
Dialectometria

www.colinguistica.org

Alguns problemas da dialectologia *clássica*

- Não ha fronteiras bem definidas entre áreas
 - Progressividade das fronteiras lineares
 - Problemas dos feixes de isoglossas
 - A subjectividade nas classificações dialectais *clássicas*
 - A escolha das isoglossas é (relativamente) subjectiva
 - Existem diferentes classificações
-

Galician speaking areas

The exact limits between Galician and asturleonese languages have been agreed internationally since the late 19th century by almost all dialectologists, the discerning isogloss being that where short latin E and O get transformed into diphthongs (pedra / piedra, roda / rueda).

Based on F. Fernandez Rei's outline as shown in his book *Dialectoloxía da lingua galega*, a well trusted and reputed source on Galician dialectology. Source data were taken from the *Atlas lingüístico galego*.

Some boundaries proposed by F. Fernandez Rei were slightly altered using newer data as shown on scholar research on particular areas.

Created by Susana Freixeiro
All rights released / public domain

- Linguistic block
- Linguistic area
- Linguistic subarea (eventually linguistic macro-subarea)

(I) WESTERN GALICIAN BLOCK

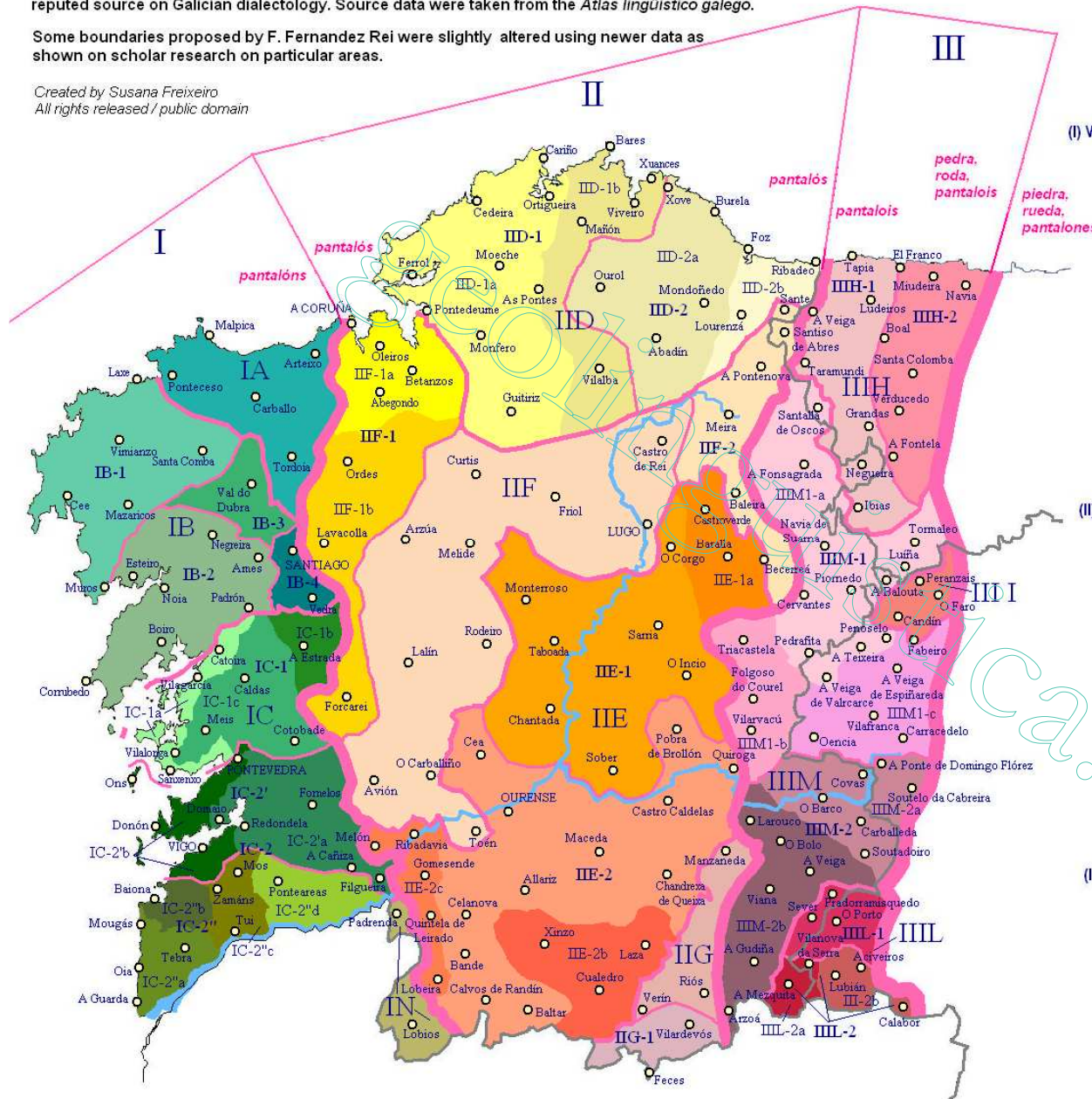
- (IA) Bergantiños area
- (IB) Finisterre area
 - (IB-1) Xallas subarea
 - (IB-2) Tambre subarea
 - (IB-3) Benibibre subarea
 - (IB-4) Santiago de Compostela subarea
- (IC) Pontevedra area
 - (IC-1) Ulla and Umia subarea
 - (IC-1a) Salnés micro-subarea
 - (IC-1b) Ulla micro-subarea
 - (IC-1c) Moraña and Cotobade micro-subarea
 - (IC-2) Verdugo and Minho macro-subarea
 - (IC-2') Verdugo subarea
 - (IC-2'a) Redondela and Mondariz micro-subarea
 - (IC-2'b) Morrazo and Frago micro-subarea
 - (IC-2'') Minho subarea
 - (IC-2'a') Lower Minho micro-subarea
 - (IC-2'b') Miñor valley micro-subarea
 - (IC-2'c') Louíña valley micro-subarea
 - (IC-2'd') Condado micro-subarea
- (IN) Lower Limia area

