualizzazione utilizzati dalla Scuola Dialettometrica di Salisburgo (DM-S)

² Cfr. in merito soprattutto Goebel (1981; 1982; 1984a).

³ Cfr. in merito i nostri lavori (Goebel, 2002; 2003a).

⁴ Si vedano le nostre relazioni di lavoro: Goebel (2007); Goebel (2008a).

Il rilevamento della variabilità linguistica reperibile sulle carte dell’AIS – chiamata nel gergo della DM-S “tassazione” – veniva effettuato secondo i principi e metodi della grammatica storica tradizionale (per la fonetica, la morfologia e la sintassi) nonché dell’onomasiologia romanza classica. A causa della lacunosità di molte carte dell’AIS e dell’uso di un questionario ridotto da parte degli ideatori dell’AIS in circa trenta località della rete esaminata, non era possibile estendere la nostra analisi tassatoria a tutte le carte dell’AIS.

Per ottimizzare la completezza della matrice dei dati dovevamo quindi scartare – almeno per le analisi qui presentate – circa un quarto delle carte AIS originali e 29 dei 407 punti di rilevamento. I restanti 378 punti venivano però completati da due punti artificiali corrispondenti all’italiano e al francese standard, e dalle seconde inchieste fatte da Paul Scheuermeier nelle città di Torino, Milano, Venezia, Bologna e Firenze. Per ovvie ragioni abbiamo anche cancellato i tre punti alloggiati nell’Italia meridionale (pp. 792, 748 e 751). Così è stata stabilita una rete con 382 punti che sta alla base di tutte le visualizzazioni di questo contributo. Ecco il bilancio consuntivo dei nostri lavori di tassazione (Tab. 1)⁵:

	carte AIS originali prese in considerazione	carte di lavoro (CL) ottenute tramite tassazione
fonetica	257	1 766
lessico	1 087	1 225
dati integrali	1 310	3 991

Tab. 1. Riassunto delle nostre attività di tassazione

La Tab. 1 dimostra la successione delle diverse tappe del trattamento-DM dei dati. A ciascuna delle tappe ivi raffigurate corrispondono presupposti, metodi e finalità ben definiti. Si badi al fatto che l’ultima tappa della catena-DM – che non compare esplicitamente sullo schema della Tab. 1 – corrisponde alla valutazione teorica delle visualizzazioni-DM da parte del dialettometra.

5. Dalla matrice dei dati alle matrici di similarità e di distanza, nonché al loro spoglio iconico

Si vedano le tavole 1 e 2: esse rappresentano due “carte di lavoro” (CL). La prima (tav. 1) è di carattere *lessicale*, la seconda (tav. 2) di carattere *morfologico*. I tipi

⁵ Si tratta dei risultati ottenuti alla fine della dialettometrizzazione dell’AIS. Le nostre relazioni di lavoro anteriori – Goebel (2007) e (2008) – poggiavano ancora su effettivi-CL minori: Goebel (2007: 1 746 CL); (2008a: 3 076 CL).

lessicali (o “tassati”) *madre*, *mamma* e *mare* che compaiono sulla tavola 1 occupano tre aree con estensioni (ed anche configurazioni geografiche) molto diverse: l’area del tassato *madre* abbraccia 62 punti, quella del tassato *mare* 19 e quella del tassato *mamma* 228 punti. Si noti inoltre la presenza di due dati mancanti. La variabilità interna (o “polinimia”) delle CL è molto alta. Mentre la tavola 1 mostra una CL 3-nima (*tri*-nima), la tavola 2 rinvia ad una CL 4-nima. Ecco il bilancio consuntivo della frammentazione tassatoria dei nostri dati (Tab. 2):

	polinimia minima	polinimia massima	numero complessivo dei tassati analizzati (rispettivamente delle loro aree)
fonetica	2	48	15 124
lessico	2	153	21 989
dati integrali	2	153	43 564

Tab. 2. Frammentazione tassatoria dei dati

La matrice dei dati da noi utilizzata in questa sede dispone quindi delle dimensioni seguenti: N (punti di rilevamento) = 382, p (carte di lavoro) = 3911. Essa costituisce la raffigurazione nominale di 43564 aree tassatorie di carattere fonetico, morfologico, sintattico e lessicale. Aggiungiamo che la definizione classificatoria delle testè menzionate 43564 aree (o tassati) è stata realizzata con metodi meramente linguistici e poggia quindi su un ingente stock di *expert knowledge* romanistico. L’elaborazione successiva delle informazioni accumulate nella matrice dei dati mira alla misurazione delle similarità e distanze reciproche tra i 382 punti di rilevamento della nostra rete. In genere tale misurazione viene eseguita mediante indici di similarità appositamente selezionati dal punto di vista linguistico fra la vasta gamma degli indici di similarità e di distanza offerti dalla tassonomia numerica. In questa sede si tratta dell’“Indice Relativo di Identità” (IRI_{jk}) e dell’“Indice Relativo di Distanza” (IRD_{jk}) che, da molti anni, rappresentano gli indici standard della DM-S. Tra l’ IRI_{jk} e l’ IRD_{jk} vige la relazione seguente: $IRI_{jk} + IRD_{jk} = 100$.

Il contenuto quantitativo delle matrici (quadrato) di similarità e di distanza deve essere quindi appositamente visualizzato, sempre in base alle necessità teoriche, analitiche e cognitive del dialettometra. A questo proposito il programma salisburghese VDM offre molteplici possibilità, sempre avvalendosi del mondo sfumato dei colori.

6. Presentazione sommaria di cinque analisi dialettometriche

Le cinque analisi presentate qui di seguito costituiscono solo un ristretto campione delle molteplici possibilità analitiche ed esplorative offerte dal programma VDM.

6.1. Presentazione ed interpretazione di tre “carte di similarità” (CDS)

Si vedano le tavole 3-5: la carta di similarità è lo strumento di lavoro più importante della DM-S. Ciascuna CDS dispone – oltre al consueto fondo di carta poligonizzato – di un punto di riferimento preselezionato e visualizza, tramite lo spiegamento spaziale di una serie di retini (per gli intervalli 4-6) e tratteggi (per gli intervalli 1-3) graduati, il calo progressivo dei valori della similarità linguistica rispetto al valore supremo di 100% del punto di riferimento in questione. Accenniamo, per una miglior comprensione delle CDS, a tre fatti cartografici di rilievo:

- Il poligono del punto di riferimento è sempre in bianco.
- In genere, il poligono bianco del punto di riferimento si trova nel centro o al margine della zona contrassegnata dal retino più scuro (intervallo 6), la quale corrisponde ai punti-AIS provvisti dei più alti tassi di similarità.
- La configurazione spaziale del profilo coropletico di una CDS (ed anche di altre visualizzazioni-DM) dipende in larga misura dalla natura di due strumenti euristici molto importanti: dall'indice di similarità nonché dall'algoritmo di visualizzazione. In questo contributo si tratta di due strumenti standard della DM-S: l'indice IRI_{jk} e l'algoritmo di visualizzazione MINMWMAX⁶. Ricordiamo inoltre che – considerate le molteplici possibilità di calcolo e di cartografazione di VDM – la generazione e la successiva comparazione di un gran numero di profili coropletici di origine diversa è facilmente realizzabile.

Quanto alla diminuzione dei tassi di similarità nello spazio, è ovvio che essa sta in stretta relazione, benché in varia misura, con l'aumento delle rispettive distanze geografiche.

Mediante i loro profili coropletici molto netti, le CDS evidenziano le seguenti proprietà geolinguistiche del rispettivo dia- o basiletto locale:

- la sua posizione relazionale all'interno della rete esaminata,
- il tasso di diffusione (penetrazione, interazione, ecc.) della sua dialettalità all'interno della rete esaminata,

⁶ Precisiamo che l'algoritmo MINMWMAX distribuisce, da ambo i lati della media aritmetica, i N-1 valori di similarità in intervalli di uguale spessore numerico: per ulteriori dettagli cfr. Goebel (1983, *passim*); (1984a, I: 90-92); (1984b: 19-20).

— l'efficacia comunicativa del rispettivo dialetto.

Le CDS si prestano inoltre a diversi modi d'interpretazione⁷. Tra le molteplici possibilità spiccano:

— 1: l'interpretazione diacronica;

— 2: l'interpretazione sociale e/o sociologica.

Ad 1: Poiché i profili spaziali delle CDS risalgono a dati geolinguistici originati diacronicamente, è lecito procedere alla loro interpretazione anche dal punto di vista storico. Questa possibilità è oltremodo proficua quando si tratta di applicare i metodi della DM-S a dati medievali: cfr. a questo proposito le nostre ricerche-DM relative all'atlante scrittologico di Anthonij Dees del 1980⁸.

Ad 2: In certi contesti giova applicare un'interpretazione sociologica *per analogiam* ammettendo che la rete esaminata si riferisca a N persone in atto di interagire tra di loro. In tale prospettiva una CDS diventa il bilancio visivo del successo (o meno) degli sforzi divulgativi (missionari, interazionali, ecc.) di uno degli N membri del gruppo fittizio esaminato.

Di particolar rilievo è inoltre la grande stabilità dei profili (coropletici) di similarità all'interno di una zona (geo)tipologicamente coerente. Per questo motivo è molto informativa la considerazione comparativa di un certo numero di profili di similarità, i cui punti di riferimento sono allineati lungo un percorso appositamente preselezionato. Scaturiscono così effetti cinematografici altamente suggestivi⁹ che gettano una nuova luce sulla sovramenzionata "gestione basilettale dello spazio da parte dell'Homo loquens".

La tavola 3, relativa alla posizione relazionale del dialetto di Milano, evidenzia molto chiaramente l'inserzione del tipo linguistico lombardo-milanese nella Padania ed il calo delle rispettive similarità verso nord (Grigioni), ovest (le valli occitaniche del Piemonte) e sud (a partire dal versante sud dell'Appennino tosco-emiliano). Si noti la compattezza areale delle zone appartenenti agli intervalli 6 e 5 nonché di quelle appartenenti all'intervallo 2 (nel nord: francoprovenzale, romancio dei Grigioni, ladino del Sudtirolo; nel sud: praticamente tutti i dialetti situati a sud della linea Gaeta-Pe-

⁷ Per una discussione circostanziata cfr. Goebel (1984a, I: 100-113); (1984b: 24-28); nonché Goebel (1981: 369-381).

