

Mare Loquens

Études d'étymologie et de géolinguistique romanes
à la mémoire de Vojmir Vinja (1921-2007)



Mare Loquens. Études d'étymologie et de géolinguistique
romanes à la mémoire de Vojmir Vinja (1921-2007)

Posebna izdanja Centra za jadranska
onomastička istraživanja – Knjiga 3

UREDNICI
José Enrique Gargallo Gil
Nikola Vuletić

NAKLADNIK
Sveučilište u Zadru

ZA NAKLADNIKA
Ante Uglešić, rektor

POVJERENSTVO ZA IZDAVAČKU DJELATNOST
SVEUČILIŠTA U ZADRU
Josip Faričić (predsjednik)

RECENZENTI
Elisabetta Carpitelli
Nada Vajs Vinja

GRAFIČKO OBLIKOVANJE
Ljubica Marčetić Marinović

PRIJELOM
Denis Gospić

TISAK
Zrinski d.d., Čakovec

NAKLADA
200 primjeraka

Izdavanje ove knjige pomogli su Ministarstvo znanosti,
obrazovanja i sporta Republike Hrvatske,
Institut Ramon Llull i Xunta de Galicia

ISBN 978-953-331-044-2

CIP zapis dostupan u računalnom katalogu Znanstvene
knjižnice Zadar pod brojem 140617072

Mare Loquens

Études d'étymologie et de géolinguistique romanes
à la mémoire de Vojmir Vinja (1921-2007)

Publiées par José Enrique Gargallo Gil et Nikola Vuletić

SVEUČILIŠTE U ZADRU, Zadar 2013.

Sommaire

JOSÉ ENRIQUE GARGALLO GIL, NIKOLA VULETIĆ Présentation	7
AUGUST KOVAČEC L'apport scientifique du professeur Vinja dans le domaine de la linguistique romane	9
Études	
JOSÉ ENRIQUE GARGALLO GIL Mar y meteorología popular en la Romania	25
JOÃO SARAMAGO La nomenclature des poissons aux Açores : aires de distribution et motivation	47
XOSÉ AFONSO ÁLVAREZ PÉREZ Oltre quella "foce stretta": un viaggio geolinguistico attraverso l'Atlantico portoghese	73
JOAN VENY Vocabulaire ichtyonymique et nautique catalan dans la Romania	93
GERMÀ COLÓN De sinonímica catalana	111
MARIA PILAR PEREA Tecnologías informáticas y variación dialectal	123
MICHEL CONTINI Geolinguistica della Sardegna. Dalla Carta corografica dello Spano al progetto AliMuS	141
HANS GOEBL Introduzione ai problemi e metodi della dialettometria, applicati ai dati dell'Atlante italo-svizzero AIS (con particolare riguardo ai dialetti del Veneto e dell'Istria)	171
RADA COSSUTTA Lessico d'origine slovena nel dialetto triestino	227
METKA FURLAN Su un prestito romanzo nelle parlate dell'Istria slovena	235
SUZANA GILJANOVIČ Prestiti romanzi ibridi nelle parlate istroslovene	249
SANDRA TAMARO Noterelle etimologiche e semantiche sull'eredità romanza (veneziana) nel lessico ciacavo di Bogliuno d'Istria	263

BARBARA BURŠIĆ GIUDICI Etimologie istriote: la casa	285
NIKOLA VULETIĆ Les mones de Pasqua de l'Adriatic croat: un recorregut geolingüístic i etimològic	299
ORSAT LIGORIO Uno sguardo ai relitti dalmato-romanzi in -CULU e -TULU	317
FRANCESCO AVOLIO Dalle Marche al Salento: L'Italia peninsulare adriatica e alcune questioni geolinguistiche	323
ANTONIO ROMANO La caratterizzazione greco-romanza della fitonimia popolare salentina: spunti metodologici	333
Mises en relief	
GORAN FILIPI Come comunicano gli istriani: è possibile una koinè assoluta in un ambiente multiculturale, multilingue e multi-etnico (sull'esempio istriano)?	363
VLADIMIR SKRAČIĆ Centre de recherche onomastique de l'Adriatique	377

Présentation

Le 15 juin 2007 s'éteint à Zagreb à l'âge de 85 ans, au terme d'une implacable maladie, Vojmir Vinja. Continueur des études sur les aspects étymologiques des contacts slavo-romanes de Petar Skok (1881-1956) et des études géolinguistiques de Mirko Deanović (1980-1984), il a été, avec son collègue Žarko Muljačić (1922-2009) et son élève Pavao Tekavčić (1931-2007), le plus illustre représentant de l'école romaniste croate de la seconde moitié du xx^e siècle. Son influence a été décisive pour le développement des études romanes en Croatie, bien au-delà de la Faculté des Lettres de Zagreb, où il avait enseigné la linguistique romane jusqu'à sa retraite en 1986. C'est ainsi qu'aujourd'hui les centres d'études romanes des universités de Zadar et de Pula, avec leur orientation vers l'étude étymologique du lexique maritime, garantissent la continuité de l'enseignement de Vojmir Vinja. À l'Université de Pula quatre atlas linguistiques ont été publiés depuis l'an 1998 : *Atlas linguistique istriote*, *Atlas linguistique istro-roumain*, *Atlas linguistique istro-vénitien* et *Atlas linguistique de la terminologie maritime des parlers d'Istrie*. Le *Centre de Recherche Onomastique de l'Adriatique* de l'Université de Zadar a publié une collection dédiée à la toponymie de l'archipel de la Dalmatie septentrionale, extrêmement riche des survivances de l'ancien dalmate et de différentes influences italo-romanes.

Avec ce volume, l'Université de Zadar veut rendre hommage au maître qui, dans les années 50 du siècle passé, fut aussi un des fondateurs du « Séminaire roman » de la Faculté des Lettres de Zadar, dans une ville ravagée par les bombardements de 1943-1944. Ce volume réunit des contributions autour des deux thèmes préférés de Vinja – l'étude étymologique et géolinguistique du lexique des côtes de la Méditerranée, mais aussi des autres côtes du monde roman, et l'étude des contacts linguistiques slavo-romans.

Nous sommes reconnaissants aux romanistes de renommée internationale qui ont bien voulu accepter de participer dans ce volume, à côté des jeunes romanistes des deux bords de la Romania européenne, la côte atlantique et la côte adriatique. Nous remercions également l'*Institut Ramon Lull* et la *Xunta de Galicia* pour leur généreux soutien.

Pour terminer cette brève présentation, nous aimerions évoquer les mots du regretté Joan Bastardas (*Diàlegs sobre la meravellosa història dels nostres mots*, Barcelona, 1996, p. 47):

Els diccionaris i atles lingüístics que parteixen no dels mots, sinó de les coses i dels conceptes, ens diuen moltes coses sobre com és de gran la diversitat amb què els homes analitzen i contemplen tot allò que és objecte de comunicació. I també quin és el pes de tota mena de circumstàncies i condicionaments, però rarament ens mostraran què hi ha de sistemàtic en les maneres de dir. Els estudis històrics i comparatius del lèxic fet partint de camps semàntics solament ens diran alguna cosa sobre el caràcter sistemàtic de les denominacions quan aquest caràcter veritablement existeix.

On est en droit de dire que Vojmir Vinja, avec son ouvrage *Jadranska fauna : etimologija i struktura naziva* (*La faune adriatique : étymologie et structure des dénominations*), avait réussi à produire une formidable étude de l'extrême diversité géolinguistique du lexique ichtyonymique et en même temps à montrer « què hi ha de sistemàtic », au-delà des limites de son bien-aimé Adriatique oriental. La traduction de cet ouvrage à une des langues de diffusion internationale reste un desideratum de la linguistique moderne.

Barcelone / Zadar, juillet 2013

José Enrique GARGALLO GIL (Université de Barcelone)
Nikola VULETIĆ (Université de Zadar)

Hans Goebel
Universität Salzburg

Introduzione ai problemi e metodi della dialettometria, applicati ai dati dell'Atlante italo-svizzero AIS (con particolare riguardo ai dialetti del Veneto e dell'Istria)

1. Brevissima presentazione delle impostazioni teorica, empirica e metodica della dialettometria (DM)

Alla pari di molte altre scienze provviste del suffisso *-metria*, la DM costituisce una disciplina di stampo induttivo che attraverso la sintesi quantitativa dei dati di atlanti linguistici (i.e. *dialettali*) tradizionali si propone di individuare e studiare ordinamenti – anzi leggi – spaziali nascosti nella massa dei dati degli stessi atlanti linguistici. Ricordiamo che in genere le singole carte degli atlanti linguistici offrono spesso strutturazioni altamente complesse, per non dire caotiche, sicché, negli ultimi cento anni, si è instaurata – da parte di non pochi filologi e linguisti – la credenza, anzi convinzione, che la distribuzione spaziale dei fatti linguistici sia esente da regolarità intrinseche o di maggior rilievo. Questa falsa credenza è stata progressivamente smentita, dagli anni 70 del secolo scorso, dalla ricerca-DM allora in gestazione, tanto per la geolinguistica sincronica quanto per la sua sorella diacronica¹.

Nel quadro della “Scuola Dialettometrica di Salisburgo”² (SDMS) si dà per assunto che le suddette leggi siano l’emanazione diretta di un atteggiamento speciale dei locutori *dia-* o *basilettali* rispetto allo

¹ Per un’applicazione medievistica della DM cf. Goebel 2008b (passim).

² Quanto ad altri lavori realizzati in seno alla SDMS rinvio ai contributi dialettometrici di Roland Bauer (2009) e Paul 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.

spazio geografico da loro abitato, chiamato da un paio di anni in seno alla SDMS “gestione basilettale dello spazio da parte dell’Homo loquens”. Si dà ugualmente per assunto che le leggi individuate nella dimensione dello *spazio* rappresentino il riscontro diretto delle leggi vigenti nella dimensione del *tempo*, scoperte e ampiamente discusse dai Neogrammatici già alla fine dell’Ottocento.

Dal punto di vista meramente operativo, la DM di Salisburgo si serve, sul versante *qualitativo*, di tutta la panoplia teorica, concettuale e metodica della geografia linguistica tradizionale (a base d’atlanti) e, su quello *quantitativo*, di una ricca gamma di metodi numerici e grafici appositamente selezionati in seno alla tassonomia numerica³ ed alla geografia 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. Il programma VDM permette l’esecuzione informatizzata di pressoché tutti gli anelli della catena metodica della SDMS: 1) immissione dei dati in una banca dati-*Access*, 2) calcolo delle similarità e distanze, 3) calcolo degli schemi dendrografici, 4) calcolo di tutti i parametri cartografici e 5) realizzazione istantanea di tutte le visualizzazioni previste: si veda in merito la figura 1.

2. La dialettometrizzazione dei dati integrali dell’AIS

Dopo le nostre prime esperienze con la parte settentrionale della rete e le carte originali (di stampa prevalentemente *lessicale*) dei volumi I, II e IV dell’AIS⁴, grazie anche agli ottimi risultati ottenuti nel quadro di una dialettometrizzazione della rete integrale dell’atlante linguistico francese ALF⁵, ci siamo proposti, negli anni 2007-2009, di realizzare un progetto analogo per la totalità della rete e delle carte linguistiche dell’AIS⁶.

Come è noto, nei suoi otto volumi di formato in-folio, l’AIS dispone di 1 705 carte linguistiche i cui dati sono distribuiti su una rete di 407 punti di rilevamento (di tipo romanzo e non-romanzo). Il nostro proposito era di analizzare la variabilità geolinguistica reperibile su tutte le carte

3 Cf. in merito l’opera fondamentale di Sneath / Sokal 1973.

4 Cf. in merito soprattutto Goebel 1981; 1982 e 1984a.

5 Cf. in merito i nostri lavori del 2002 e del 2003a.

6 Si vedano le nostre relazioni di lavoro del 2007 e del 2008a. Solo i nostri contributi del 2010 e del 2011 si avvalgono dell’effettivo finale di 3 911 carte di lavoro (CL) complessive.

dell’AIS secondo le categorie linguistiche della *fonetica*, della *morfologia*, della *sintassi* e del *lessico*, considerando non solo le *carte* propriamente dette dell’AIS (carte 1-1 681), ma anche tutti i *materiali intabellati* presentati alla fine dell’ottavo volume (“Konjugationstabellen”, 1 682-1 705).

3. Dalla fonte originale (AIS) alla matrice dei dati

Si veda la figura 1.

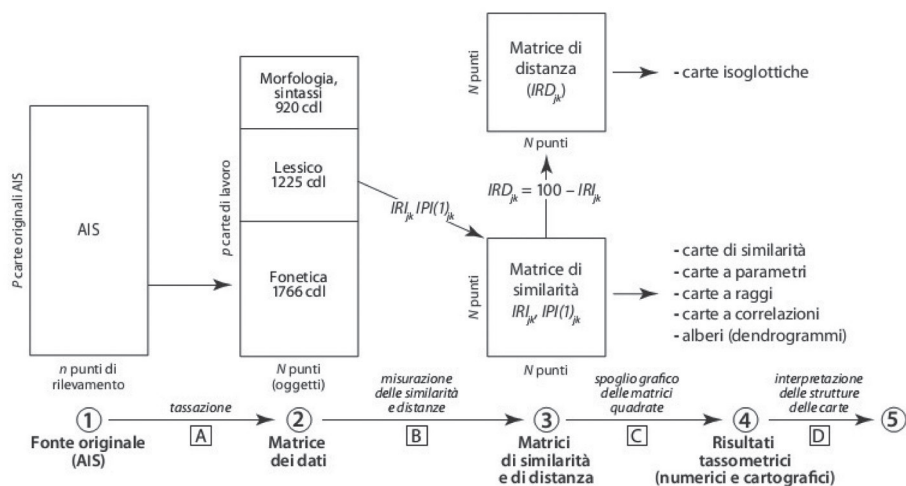


FIGURA 1. Diagramma dei metodi di misurazione e di visualizzazione utilizzati dalla “Scuola Dialettometrica di Salisburgo” (SDMS)

Il rilevamento della variabilità linguistica reperibile sulle carte dell’AIS – chiamata nel gergo della SDMS “tassazione” – è stato effettuato secondo i principi e metodi della grammatica storica tradizionale (per la *fonetica*, la *morfologia* e la *sintassi*) nonché dell’onomasiologia romanza classica (per il *lessico*). A causa della lacunosità di molte carte dell’AIS e dell’uso sistematico di un *questionario ridotto* da parte dei rilevatori dell’AIS in circa trenta località della rete esaminata, non era possibile estendere l’analisi tassatoria a tutte le carte dell’AIS.

Per assicurare la completezza della matrice dei dati si doveva quindi scartare più di un quarto delle carte AIS originali⁷, nonché 29 dei 407

7 Su 1 705 carte e tabelle linguistiche a disposizione sono state analizzate solo 1 244 unità, cioè 73% dell’effettivo totale.

punti di rilevamento. I restanti 378 punti venivano però completati da due punti artificiali corrispondenti all'italiano (P. 999) ed al francese standard (P. 998) e dalle seconde inchieste fatte da Paul Scheuermeier nelle seguenti città: Torino, Milano, Venezia, Bologna e Firenze.

Per ovvie ragioni si è proceduto alla cancellazione di tre punti non romanzi ubicati nell'Italia meridionale (greco: PP.-AIS 792 e 748; albanese: P.-AIS 751). In questo modo si è stabilita una rete con 382 punti, che rappresenta la base di tutte le visualizzazioni di questo contributo.

Ecco il bilancio consuntivo dei nostri lavori di tassazione per la *fonetica*, il *lessico*, la *morfo-sintassi* nonché il *corpus integrale*⁸:

Fonetica			
carte originali esaminate	carte di lavoro	tassati (o aree tassatorie)	tassati/carta di lavoro
288	1 766	15 124	8,56

Lessico			
carte originali esaminate	carte di lavoro	tassati (o aree tassatorie)	tassati/carta di lavoro
1 076	1 225	21 989	17,95

Morfologia, Sintassi			
carte originali esaminate	carte di lavoro	tassati (o aree tassatorie)	tassati/carta di lavoro
349	920	7 333	7,97

Corpus integrale			
carte originali esaminate	carte di lavoro	tassati (o aree tassatorie)	tassati/carta di lavoro
1 244	3 911	43 564	11,38

La figura 1 dimostra la successione delle diverse tappe del trattamento-DM dei dati di un qualsiasi atlante linguistico. A ciascuna delle quattro tappe (A, B, C e D) ivi raffigurate corrispondono presupposti, metodi e finalità ben definiti. Si badi al fatto che l'ultima tappa (D) della catena-DM corrisponde alla valutazione teorica delle visualizzazioni-DM da parte del dialettometra e rappresenta quindi la quintessenza della procedura intera.

8 Si tratta dei risultati disponibili alla fine della dialettometrizzazione dell'AIS. Le nostre relazioni del 2007 e del 2008 poggiavano ancora su effettivi-CL minori: Goebel 2007: 1 746 CL; Goebel 2008a: 3 076 CL.

Il processo della “tassazione” – che nient’altro è che una forma di tipizzazione dei dati grezzi secondo determinati criteri scientifici – può essere applicato a tutte le categorie linguistiche, dalla fonetica alla morfologia e alla sintassi. Si deve quindi distinguere tra il livello delle *carte originali* dell’AIS e quello delle *carte tassate* (o tipizzate), battezzate da noi, già negli anni 70 del secolo scorso “*carte di lavoro*” (CL).

Volgiamo ora lo sguardo alle tavole 1 e 2: esse rappresentano due delle 3 911 CL da noi analizzate. La prima CL (tavola 1) è di carattere *fonetico*, la seconda CL (tavola 2) di carattere *lessicale*. I diversi tipi fonetici o lessicali (o “tassati” secondo la terminologia della SDMS) che compaiono sulle tavole 1 e 2, occupano aree provviste di estensioni ed anche di configurazioni geografiche molto diverse. Si veda, a titolo d’esempio, la tavola 2 dove l’area del tassato *zio* abbraccia 225 punti (o poligoni), quella del tassato *barba* 108 e quella del tassato *oncle* 9 punti. Si noti inoltre la presenza di dati mancanti (lacune): si evidenziano otto lacune sulla tavola 1 ed una sola lacuna sulla tavola 2.

Il contenuto *qualitativo* della matrice dei dati può quindi esser caratterizzato – in termini *quantitativi* – secondo il numero dei *tassati* analizzati, il quale ovviamente corrisponde a quello delle rispettive *aree tassatorie*.

È degno di nota ancora un altro fatto di carattere *quantitativo*: la variabilità interna (detta anche “polinimia”) delle CL secondo la frequenza dei rispettivi tassati. Mentre la tavola 1 mostra una CL 10-nima (*deca-nima*), la tavola 2 rinvia ad una CL 8-nima (*octo-nima*). Ovviamente il nostro corpus esclude le CL *mono-nime* che, essendo prive di *variazione* interna, non servono per analisi *variazionali*.