(II) CENTRAL GALICIAN BLOCK

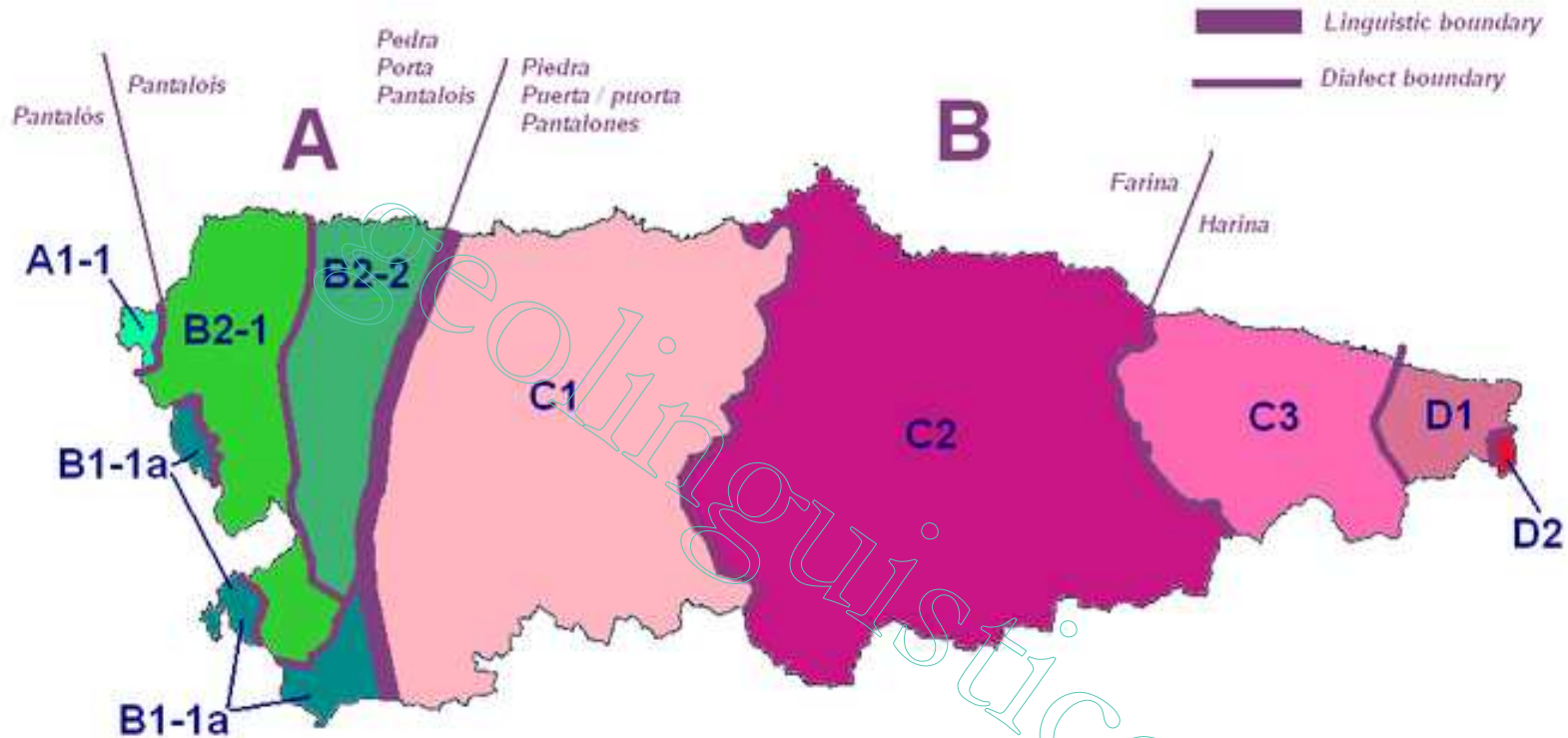
- (II) Mindoniensis area
 - (II-1) Western Mindoniensis area
 - (II-1a) Capelada, Ferrol and Guitiriz micro-subarea
 - (II-1b) Viveiro and Xermade micro-subarea
 - (II-2) Eastern Mindoniensis area
 - (II-2a) A Mariña-Xistral micro-subarea
 - (II-2b) A Mariña-Lourenzá micro-subarea
- (IE) Lucu-aurensis area
 - (IE-1) Lucensis subarea
 - (IE-1a) Berceréa micro-subarea
 - (IE-2) Aurensis subarea
 - (IE-2b) Upper Limia micro-subarea
 - (IE-2c) O Ribeiro micro-subarea
- (II) Central transitional area
 - (II-1) Western subarea
 - (II-1a) As Mariñas micro-subarea
 - (II-1b) Ordes and Terra de Montes micro-subarea
 - (II-2) Meira subarea
- (II) Eastern transitional area
 - (II-1) Ombra and Vilardevós subarea