⁸ Cfr. Goebel (2008b).

⁹ Si tratta di un'applicazione della "cartografia dinamica": cfr. a questo proposito Goebel (1982: 38-40; 78-81) nonché (1984b: 39-31). Oggigiorno la confezione di tali suggestive sequenze cinematografiche è facilmente realizzabile mediante il programma PowerPoint. Ovviamente anche in questa sede l'uso di colori è indispensabile.

scara. Si badi inoltre al fatto che la più grande dissimilarità (“antipodi”: intervallo 1) si manifesta solo in Sardegna.

La tavola 4 esibisce un profilo d’identità “tipicamente meridionale”. Allorché la localizzazione geografica delle zone con la più alta similarità (intervallo 6) è completamente cambiata, rimangono inalterate la grande compattezza delle zone raggruppate nei sei intervalli visualizzati e la diminuzione graduata della similarità linguistica coll’aumento della distanza geografica. Si noti che gli “antipodi” del dialetto di Napoli si trovano in Valle d’Aosta, nei Grigioni e nella Ladinia sudtirolese.

La tavola 5 (relativa all’italiano standard) rivela, oltre all’armonica inserzione relazionale dell’italiano standard nella parte centrale dello “stivale”¹⁰, anche una marcata presenza – con una chiara discontinuità territoriale rispetto all’Emilia-Romagna – della “tipicità italiana” nel Veneto, la cui preminenza – a dire il vero – ci ha sempre enormemente meravigliato. Si tratta, in effetti, di una situazione molto particolare che con ogni probabilità rinvia a effetti sostratistici risalenti al periodo della romanizzazione del Veneto nel terzo secolo a. C. Si noti inoltre l’ubicazione degli antipodi tipologici dell’italiano, contrassegnati tutti dai tratteggi 1 e 2, e rintracciabili ovviamente alle diverse periferie della nostra rete: Sardegna centrale e meridionale, Piemonte occidentale e settentrionale, Valle d’Aosta, nonché le tre zone ladine (o retoromanze).

6.2. *Presentazione ed interpretazione di una carta a parametro: la sinossi dei coefficienti di asimmetria di Fisher (CAF) delle distribuzioni di similarità*

Si veda la tavola 6: la natura statistica molto variabile delle N distribuzioni di similarità di cui consiste ciascuna matrice di similarità (colle dimensioni N per N), ci ha ben presto indotti ad esaminare più da vicino – sempre in chiave comparativa – i valori dei diversi parametri statistici (come il “minimo”, il “massimo”, la “media aritmetica”, ecc.)¹¹. Si è così rivelata la grande utilità geolinguistica della sinossi di alcuni parametri relativi alla simmetria della rispettiva distribuzione di similarità tra i quali spicca il “coefficiente di asimmetria di Fisher (CAF)”¹² che ci permette di studiare più da vicino il cosiddetto *Sprachausgleich* (‘compromesso linguistico’).

Per *Sprachausgleich* – che ovviamente ha diversi gradi di intensità – si intende la più o meno grande integrazione connettiva di un basiletto particolare nell’insieme della

¹⁰ La tavola 5 visualizza inoltre i diversi gradi di italianizzazione dei dialetti dell’Italo-, Reto- e Sardo-Romania.

¹¹ Per ulteriori dettagli cfr. Goebel (1984a, I: 136-164).

¹² Al CAF *italiano* (ed anche *francese*) corrispondono i termini inglese *skewness* e tedesco *Schiefte*.

rispettiva rete geolinguistica. Pare logico – e corrisponde inoltre alla nostra esperienza empirica – che tra i diversi basiletti di una rete geolinguistica data ce ne siano alcuni che assumono un rilievo molto isolato rispetto ai loro omologhi ed altri i cui legami con la maggioranza delle componenti della corrispondente rete sono eccellenti.

Ecco una riflessione atta a facilitare la comprensione di questo fenomeno. Ciascuno dei 382 vettori delle altrettante località della nostra rete è contrassegnato da 3911 attributi (o tassati) geolinguistici. Orbene, ciascuno di questi 3911 tassati corrisponde ad un'area geolinguistica la cui estensione può variare – teoricamente – tra 1 e 381 punti (o poligoni)¹³. È quindi di somma importanza per il livello d'integrazione di un vettore(-punto) particolare nella totalità della rispettiva rete, la ripartizione percentuale di aree con estensioni minori, medie o maggiori. Da questa considerazione si evidenzia che un punto-vettore che dispone prevalentemente di aree con un'estensione molto ridotta non può che giocare un ruolo comunicativo molto limitato in seno alla rete intera. E viceversa.

Che la simmetria molto variabile di una distribuzione di similarità data possa avere un particolare significato linguistico, si può evidenziare tramite un raffronto delle leggende numeriche di due carte di similarità. La leggenda della tavola 3 (relativa al dialetto lombardo di Milano, P. 261) ci ricorda che dei 381 valori di similarità cartografati 194 (= 104 + 78 + 12 punti su 381: = 50.91%) – ossia la maggioranza – si trovano *sopra* la media aritmetica di 61.42%, laddove 187 (= 18 + 95 + 74 punti su 381: = 49.08%) si accumulano *sotto* la media aritmetica.

Una situazione diversa risulta dalla tavola 4 relativa al dialetto di Napoli (P. 721) dove la stessa ripartizione si presenta come segue: punti situati *sopra* la media aritmetica di 58.58%: 150 (= 78 + 56 + 16 punti su 381: = 39.37%), ossia la minoranza; numero dei punti situati *sotto* la media aritmetica: 231 (= 134 + 74 + 23 punti su 381: = 60.62%).

La connessione comunicativa (relativa) del dialetto di Milano (P. 261) col resto della rete è quindi superiore a quella del dialetto di Napoli (P. 721).

Il coefficiente di asimmetria di Fisher (CAF)¹⁴ assume il valore 0 per una simmetria totale della rispettiva distribuzione di frequenza, valori negativi nel caso di una dislocazione (asimmetrica) verso *destra* (maggioranza dei valori *sopra* la media aritmetica) e valori positivi nel caso di una dislocazione (asimmetrica) verso *sinistra* (maggioranza dei valori *sotto* la media aritmetica).

Sulla tavola 6, ai valori *negativi* corrispondono i poligoni con i tratteggi 1 e 2,

¹³ L'estensione massima di N, cioè di 382, punti non può essere ammessa in questa sede perché significherebbe un azzeramento totale della variabilità geolinguistica.

¹⁴ Per la formula del CAF cfr. Goebel (1981: 394) e (1984a, I: 150).

mentre i valori *positivi* vengono marcati dai poligoni con i retini 5 e 6. Ecco il significato linguistico di tratteggi e retini:

— Poligoni con i tratteggi 1 e 2: zone con un tasso molto alto del compromesso linguistico (*Sprachausgleich*), contrassegnate – tanto in dia- quanto in sincronia – da un grande dinamismo interattivo rispetto alla totalità delle rete dell’AIS¹⁵. Si tratta dunque di zone molto attive tanto dal punto di vista espansivo quanto ricettivo.

— Poligoni con i retini 5 e 6: zone con un tasso molto basso del compromesso linguistico (*Sprachausgleich*), caratterizzate quindi – tanto in dia- quanto in sincronia – da un gran conservatorismo linguistico rispetto alla totalità delle rete esaminata. Si tratta pertanto di zone la cui tipicità linguistica è stata preservata per molto tempo ed in larga misura da intrusioni straniere di qualsiasi origine.

Uno sguardo alla tavola 6 ci rivela – non senza difficoltà per ovvi motivi cartografici – che la ripartizione spaziale dei retini 5 e 6 è molto chiara. Lo stesso dicasi dei tratteggi: nella Padania i tratteggi 1 e 2 creano una dipartimentalizzazione ben definita in cui emergono, come zone divisorie provviste però di altissimi tassi interattivi, il corso dell’Adige e del Po nonché lunghi tratti della catena appenninica dalla Liguria fino alle Marche. Un altro solco vallivo contrassegnato in parte da tratteggi della classe 1 (e quindi provvisto di un alto potenziale interattivo) s’incunea tra l’Italia mediana (Lazio, Umbria, Marche) e quella del sud (Abruzzo e Molise, Campania, ecc.).

Le zone nei retini 6 (ecc.) rappresentano invece configurazioni spiccatamente puntuali e provviste quindi di un’estensione corica geograficamente coerente.

È ovvio che le isole linguistiche situate nel sud (pp. 715, 760) ed in Sicilia (pp. 865, 836, 817) e soggette sin dalla loro genesi a perpetui interscambi linguistici rientrano negli intervalli 1 e 2. Lo stesso vale per la parte settentrionale della Sardegna, la quale, considerando la sua rampante toscanizzazione plurisecolare tramite la Corsica, dispone, in ultima analisi, di molte caratteristiche di un’isola linguistica vera e propria.

6.2. *Presentazione ed interpretazione di due carte interpuntuali (carta isoglottica e carta a raggi)*

Si vedano le tavole 7 e 8: dal punto di vista cartografico le carte interpuntuali rientrano nella classe delle carte “isaritmiche”, i cui elementi iconici di base non sono più

¹⁵ Il rispettivo dinamismo interattivo non si limita dunque alla vicinanza immediata dei punti della rete, bensì abbraccia sempre la totalità di tutti i punti d’inchiesta della rete.

le aree dei poligoni, bensì i lati di *triangolo* o i lati di *poligono* ambedue generati tramite un previo trattamento geometrico della rete esaminata. La confezione delle tavole 7 e 8 si basa quindi sulle tappe seguenti¹⁶:

— Triangolazione delle rete-AIS secondo i principi della geometria di Delaunay¹⁷. Ne risultano – per la rete della tavola 8 – 970 lati di *triangolo*.

— Poligonizzazione successiva della rete triangolata secondo i principi della geometria di Voronoi¹⁸: ne risultano – per la rete della tavola 7 – 970 lati di *poligono*.