Ecco il bilancio consuntivo della frammentazione tassatoria dei nostri dati:

	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

Le regolarità inerenti ai dati atlantistici, menzionate già all’inizio di questo contributo, si manifestano anche a livello delle CL: si veda in merito la figura 2.

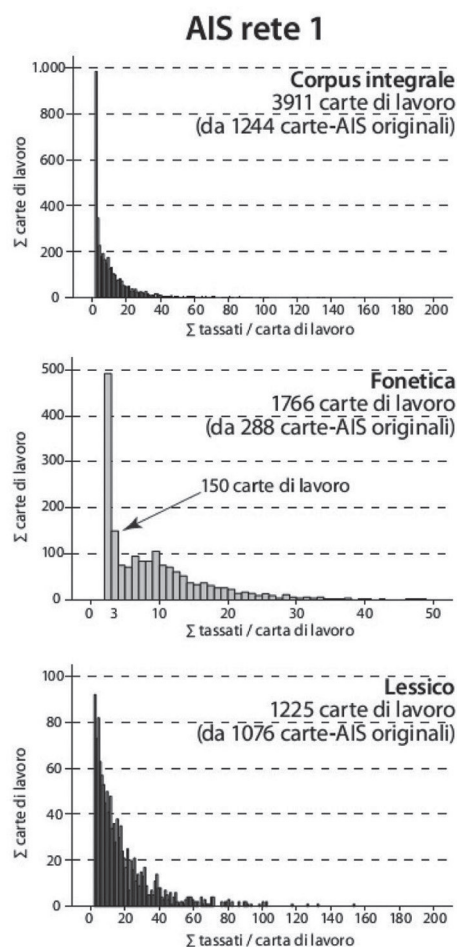


FIGURA 2. Istogrammi relativi alla relazione tra la polinimia interna ed il numero delle carte di lavoro (CL) analizzate: corpus *integrale* (3 911 CL), corpus *fonetica* (1 766 CL), corpus *lessicale* (1 225 CL)

Esempio di intelligenza: nel corpus *fonetica* (si vedano la freccia nonché le rispettive indicazioni lungo gli assi x e y) ci sono 150 CL che dispongono di 3 tassati (e sono quindi *tri-nime*).

Dalle tre curve esponenziali si evidenzia chiaramente che il numero delle CL *oligo-nime* è di gran lunga superiore a quello delle CL *poli-nime* e che tra i due poli estremi si instaura un decremento quantitativo provvisto di una sorprendente gradualità⁹.

La matrice dei dati da noi utilizzata dispone quindi delle dimensioni seguenti: N (= \sum punti di rilevamento) = 382, p (= \sum carte di lavoro) = 3 911. Essa costituisce la raffigurazione nominale di 43 564 aree tassatorie

⁹ Questo fatto è stato studiato più da vicino dal linguista quantitativo tedesco Gabriel Altmann nel 1985 che ne ha derivato una legge linguistica particolare ("legge di Goebel").

di carattere fonetico, morfologico, sintattico e lessicale intrecciate tra loro in maniera altamente complessa.

Aggiungiamo ancora il fatto che la definizione classificatoria dei testè menzionati 43 564 tassati (o aree) è stata realizzata con metodi meramente *linguistici* e poggia quindi su un ingente stock di “expert knowledge” romanistico.¹⁰

4. La misurazione delle *similarità* e *distanze* e la costituzione delle matrici di *similarità* e di *distanza*

L’elaborazione successiva delle informazioni accumulate nella matrice dei dati (tappa B; si veda la figura 1) 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 internazionale¹¹. In questa sede si tratta dell’“Indice Relativo d’Identità” (IRI_{jk}) e dell’“Indice Relativo di Distanza” (IRD_{jk}) i quali, da molti anni, rappresentano gli indici standard della SDMS¹². Tra l’ IRI_{jk} e l’ IRD_{jk} vige la relazione seguente: $IRI_{jk} + IRD_{jk} = 100$.

La figura 3 visualizza il processo della misurazione delle similarità rilevabili tra due vettori-attributi accoppiati. Si veda in merito l’esempio di misurazione (relativo ai punti d’inchiesta 3 e 4) descritto più in dettaglio nella leggenda della figura 3.

In seguito a ciò il contenuto quantitativo delle matrici (quadrate) di *similarità* e di *distanza* deve essere appositamente visualizzato, sempre in base alle necessità teoriche, analitiche e cognitive del dialettometra. Si tratta quindi di spogliare debitamente il contenuto delle due matrici quadrate. A questo proposito il programma salisburghese VDM offre molteplici possibilità, sempre avvalendosi del mondo sfumato dei colori: rinviamo in merito alla tappa C sulla figura 1.

10 Non condivido l’ottimismo procedurale, manifestato e praticato in altra sede, di sostituire il lavoro di tassazione “fatto a mano” con metodi automatizzati di tipo “data mining”.

11 Cito in merito ancora una volta il manuale classico di Sneath / Sokal 1973.

12 Sulla figura 1 viene menzionato anche l’indice ponderato $IPI(1)_{jk}$: si veda in merito l’interstizio tra la matrice dei dati e quella di similarità nonché l’interno del simbolo della matrice di similarità. La sigla $IPI(1)$ significa “Indice Ponderato d’Identità col peso [numerico] 1”. Il peso numerico può variare tra 1 e infinito. L’uso dell’ $IPI(x)$ è consigliabile in certi contesti-DM: per una descrizione dettagliata cf. Goebel 1987.

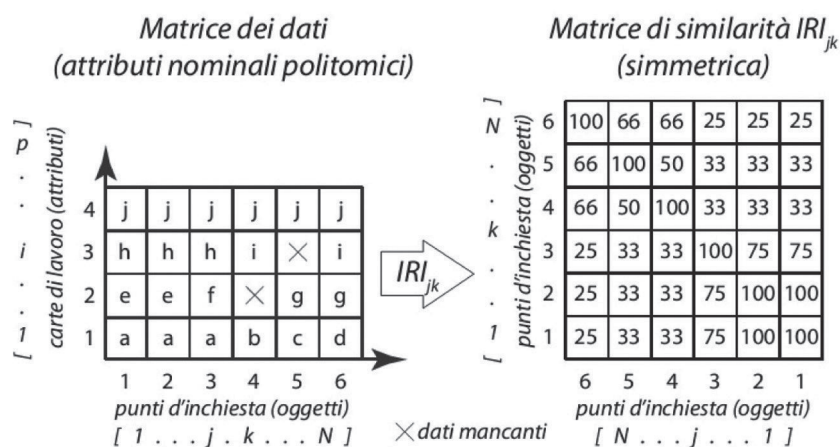


FIGURA 3. Schema di misurazione dialettometrica: matrice dei dati e matrice di similarità
Matrice dei dati: rappresentazione di un atlante linguistico modello dopo tassazione avvenuta
Matrice di similarità: contiene 36 valori di similarità calcolati tramite l'indice di similarità standard della SDMS IRI_{jk} .

Esempio d'intelligenza: secondo la logica algoritmica dell' IRI_{jk} , la similarità (linguistica), rilevabile tra i vettori-oggetto dei punti d'inchiesta 3 e 4, s'ammonta a 33%. Questa percentuale risulta dalla relazione tra una co-identità (j/j) e tre co-presenze (a/b , b/i , j/j) (di cui due sono co-differenze [a/b , b/i] ed una co-identità [j/j]), ricavabili dai vettori-oggetto accoppiati dei punti d'inchiesta 3 e 4. La coppia f/X non serve ad una misurazione e viene perciò tralasciata.

5. Spoglio A della matrice di similarità: presentazione di sette carte di similarità

Si vedano le tavole 3-9.

La carta di similarità (CDS) rappresenta lo strumento di lavoro più importante della SDMS. 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 tinte cromatiche graduate secondo la logica dello spettro solare (ossia di quello dell'arcobaleno), il calo progressivo dei valori della similarità linguistica rispetto al valore supremo di 100% valevole per il punto di riferimento stesso¹³ ($IRI_{jj} = 100$).

Accenniamo, per una miglior comprensione delle CDS, a quattro fatti cartografici di rilievo:

¹³ Una delle proprietà matematiche dell'indice IRI_{jk} è la condizione di *riflessibilità* che fa sì che una misurazione riflessa (punto j con punto j) conduca al risultato 100 ($IRI_{jj} = 100$).

a) Il poligono del punto di riferimento rimane sempre in bianco e si trova, di solito, nel centro o al margine della zona contrassegnata dal color rosso (intervallo 6), la quale corrisponde ai punti-AIS provvisti dei più alti tassi di similarità.

b) In genere il profilo della strutturazione spaziale complessiva è molto regolare e corrisponde, metaforicamente parlando, ad una “montagna” provvista di una sola cima e di molte falde, pendii e vallate.

c) La configurazione spaziale del profilo coropletico di una CDS (ed anche di altre visualizzazioni-DM) dipende – prescindendo dal rispettivo corpus di base – in larga misura dalla natura di due strumenti euristici molto importanti: dall’indice di similarità nonché dall’algoritmo di visualizzazione. In questo contributo vengono applicati tre strumenti standard della SDMS: l’indice IRI_{jk} ed i due algoritmi di visualizzazione MINMWMAX e MEDMW¹⁴.

d) Ciascuna delle nostre carte coropletiche¹⁵ e isaritmiche¹⁶ consta di tre parti: il cartogramma propriamente detto, la leggenda numerica (sempre a sinistra in basso) e un istogramma (sempre a destra in basso). Mentre il cartogramma è del più alto interesse per il *geolinguista*, le informazioni deducibili dalla leggenda e dall’istogramma coinvolgono piuttosto lo *statistico* o il *linguista quantitativo*.

Ricordiamo inoltre che – considerate le molteplici possibilità di calcolo e di cartografazione offerte da VDM – la generazione e la successiva comparazione di un gran numero di profili coropletici di origine diversa è realizzabile in poco tempo e con grande facilità.

Quanto alla diminuzione dei tassi di similarità nello spazio, è ovvio che essa si pone in stretta relazione, benché in varia misura, con l’aumento delle rispettive distanze geografiche. Utilizzeremo questo fatto – che rappresenta d’altronde un *universale linguistico* di primo rango – nel quadro della dialettometria correlativa: si veda il paragrafo 10.

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

¹⁴ 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. L’algoritmo MEDMW invece – pur rispettando la posizione centrale della media aritmetica – crea, da ambedue le parti di quest’ultima, intervalli i più uguali possibile quanto al numero dei rispettivi poligoni (o punti d’atlante): per ulteriori dettagli cf. Goebel 1983: *passim*; 1984a, I: 90-92 e 1984b: 19-20.

¹⁵ La cartografia internazionale intende per “carta coropletica” un cartogramma composto da tasselli corici disgiunti provvisti di apposite colorazioni sfumate.

¹⁶ Il predicato „isaritmico“ si riferisce esclusivamente alle tavole 13 e 14. S’intende per “carta isaritmica” un cartogramma basata su una simbologia lineiforme o decisamente lineare.

- la sua posizione relazionale all'interno della rete esaminata,
- il tasso di diffusione (penetrazione, interazione, ecc.) della sua dialettalità in seno alla rete esaminata,
- l'efficacia "comunicativa" del rispettivo dialetto.

Le CDS si prestano inoltre a diversi modi d'interpretazione¹⁷. 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 *in tempi trascorsi*, 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 SDMS a dati medievali: cf. a questo proposito le nostre ricerche-DM relative all'atlante scrittologico di Anthonij Dees del 1980¹⁸.

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 suggestivi¹⁹ che gettano una nuova luce sulla sovramenzionata "gestione basilettale dello spazio da parte dell'Homo loquens".

5.1. Presentazione e interpretazione della tavola 3: una CDS relativa al P.-AIS 544 (Arezzo, Toscana)

La tavola 3 (relativa ad un basiletto toscano "tipico") rivela, oltre all'armonica inserzione relazionale del dialetto di Arezzo nella parte

¹⁷ Per una discussione circostanziata cf. Goebel 1984a, I: 100-113 e 1984b: 24-28 nonché Goebel 1981: 369-381.

¹⁸ Cf. Goebel 2008b e Dees 1980.

¹⁹ Si tratta di un'applicazione della "cartografia dinamica": cf. a questo proposito Goebel 1982: 38-40 e 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 dei colori rimane indispensabile.

centrale dello “stivale”, anche una marcata presenza – con una chiara discontinuità territoriale rispetto all’Emilia-Romagna – della “tipicità mediana” 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 avvenuta alla fine del terzo secolo a. C²⁰.

Si noti inoltre l’ubicazione degli „antipodi” tipologici del toscano di Arezzo, contrassegnati tutti in blu scuro e blu chiaro (equivalenti agli intervalli 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).

I tratteggi bianchi su fondo *rosso* (si veda il P.-AIS 535, vicino al P.-AIS 544) e *blu scuro* (si veda il P.-AIS 121, situato in Valle d’Aosta) rinviano rispettivamente ai valori *massimo* (86,64) e *minimo* (46,22) della rispettiva distribuzione di similarità. Metaforicamente parlando si potrebbe anche qualificare la funzione dei rispettivi poligoni come quelle del “miglior amico” e del “peggiore nemico” (o “capro espiatorio”).

5.2. Presentazione e interpretazione della tavola 4: una CDS relativa al P.-AIS 376 (Venezia)

Il potenziale linguistico esaminato si riferisce al primo (*carpentiere di ferro*) dei due “sujets”, intervistato da Paul Scheuermeier a Venezia nel 1922.

La zona coperta dai poligoni con colori *caldi* (rosso, arancione, giallo) si stende dai piedi delle Alpi fino alla “linea Roma-Ancona”, coinvolgendo quindi anche l’Italia mediana. Tale non sarebbe il caso con un profilo di similarità lombardo, piemontese o emiliano-romagnolo dove i poligoni contrassegnati con colori caldi non trasgredirebbero mai la “linea La Spezia-Rimini”. Questo fatto si spiega per la posizione geotipologica dei dialetti veneti caratterizzati in genere da una fortissima “medianizzazione” intrinseca che, come pare, risale già al periodo della romanizzazione²¹.

E ovvio che i dialetti dell’Istria sono molto vicini alla dialettalità della Serenissima e figurano quindi negli intervalli 6 (rosso) e 5 (arancione). Si osservino inoltre la posizione laterale della totalità del sud d’Italia e della Sardegna nonché la perifericità dei dialetti dell’arco alpino occidentale e dei dialetti ladini (retoromanzi).

20 Cf. in merito le nostre riflessioni in Goebel 2008a: 58-61.

21 Cf. la nota precedente.

5.3. Presentazione e interpretazione della tavola 5: una CDS relativa al P.-AIS 368 (Pirano/Piran, Slovenia)

Il profilo coropletico della tavola 5 è tipicamente “veneto” e combacia – tranne poche eccezioni – con quello della tavola 4.

5.4. Presentazione e interpretazione della tavola 6: una CDS relativa al P.-AIS 378 (Montona/Montovun, Croazia)

Il profilo coropletico della tavola 6 corrisponde praticamente a quello della tavola 5. Si osservi l'ubicazione del valore massimo della distribuzione di similarità (83,21): esso si trova al posto del P.-AIS 368 (Pirano/Piran), cioè nella vicinanza immediata del punto di riferimento 378. Lo stesso dicasi anche della tavola 5, dove il valore massimo (ugualmente 83,21) si trova al posto del P.-AIS 378.

5.5. Presentazione e interpretazione delle tavole 7 e 8: due CDS relative al P.-AIS 397 (Rovigno/Rovinj, Croazia) ed al P.-AIS 398 (Dignano/Vodnjan, Croazia)

È risaputo che ai tempi delle inchieste di Paul Scheuermeier (fatte nel 1922), gli abitanti delle località di Rovigno/Rovinj e Dignano/Vodnjan disponevano – e tuttora dispongono – di competenze linguistiche abbastanza variegata che spaziano dal *croato* ed *italiano* al *veneziano coloniale* nonché all'*istrioto* (o *istroromanzo*) antico. Vista la debolezza sociolinguistica dell'*istrioto* già ai tempi di P. Scheuermeier e la sua progressiva interferenza col *veneziano coloniale* è difficile dire in quale misura le risposte rilevate da P. Scheuermeier riflettano l'*istrioto* da una parte o il *veneziano coloniale* dall'altra²².

Sta di fatto che i profili coropletici delle tavole 7 e 8 sono strettamente apparentati tra di loro staccandosi però leggermente da quelli delle tavole 4-6. Accenno in merito soprattutto alla posizione più vicina dei dialetti friulani, tutti contrassegnati in giallo (intervallo 4) ed alla dislocazione dell'ubicazione del “peggior nemico” (tratteggio bianco su blu scuro) dalla Valle d'Aosta nella Sardegna centrale. Ciò non impedisce che l'estensione delle zone colorate in rosso ed arancione rimanga la stessa di prima (si vedano le tavole 4-6) coinvolgendo cioè i basiletti di Rovigno/Rovinj e Dignano/Vodnjan nel geotipo veneto *tout court*.

22 Cf. in merito anche Goebel 1988: 91-93.

5.6. Presentazione e interpretazione della tavola 9: una CDS relativa al P.-AIS 29 (Santa Maria, Grigioni, Svizzera)

Si tratta di un profilo coropletico tipicamente “padano” che si manifesterebbe in forma molto simile anche da altri punti-AIS *ladini* (*retoromanzi*) e *alto-italiani* (*padani*). I poligoni gialli si fermano alla “linea La Spezia-Rimini” escludendo la zona del ligure da una parte ed alcuni angoli del veneto dall’altra. È chiaro che i poligoni in rosso ed arancione coprono la Rumantschia intera verso *ovest*; non meno evidente è che la Ladinia dolomitica ed il Friuli occupino posizioni similmente privilegiate verso *est*.

Paragonando tra di loro le strutture coropletiche delle tavole 4-9 si noti che i loro fondamenti, rappresentati dagli intervalli 1-3 (blu scuro, blu chiaro e verde) rimangono praticamente inalterati e che le differenze si evidenziano soprattutto nel quadro degli intervalli 5 (arancione) e 6 (rosso).

6. Presentazione di una carta di prossimità

Si veda la tavola 10 (relativa al P.-AIS 29, Santa Maria, Grigioni, Svizzera)

Mentre le CDS servono, in ultima analisi, a visualizzare il calo delle similarità linguistiche coll’aumento della distanza geografica rispetto al punto di riferimento, e dimostrano così gli effetti della “gestione basilettale dello spazio da parte dell’Homo loquens”, la “carta di prossimità” serve a visualizzare la stessa costellazione dal punto di vista della geometria euclidea.

In effetti, è possibile calcolare, mediante il noto teorema di Pitagora, le *distanze* reciproche (*d*) tra tutti i punti della rete dell’AIS e convertirle dopo in valori di *prossimità* (*s*) mediante la formula $d + s = 1$ (o 100%). I valori di *prossimità* vengono di seguito inseriti in una matrice (quadrata) di *prossimità* la quale si presta alle stesse visualizzazioni come le CDS: si veda in merito la tavola 10. Non c’è da stupirsi che il suo profilo coropletico sia contrassegnato da una regolarità geometrica (“euclidea”) perfetta.