(III) EASTERN GALICIAN BLOCK

- (III) Asturian area
 - (III-1) A Veiga, Os Oscos and Ibias subarea
 - (III-2) Salave, Coaña and Boal subarea
- (III) Ancares area
- (III) Zamora area
 - (III-1) O Porto subarea
 - (III-2) A Mezquita and Lubián subarea
 - (III-2a) A Mezquita micro-subarea
 - (III-2b) Lubián and Hermisende micro-subarea
- (III) Central Western area
 - (III-1) A Fonsagrada, Courel and Bierzo subarea
 - (III-1a) A Fonsagrada and Cervantes micro-subarea
 - (III-1b) Courel micro-subarea
 - (III-1c) Bierzo micro-subarea
 - (III-2) Valdeorras and Viana subarea
 - (III-2a) Valdeorras micro-subarea
 - (III-2b) O Bolo and Viana micro-subarea



LANGUAGES AND DIALECTS OF ASTURIAS



A. GALICIAN GROUP

A1-1. Central Galician (Meira subarea)

B1-1a. Eastern Galician (Fonsagrada, Courel and Bierzo subarea. Fonsagrada and Cervantes micro-subarea)

EONAVIAN:

B2-1. Eastern Galician (A Veiga, Ozcos and Ibias subarea)

B2-2. Eastern Galician (Salave, Coaña and Boal subarea)