— Allestimento della simbologia cartografica per conferire parallelamente ai lati di triangolo e di poligono uno spessore ed una colorazione variabile a seconda della grandezza (o meno) dei rispettivi valori interpuntuali da visualizzare. Il programma VDM consente la realizzazione rapida di ambedue le possibilità (ovviamente a colori).

Dal punto di vista tasso- o dialettometrico la carta isoglottica della tavola 7 è basata su 970 valori di *distanza* dell'IRD_{jk} oscillanti tra 10.19 e 47.52%. La visualizzazione dei suddetti 970 valori viene effettuata lungo altrettanti lati di poligoni secondo il principio cartografico seguente: più i valori dell'IRD sono alti, più i rispettivi lati di poligono sono spessi e scuri. E viceversa. In questa maniera spuntano, in non poche zone della nostra carta, fenomeni lineiformi equivalenti ai ben noti fasci d'isoglosse. Questo accade al sud della Rumantschia, lungo l'arco alpino occidentale, il Po e l'Adige, tra il Veneto ed il Friuli, lungo l'Appennino, nella Sardegna settentrionale ed intorno a tutte le isole linguistiche del sud e della Sicilia.

Mentre la configurazione lineiforme spessa e scura situata lungo l'Appennino tosco-emiliano evoca la ben nota "linea La Spezia-Rimini" di memoria wartburghiana¹⁹, la sua equivalente insediata tra Roma ed Ancona offre un assetto tutt'altro che lineare. Si tratta piuttosto di una larga zona di passaggio con effetti di compartimentalizzazione ben marcati.

I paraggi dove si accumulano i poligoni sottili marcati in grigio chiaro rimandano al contrario a zone linguisticamente coerenti (o nuclei dialettali).

Mentre sulla *carta isoglottica* della tavola 7 si evidenziano fenomeni di confinazione e di distanziamento interpuntuale, dalla *carta a raggi* della tavola 8 si ricavano fatti inversi: collegamenti linguistici a brevissima distanza. Per questo

¹⁶ Per il problema delle carte interpuntuali cfr. soprattutto Goebel (1983 *passim*) nonché (1984a, I: 183-196).

¹⁷ Cfr. Okabe *et al.* (1992: 72 ss.).

¹⁸ Cfr. Okabe *et al.* (1992: 65 ss.).

¹⁹ Per la storia dei concetti delle linee La Spezia-Rimini e Roma-Ancona cfr. Goebel (2003b: 634 s.).

motivo la variabile-DM visualizzata è un indice di *similarità* (IRI_{jk}). I rispettivi valori oscillano tra 52.48% e 89.81%. Si badi al fatto che la somma tra il valore minimo della tavola 7 (= 10.19%) ed il valore massimo della tavola 8 (= 89.81%) ammonta a 100%. È questa la conseguenza della già menzionata relazione tra similarità (IRI_{jk}) e distanza (IRD_{jk}) secondo la formula $IRI_{jk} + IRD_{jk} = 100$.

Tramite lo spiccato addensamento spaziale di molti triangoli spessi e scuri il profilo isaritmico della tavola 8 mostra l'ubicazione di nuclei dialettali. Per una miglior comprensione di questa carta giova anche l'idea metaforica che ciascuno dei 970 triangoli funge da "tubo" posto tra due località contigue, nel quale corrono "flussi" di similarità (o "amicizia") di varia intensità.

Bisogna però aggiungere che, dal punto di vista tassometrico, l'analisi interpuntuale è piuttosto superficiale. In effetti, i 970 valori interpuntuali rappresentano, sui 72771 valori delle rispettive matrici di distanza o di similarità, solo il 1.33% della totalità delle informazioni quantitative disponibili.

6.4. *Presentazione dell'analisi dendrografica tramite l'algoritmo di Joe Ward, Jr.*

Si vedano le tavole 9 e 10: uno dei mezzi più efficaci di classificazione numerica in genere e della DM-S in particolare è, da oltre 30 anni, la "classificazione gerarchica agglomerativa" (CGA) (così denominata a livello internazionale). La CGA fornisce, in quanto *output* euristico, schemi dendrografici ("alberi") di qualsiasi natura. Ecco i sommi capi metodici per un buon intendimento della CGA²⁰:

- Il punto di partenza – la matrice di similarità.
- Si tratta di nuovo di una classificazione quantitativa della totalità dei dati compresi nella rispettiva matrice di similarità (N per N).
- L'agglomerazione algoritmica.

I suddetti alberi – che per una matrice di similarità colle dimensioni N per N dispongono sempre di N foglie – nascono da un processo reiterato di fusioni binarie retto da un peculiare algoritmo agglomerativo. Tale processo inizia sempre a partire dalle foglie: sulla tavola 9 si veda in merito la freccia inferiore ("creazione statistica dell'albero"). La classificazione (o tassometria) numerica internazionale si serve di un gran numero di appositi algoritmi, al quale corrisponde una molteplicità analoga di alberi. L'utilità di questi alberi per un dato progetto di ricerca deve essere accertata

²⁰ Per una breve presentazione della tecnica della CGA in lingua italiana cfr. Goebel (1992, passim).

caso per caso tramite una valutazione combinata delle necessità della statistica e della rispettiva disciplina empirica (in questa sede: della geolinguistica romanza).

— Le proprietà formali (= statistiche) degli alberi generati dalla CGA: gli alberi della CGA costituiscono una gerarchia con ramificazioni – esclusivamente binarie – di classi (*clusters*, ecc.) disgiunte, la cui variabilità quantitativa interna cresce a partire dalle foglie (= chioma) fino alla radice (= tronco) dell’albero. Soprattutto per l’interpretazione geolinguistica dei diversi piani della struttura ramificata dell’albero la presa in considerazione della variabilità cangiante delle diverse classi è essenziale.

— Il punto di arrivo – la spazializzazione degli alberi calcolati.

Mentre per l’uso esclusivamente diacronico di schemi dendrografici (come viene praticato spesso dalla lessicostatistica²¹) – si veda in merito la freccia superiore (“interpretazione diacronica dell’albero”) – bastava l’analisi puntuale dell’albero stabilito, per la DM la situazione è diversa. In sede di DM bisogna sempre procedere ad una spazializzazione dei risultati dell’elaborazione classificatoria dell’albero: rimando alle tavole 9 e 10 che ne mostrano il principio. Grazie all’efficacia del programma VDM la spazializzazione selettiva degli alberi calcolati può essere effettuata in maniera molto comoda, veloce e – ovviamente – coll’aiuto di colori. Il calcolo e la generazione degli schemi dendrografici nonché spaziali delle tavole 9 e 10 sono stati realizzati tramite VDM in pochi minuti (escluso il tempo necessario per la trasposizione dei colori in retini grigi).

L’algoritmo dendrografico applicato è stato proposto dallo statistico americano Joe Ward, Jr. La sua utilità per le finalità della geolinguistica è stata spesso comprovata in base a dati geolinguistici di origine assai diversa²². La selezione e definizione dei rami dell’albero (“dendremi”) (tavola 9) corrisponde a quella delle rispettive aree (“coremi”)²³ della spazializzazione (tavola 10). In questa sede si tratta di una selettività media – con solo 11 dendremi-coremi (DC) – che evidenzia gruppi dialettali universalmente conosciuti.

L’interpretazione linguistica deve sempre iniziare con le prime biforcazioni dopo la radice (ossia dopo il tronco). La scissione maggiore (nei DC 1-7 versus i DC 8-11) fa risaltare la “linea La Spezia-Rimini”. Tramite la bipartizione ulteriore

²¹ Per le filiazioni tra la DM dendrografica e la lessicostatistica cfr. Goebel (2003a: 84-89).

²² Rimando in merito alla bibliografia dei nostri contributi-DM: http://www.sbg.ac.at/rom/people/prof/goebel/dm_publi.htm

²³ La coppia terminologica *dendrema-corema* viene da noi utilizzata già da parecchio tempo: cfr. Goebel (2003a: 85, nota 13).

del DC “padano” 1-7 viene scartata la *Rumantschia* intera (DC 1). La parte restante della Padania si scompone nel dominio veneto da una parte (DC 7) ed il resto gallo-romanzo dall'altra (DC 2-6).

Per l'Italia mediana e meridionale la bipartizione del DC 8-11 finisce con il distacco del centro e sud della Sardegna (DC 11) dal resto (DC 8-10). E così via.

Vanno sottolineati però ancora due fatti: l'onnipresenza della grande compattezza spaziale dei coremi e l'incontrovertibile rilevanza divisoria della “linea La Spezia-Rimini” anche in sede di classificazione dendrografica.

Ovviamente ciascuno degli 11 DC delle tavole 9 e 10 richiederebbe (e meriterebbe) un'analisi più dettagliata, alla quale dobbiamo però rinunciare per motivi di spazio.

6.5. *La dialettometria correlativa*

La DM correlativa, nata dallo studio approfondito delle relazioni che, senza ombra di dubbio, intercorrono tra la distanza (o prossimità) geografica e quella linguistica, è stata aggiunta solo nel 2004 all'armamentario metodico della DM-S²⁴. Questo si evidenzia subito da una comparazione tra una carta di *similarità* normale – calcolata cioè con l'indice IRI_{jk} – ed una carta analoga basata su dei valori di prossimità calcolati secondo il noto teorema di Pitagora e quindi secondo i principi della geometria euclidea (carta di *prossimità*)²⁵. Essendo già a disposizione del programma VDM le coordinate geografiche (x e y) dei punti dell'AIS, l'applicazione successiva del teorema di Pitagora risulta molto facile.

Per l'accertamento delle (cor)relazioni esistenti tra la variabilità spaziale della lingua e quello dello spazio euclideo sarebbe possibile, almeno in teoria, esaminare *de visu*, nel quadro di un qualsiasi progetto-DM, N carte di *similarità* con altrettante carte di *prossimità*. Ovviamente questa considerazione meramente visuale, e quindi impressionistica, non condurrebbe a risultati scientificamente “probanti”. Siccome la statistica moderna offre la possibilità di misurare la relazione tra due serie di valori empiricamente rilevati mediante appositi indici di correlazione, l'applicazione di tali misure e la successiva visualizzazione sinottica dei rispettivi risultati potrebbero costituire una soluzione di gran lunga preferibile alla suddetta comparazione visiva reiterata di profili coropletici accoppiati.