Anticipando sul paragrafo 10, è lecito chiedersi in quale maniera i profili coropletici delle tavole 9 e 10 possano paragonarsi tra di loro: a) tramite un’introspezione comparativa meramente visiva, b) tramite una misurazione correlativa con mezzi statistici. Utilizzando quest’ultimo metodo tramite l’applicazione di uno degli indici di correlazione più noti in sede di statistica descrittiva [$r(\text{BP})$], si ottiene

il valore +0,744. Ovviamente questo valore *isolato* non ha alcuna importanza *linguistica*. Riuscendo però a procurarci tutti i 382 valori-*r*(BP) possibili ed a visualizzarli nella consueta maniera, si aprono nuovi orizzonti euristici anche per il *geolinguista-dialettometra*. Per ulteriori dettagli rinviamo al paragrafo 10.

7. Spoglio B della matrice di similarità: presentazione e interpretazione di due carte a parametri: le sinossi dei CAF (“Coefficienti di Asimmetria di Fisher”) e dei QuoCo (“Quozienti di Comunicazione”) delle 382 distribuzioni di similarità

Si vedano le tavole 11 e 12.

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 – non solo i valori dei diversi parametri statistici (come il “minimo”, il “massimo”, la “media aritmetica”, ecc.)²³ ma anche lo loro importanza geolinguistica. Si è così rivelata la grande utilità geolinguistica della sinossi di alcuni parametri relativi alla *simmetria* della rispettiva distribuzione di similarità, tra i quali spiccano due indici che misurano il grado (o meno) della simmetria di una distribuzione di similarità: si tratta, da una parte, del “Coefficiente di Asimmetria di Fisher” (CAF)²⁴ e, dall'altra, del “Quoziente di Comunicazione” (QuoCo). Ambedue gli indici ci permettono di studiare più da vicino il fenomeno del cosiddetto *Sprachausgleich* (“compromesso linguistico”) che è strettamente legato all'embricazione spesso molto complessa delle aree tassatorie contenute nella matrice stessa dei dati.

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 rispettiva rete geolinguistica. Tale integrazione si fa attraverso l'intreccio graduato di aree tassatorie di *piccole* (*oligo*-coriche), *medie* (*meso*-coriche) e *grandi* (*mega*-coriche) dimensioni col resto dei dati atlantistici.

La considerazione quantitativa dell'occorrenza di aree *oligo*-, *meso*- e *mega*-coriche nei diversi vettori-attributi di una matrice dei dati può essere realizzata efficacemente tramite la misura della simmetria delle diverse distribuzioni di similarità calcolate a base della matrice dei dati.

²³ Per ulteriori dettagli cf. Goebel 1984a, I: 136-164.

²⁴ Al CAF *italiano* (ed anche *francese* e *spagnolo*) corrispondono i termini inglese *skewness* e tedesco *Schiefte*. Il termine catalano è *biaix*.

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-attributi delle altrettante località della nostra rete-AIS è contrassegnato da 3 911 attributi (o tassati) geolinguistici. Orbene, ciascuno di questi 3 911 tassati corrisponde ad un'area geolinguistica la cui estensione può variare – teoricamente – tra 1 e 381 punti (o poligoni)²⁵. È quindi di somma importanza per il livello d'integrazione di un vettore-attributi (o punto d'inchiesta) particolare nella totalità della rispettiva rete, la ripartizione percentuale di aree *oligo-*, *meso-* e *mega-*coriche. Da questa considerazione si evidenzia che un vettore-attributi che dispone prevalentemente di aree tassatorie con un'estensione abbastanza *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 toscano di Arezzo, P.-AIS 544) ci ricorda che dei 381 valori di similarità cartografati 154 (= 96 + 20 + 38 punti su 381: = 40,41%) – ossia la minoranza – si trovano sopra la media aritmetica di 63,70, laddove 227 (= 26 + 49 + 152 punti su 381: = 59,58%) si accumulano sotto la media aritmetica.

Una situazione diversa risulta dalla tavola 4 relativa al dialetto di Venezia (P.-AIS 376) dove la stessa ripartizione si presenta come segue: punti situati sopra la media aritmetica di 62,55: 212 (= 161 + 34 + 17 punti su 381: = 55,64%), ossia la maggioranza; numero dei punti situati sotto la media aritmetica: 169 (= 33 + 49 + 87 punti su 381: = 44,35%).

La connessione comunicativa (relativa) del dialetto di Venezia (P.-AIS 376) col resto della rete è quindi superiore a quella del dialetto di Arezzo (P.-AIS 544).

È precisamente la percentuale dei valori di similarità situati sopra la media aritmetica che viene misurata dal QuoCo (“Quoziente di Comunicazione”)²⁶.

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

26 Per la formula ed una descrizione dettagliata del QuoCo cf. Goebel 1984a, I: 154-157.

L'altro indice di simmetria è statisticamente più complesso. Il “coefficiente di asimmetria di Fisher” (CAF)²⁷ 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).

7.1. Presentazione della tavola 11: la sinossi dei “Coefficienti di Asimmetria di Fisher” (CAF) di 382 distribuzioni di similarità

Sulla tavola 11, ai valori negativi corrisponde la maggioranza dei poligoni in blu scuro e blu chiaro (intervalli 1 e 2), mentre i valori positivi vengono marcati dai poligoni situati negli intervalli 3-6. Ecco il significato linguistico dei colori:

- poligoni negli intervalli 1 e 2 (con colorazione azzurra): 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’AIS²⁸. Si tratta dunque di zone molto attive e provviste di molti contatti linguistici, tanto dal punto di vista *espansivo* quanto da quello *ricettivo*.
- poligoni situati prevalentemente negli intervalli 4-6 (contrassegnati cioè con i colori caldi): zone con un tasso molto basso del compromesso linguistico (*Sprachausgleich*), caratterizzate quindi – tanto in dia- quanto in sincronia – da un gran conservatorismo (o immobilismo) linguistico rispetto alla totalità delle rete esaminata. Si tratta pertanto di zone la cui tipicità linguistica originale è stata preservata per molto tempo ed in larga misura da intrusioni straniere di qualsiasi origine.

Uno sguardo alla tavola 11 ci rivela che la ripartizione spaziale dei poligoni marcati con colori caldi è molto chiara. Lo stesso dicasi dei poligoni collocati negli intervalli 1 e 2: nella Padania i poligoni in blu scuro e blu chiaro creano una dipartimentalizzazione ben definita in cui emergono, come solchi divisorii provvisti però di altissimi tassi interattivi,

²⁷ Per la formula del CAF cfr. Goebel 1981: 394 e 1984a, I: 150.

²⁸ 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 esaminata.

il corso dell'Adige e del Po nonché lunghi tratti della catena appenninica dalla Liguria fino alle Marche.

Un'altra incavatura contrassegnata in parte da poligoni delle classi 1 e 2 (e quindi provvista di un alto potenziale interattivo) s'incunea tra l'Italia mediana (Lazio, Umbria, Marche) e quella del sud (Abruzzo e Molise, Campania, ecc.).

Si osservi ancora che alle zone cromaticamente "calde" corrispondono invece configurazioni spiccatamente *coriche* o *puntuali* e dotate quindi di un'estensione corica geograficamente coerente, mentre le zone contrassegnate da tinte azzurre assumono configurazioni *lineiformi* o addirittura *circolari*.

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

7.2. Presentazione della tavola 12: la sinossi dei "Quozienti di Comunicazione" (QuoCo) di 382 distribuzioni di similarità

Mentre sulla tavola 11 i colori blu scuro e blu chiaro equivalevano ad un alto tasso del compromesso linguistico ("Sprachausgleich"), sulla tavola 12 il significato linguistico dei colori è decisamente opposto: laddove si evidenziano poligoni in rosso e arancione, siamo in presenza di un alto tasso di intercambio linguistico.

In linea di massima, i messaggi delle due tavole sono complementari; affiorano però alcune divergenze, soprattutto nella zona di contatto tra la Lombardia ed il Piemonte.

8. Spoglio della matrice di distanza: presentazione e interpretazione di due carte isoglottiche

Si vedano le tavole 13 e 14.

Dal punto di vista cartografico le carte isoglottiche rientrano nella classe delle carte "isaritmiche", i cui elementi iconici di base non sono più le *aree* dei poligoni, bensì i *lati* di poligono generati tramite un previo

trattamento geometrico della rete esaminata²⁹. La confezione del fondo di carta delle tavole 13 e 14 si basa quindi sulle tappe seguenti³⁰:

- triangolazione delle rete-AIS secondo i principi della geometria di Delaunay³¹. Ne risultano – per la rete delle tavole 13 e 14 – 970 lati di *triangolo*.
- poligonizzazione successiva della rete triangolata secondo i principi della geometria di Voronoi³²: ne risultano 970 lati di *poligono*.
- allestimento della simbologia cartografica per conferire ai lati 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 proprietà (*spessore* e *colorazione*).

Dal punto di vista tasso- o dialettometrico la carta isoglottica della tavola 13 (relativa alla *fonetica*) è basata su 970 valori di distanza dell'IRD_{jk} oscillanti tra 5,34 e 47,34%, mentre sulla tavola 14 (relativa al *lessico*) la stessa variazione si evidenzia tra i valori-IRD_{jk} 11,19 e 57,49%.

La visualizzazione dei suddetti 970 valori viene effettuata lungo altrettanti lati di poligoni secondo il principio cartografico seguente: più i valori dell'IRD_{jk} sono alti, più i rispettivi lati di poligono sono *spessi* e *azzurri*. E viceversa.

In questa maniera spuntano, in non poche zone delle nostre due carte, 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 in *blu scuro* situata lungo l'Appennino tosco-emiliano evoca la ben nota "linea La Spezia-Rimini" di memoria wartburghiana³³, 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.

29 Si detto tra parentesi che esiste, accanto alla "carta isoglottica", ancora un altro tipo di visualizzazione interpuntuale di estrazione "isaritmica", chiamata "carta a raggi". Essa è basata sulla *triangolazione* della rete e visualizza, tramite la variazione dello *spessore* e del *colore* dei *lati di triangolo*, non delle *distanze* bensì delle *similarità* interpuntuali. Il suo messaggio cartografico è quindi complementare a quello della carta isoglottica. Per esempi concreti si veda Goebel 1983: *passim*, 2008a: 93, 2010: 33 e 2011: 99.

30 Per il problema delle carte *interpuntuali* in genere cfr. soprattutto Goebel 1983: *passim* nonché 1984a, I: 183-196.

31 Cf. Okabe *et al.* 1992: 72 ss.

32 Cf. Okabe *et al.* 1992: 65 ss.

33 Per la storia dei concetti delle "linee" *La Spezia-Rimini* e *Roma-Ancona* cf. Goebel 2003b: 634 s.

I settori dove si accumulano i lati di poligono *sottili* marcati in rosso rimandano al contrario a zone linguisticamente *coerenti* (o *nuclei dialettali*).

Si noti il fatto che la tavola 13 dimostra, sulla base di 1 766 CL di tipo *fonetico*, la strutturazione isoglottica dell' AIS dal punto di vista *fonetico*, mentre la tavola 14 ci rivela la strutturazione isoglottica dell' AIS dal punto di vista *lessicale*. Il rispettivo corpus consta di 1 225 CL di tipo *lessicale*. Le divergenze tra queste due analisi interpuntuali sono poche. Questo fatto prova che l'impatto delle due categorie linguistiche sulla "gestione basilettale dello spazio" è molto simile.

Bisogna però aggiungere che, dal punto di vista tassometrico, l'analisi interpuntuale è piuttosto superficiale. In effetti, i 970 valori interpuntuali rappresentano, sui 72 771 valori³⁴ delle rispettive matrici di distanza o di similarità, solo l'1,33% della totalità delle informazioni quantitative disponibili, mentre altre analisi-DM poggiano su tutti i dati contenuti nelle matrici di similarità o di distanza.

9. Spoglio C della matrice di similarità: le molteplici possibilità della dialettometria correlativa

La DM correlativa – nata dallo studio approfondito delle relazioni che, come risulta dalla considerazione delle tavole 4-9, intercorrono tra la similarità linguistica e la prossimità geografica – è stata aggiunta solo nel 2004 al repertorio metodico della SDMS³⁵. Questa stretta interdipendenza si evidenzia subito da una comparazione visiva³⁶ tra una *carta di similarità* normale – calcolata cioè con l'indice IRI_{jk} – ed una carta analoga (*carta di prossimità*) basata su valori di prossimità calcolati secondo il noto teorema di Pitagora e quindi secondo i principi della geometria euclidea: si vedano in merito le tavole 9 e 10.

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 del linguaggio e quella delle relazioni euclidee sarebbe possibile, almeno in teoria, esaminare *de visu*, nel quadro di un qualsiasi progetto-

34 In questa sede si deve applicare la nota formula combinatoria $N/2 (N - 1)$. Ne deriva, per $N = 382$, l'espressione seguente: $382/2 \cdot 381 = 72 771$.

35 Per una prima presentazione delle molteplici possibilità della DM correlativa rinviamo alla nostra relazione di lavoro del 2005.

36 Per altri due campioni di una carta di prossimità cf. Goebel 2007: 263 e 2008a: 95.

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 descrittiva 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 quadrate (di *similarità* o di *prossimità*) 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 SDMS: si veda la figura 4.

La misura di correlazione statisticamente più adatta nell'ambito delle nostre esperienze geolinguistiche è il coefficiente di correlazione di [Auguste] Bravais e [Karl] Pearson $r(\text{BP})$ ³⁷. Il $r(\text{BP})$ misura la correlazione lineare tra due variabili quantitative assumendo sempre 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 affiancare ad ambedue le possibilità un'adeguata interpretazione linguistica, sempre in chiave dia- e sincronica.

9.1. La correlazione spaziale tra “linguaggio” e spazio (euclideo)

Si veda la tavola 15.

La tavola 15 rivela una strutturazione spaziale ben interpretabile dal punto di vista linguistico.

Il significato dei rispettivi colori è questo:

- zone contrassegnate in *rosso* ed in *arancione*: settori della rete-AIS dove persiste un'armonia originaria tra “linguaggio” e spazio, esenti da trasferimenti di elementi linguistici in contrasto con lo svolgimento

³⁷ Per la formula del $r(\text{BP})$ cf. Goebel 2005: 329.

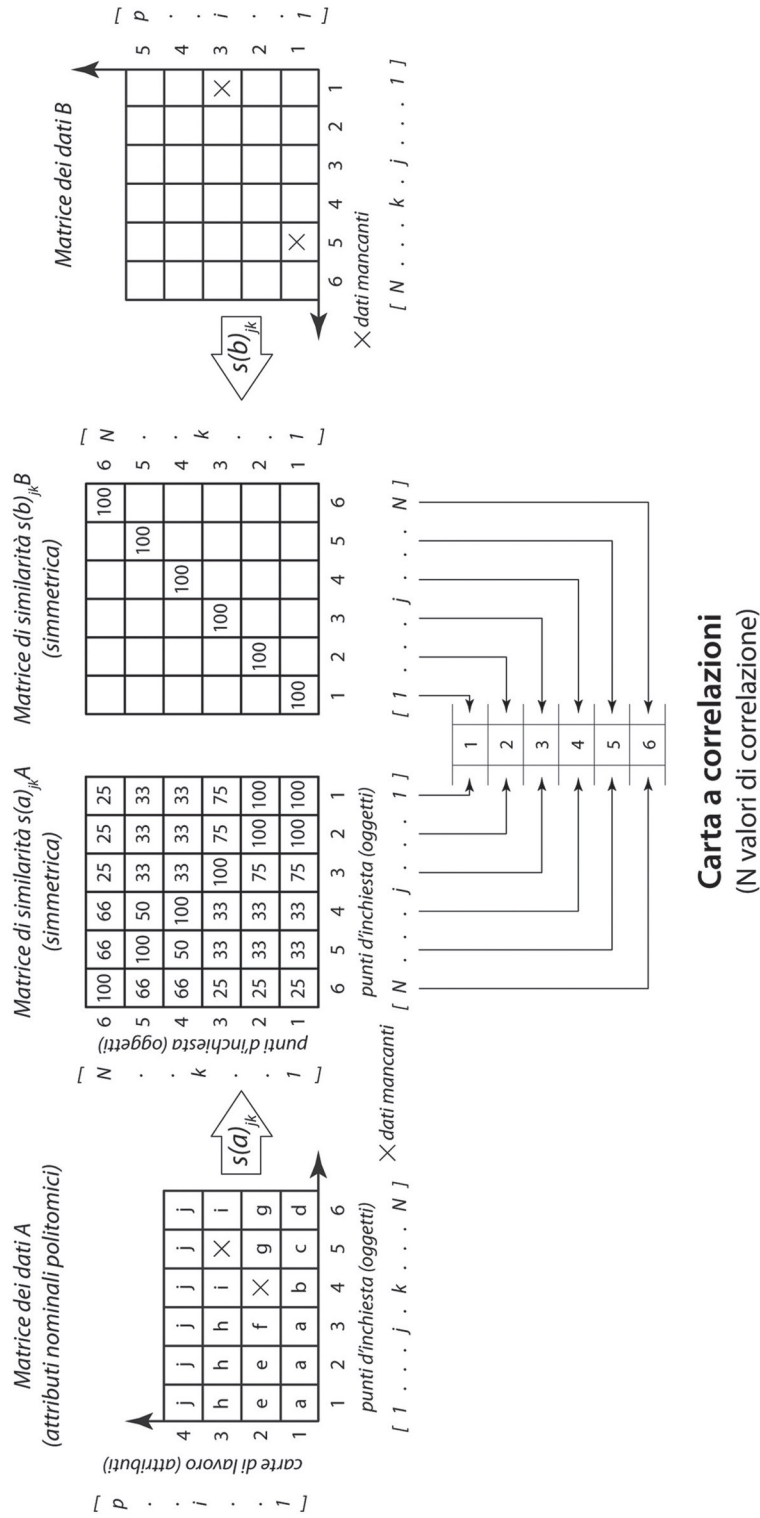


FIGURA 4. Schema di calcolo per la generazione di una carta a correlazioni (con 6 valori) mediante due matrici dei dati e le rispettive matrici di similarità (o di prossimità), tutte provviste di 6 oggetti (elementi o punti d'inchiesta)

“naturale” dello scambio linguistico. Si tratta cioè di regioni conservatrici e rimaste linguisticamente piuttosto incontaminate.