B. ASTURIAN GROUP

C1. Western Asturian

C2. Central Asturian

C3. Eastern Asturian

D1. Cantabrian

D2. Montañese

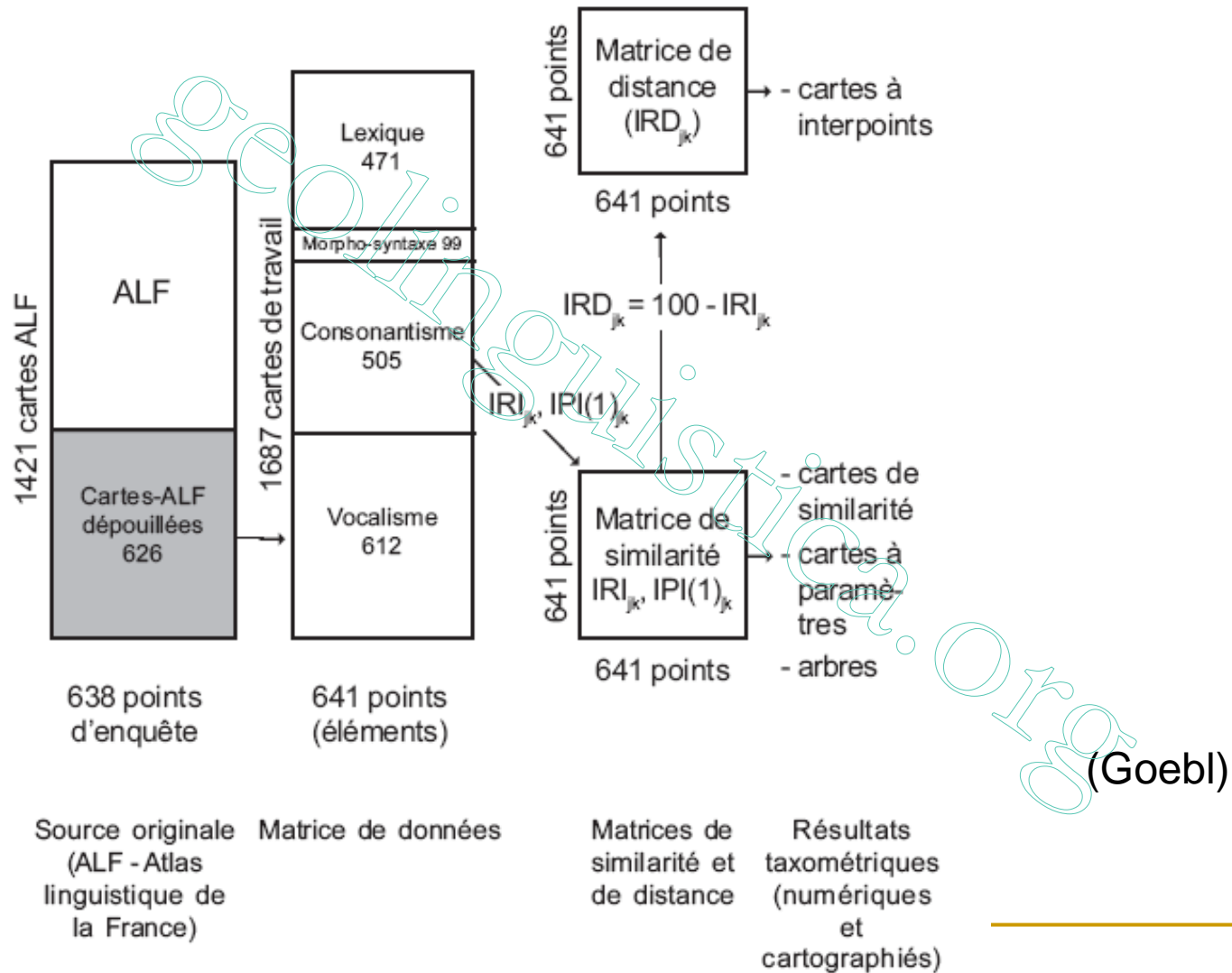
Created with data of Zamora Vicente's Dialectología Española, Fernández Rei's Dialectoloxía da lingua galega, ALPI and ALGa

*Susana Freixeiro
 All rights released / public domain*

Os princípios da dialectometria

- Avaliação da distância linguística entre pontos. Da análise qualitativa à análise **quantitativa**
 - ❑ Cúmulo das diferenças linguísticas: exame de grande número de dados, possivelmente de diferentes aspectos da língua
 - Diferentes escolas
 - ❑ Origens: Séguy e Guiter
 - ❑ Actualidade: Salzburgo (Goebel e Bauer); Holanda (Nerbonne); Estados Unidos (Kretzschmar)
-

Chaîne dialectométrique



Os principais passos (I)

- **Elaboração de cartas**
(mosaicos triangulares e poligonais)

SLC@Linguistica.Org

Os principais passos (II)

- **Seleccção dos dados**
 - Número mínimo de 100 mapas
 - Depende também das perspectivas desejadas
- **Taxonomia dos dados: classificação e obtenção de matrizes.**



Atributos (mapas	2	1	2	1	3
de atlas)	1	1	1	1	1
		1	2	3	4
		Objectos (pontos da rede)			

Matriz de dados

			Corpus: TOT: 641 PP.	Corpus: OC: 220 PP.	Corpus: NORD: 421 PP.		
1	2		3	4	5	6	
Catégorie linguistique	Cartes - <i>ALF</i>		Cartes de travail	Taxats	Taxats	Taxats	
	Décompte <i>partiel</i>	Décompte intégral	Décompte intégral	Décompte intégral	Décompte intégral	Décompte intégral	
1	Phonétique	247	—	1.117	12.023	7.507	8.948
2	<i>Vocalisme</i>	242	—	612	8.657	5.218	6.518
3	<i>Consonantisme</i>	227	—	505	3.366	2.289	2.430
4	Morpho-syntaxe	84	—	99	989	634	625
5	Lexique	463	—	471	7.031	3.680	4.691
	Total	—	626	1.687	20.043	11.821	14.264