Si tratta in particolare dell'allestimento di due matrici di similarità/prossimità

²⁴ Per una prima presentazione delle molteplici possibilità della DM correlativa rinvio alla nostra relazione di lavoro, Goebel (2005).

²⁵ Per due campioni di una carta di prossimità cfr. Goebel (2007: 263) e (2008a: 95).

provviste delle stesse dimensioni N per N , i cui vettori debitamente accoppiati si prestano ad altrettanti calcoli della rispettiva correlazione numerica. Dopo l'esecuzione degli N calcoli di correlazione, gli N valori calcolati vengono cartografati secondo le norme abituali della DM-S.

La misura di correlazione statisticamente più adatta nell'ambito delle nostre esperienze geolinguistiche è il coefficiente di correlazione di Bravais e Pearson $r(\text{BP})^{26}$. Il $r(\text{BP})$ misura la correlazione lineare tra due variabili quantitative assumendo valori compresi tra -1 e $+1$.

Ai fini dell'interpretazione e di una buona comprensione delle carte a correlazione giova l'assunto che le due variabili correlate tra di loro siano "forze" o "energie" in procinto di diffondersi nello spazio. In questa prospettiva ci si può chiedere se tale diffusione avvenga, da parte delle variabili, in perfetta *sintonia* o in *disarmonia* più o meno marcata. Ovviamente si deve attribuire ad ambedue le possibilità un'adeguata interpretazione linguistica, sempre in chiave dia- e sincronica.

6.5.1. La correlazione spaziale tra "lingua" e "geografia"

Si veda la tavola 11: essa rivela una strutturazione spaziale ben interpretabile dal punto di vista linguistico. Il significato dei rispettivi retini e tratteggi è questo:

— Classi 5 e 6 (retini): zone dove persiste un'armonia originaria tra "lingua" e "spazio", esenti da trasferimenti di fonemi o lessemi in contrasto con lo scambio linguistico di stampo "naturale". Si tratta cioè di zone conservatrici e rimaste linguisticamente piuttosto incontaminate.

— Classi 1 e 2 (tratteggi): zone dove un'eventuale armonia originaria tra "lingua" e "spazio" è stata profondamente perturbata "da fuori", tramite l'ingerenza di dinamismi culturali, sociali o politici di varia origine. Si tratta di zone linguisticamente inquiete, pervase da dinamismi (socio)linguistici anche di lunga durata.

Un rapido sguardo alla tavola 11 pone in evidenza che le zone "inquiete" (intervallo 1) si collocano soprattutto nella parte settentrionale della Toscana, in Liguria e nel Veneto (con inclusione dell'Istria) e nella Sardegna settentrionale. Dalla tavola 11 si deduce inoltre il fatto – d'altronde molto plausibile – che le cinque isole linguistiche nel sud (PP: 715 e 760) ed in Sicilia (PP: 865, 836 e 817) costituiscono zone linguisticamente movimentate. La fondazione di un'isola linguistica rappresenta, in effetti, la rottura totale di un'eventuale armonia originaria tra lingua e spazio.

²⁶ Per la formula del $r(\text{BP})$ cfr. Goebel (2005: 329).

La posizione “movimentata” del nord della Toscana e del Veneto è altamente emblematica soprattutto alla luce della “medianizzazione” del Veneto²⁷ e dei rispettivi processi di esportazione (provenienti dalla Toscana) e di importazione (da parte dei dialetti veneti) di materia linguistica di varia natura.

Si noti anche la bipartizione del dominio retoromancio grigionese, dove la Surselva assume il ruolo di chi ha abbandonato in larga misura il “passo cadenzato” col resto della rete, mentre i dialetti dell’Engadina sono rimasti fedeli alle possibilità dinamiche offerte loro in maniera “naturale” dalla geografia dello spazio padano. Lo stesso dicasi per la Ladinia dolomitica ed il Friuli.

6.5.2. *La correlazione spaziale tra due categorie linguistiche diverse (fonetica e lessico)*

Si veda la tavola 12: il modulo della DM correlativa si presta anche alla comparazione delle gestioni spaziali di due categorie linguistiche diverse. Si veda a questo proposito la tavola 12 dove si evidenzia che il “passo cadenzato” spaziale della *fonetica* e del *lessico* segue lo stesso ritmo tanto nei settori centrali della Padania quanto nel sud, mentre ci troviamo di fronte ad una vera e propria disarmonia delle due categorie sui versanti occidentali e settentrionali del Piemonte, nonché alla periferia della Toscana e dell’Umbria, in Calabria, nel sud della Puglia e nelle isole (Sicilia e Sardegna). Lo stesso divario si manifesta anche nelle cinque isole linguistiche. Questi fatti si spiegano con tensioni esistenti all’interno delle grammatiche dei rispettivi basiletti tra le categorie della *fonetica* e del *lessico*.

Ovviamente la comparazione correlativa potrebbe essere estesa anche ad altre categorie linguistiche. Per mancanza di spazio dobbiamo limitarci, in questa sede, al semplice accenno che queste comparazioni forniscono dei risultati molto interessanti, suscettibili di aprire – in sede di geolinguistica (o di tipologia areale) – nuovi orizzonti per la grammatica variazionale

Riassumendo, si può dire che la DM correlativa rappresenta uno strumento euristico davvero prezioso che si affianca in maniera proficua agli altri metodi della DM salisburghese.

²⁷ Si veda in merito il § 6.1.

7. Conclusione

Arrivati al termine di questa sommaria presentazione, ci preme riaffermare alcune caratteristiche salienti della DM salisburghese:

- si tratta di un metodo induttivo che, tramite la sintesi numerica di dati empirici molto variegati, porta alla scoperta di regolarità e leggi geolinguistiche, nascoste a prima vista nei dati analizzati;
- si tratta di un metodo orientato verso l'analisi quantitativa dello spazio, che proprio per questo motivo considera la cartografazione (o visualizzazione) generale dei suoi risultati numerici come *via regia* della propria euristica;
- si tratta di un metodo che, per via della sua impostazione comparativa intrinseca, rimane sempre aperto a cooperazioni interdisciplinari di qualsiasi natura;
- si tratta di un metodo, la cui base empirica è costituita dagli atlanti linguistici propriamente detti ed anche da analoghe raccolte empiriche;
- si tratta di un metodo che, nato in seno alla romanistica, rientra nella tradizione della geografia linguistica classica e si vanta di conciliare armonicamente saperi tradizionali ed innovazioni metodiche nonché tecnologiche di recente data.

L'abbondanza euristica dell'*output* iconico della DM impone, ai suoi seguaci di fare ricorso ad una filosofia mediatica nuova per rimediare alle insufficienze materiali della diffusione dei suoi risultati in forma meramente cartacea. Dal momento che questi problemi possono ora essere facilmente risolti tramite il nostro programma VDM ("Visual DialectoMetry"), faccio, a guisa di conclusione, la duplice offerta di spedire, da un lato, un CD-ROM con il programma VDM ed i dati del progetto sopra descritto, nonché d'impartire, dall'altro, nella stessa Salisburgo un tirocinio-VDM di un paio di giorni a tutti i colleghi interessati.

8. Ringraziamenti

- Sussidio finanziario: "Fonds zur Förderung der wissenschaftlichen Forschung in Österreich" (FWF), Vienna [Fondazione austriaca per la ricerca scientifica] (progetto di ricerca no. 18365).
- La tassazione ed immissione informatica delle 3 911 carte di lavoro (CL) è stata eseguita dai seguenti romanisti salisburghesi: Emese Lörincz, Christine Greil, Julia Hattinger, Daniel Rötzer, Mirjam Högl, Gertraud Klingler e Xavier Casassas.
- Ideazione, creazione e aggiornamento continuo del programma VDM (Visual DialectoMetry): Edgar Haimerl, Seattle (Stati Uniti d'America).

- Realizzazione dei grafici in bianco nero: Werner Goebel, Vienna.
- Revisione stilistica del testo italiano: Adelaide Fiocchi Baehr, Salisburgo.

A tutte le persone citate, nonché all'ente di ricerca FWF, vada l'espressione della mia sincera e profonda gratitudine.

Hans Goebel
e-mail <Hans.Goebel@sbg.ac.at>

BIBLIOGRAFIA

- AIS: JABERG K. & JUD J. (eds.), *Sprach- und Sachatlas Italiens und der Südschweiz*, Zofingen, Ringier, 1928-1940, 8 vol. (ristampa: Nendeln [Liechtenstein], Kraus, 1971).
- ALF: GILLIÉRON J. & EDMOND E. (eds.), *Atlas linguistique de la France*, Paris, Champion, 1902-1910, 10 vol. (ristampa: Bologna, Forni, 1968).
- BAUER R., *Dialektometrische Einsichten. Sprachklassifikatorische Oberflächenmuster und Tiefenstrukturen im lombardisch-venedischen Dialektraum und in der Rätoromania*, San Martin de Tor, Istitut ladin Micurà de Rü, 2009 (Ladinia Monographica 01).
- GOEBL H., Éléments d'analyse dialectométrique (avec application à l'AIS), in *Revue de Linguistique Romane* 45, 1981, 349-420.
- GOEBL H., *Dialektometrie. Prinzipien und Methoden des Einsatzes der Numerischen Taxonomie im Bereich der Dialektgeographie*, Österreichische Akademie der Wissenschaften, Denkschriften, philosophisch-historische Klasse, vol. 157, Wien, 1982, 1-123 [41 tavole].
- GOEBL H., Parquet polygonal et treillis triangulaire: les deux versants de la dialectométrie interponctuelle, in *Revue de Linguistique Romane* 47, 1983, 353-412.
- GOEBL H., *Dialektometrische Studien. Anhand italo-romanischer, rätoromanischer und gallo-romanischer Sprachmaterialien aus AIS und ALF*, Tübingen, Niemeyer, 1984a, (Beihefte zur Zeitschrift für Romanische Philologie, vol. 191-193).
- GOEBL H., Lineamenti di dialettometria (con applicazione all'AIS), in *Guida ai dialetti veneti* VI, 1984b, 7-53.
- GOEBL H., Una classificazione gerarchica di dati geolinguistici tratti dall'AIS. Saggio di dialettometria dendrografica, in *Rivista di dialettologia italiana* 16, 1992, 67-80.
- GOEBL H., Analyse dialectométrique des structures de profondeur de l'ALF, in *Revue de Linguistique Romane* 66, 2002, 5-63.
- GOEBL H., Regards dialectométriques sur les données de l'Atlas linguistique de la France (ALF): relations quantitatives et structures de profondeur, in *Estudis Romànics* 25, 2003a, 59-120.