- zone contrassegnate in *blu scuro* ed in *blu chiaro*: settori della rete-AIS dove un’eventuale armonia originaria tra “linguaggio” 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 15 pone in evidenza che le zone “movimentate” (in blu scuro: 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 15 si deduce inoltre il fatto – d’altronde molto plausibile – che le cinque isole linguistiche situate nel sud (PP.-AIS 715 e 760) ed in Sicilia (PP.-AIS 865, 836 e 817) costituiscono nuclei basilettali linguisticamente movimentati. Sta di fatto che la fondazione di un’isola linguistica rappresenta la rottura totale di un’eventuale armonia originaria tra linguaggio e spazio.

La posizione “movimentata” del nord della Toscana e del Veneto è altamente emblematica soprattutto alla luce della già menzionata “medianizzazione” del Veneto³⁸ 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 dell’AIS, 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.

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

Si veda la tavola 16.

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 16 dove si evidenzia che il “passo cadenzato” spaziale della *fonetica* e del *lessico* segue lo stesso ritmo tanto nei settori

³⁸ Si veda in merito il paragrafo 6.1.

centrali della Padania quanto nel sud, mentre ci troviamo di fronte ad una vera e propria disarmonia delle due categorie sulle frontiere 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 situate nel sud ed in Sicilia.

Questi fatti si spiegano con tensioni esistenti all'interno delle strutture grammaticali dei rispettivi basiletti tra le categorie della *fonetica* e del *lessico*.

9.3. La correlazione spaziale particolare tra due categorie linguistiche diverse (fonetica e lessico) e lo spazio (euclideo)

Si vedano le tavole 17 e 18.

Ovviamente è possibile ripetere l'esperienza-DM fatta nel paragrafo 10.1. correlando tra di loro, da una parte, le dimensioni della *fonetica* e dello *spazio euclideo* e, dall'altra, quelle del *lessico* e dello *spazio euclideo*. I rispettivi risultati cartografici sono ben lungi dal combaciare perfettamente. Mentre possiamo constatare una certa convergenza per il sud della rete dell'AIS – dove le due categorie linguistiche si sono sviluppate il più delle volte seguendo la “logica” dello spazio euclideo –, dobbiamo rassegnarci a prendere atto di divergenze profonde per l'Italia settentrionale ed anche per alcune delle isole linguistiche del sud.

Sulla tavola 17 si palesano alcune zone “motrici”, come la Toscana settentrionale, la Liguria ed il Veneto, le quali avevano già fatto capolino sulla tavola 15, mentre nel resto della Padania l'evoluzione fonetica sembra essersi svolta secondo ritmi piuttosto “naturali”.

Quanto all'evoluzione diacronica del lessico, le zone più marcate da dinamismi *espansivi* e *ricettivi* si allargano dalla Toscana settentrionale alla Rumantschia, attraversando la Liguria, il Piemonte nonché la Lombardia occidentale.

In ogni caso stupisce l'andamento molto regolare dei due profili coropletici.

Visto lo stretto legame tra spazio e tempo si può ipotizzare che studi diacronici condotti sulle regioni testè menzionate potrebbero fornire ulteriori chiarimenti sui dinamismi e contatti linguistici evocati.

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

esibiscono dei risultati molto interessanti, suscettibili di aprire – in sede di *geolinguistica* o di *tipologia areale* – nuove prospettive per la ricerca sul *variazionismo grammaticale*.

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.

10. Spoglio D della matrice di similarità: l'analisi dendrografica tramite l'algoritmo di Joe Ward, Jr.

Si vedano le tavole 19 e 20.

Uno dei mezzi più efficaci in sede di classificazione numerica in genere e della SDMS in particolare è, da oltre 40 anni, la “Classificazione Gerarchica Agglomerativa” (CGA) (così denominata a livello internazionale). La CGA fornisce, in quanto output euristico, schemi dendrografici (“alberi”, dendrogrammi) di qualsiasi natura. Ecco i sommi capi metodici per un buon intendimento della CGA:

- Il punto di partenza dei calcoli-CGA è la matrice di similarità stessa.
- Si tratta di nuovo di una classificazione quantitativa della *totalità* dei dati compresi nella rispettiva matrice di similarità (N per N).
- 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 19 si veda in merito la freccia inferiore (“senso dell’agglomerazione”). Gli alberi generati in seno della CGA sono perciò provvisti di ramificazioni binarie gerarchizzate.

La *classificazione numerica* (o *tassometria*) internazionale³⁹ si serve di un gran numero di appositi algoritmi, al quale corrisponde ovviamente una molteplicità analoga di dendrogrammi. L'utilità di questi alberi per un dato progetto di ricerca deve essere accertata caso per caso tramite una valutazione combinata delle necessità della *statistica* e della rispettiva *disciplina empirica* (in questa sede: della *geolinguistica romanza*).

Si osservi che gli alberi-CGA costituiscono una gerarchia con ramificazioni – esclusivamente binarie – di classi (clusters, ecc.) disgiunte, la cui variabilità quantitativa interna cresce a partire dalle

39 Rimando di nuovo al manuale classico di Sneath / Sokal 1973.

foglie (= “chioma”) fino alla radice (= “tronco”) dell’albero. Soprattutto per l’interpretazione geolinguistica dei diversi piani della struttura ramificata dell’albero si rende indispensabile la presa in considerazione della variabilità cangiante delle diverse classi.

Il punto più importante in sede di DM è però la proiezione finale delle strutture degli alberi calcolati nello spazio (“spazializzazione”). Mentre per l’uso esclusivamente *diacronico* di schemi dendrografici (come viene praticato spesso dalla lessicostatistica⁴⁰) – si veda in merito la freccia superiore (intitolata “frammentazione diacronica”) – bastava l’analisi dell’albero stesso, per la DM al contrario la situazione è diversa. In sede di DM bisogna sempre procedere ad una spazializzazione dei risultati dell’elaborazione classificatoria (selettiva) dell’albero: rimando alle tavole 19 e 20 (nonché 23 e 24) che ne dimostrano il principio.

Grazie alle possibilità informatiche del programma VDM la spazializzazione selettiva degli alberi calcolati può essere effettuata in maniera molto comoda, veloce e – ovviamente – coll’aiuto di colori. A prova di ciò rammento che il calcolo e la generazione degli schemi *dendrografici* e *spaziali* delle tavole 19 e 20 sono stati realizzati tramite VDM in pochi minuti.

L’algoritmo dendrografico applicato è stato proposto dallo statistico americano Joe Ward, Jr. nel 1963⁴¹. La sua utilità per le finalità della geolinguistica è stata spesso comprovata in base a dati geolinguistici di origine assai diversa⁴². La selezione e definizione dei *rami* dell’albero (“dendremi”) (tavola 19) corrisponde a quella delle rispettive *aree* (“coremi”)⁴³ della *spazializzazione* (tavola 20). In questa sede si tratta di una selettività media – con 18 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 prima biforcazione (nei DC 1-10 versus i DC 11-18) fa risaltare la ben nota “linea La Spezia-Rimini”. Tramite la bipartizione ulteriore del macro-DC “padano” 1-10 viene scartata la Rumantschia intera (DC 1). La parte restante della Padania si scompone nel dominio veneto da una parte (DC 8-10) ed il resto gallo-romanzo dall’altra (DC 2-7).

40 Per le relazioni tra la DM dendrografica e la lessicostatistica cf. Goebel 2003a: 84-89.

41 Per una descrizione e valutazione dell’algoritmo di Ward cf. anche Sneath / Sokal 1973: 204s. e 241s.

42 Rinvio in merito alla bibliografia (sempre aggiornata) dei nostri contributi-DM: <http://www.sbg.ac.at/rom/people/prof/goebl/dm_publi.htm>

43 La coppia terminologica *dendrema-corema* viene da noi utilizzata già da parecchio tempo: cf. Goebel 2003a: 85, nota 13.

Per l'Italia mediana e meridionale la bipartizione ulteriore del macro-DC 11-18 continua con il distacco dell'Italia mediana (DC 11-13) dal sud e dalla Sicilia (DC 14-17) nonché dalla Sardegna centrale e meridionale (DC 18). 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 dei 18 DC delle tavole 19 e 20 richiederebbe (e meriterebbe) un'analisi più dettagliata, alla quale dobbiamo però rinunciare per motivi di spazio.

11. Praticando la DM con reti diluite

Si vedano le tavole 21-24.

Lo scopo principale di questo paragrafo è di carattere prevalentemente *cartografico*. Si tratta di evidenziare, tramite la riduzione della rete di ricerca dell'AIS alle zone costiere, la persistenza (o meno) dell'*utilità euristica* di alcune procedure standard della DM.

Ovviamente il successo euristico dei progetti-DM finora realizzati dipende in larga misura dalla configurazione spaziale delle rispettive reti. La compattezza corica della rete dell'ALF si presta ottimamente alla produzione di profili quantitativi graduati. Lo stesso dicasi, benché con alcune riserve, anche della rete oblunga dell'AIS. In ambedue i casi l'ampiezza della rete garantisce la *strutturazione accidentata* dei risultati della classificazione-DM praticata, ed anche la loro *chiara percezione* da parte del dialettometra.

Quando però ci si accinge a diluire la compattezza della rete originale ed ad estrarne una sequenza lineiforme di poligoni, la buona percezione dei risultati della classificazione-DM viene limitata da una parte o chiaramente messa in dubbio dall'altra.

Gli esempi seguenti si basano su una riduzione drastica del numero dei punti d'inchiesta da 382 a 142; rimane invece intatto il numero delle CL sintetizzate. Questo ridimensionamento ha condotto alla definizione di una nuova matrice dei dati, che è stata sottoposta successivamente ai soliti calcoli delle similarità e distanze.

Presentiamo in merito due CDS ed un'analisi dendrografica.

I I.1. Presentazione e interpretazione delle tavole 21 e 22: due CDS relative ai PP.-AIS 376 (Venezia) e 721 (Napoli)

Si vedano le tavole 21 e 22.

È molto utile comparare *de visu* la tavola 21 con la tavola 4 visto che ambedue dispongono dello stesso punto di riferimento: P.-AIS 376, Venezia. Ci sono, in effetti, molte convergenze tra la strutturazione dei due profili coropletici. Le divergenze si spiegano per le strutture “appiattite” all’interno della matrice dei dati diluita. È però degno di nota che la distribuzione dei colori lungo la linea costiera risulta praticamente ovunque identica, salvo alcune eccezioni ubicate soprattutto nel Friuli. Si noti inoltre il fatto che i valori *minimo* e *massimo* delle due leggende numeriche sono identici.

È ovvio che il profilo coropletico della tavola 22 (relativo al P.-AIS 721, Napoli) è diverso da quello della tavola 21 ed è evidente che i poligoni con colori caldi occorrono solo a sud della “linea La Spezia-Rimini”.

I I.2. Presentazione e interpretazione delle tavole 23 e 24: l’analisi dendrografica tramite l’algoritmo di Joe Ward, Jr.

Le norme tassometriche e cartografiche qui applicate corrispondono perfettamente a quelle del paragrafo 11. L’algoritmo-Ward fornisce di nuovo, da un lato, un albero binario ben strutturato che consente, dall’altro, una spazializzazione coerente e chiara. Il macro-DC 1-7 si stende fino alla linea “La Spezia-Rimini”, mentre i restanti DC (8-11) coinvolgono il sud e le isole. Si noti che l’isola linguistica occitana di Guardia Piemontese (P.-AIS 760) e la sua omologa gallo-italica di San Fratello (P.-AIS 817) appartengono ambedue al DC 8 che rappresenta, tra i DC situati al sud della “linea La Spezia-Rimini”, il gruppo quantitativamente più eterogeneo e quindi tipologicamente meno “marcato”. Di nuovo risalta la grande coesione spaziale dei coremi calcolati.

12. Conclusione

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

- si tratta di un metodo induttivo nonché esploratorio che, tramite la sintesi numerica di dati empirici molto variegati, porta alla *scoperta* *elo all'esplorazione* di *regolarità e leggi geolinguistiche*, nascoste a prima vista nei dati analizzati;
- si tratta di un metodo che, dato il suo orientamento verso lo studio quantitativo dello spazio, considera l'uso generalizzato di *cartografazioni e visualizzazioni* dei suoi risultati numerici come *via regia* della propria euristica;
- si tratta di un metodo che, per via della sua impostazione geografica e comparativa intrinseca, rimane sempre aperto a *cooperazioni interdisciplinari* di qualsiasi natura⁴⁴;
- si tratta di un metodo, la cui base empirica è costituita dagli atlanti linguistici propriamente detti ed anche da analoghe raccolte empiriche, fermo restando il principio che quest'ultime siano di natura *basilettale*;
- si tratta di un metodo che, nato in seno alla Romanistica, rientra nella tradizione della geografia linguistica romanza 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 particolare al fine di rimediare alle eventuali insufficienze materiali della diffusione dei suoi risultati in forma meramente cartacea. Questo si riferisce soprattutto alle molteplici resistenze amministrative e logistiche nonché alle spesso insormontabili difficoltà finanziarie che ostacolano l'uso dei *colori* per la riproduzione degli irrinunciabili grafici.

Dal momento che questi problemi possono ora essere facilmente risolti – almeno sul piano informatico – tramite il nostro programma VDM (“Visual DialectoMetry”), vi propongo, per concludere, la nostra duplice offerta di spedire da un lato, a chi ne fa richiesta, 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.

44 Rimando in merito alla nostra cooperazione interdisciplinare con un gruppo di genetisti italiani pubblicata nel 2005: Scapoli / Goebel / Sobota / Mamolini / Rodriguez-Laralde / Barrai 2005.

13. Ringraziamenti

- Sussidio finanziario: “Fonds zur Förderung der wissenschaftlichen Forschung in Österreich” (FWF), Vienna [Fondazione austriaca per la ricerca scientifica] (progetto di ricerca 18 365)
- La tassazione ed immissione informatica delle 3 911 carte di lavoro (CL) è stata eseguita dai seguenti romanisti salisburghesi: Eme-se 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 delle 4 figure e 24 tavole: 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.

Riferimenti bibliografici e abbreviazioni

- AIS = JABERG, Karl / JUD, Jakob (1928-1940): *Sprach- und Sachatlas Italiens und der Südschweiz*. Zofingen: Ringier (ristampa: Nendeln: Kraus 1971). 8 voll.
- ALF = GILLIÉRON, Jules / EDMONT, Edmond (1902-1910): *Atlas linguistique de la France*. 10 vols. Paris: Champion (ristampa: Bologna: Forni, 1968).
- ALTMANN, Gabriel (1985): «Die Entstehung diatopischer Varianten. Ein stochastisches Modell». *Zeitschrift für Sprachwissenschaft* 4, p. 139-145.
- BAUER, Roland (2009): *Dialektometrische Einsichten. Sprachklassifikatorische Oberflächenmuster und Tiefenstrukturen im lombardisch-venedischen Dialektraum und in der Rätoromania*. San Martin de Tor: Istitut ladin Micurà de Rü (Ladinia Monographica 01).
- CAF = Coefficiente di Asimmetria di Fisher
- CDS = carta di similarità
- CGA = Classificazione Gerarchica Agglomerativa
- CL = carta di lavoro
- DC = dendrema-corema
- DEES, Anthonij (1980): *Atlas des formes et des constructions des chartes françaises du 13^e siècle*. Tübingen: Niemeyer (Beihefte zur Zeitschrift für Romanische Philologie, vol. 178).
- DM = dialettometria, relativo alla dialettometria, dialettometrico
- GOEBL, Hans (1981): «Éléments d’analyse dialectométrique (avec application à l’AIS)». *RLiR* 45, p. 349-420.
- (1982): «Dialektometrie. Prinzipien und Methoden des Einsatzes der Numerischen Taxonomie im Bereich der Dialektgeographie». *Denkschriften [Memorie], philosophisch-historische Klasse, vol. 157, Wien: Österreichische Akademie der Wissenschaften*, p. 1-123 [41 tavole].

- (1983): «Parquet polygonal et treillis triangulaire: les deux versants de la dialectométrie interponctuelle». *RLiR* 47, p. 353-412.
- (1984a): *Dialektometrische Studien. Anband italo-romanischer, rätoromanischer und galloromanischer Sprachmaterialien aus AIS und ALF*. Tübingen: Niemeyer (Beihefte zur Zeitschrift für Romanische Philologie, vol. 191-193). 3 voll.
- (1984b): «Lineamenti di dialettometria (con applicazione all'AIS)». *Guida ai dialetti veneti* VI, p. 7-53.
- (1987): «Points chauds de l'analyse dialectométrique: pondération et visualisation». *RLiR* 51, p. 63-118.
- (1988): «Il posto dialettometrico che spetta ai punti-AIS 338 (Adornano, Friuli), 398 (Dignano/Vodnjan, Istria) e 367 (Grado, Friuli). Presentazione di tre carte di similarità». *Linguistica* 28, p. 75-103.
- (2002): «Analyse dialectométrique des structures de profondeur de l'ALF». *RLiR* 66, p. 5-63.
- (2003a): «Regards dialectométriques sur les données de l'Atlas linguistique de la France (ALF): relations quantitatives et structures de profondeur». *Estudis Romànics* 25, p. 59-120.
- (2003b): «Entwicklung der romanischen Sprachgrenzen: Italo-romania und Ostalpenraum». ERNST, Gerhard / GLEßGEN, Martin-Dietrich / SCHMITT, Christian / SCHWEICKARD, Wolfgang (eds.): *Romanische Sprachgeschichte. Histoire linguistique de la Romania. Ein internationales Handbuch zur Geschichte der romanischen Sprachen. Manuel international d'histoire linguistique de la Romania*. Vol. 1. Berlin / New York: de Gruyter, p. 631-645.
- (2005): «La dialectométrie corrélatrice: un nouvel outil pour l'étude de l'aménagement dialectal de l'espace par l'homme». *RLiR* 69, p. 321-367.
- (2007): «Sprachgeographische Streifzüge durch das Netz des Sprachatlases AIS». *Ladinia* XXXI, p. 187-271.
- (2008a): «La dialettometrizzazione integrale dell'AIS. Presentazione dei primi risultati». *RLiR* 72, p. 25-113.
- (2008b): «Sur le changement macrolinguistique survenu entre 1300 et 1900 dans le domaine d'Oil. Une étude diachronique d'inspiration dialectométrique». *Dialectologia* 1, p. 3-43.
- (2010): «Introducción a los problemas y métodos según los principios de la Escuela Dialectométrica de Salzburgo (con ejemplos sacados del "Atlante italo-svizzer", AIS)». AURREKOETXEA, Gotzon / ORTMAETXEA, Jose Luis (eds.): *Tools for Linguistic Variation*. Bilbao / Bilbo: Universidad del País Vasco / Euskal Herriko Unibertsitatea, p. 3-39.
- (2011): «Brevissima presentazione della dialettometria (Con esempi tratti dall'AIS)». BOMBI, Raffaella / D'AGOSTINO, Mari / DAL NEGRO, Silvia / FRANCESCHINI, Rita (eds.): *Lingue e culture in contatto. In memoria di Roberto GUSMANI, Bolzano, 18-19 febbraio 2010. Atti del 10° Congresso dell'Associazione Italiana di Linguistica Applicata*. Perugia: Guerra edizioni (Pubblicazioni della Associazione Italiana di Linguistica Applicata), p. 71-103.