- índice relativo de identidade (IRI)

$$IRI_{1,2} = 100 \cdot \frac{1}{1+1} = 50\%$$

$$IRI_{1,3} = 100 \cdot \frac{2}{2+0} = 100\%$$

$$IRI_{1,4} = 100 \cdot \frac{1}{1+1} = 50\%$$

	4	50	50	50	100
	3	100	50	100	50
objectos (pontos da rede)	2	50	100	50	50
	1	100	50	100	50
		1	2	3	4
		Objectos (pontos da rede)			

Matriz simétrica dos IRI

- **índice relativo de distância (IRD)**

- Complementar do IRI: $IRI_{1,2} + IRD_{2,1} = 100$

$$IRI_{1,2} = 50 \text{ logo } IRD_{2,1} = 50$$

$$IRI_{1,3} = 100, \text{ logo } IRD_{3,1} = 0$$

$$IRI_{1,4} = 50, \text{ logo } IRD_{4,1} = 50$$

$$IRI_{2,3} = 50, \text{ logo } IRD_{3,2} = 50$$

$$IRI_{2,4} = 50, \text{ logo } IRD_{4,2} = 50$$

$$IRI_{3,4} = 50, \text{ logo } IRD_{4,3} = 50$$

	4	50	50	50	0
	3	0	50	0	50
objectos (pontos	2	50	0	50	50
da rede)	1	0	50	0	50
		1	2	3	4
		Objectos (pontos da rede)			

Matriz simétrica dos IRD

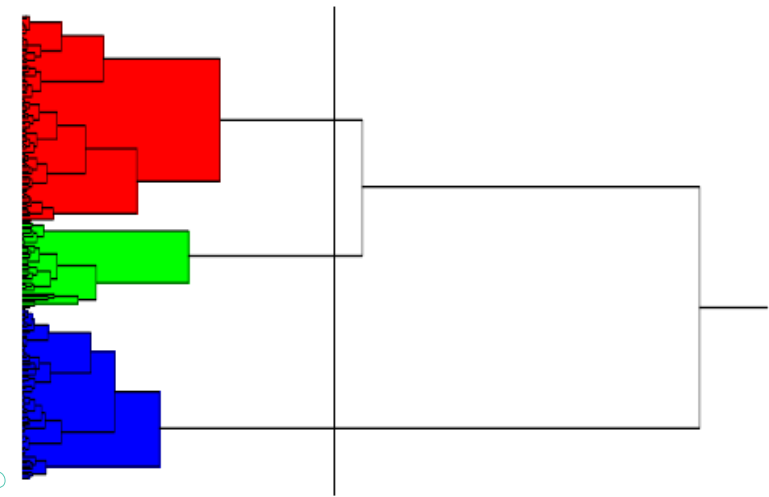
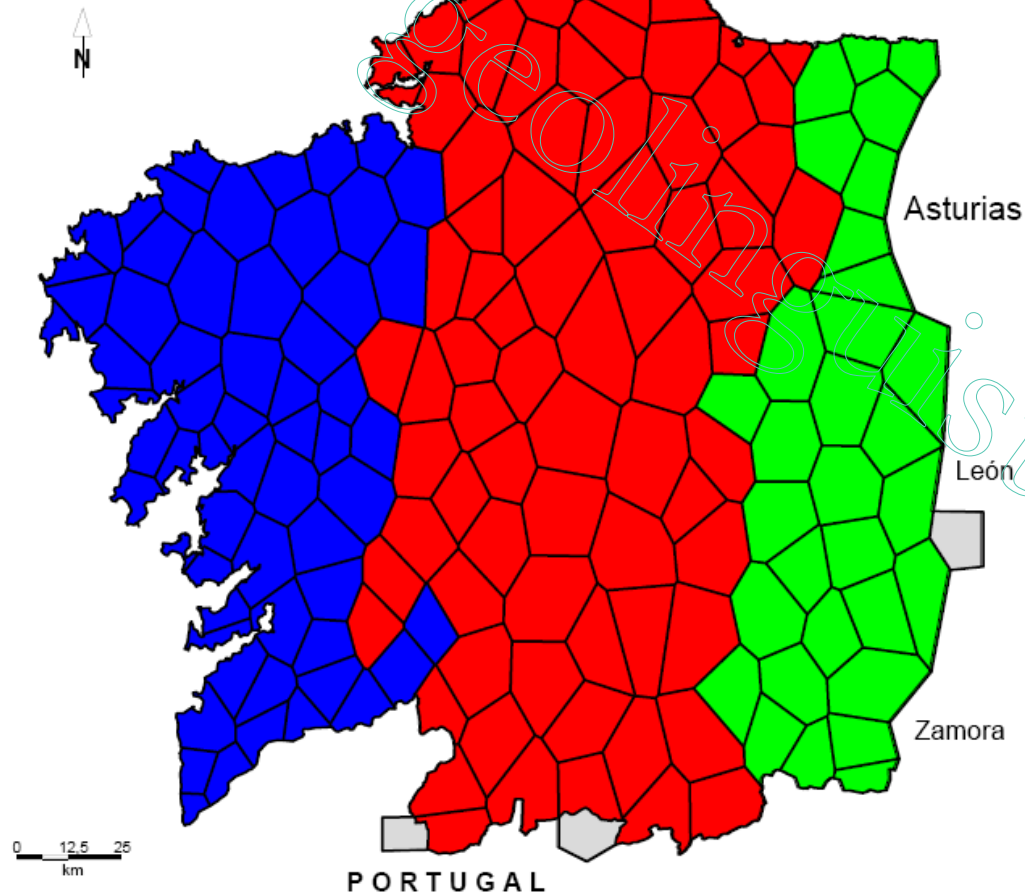
Estabelecimento de intervalos