GOEBL H., Entwicklung der romanischen Sprachgrenzen: Italomania und Ostalpenraum, in G. ERNST, M.-D. GLESSGEN, Chr. SCHMITT & W. SCHWEICKARD (eds.), *Romanische Sprachgeschichte. Histoire linguistique de la Romania. Ein internationales Handbuch zur Geschichte der romanischen Sprachen. Manuel international d'histoire linguistique de la Romania*, Berlin, New York, de Gruyter, 2003b, vol. 1, 631-645.

GOEBL H., La dialectométrie corrélatrice: un nouvel outil pour l'étude de l'aménagement dialectal de l'espace par l'homme, in *Revue de Linguistique Romane* 69, 2005, 321-367.

GOEBL H., Sprachgeographische Streifzüge durch das Netz des Sprachatlasses AIS, in *Ladina* XXXI, 2007, 187-271.

GOEBL, H., La dialettometrizzazione integrale dell'AIS. Presentazione dei primi risultati, in *Revue de Linguistique Romane* 72, 2008a, 25-113.

GOEBL H., Sur le changement macrolinguistique survenu entre 1300 et 1900 dans le domaine d'Oïl. Une étude diachronique d'inspiration dialectométrique, in *Dialectologia* 1, 2008b, 3-43.

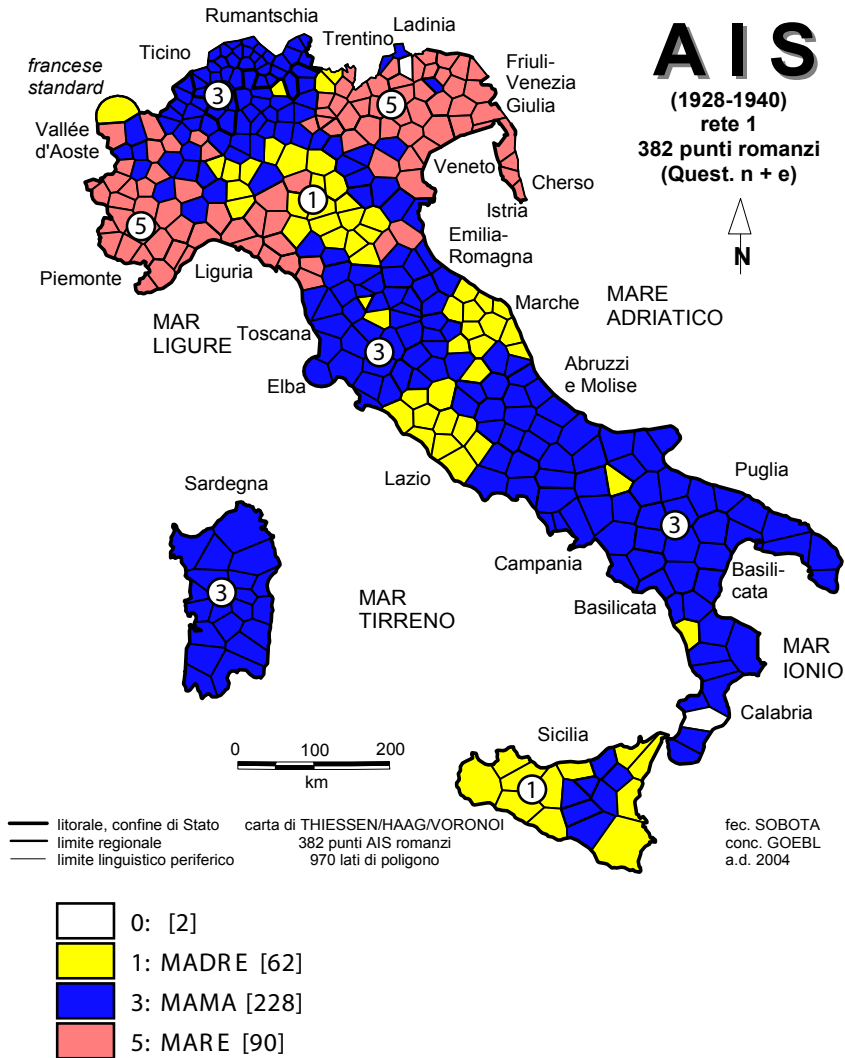
OKABE A., BOOTS B. & SUGIHARA K., *Spatial Tessellations. Concepts and Applications of Voronoi Diagrams*, Chichester, New York, Brisbane, Toronto, Singapore, Wiley, 1992.

VIDESOTT P., *Padania scrittologica: analisi scrittologiche e scrittometriche di testi in italiano settentrionale antico dalle origini al 1525*, Tübingen, Niemeyer, 2009 (Beihefte zur Zeitschrift für Romanische Philologie, vol. 353).

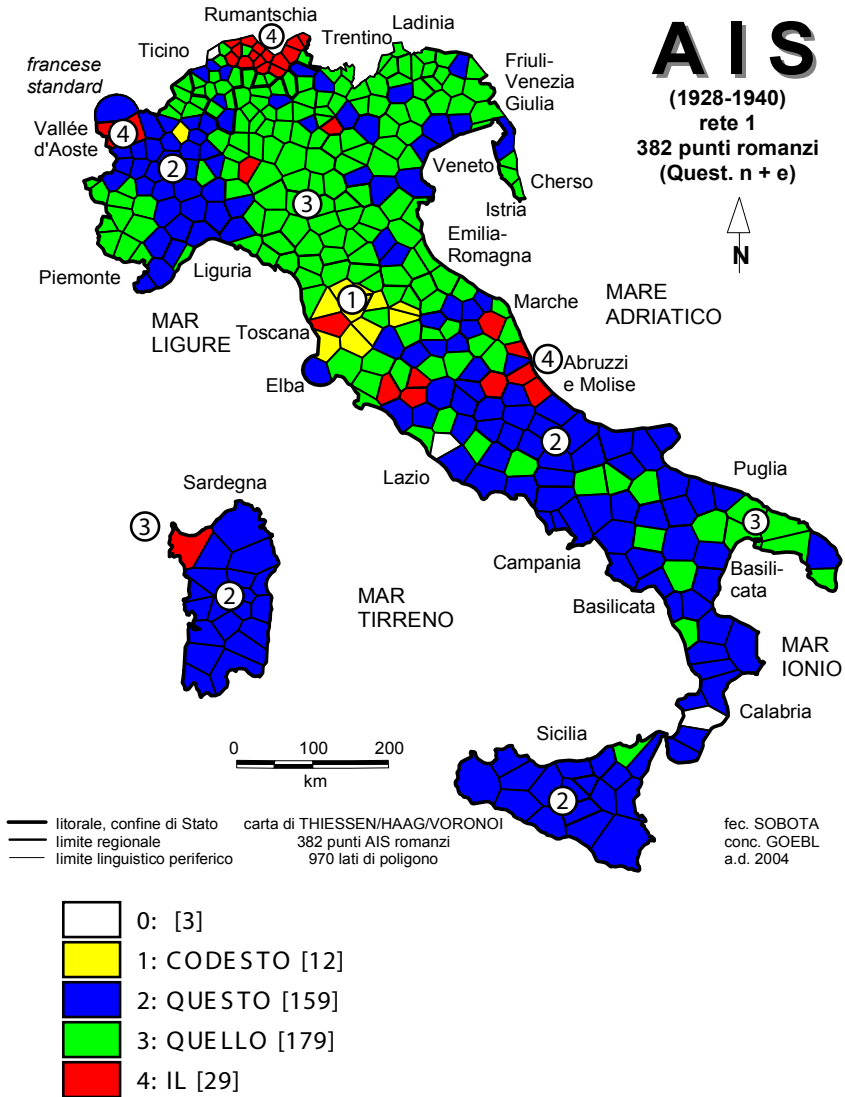
ABBREVIAZIONI

- CAF: coefficiente di asimmetria di Fisher
- CDS: carta di similarità
- CL: carta di lavoro
- DC: dendrema-corema
- DM-S: dialettometria praticata secondo i principi della “Scuola dialettometrica di Salisburgo”
- IRD: Indice Relativo di Distanza
- IRI: Indice Relativo di Identità