IRD_{jk} = Indice Relativo di Distanza

IRI_{jk} = Indice Relativo d'Identità

OKABE, Atsuyuki / BOOTS, Barry / SUGIHARA, Kokichi (1992): *Spatial Tesselations. Concepts and Applications of Voronoi Diagrams*. Chichester, New York, Brisbane, Toronto, Singapore: Wiley.

QuoCo = Quoziente di Comunicazione

r(BP) = coefficiente di correlazione di [Auguste] Bravais e [Karl] Pearson

- SCAPOLI, Chiara / GOEBL, Hans / SOBOTA, Slawomir / MAMOLINI, Elisabetta / RODRIGUEZ-LARRALDE, Alvaro / BARRAI, Italo (2005): «Surnames and Dialects in France: Population Structure and cultural evolution». *Journal of Theoretical Biology* 237, p. 75-86.
- SDMS = “Scuola Dialettometrica di Salisburgo”
- SNEATH, Peter A. / SOKAL, Robert R. (1973): *Numerical Taxonomy. The Principles and Practice of Numerical Classification*. San Francisco: Freeman.
- VDM = “Visual DialectoMetry” (software dialettometrico, utilizzato a Salisburgo)
- VIDESOTT, Paul (2009): *Padania scrittologica: analisi scrittologiche e scrittometriche di testi in italiano settentrionale antico dalle origini al 1525*. Tübingen: Niemeyer (Beihefte zur Zeitschrift für Romanische Philologie, vol. 353).
- WARD, Joe H., Jr. (1963): «Hierarchical Grouping to optimize an Objective Function». *Journal of the American Statistical Association* 58, p. 236-244.

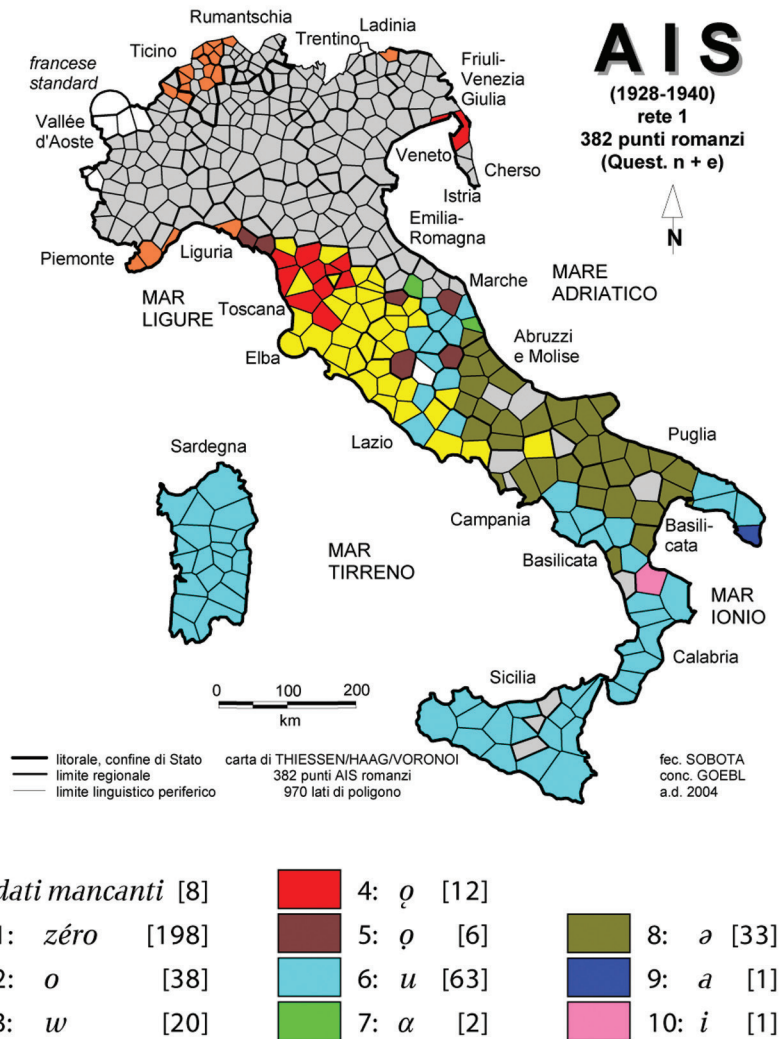


TAVOLA 1. Una delle 1 766 carte di lavoro di tipo *fonetico*: la distribuzione geografica dei risultati della -U finale nell'etimo latino COGNÁTU (secondo AIS 27 *il suo cognato*)

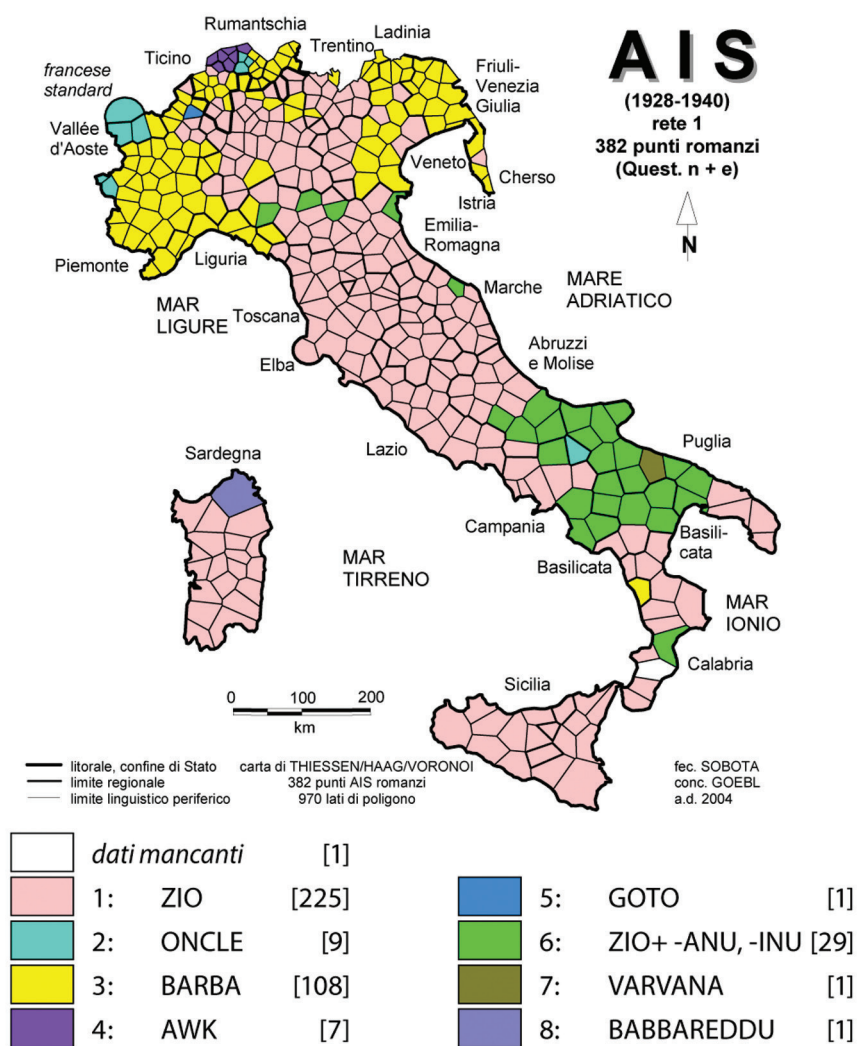
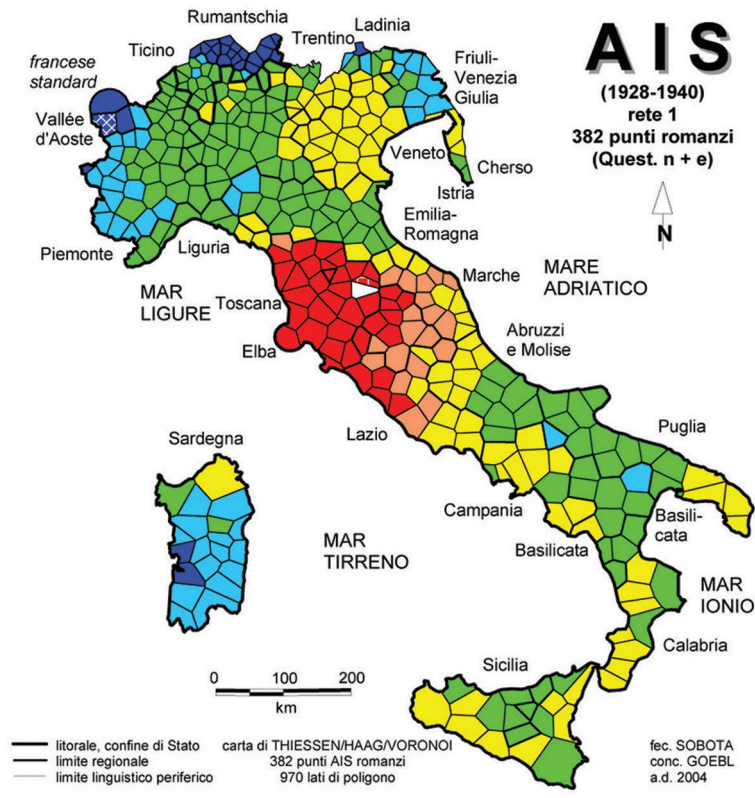


TAVOLA 2. Una delle 1 225 carte di lavoro di tipo lessicale: la distribuzione geografica delle denominazioni del concetto "zio" (secondo AIS 19 il loro zio)



Visualizzazione MINMWMAX 6-tuplo

[1]	46.22 – 52.05	(n = 26)
[2]	– 57.87	(n = 49)
[3]	– 63.70	(n = 152)
[4]	– 71.35	(n = 96)
[5]	– 78.99	(n = 20)
[6]	– 86.64	(n = 38)
		(Σ = 381)

Distribuzione di similarità MINMWMAX 6-tuplo

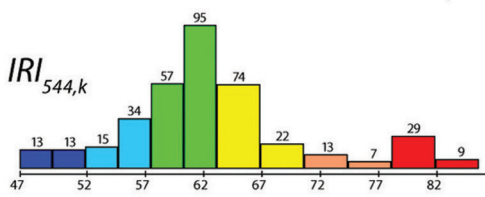
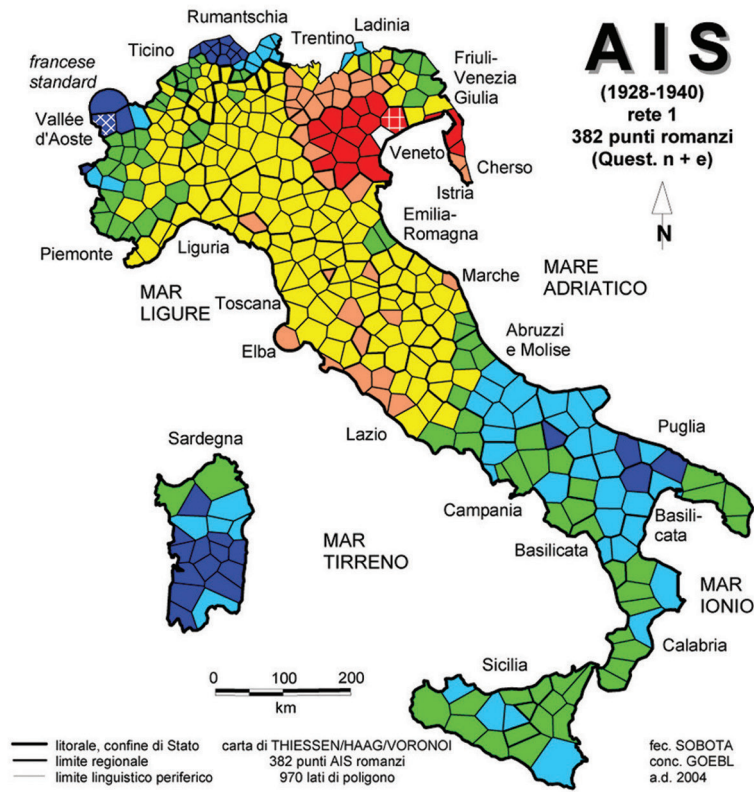


TAVOLA 3. Carta di similarità relativa al punto di riferimento 544 (Arezzo, Toscana)
Indice di similarità: IRI_{jk}
Corpus: 3 911 carte di lavoro (di tutte le categorie linguistiche)
Visualizzazione: MINMWMAX 6-tuplo



Visualizzazione MINMWMAX 6-tuplo

[1]	46.52 – 51.86	(n = 33)
[2]	– 57.21	(n = 49)
[3]	– 62.55	(n = 87)
[4]	– 68.91	(n = 161)
[5]	– 75.27	(n = 34)
[6]	– 81.63	(n = 17)
		(Σ = 381)

Distribuzione di similarità MINMWMAX 6-tuplo

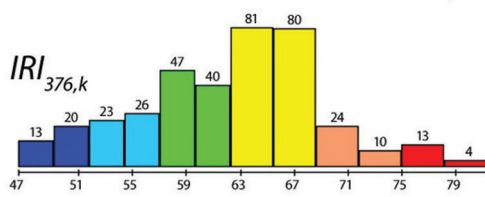
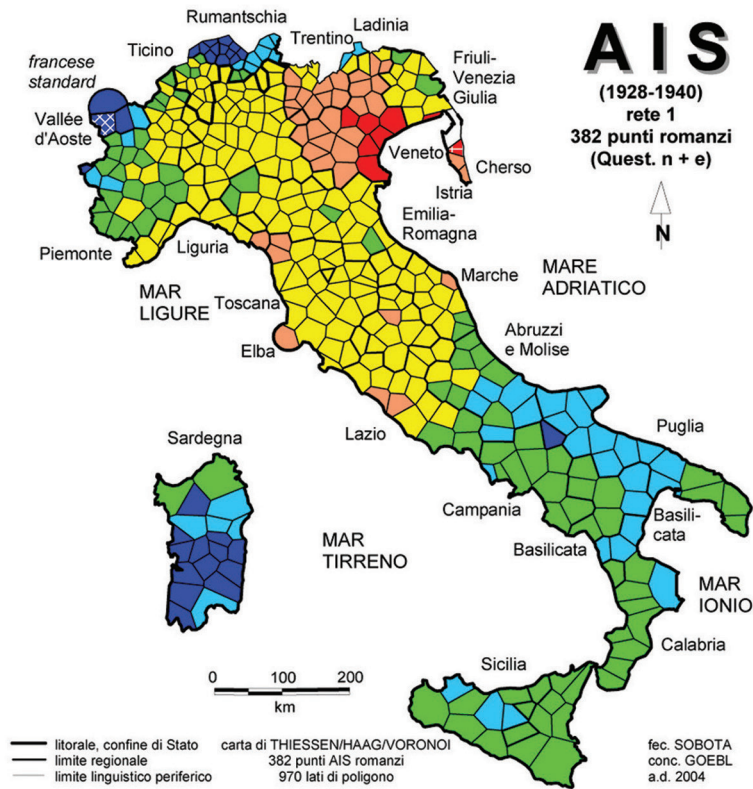


TAVOLA 4. Carta di similarità relativa al punto di riferimento 376 (Venezia, *sujet 1*; Veneto)
 Indice di similarità: IRI_{jk}
 Corpus: 3 911 carte di lavoro (di tutte le categorie linguistiche)
 Visualizzazione: MINMWMAX 6-tuplo



Visualizzazione MINMWMAX 6-tuplo

[1]	46.71 – 52.34	(n = 30)
[2]	– 57.97	(n = 44)
[3]	– 63.59	(n = 89)
[4]	– 70.13	(n = 173)
[5]	– 76.67	(n = 37)
[6]	– 83.21	(n = 8)
		($\Sigma = 381$)

Distribuzione di similarità MINMWMAX 6-tuplo

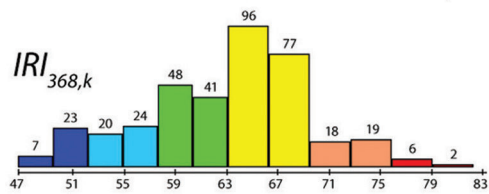
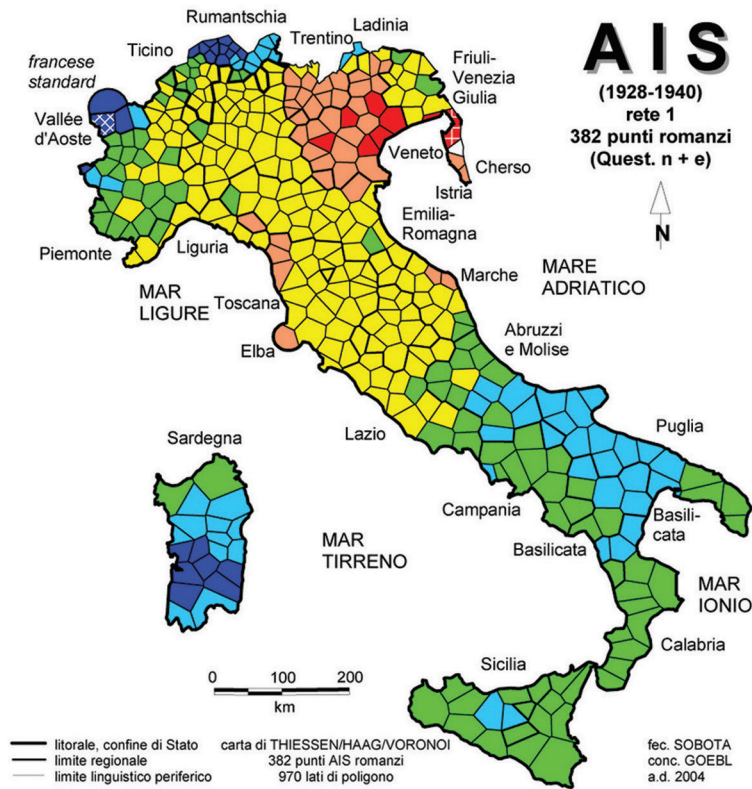


TAVOLA 5. Carta di similarità relativa al punto di riferimento 368 (Pirano/Piran, Slovenia)

Indice di similarità: IRI_k

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

Visualizzazione: MINMWMAX 6-tuplo



Visualizzazione MINMWMAX 6-tuplo

[1]	44.25 – 50.30	(n = 22)
[2]	– 56.35	(n = 51)
[3]	– 62.41	(n = 94)
[4]	– 69.34	(n = 167)
[5]	– 76.28	(n = 39)
[6]	– 83.21	(n = 8)
		(Σ = 381)