secolinguistica.Org

Saramago e Cardoso (no prelo), sobre o *Atlas do Sergipe* (Brasil)

Dendrogramas

ALGa

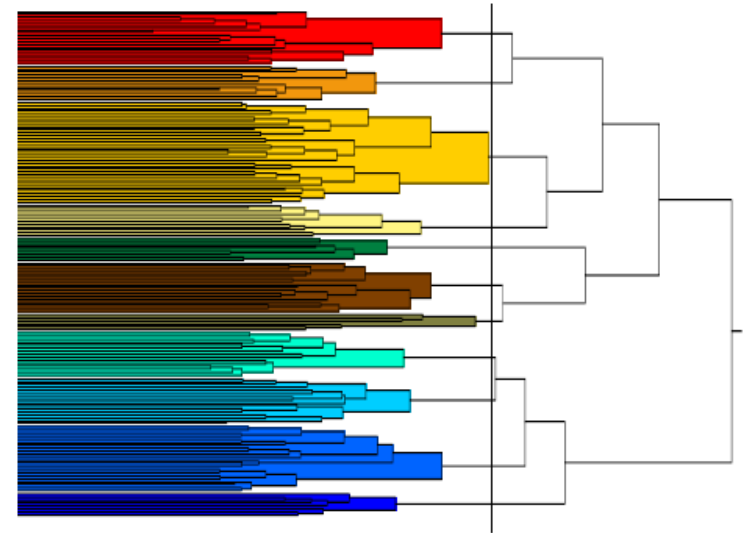
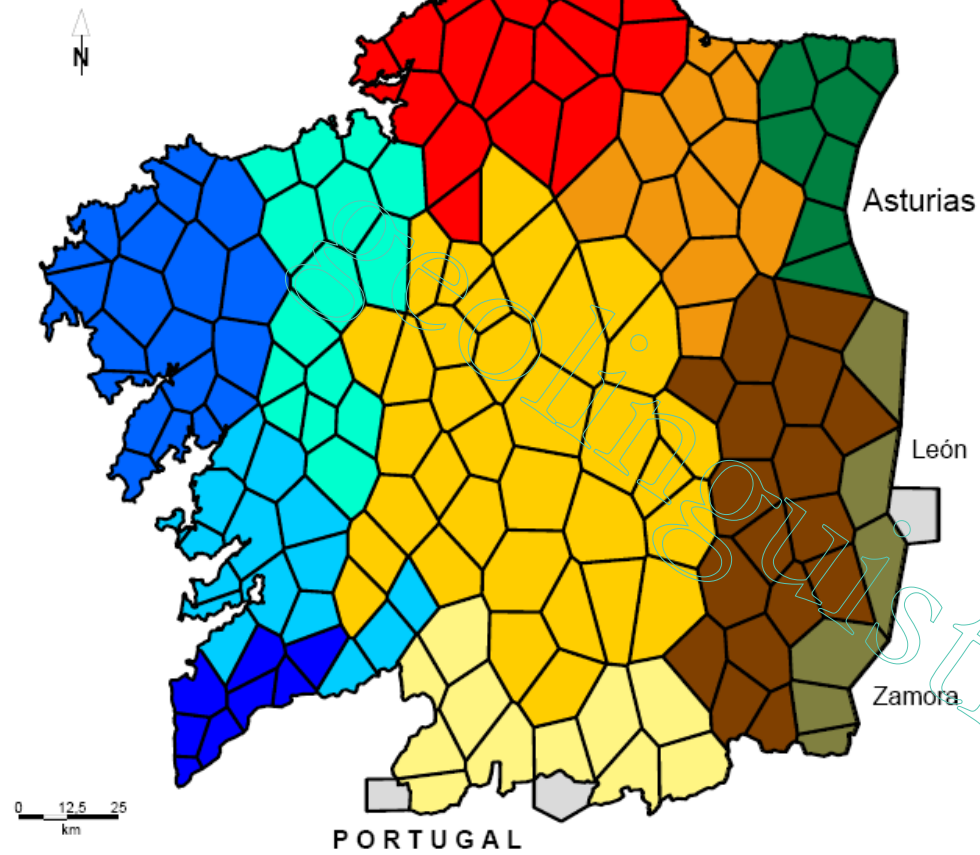