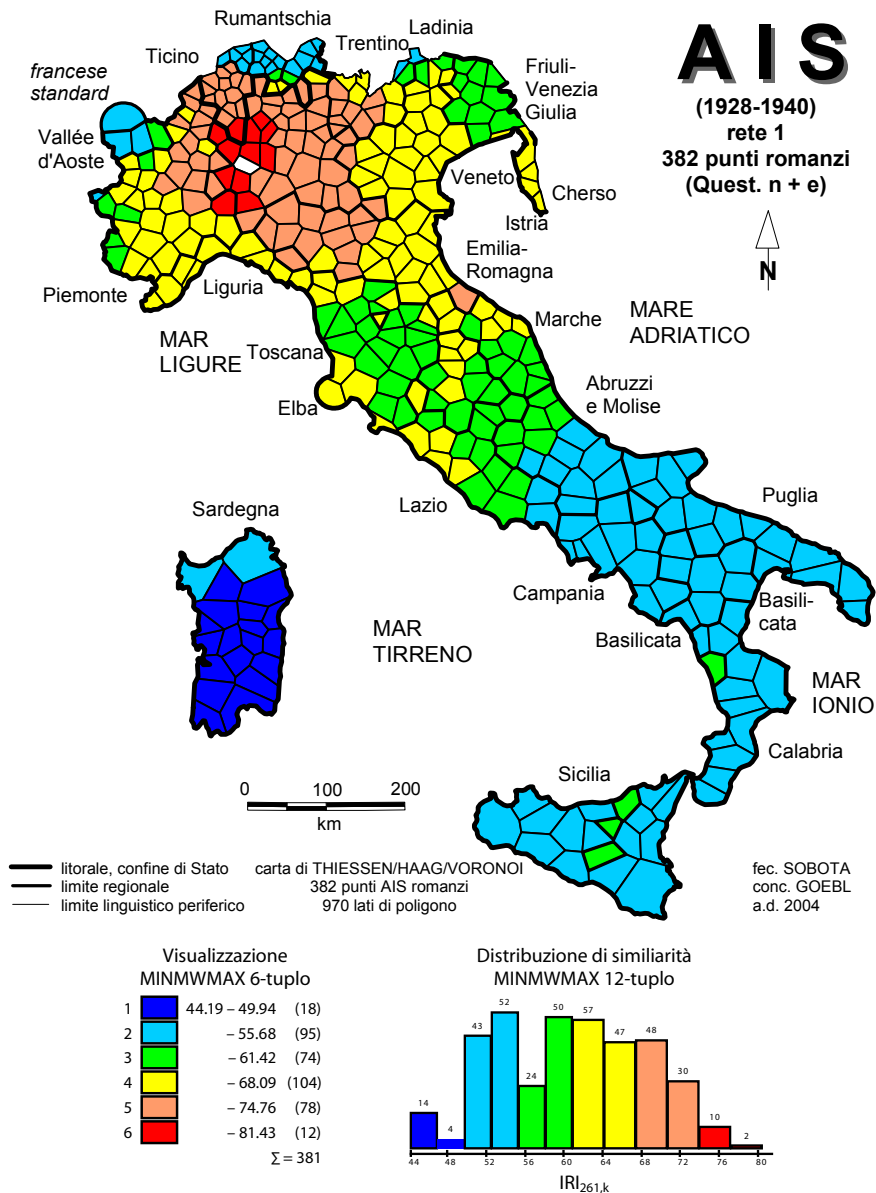
TAVOLE



Tav. 1: Una delle 1225 carte di lavoro *lessicali*: ripartizione geografica delle denominazioni della “madre” (secondo AIS 8 *sua madre*).



Tav. 2: Una delle 289 carte di lavoro *morfologiche*: ripartizione geografica delle realizzazioni dialettali dell'aggettivo dimostrativo *questo* (secondo AIS 42 *questo bambino*)

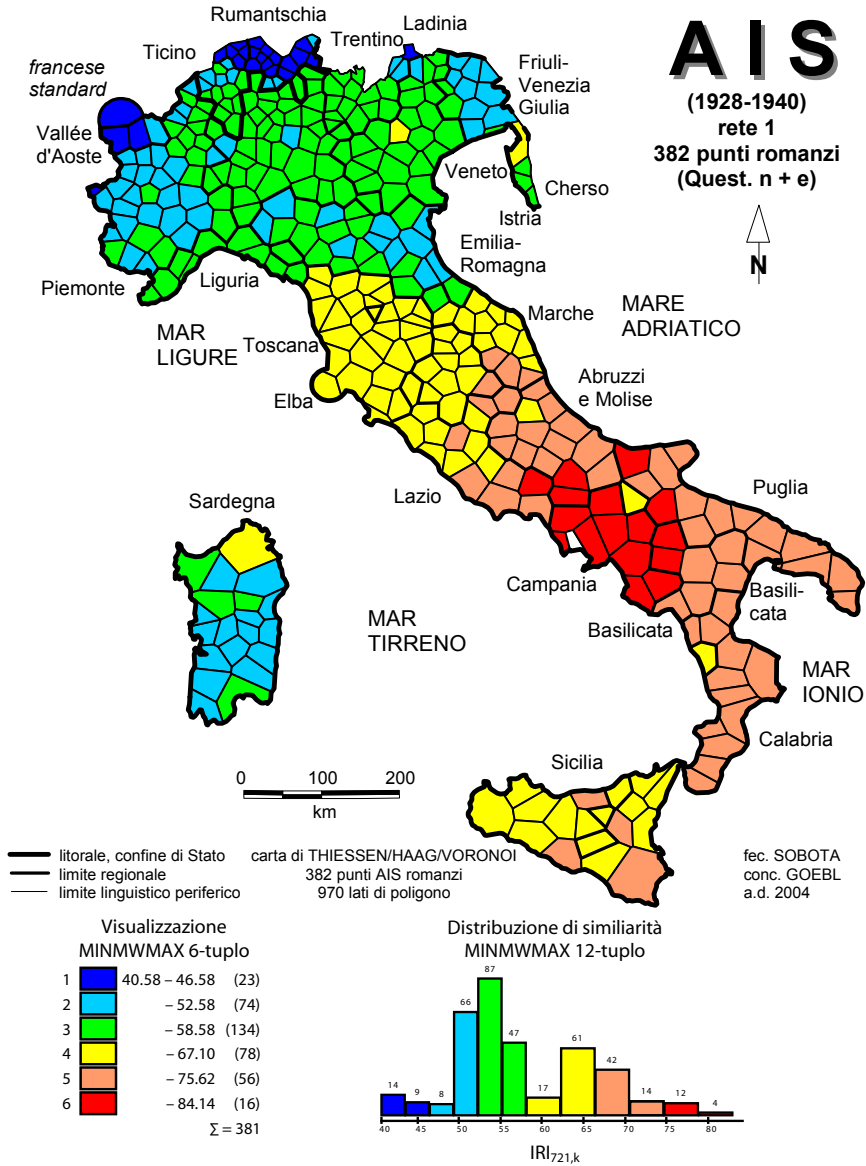


Tav. 3: Carta di similarità relativa al punto di riferimento 261 (Milano)

Indice di similarità: IRI_{jk}

Corpus: 3 911 carte di lavoro (tutte le categorie linguistiche)

Algoritmo d'intervallizzazione: MINMWMAX 6-tuplo

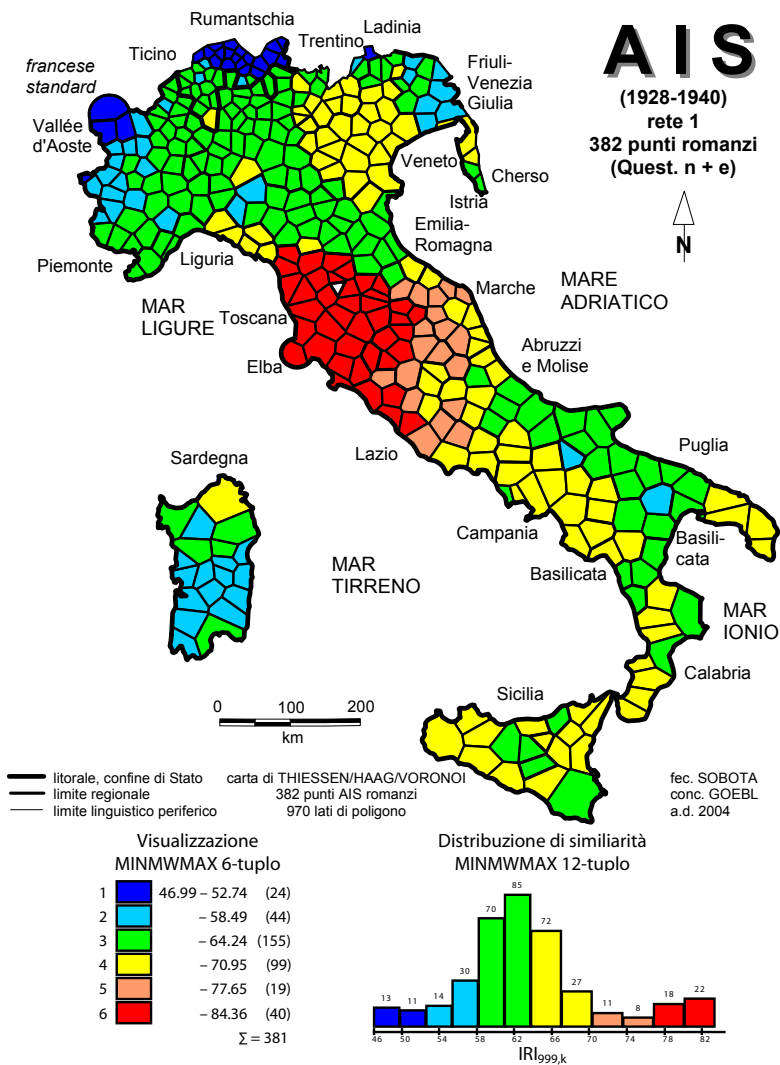


Tav. 4: Carta di similarità relativa al punto di riferimento 721 (Napoli)