Distribuzione di similarità MINMWMAX 6-tuplo

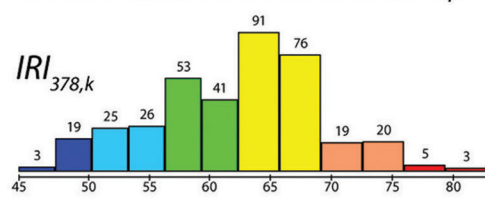


TAVOLA 6. Carta di similarità relativa al punto di riferimento 378 (Montona/Motovun, Croazia)
Indice di similarità: IRI_k
Corpus: 3 911 carte di lavoro (di tutte le categorie linguistiche)
Visualizzazione: MINMWMAX 6-tuplo

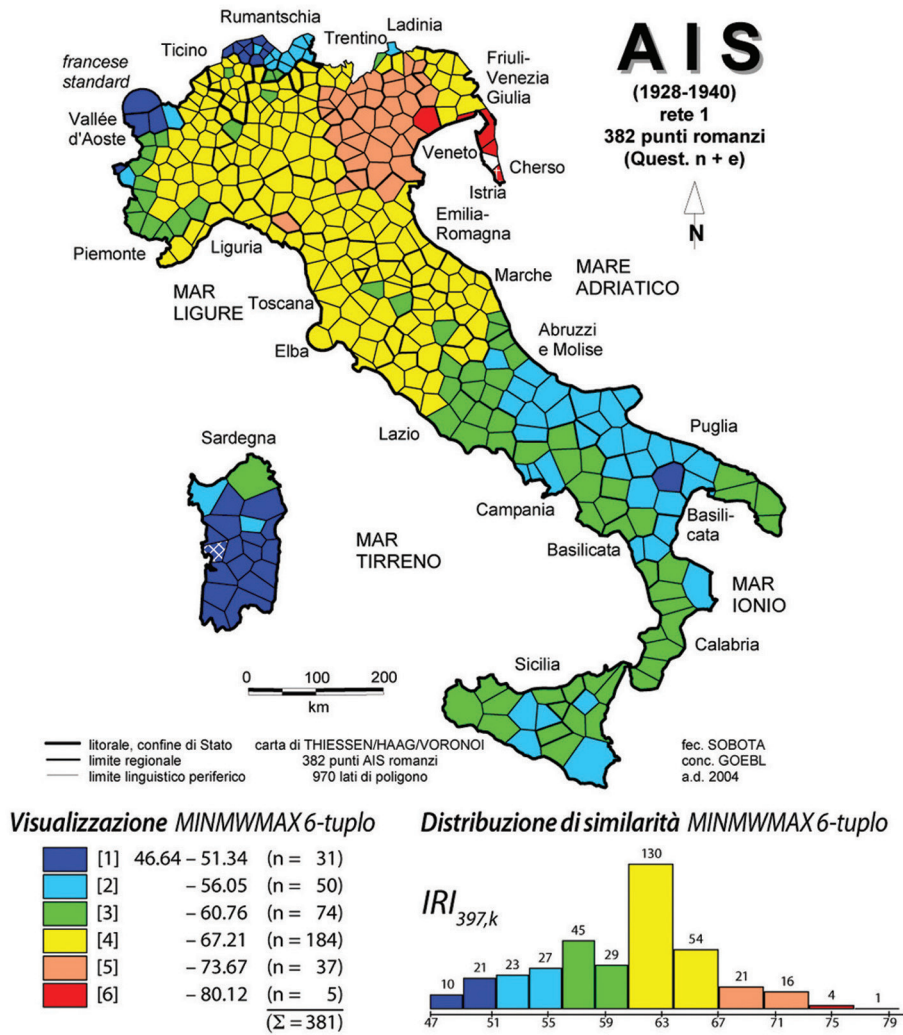
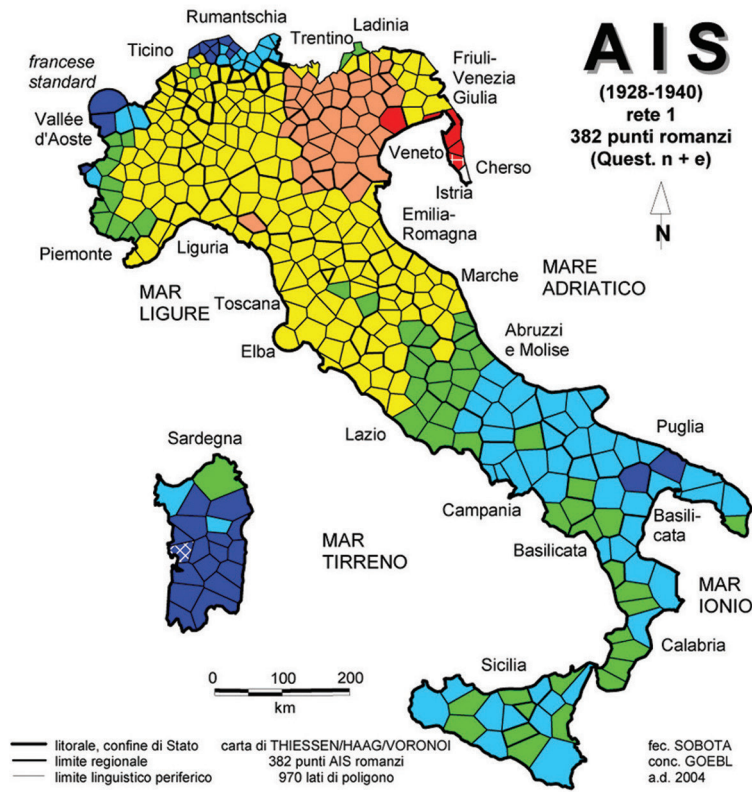


TAVOLA 7. Carta di similarità relativa al punto di riferimento 397 (Rovigno/Rovinj, Croazia)

Indice di similarità: IRI_{ik}

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

Visualizzazione: MINMWMAX 6-tuplo



Visualizzazione MINMWMAX 6-tuplo

[1]	45.79–50.75	(n = 30)
[2]	–55.71	(n = 64)
[3]	–60.68	(n = 58)
[4]	–67.16	(n = 183)
[5]	–73.64	(n = 41)
[6]	–80.12	(n = 5)
		(Σ = 381)

Distribuzione di similarità MINMWMAX 6-tuplo

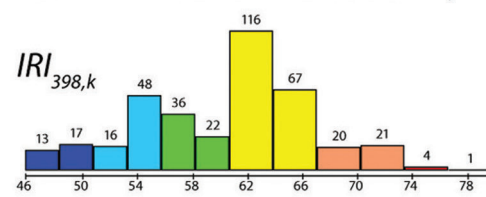


TAVOLA 8. Carta di similarità relativa al punto di riferimento 398 (Dignano/Vodnjan, Croazia)
 Indice di similarità: IRI_{ik}
 Corpus: 3 911 carte di lavoro (di tutte le categorie linguistiche)
 Visualizzazione: MINMWMAX 6-tuplo

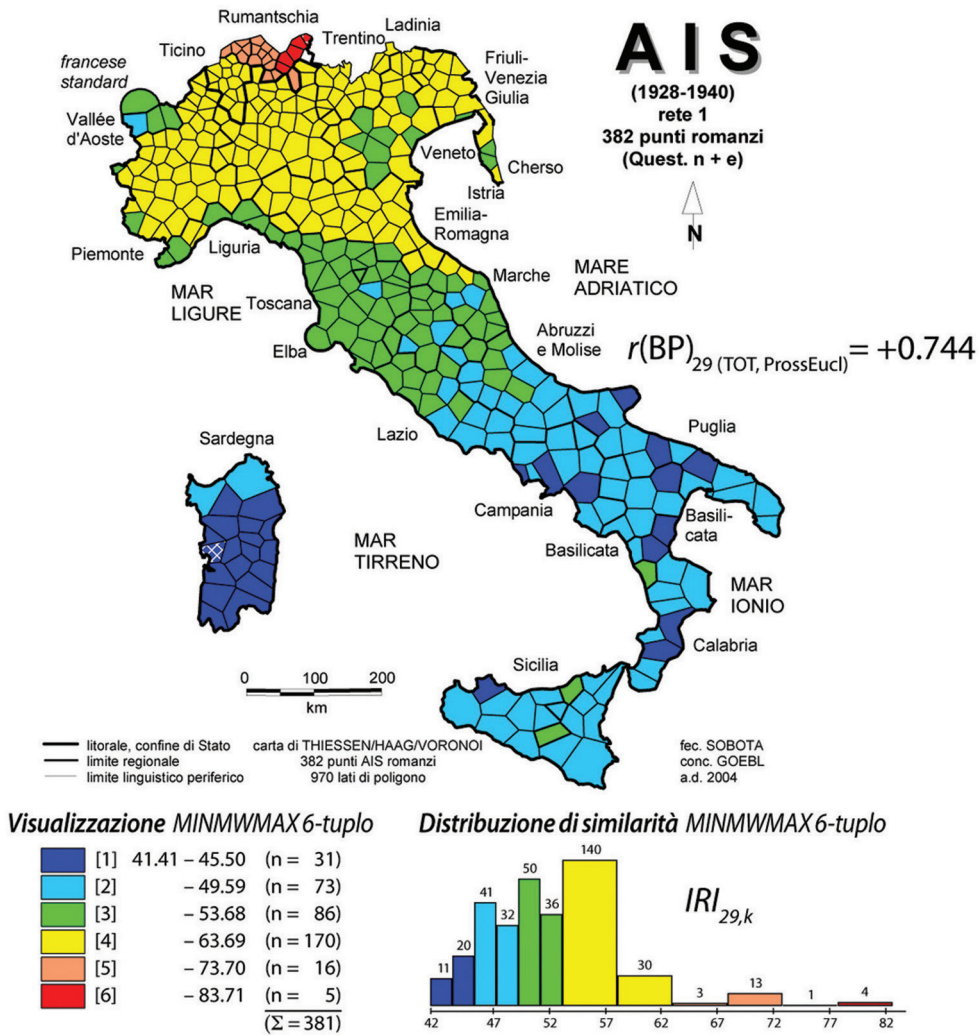


TAVOLA 9. Carta di similarità relativa al punto di riferimento 29 (Santa Maria, Val Monastero, Grigioni, Svizzera)

Indice di similarità: $IRI_{29,k}$

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

Visualizzazione: MINMWMAX 6-tuplo

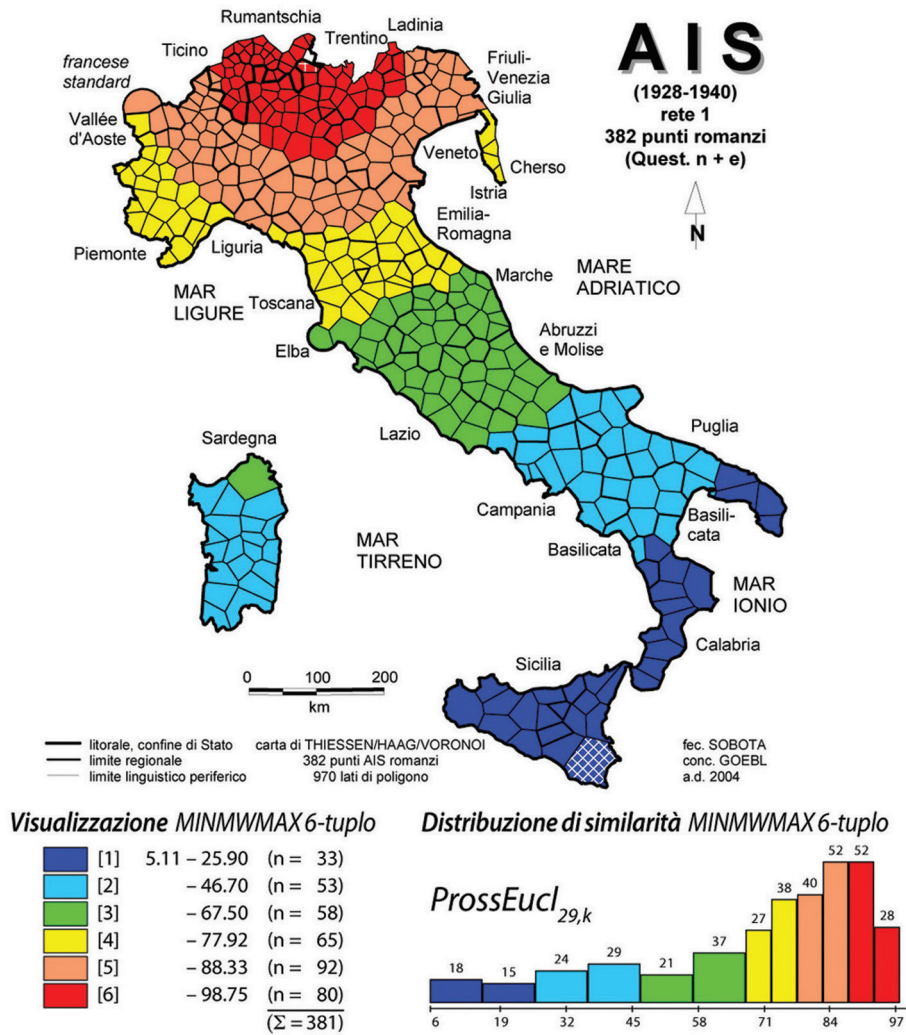


TAVOLA 10. Carta di prossimità relativa al punto di riferimento 29 (Santa Maria, Val Monastero, Grigioni, Svizzera)

Indice di prossimità: 100 - distanza euclidea (secondo il teorema di Pitagora)

Visualizzazione: MINMWMAX 6-tuplo

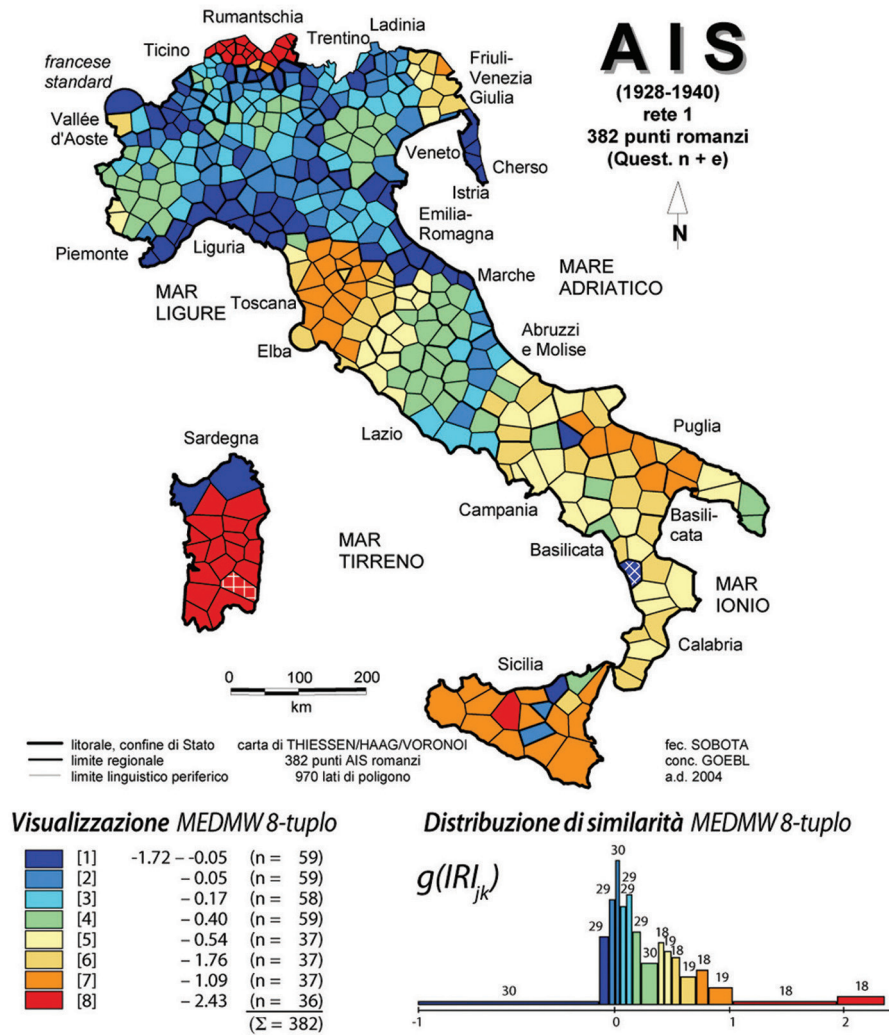


TAVOLA 11. Sinossi coropletica dei “Coefficienti di Asimmetria di Fisher” (CAF) $[g(IRI_{jk})]$ di 382 distribuzioni di similarità
Indice di similarità: IRI_{jk}
Corpus: 3 911 carte di lavoro (di tutte le categorie linguistiche)
Visualizzazione: MEDMW 8-tuplo

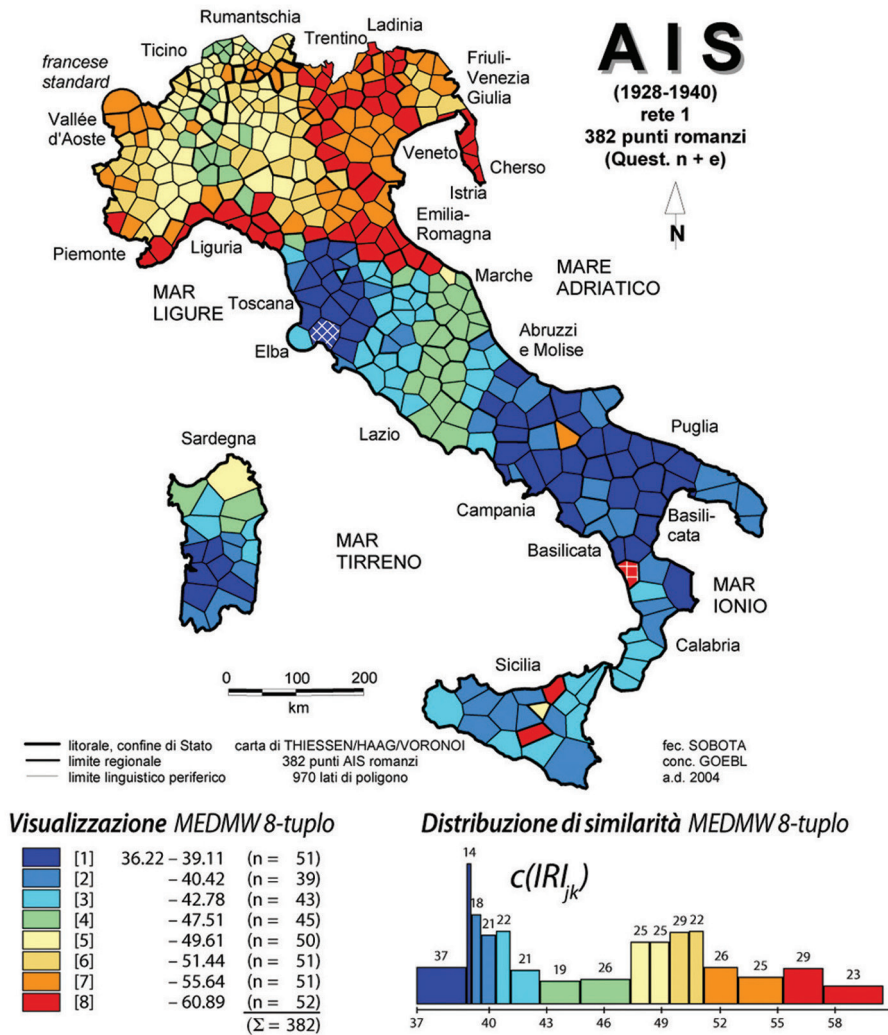


TAVOLA 12. Sinossi coropletica dei “Quozienti di Comunicazione” (QuoCo) [$c(IRI_{jk})$] di 382 distribuzioni di similarità
Indice di similarità: IRI_{jk}
Corpus: 3 911 carte di lavoro (di tutte le categorie linguistiche)
Visualizzazione: MEDMW 8-tuplo