Análise de agrupamentos
Corte: 3 agrupamentos
Número de puntos: 167
Número de mapas analizados: 189
Matriz: índice relativo de identidade
Algoritmo aglomerativo: método de Ward
Programa informático: *Visual Dialectometry* (H. Goebel / E. Haimlerl - Universidade de Salzburgo)

Figura 2

Sousa (2006)

ALGa

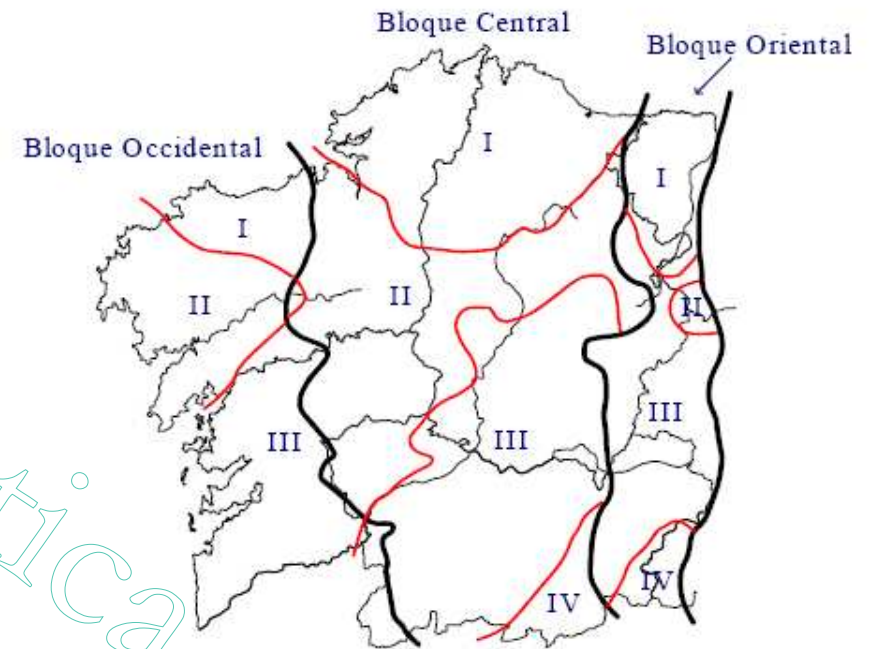
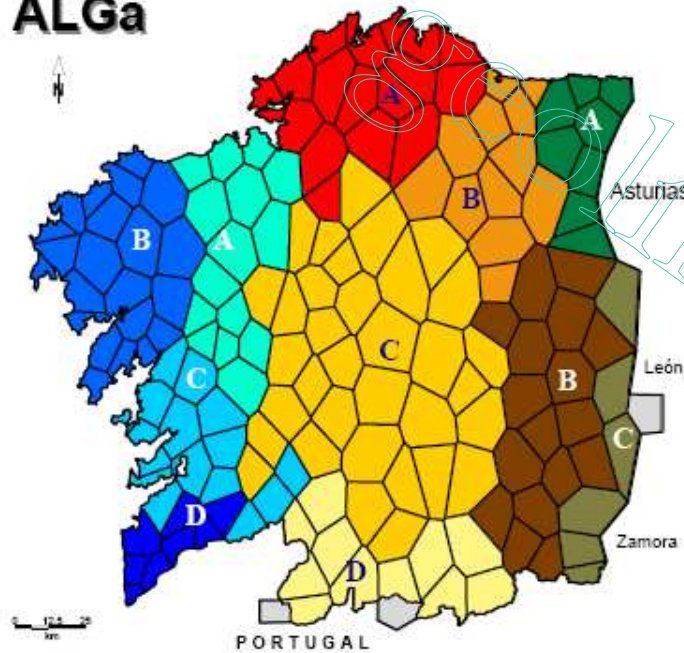