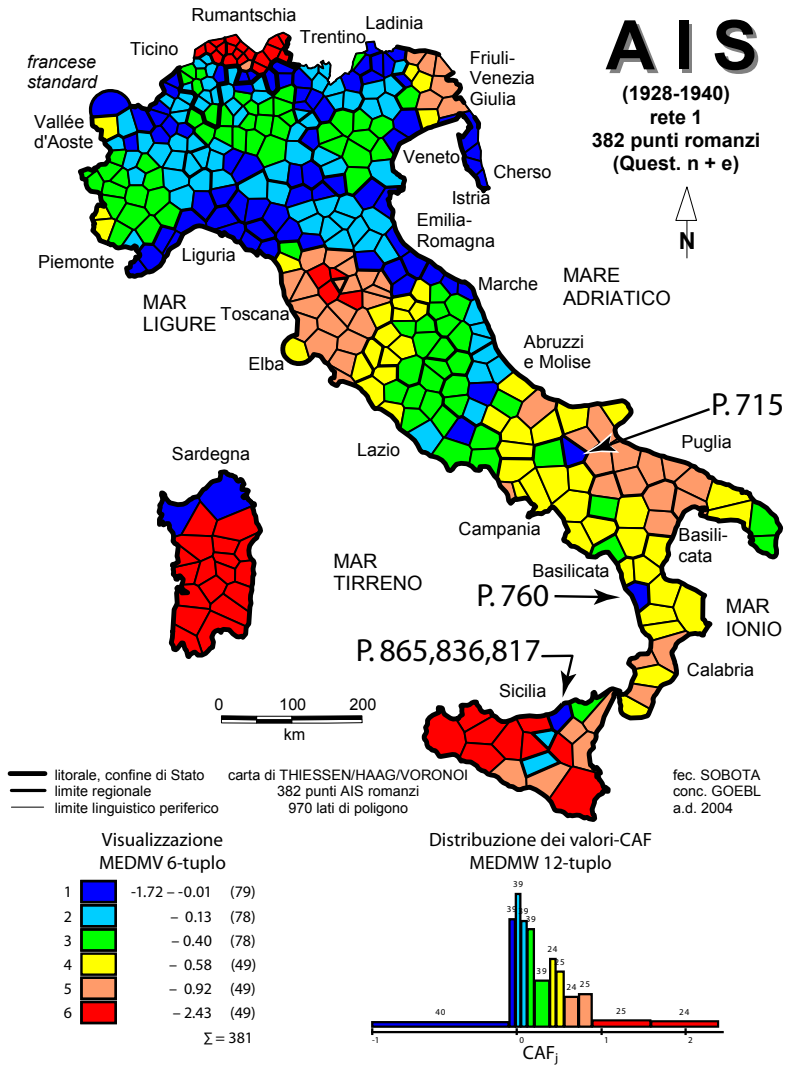
Indice di similarità: IRI_{jk}

Corpus: 3 911 carte di lavoro (tutte le categorie linguistiche)

Algoritmo d'intervallizzazione: MINMWMAX 6-tuplo



Tav. 5: Carta di similarità relativa al punto di riferimento 999 (italiano standard)
Indice di similarità: IRI_{jk}
Corpus: 3 911 carte di lavoro (tutte le categorie linguistiche)
Algoritmo d'intervallizzazione: MINMWMAX 6-tuplo

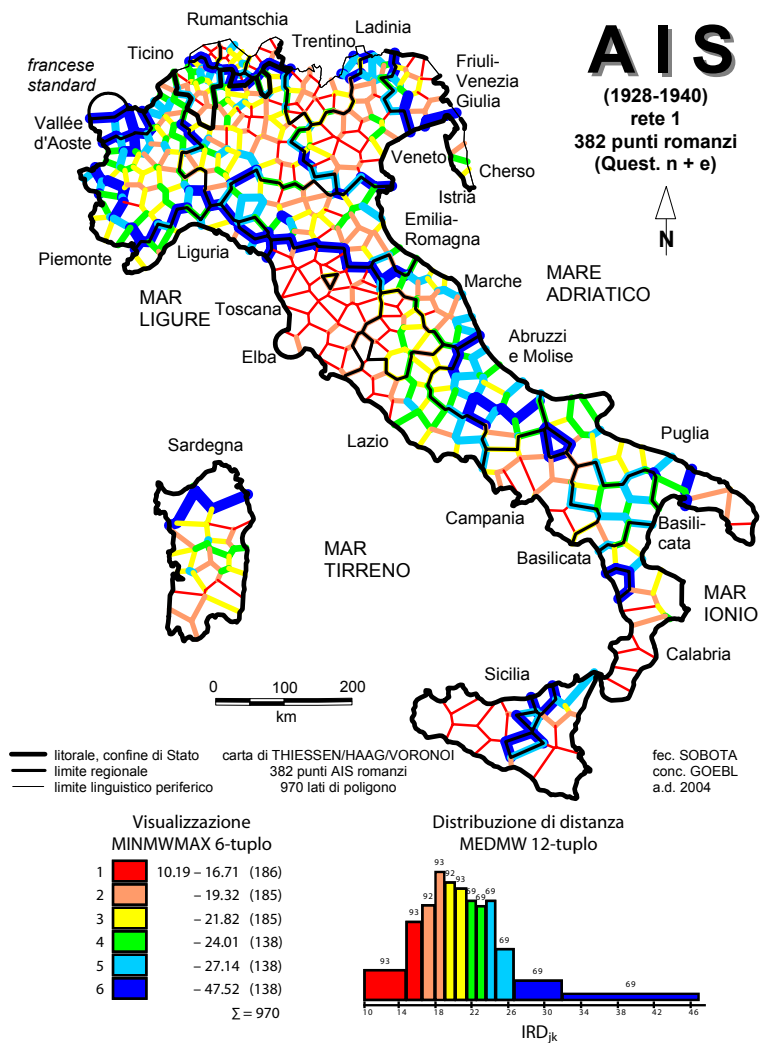


Tav. 6: Sinossi coropletrica dei coefficienti di asimmetria di Fisher (CAF) di 382 distribuzioni di similarità

Indice di similarità: IRI_{jk}

Corpus: 3 911 carte di lavoro (tutte le categorie linguistiche)

Algoritmo d'intervallizzazione: MEDMW 6-tuplo

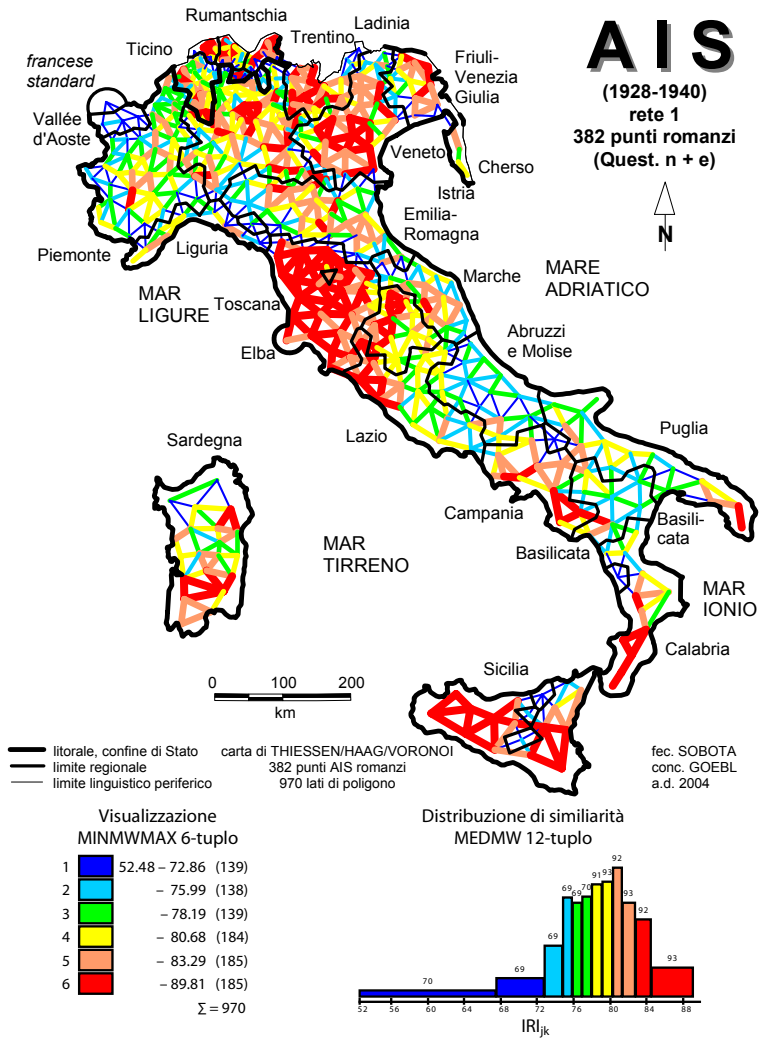


Tav. 7: Carta isoglottica

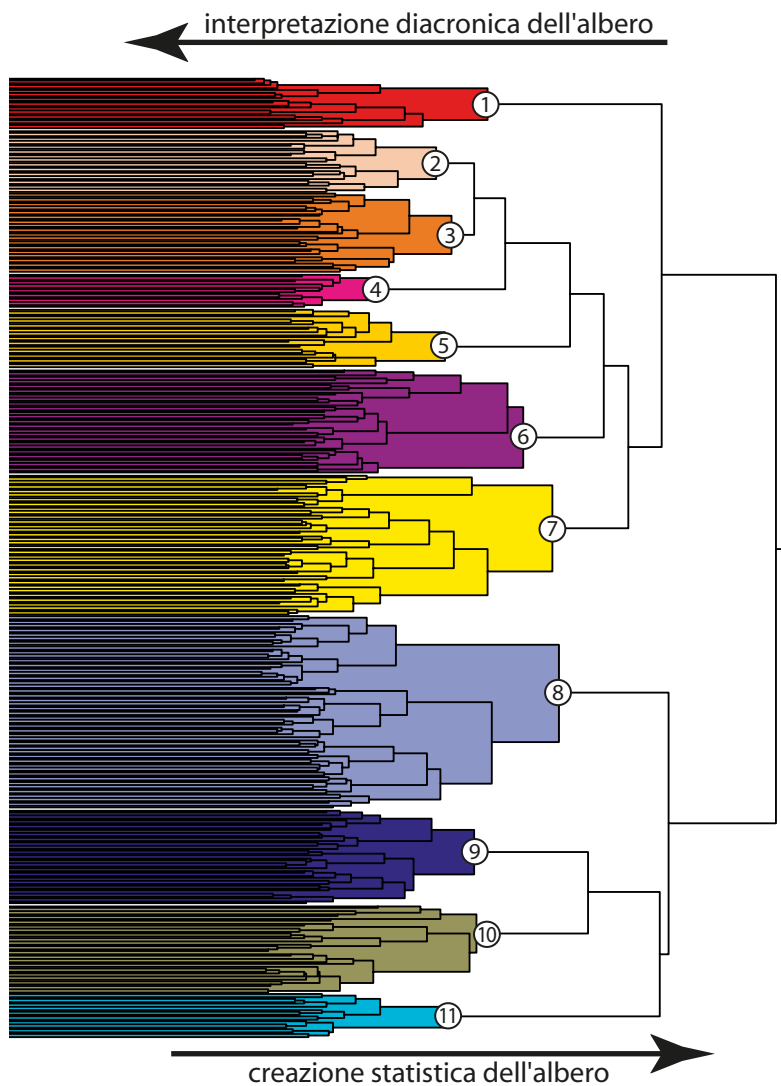
Indice di distanza: IRD_{jk}

Corpus: 3 911 carte di lavoro (tutte le categorie linguistiche)