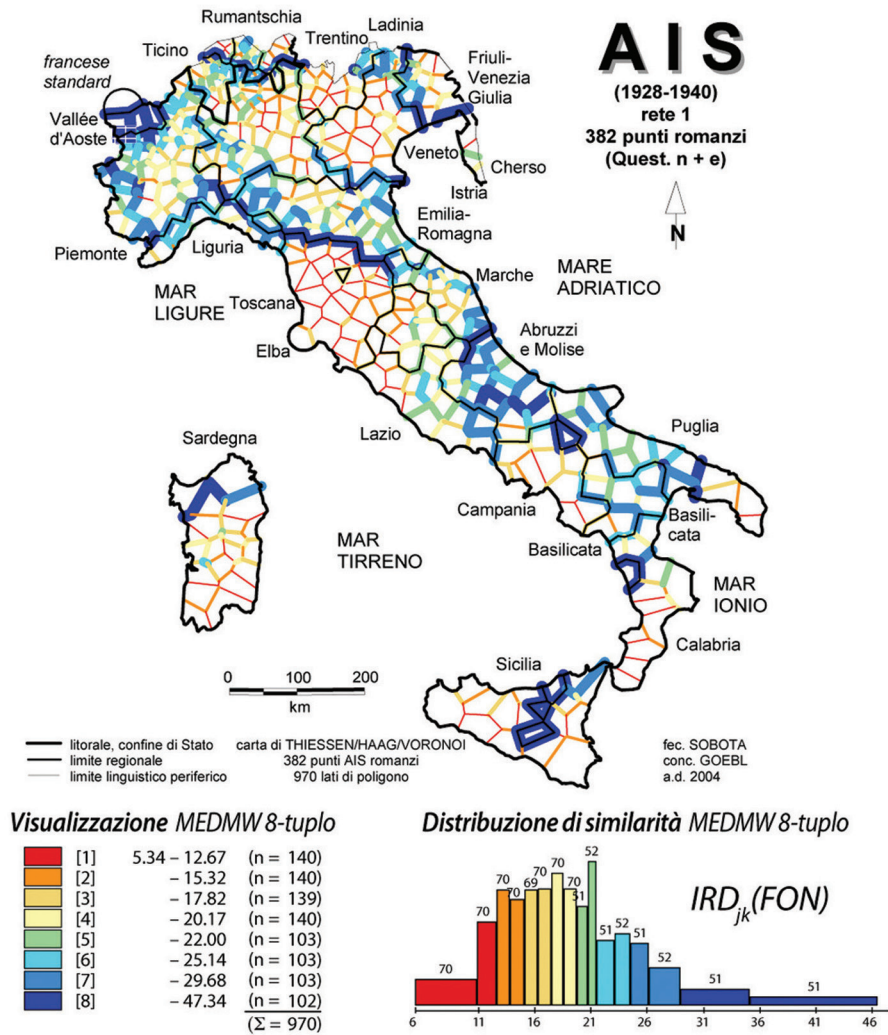


TAVOLA 13. Carta isoglottica
Indice di distanza: IRD_{jk}
Corpus: 1 766 carte di lavoro di tipo fonetico
Visualizzazione: MED 8-tuplo

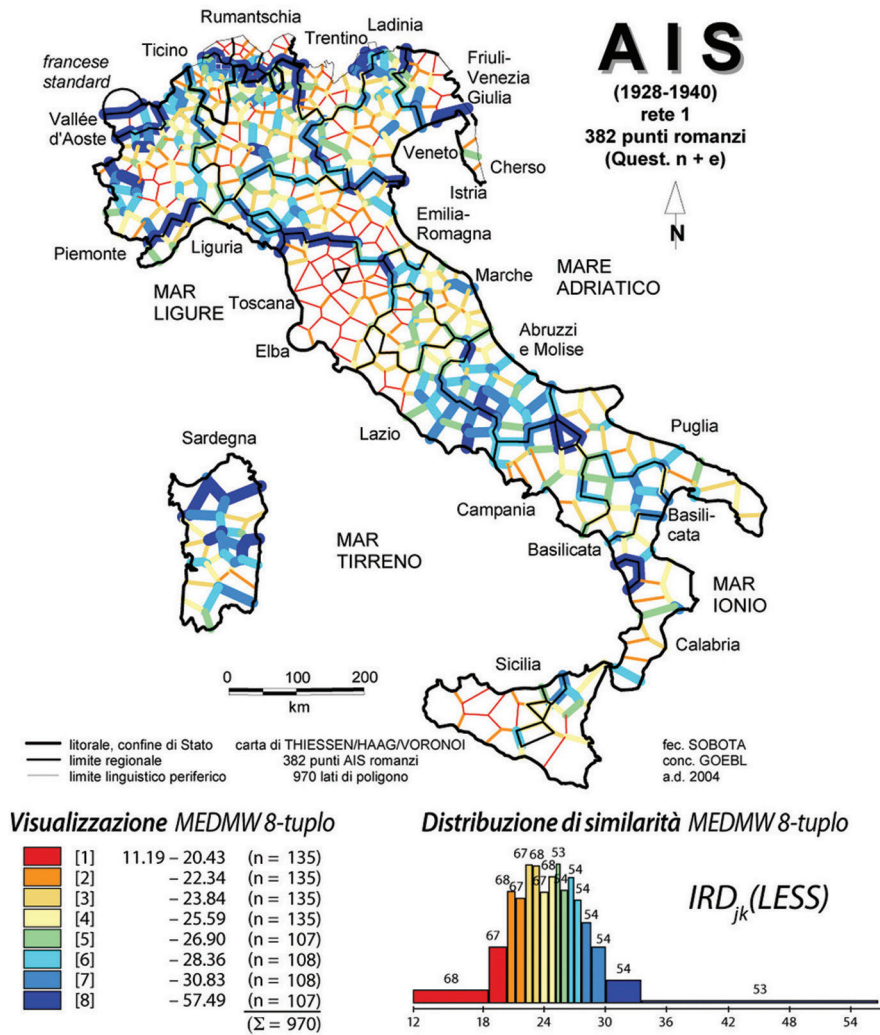


TAVOLA 14. Carta isoglottica
Indice di distanza: IRD_{jk}
Corpus: 1 225 carte di lavoro di tipo lessicale
Visualizzazione: MED 8-tuplo

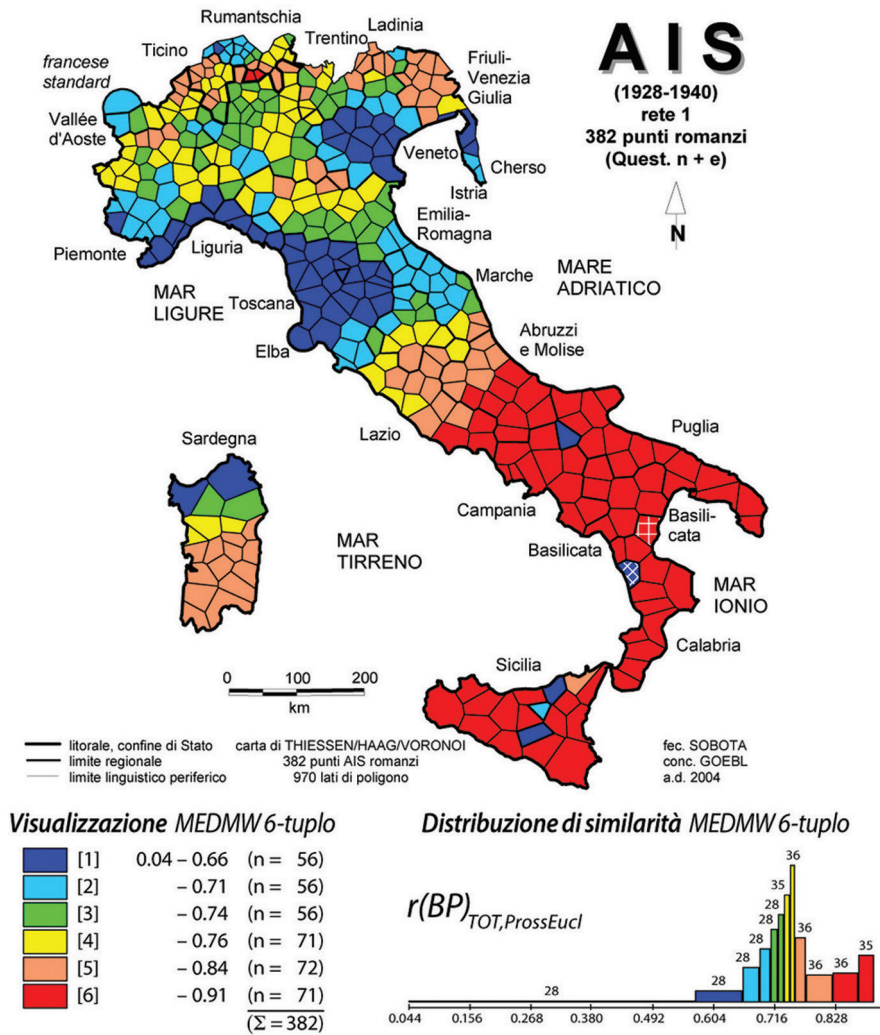


TAVOLA 15. Carta a correlazioni [mediante $r(BP)$]:
similarità linguistica (secondo IRI_{jk}) e *prossimità geografica* (secondo il teorema di Pitagora)
 Corpus linguistico: 3 911 carte di lavoro (di tutte le categorie linguistiche)
 Visualizzazione: MEDMW 6-tuplo

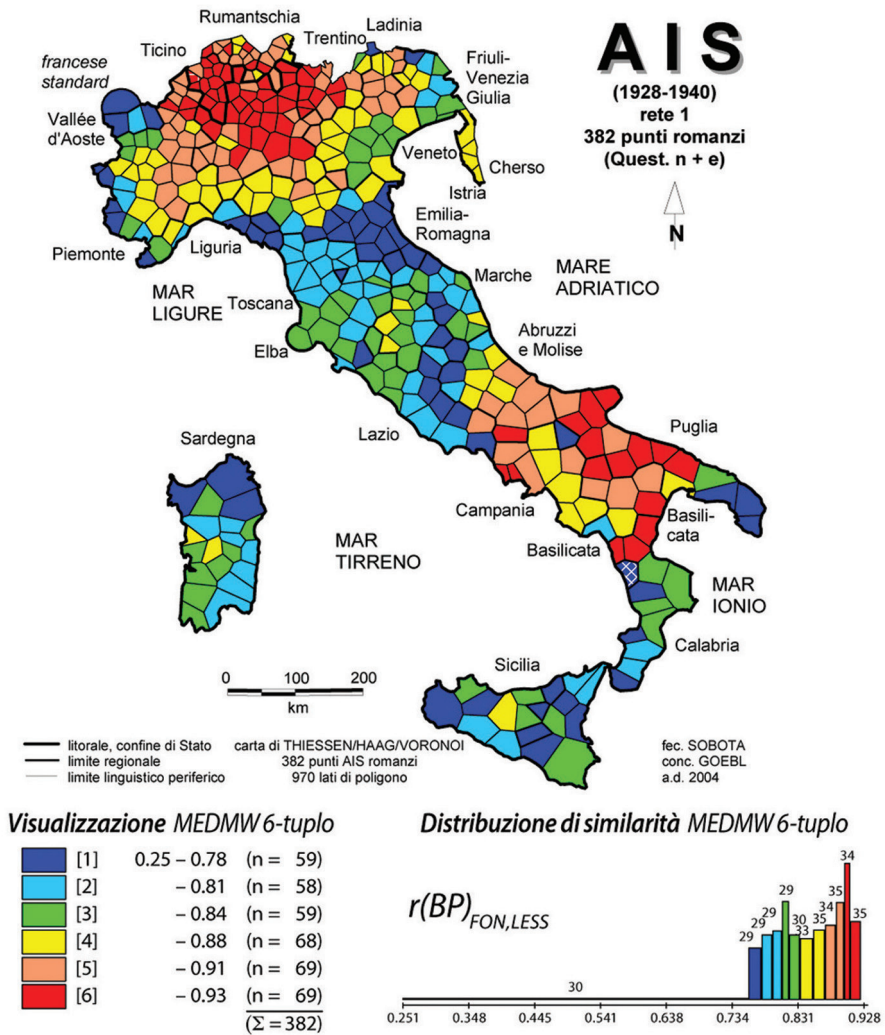


TAVOLA 16, Carta a correlazioni [mediante $r(BP)$]:
 similarità fonetica e similarità lessicale (sempre secondo IRI_{jk})
 Corpus fonetico: 1 766 carte di lavoro; corpus lessicale: 1 225 carte di lavoro
 Visualizzazione: MEDMW 6-tuplo

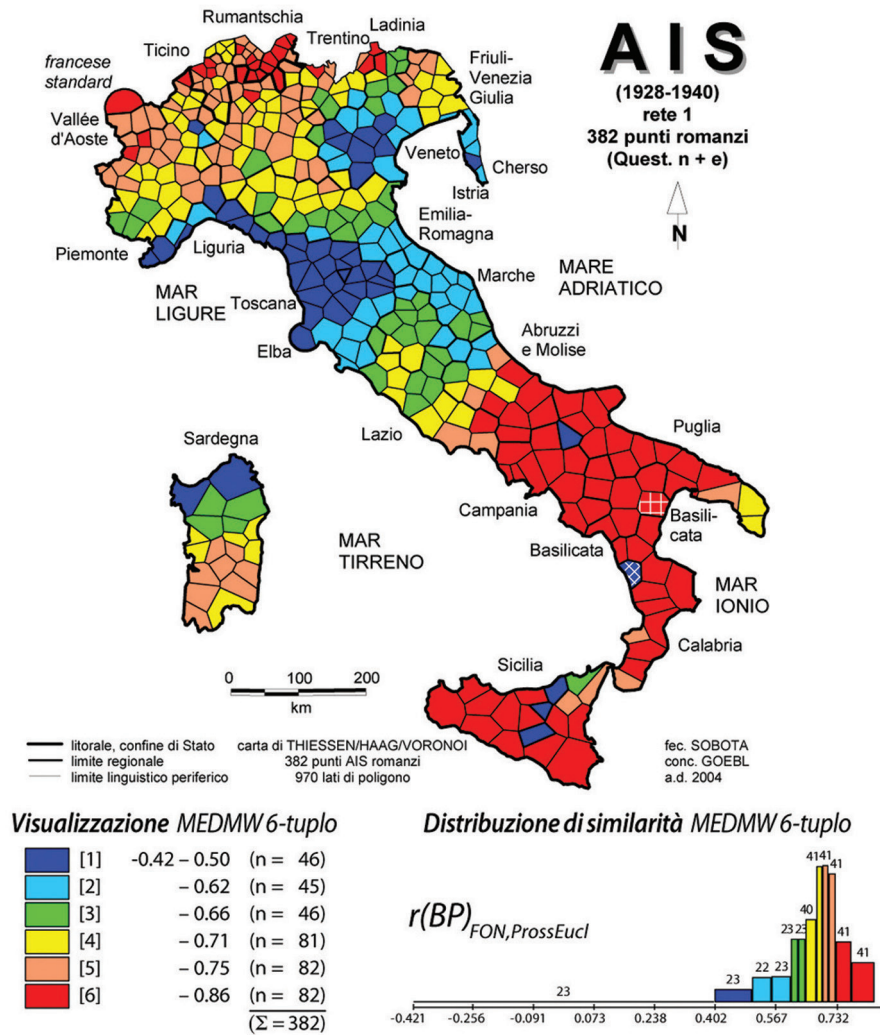


TAVOLA 17. Carta a correlazioni [mediante $r(BP)$]:
similarità fonetica (secondo IRI_k) e *prossimità geografica* (secondo il teorema di Pitagora)
 Corpus fonetico: 1 766 carte di lavoro
 Visualizzazione: MEDMW 6-tuplo

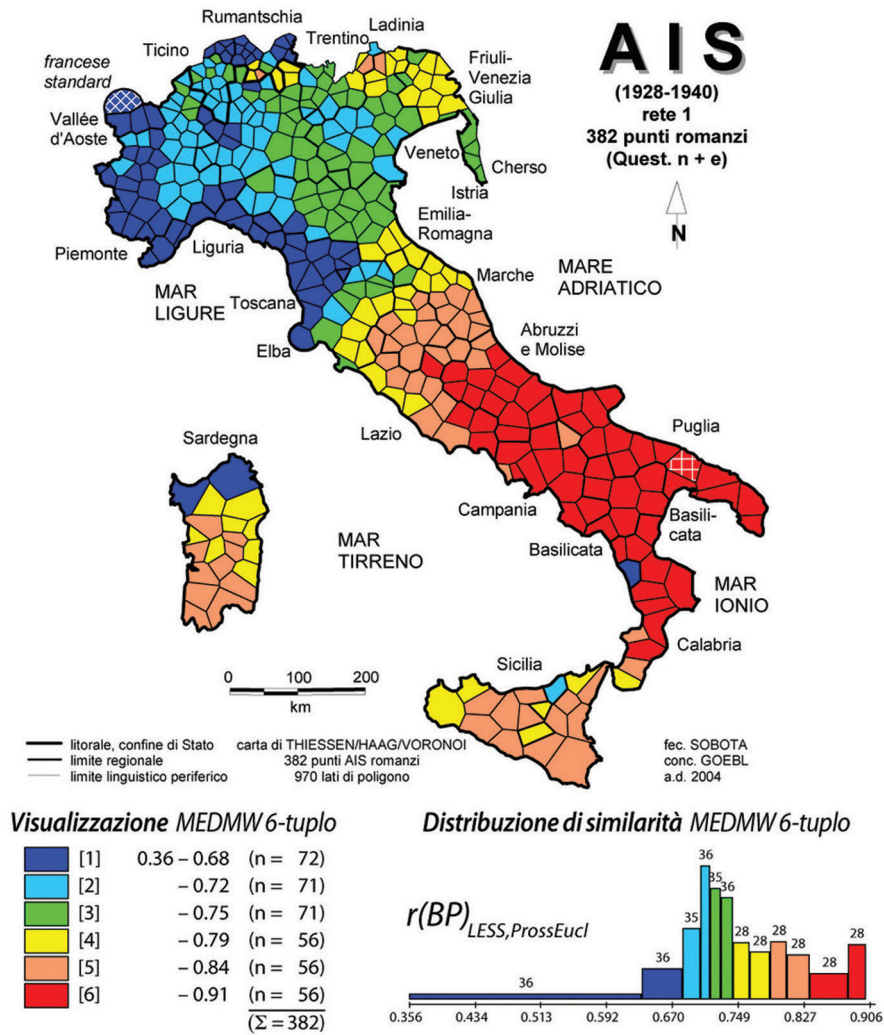


TAVOLA 18, Carta a correlazioni [mediante $r(BP)$]:
similarità lessicale (secondo IRI_k) e prossimità geografica (secondo il teorema di Pitagora)
Corpus lessicale: 1 225 carte di lavoro
Visualizzazione: MEDMW 6-tuplo