Algoritmo d'intervallizzazione: MEDMW 6-tuplo



Tav. 8: Carta a raggi
Indice di similarità: IRI_{jk}
Corpus: 3 911 carte di lavoro (tutte le categorie linguistiche)
Algoritmo d'intervallizzazione: MEDMW 6-tuplo

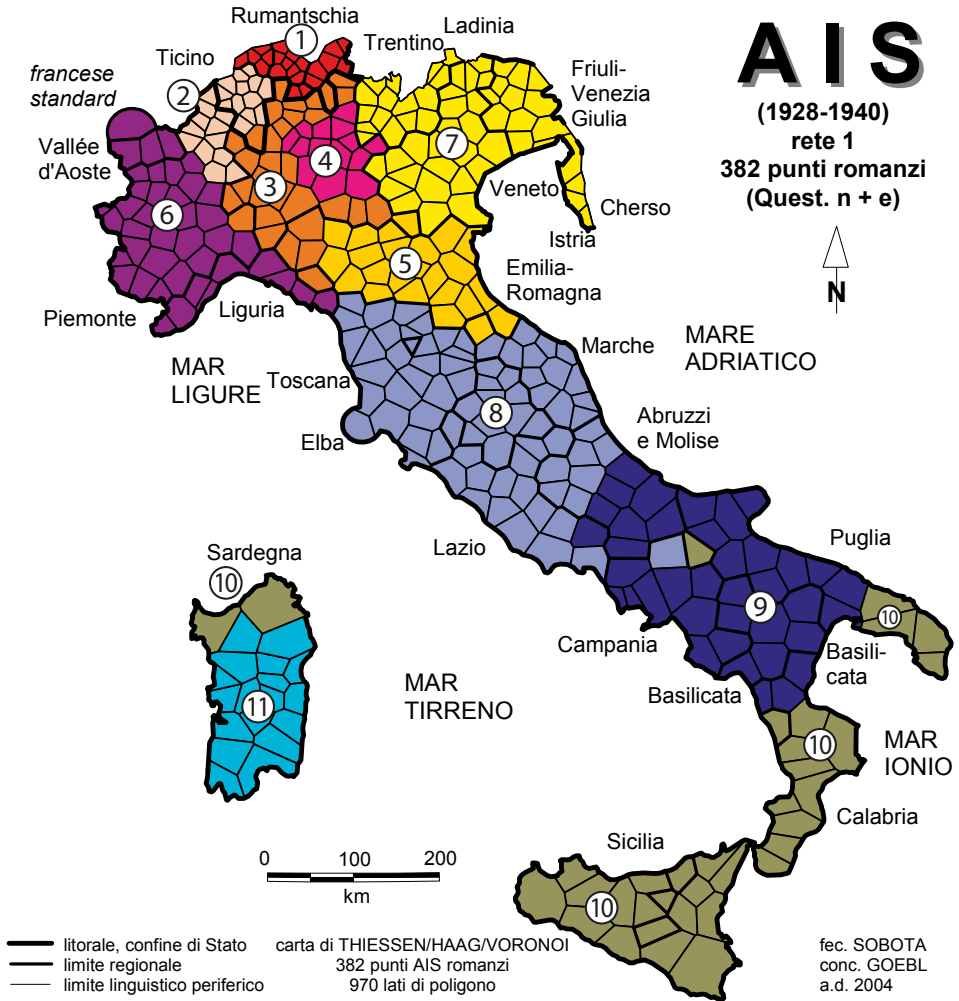


Tav. 9: Classificazione gerarchica agglomerativa (CGA) secondo il metodo proposto da Joe Ward, Jr.

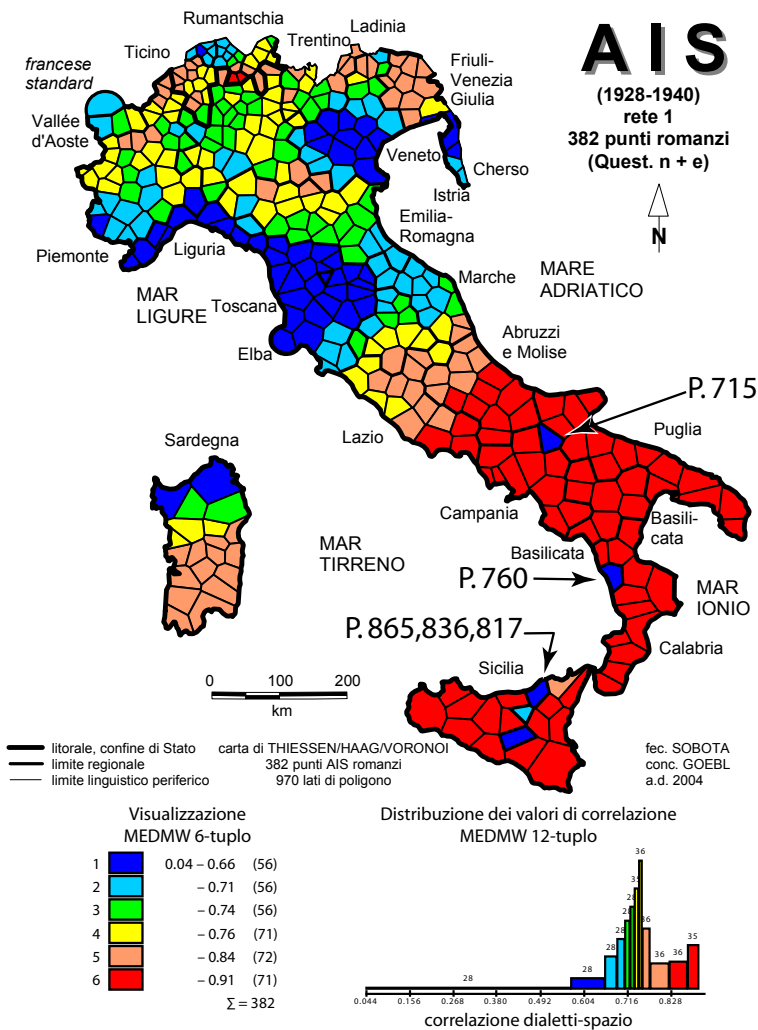
Indice di similarità: IRI_{jk}

Corpus: 3 911 carte di lavoro (tutte le categorie linguistiche)

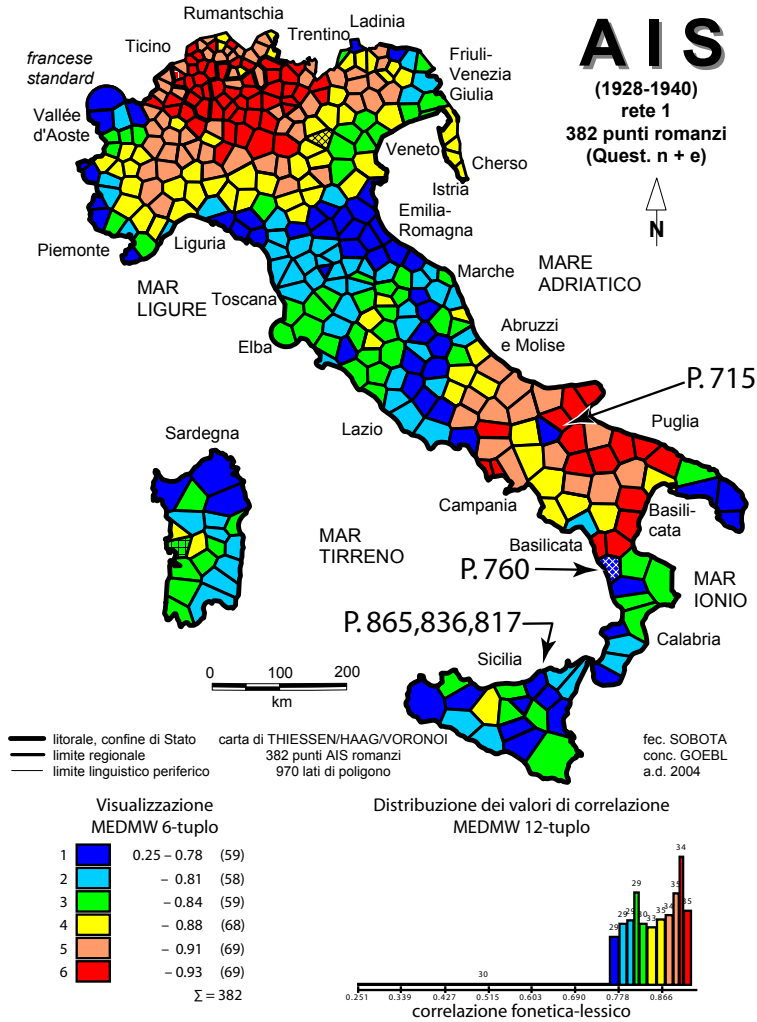
Numero dei dendremi retinati: 11



Tav. 10: Spazializzazione della struttura dendrematica della tav. 9
Numero dei coremi retinati: 11



Tav. 11: Carta a correlazioni [mediante $r(\text{BP})$]: similarità linguistica [*generale*] (IRI_{jk}) e prossimità geografica (euclidea)
 Corpus (linguistico): 3 911 carte di lavoro (tutte le categorie linguistiche)
 Algoritmo d'intervallizzazione: MEDMW 6-tuplo



Tav. 12: Carta a correlazioni [mediante $r(BP)$]: due misurazioni di similarità linguistica (IRI_{jk}): fonetica (A) e lessico (B)

Corpus A: 1766 carte di lavoro *fonetiche*

Corpus B: 1225 carte di lavoro *lessicali*

Algoritmo d'intervalizzazione: MEDMW 6-tuplo