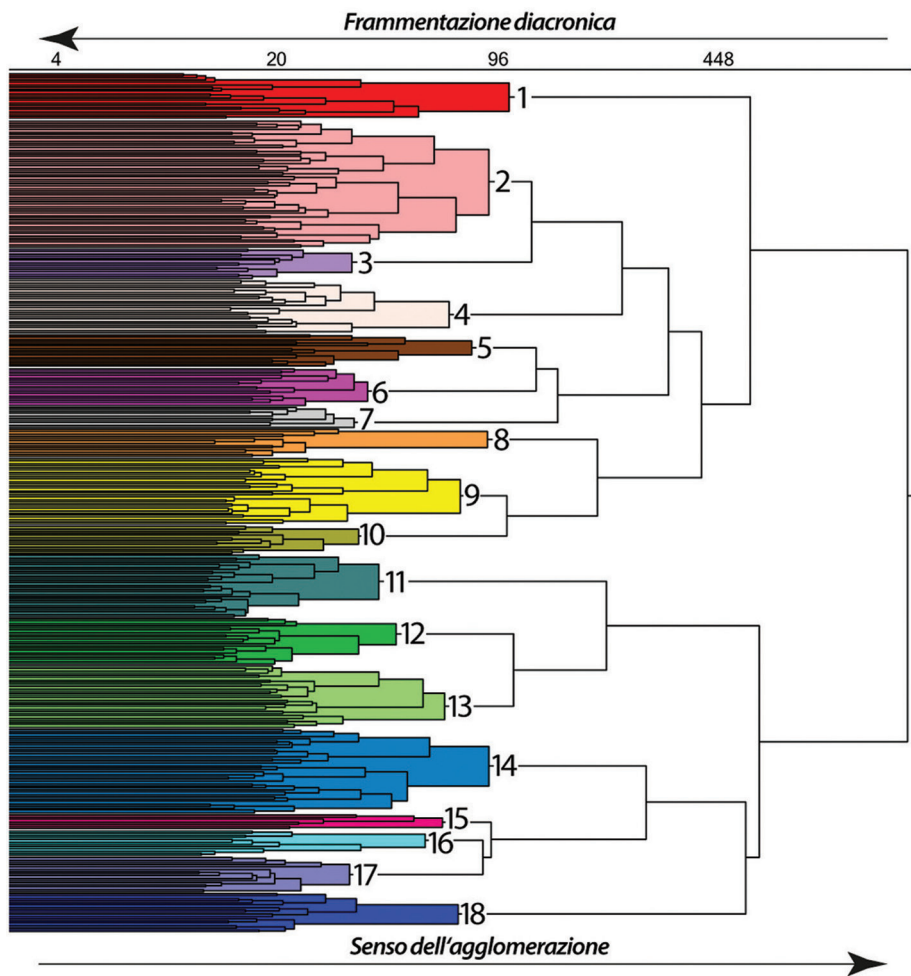


TAVOLA 19. Classificazione Gerarchica Agglomerativa: metodo di Joe Ward, Jr.
 Indice di similarità: IRI_{jk}
 Corpus: 3 911 carte di lavoro (di tutte le categorie linguistiche)
 Numero dei dendremi colorati: 18

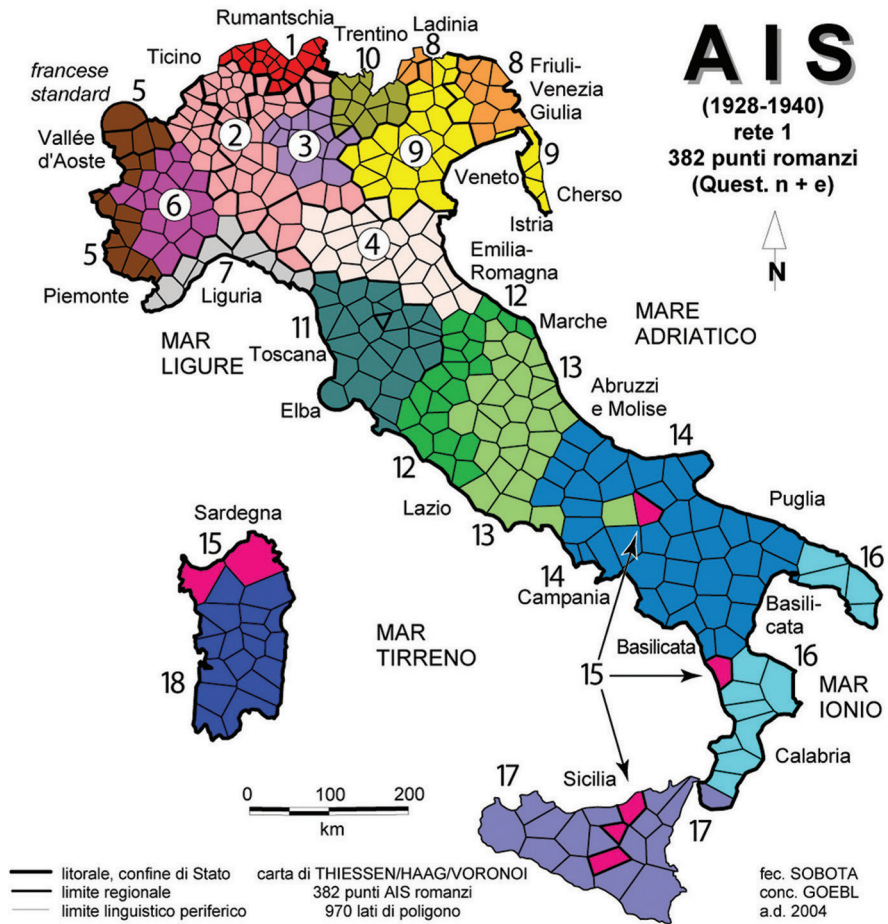
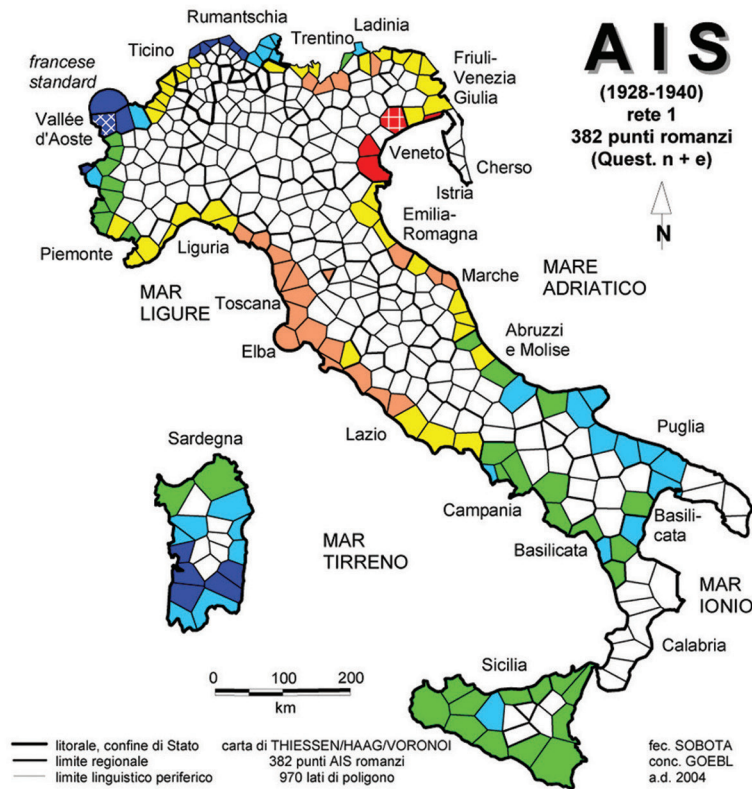


TAVOLA 20. Spazializzazione dei 18 dendremi colorati della tavola 19



Visualizzazione MINMWMAX 6-tuplo

[1]	46.52 – 50.93	(n = 15)
[2]	– 55.34	(n = 27)
[3]	– 59.75	(n = 35)
[4]	– 67.04	(n = 38)
[5]	– 74.34	(n = 22)
[6]	– 81.63	(n = 4)
		($\Sigma = 141$)

Distribuzione di similarità MINMWMAX 6-tuplo

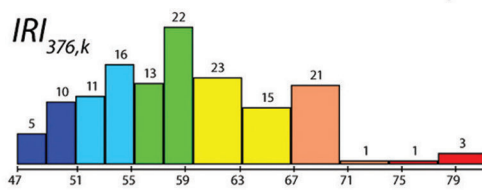


TAVOLA 21. Carta di similarità relativa al punto di riferimento 376 (Venezia, sujet 1; Veneto)

Indice di similarità: IRI_k

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

Rete diluita a 142 punti-AIS in posizione periferica

Visualizzazione: MINMWMAX 6-tuplo

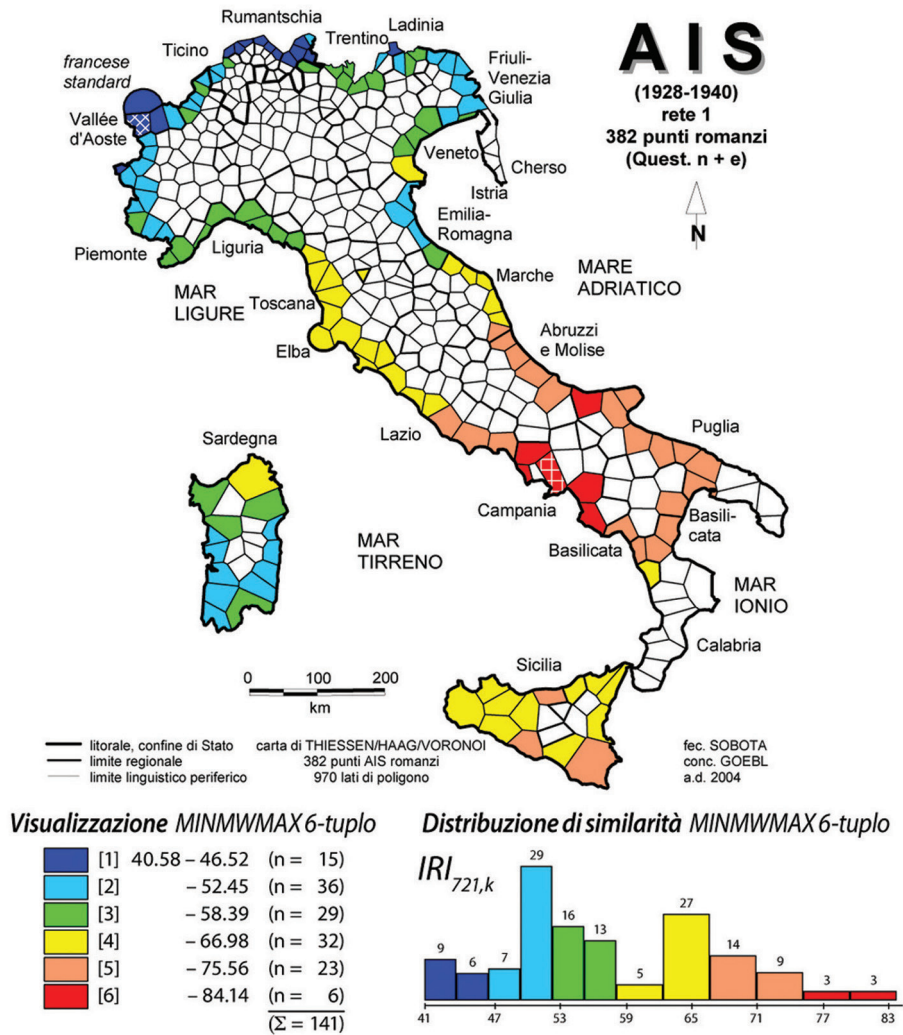


TAVOLA 22. Carta di similarità relativa al punto di riferimento 721 (Napoli, Campania)
 Indice di similarità: $IRI_{721,k}$
 Corpus: 3 911 carte di lavoro (di tutte le categorie linguistiche)
 Rete diluita a 142 punti-AIS in posizione periferica
 Visualizzazione: MINMWMAX 6-tuplo

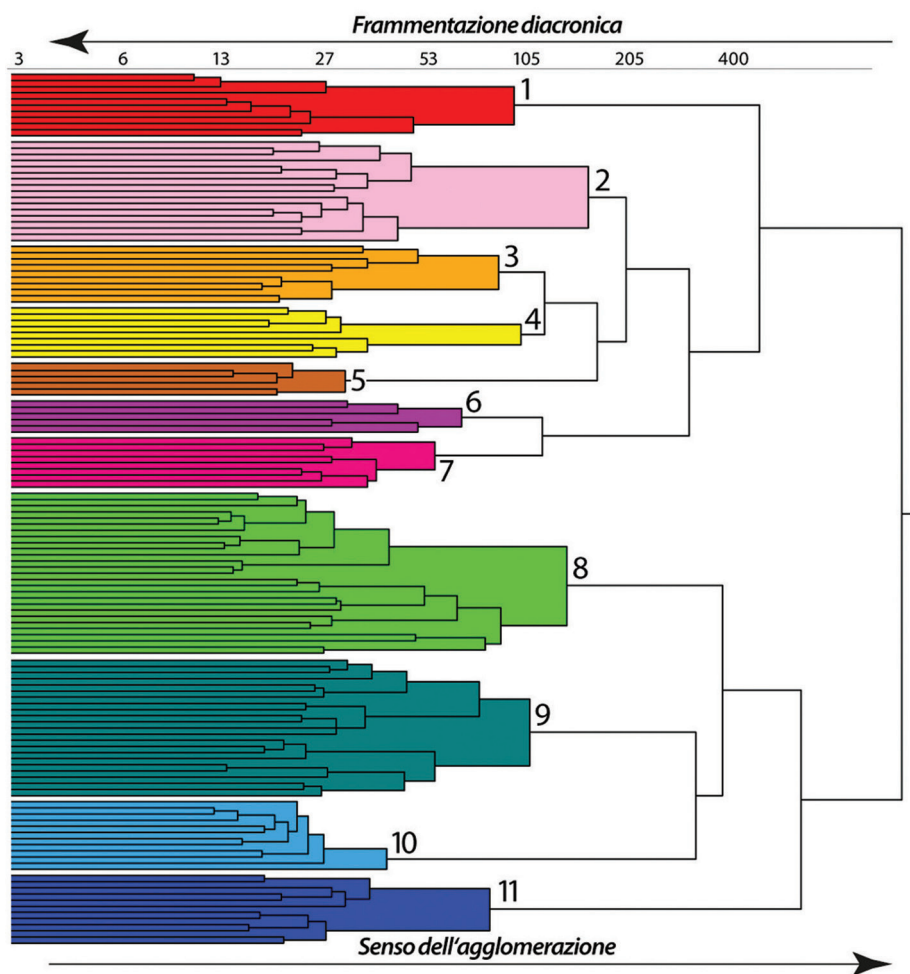


TAVOLA 23. Classificazione Gerarchica Agglomerativa (CGA): metodo di Joe Ward, Jr.
 Indice di similarità: IRI_{jk}
 Corpus: 3 911 carte di lavoro (di tutte le categorie linguistiche)
 Numero dei dendremi colorati: 11
 Rete diluita a 142 punti-AIS in posizione periferica

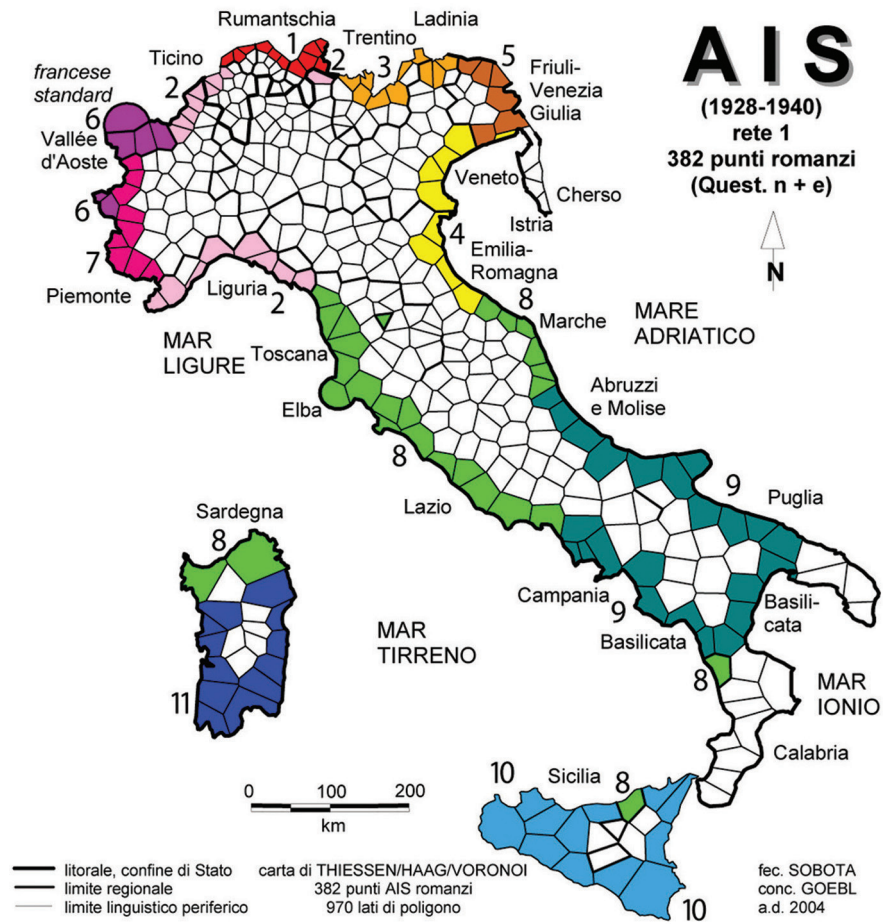


TAVOLA 24. Spazializzazione degli 11 dendremi colorati della tavola 23