Análise de agrupamentos
Corte: 11 agrupamentos
Número de puntos 167
Número de mapas analizados: 189
Matriz: índice relativo de identidade
Algoritmo aglomerativo: método de Ward
Programa informático: *Visual Dialectometry* (H. Goebel / E. Haimlerl - Universidade de Salzburgo)

Figura 3
División en 3 bloques

Alguns exemplos

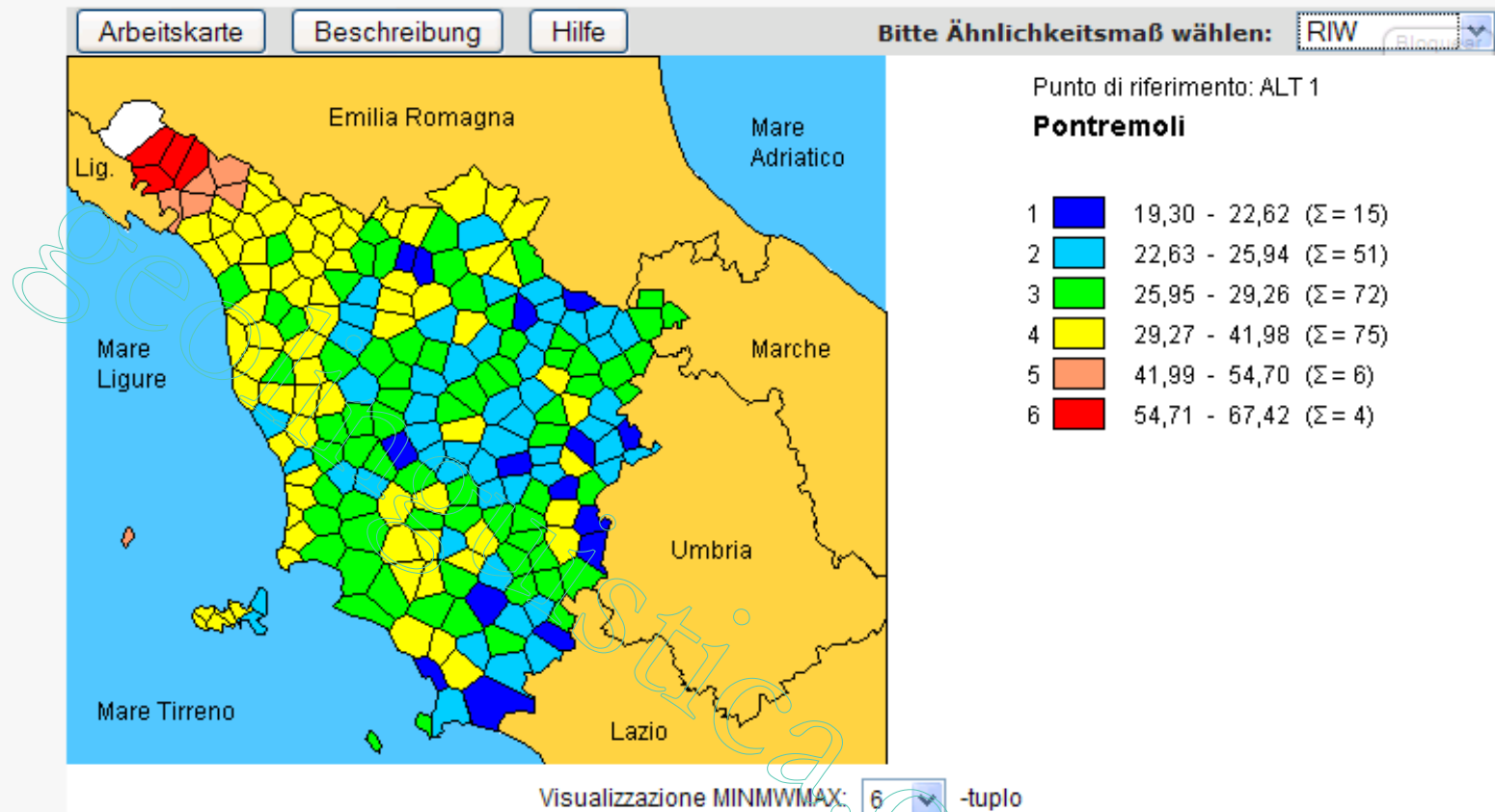
ALGa



Sousa (2006)

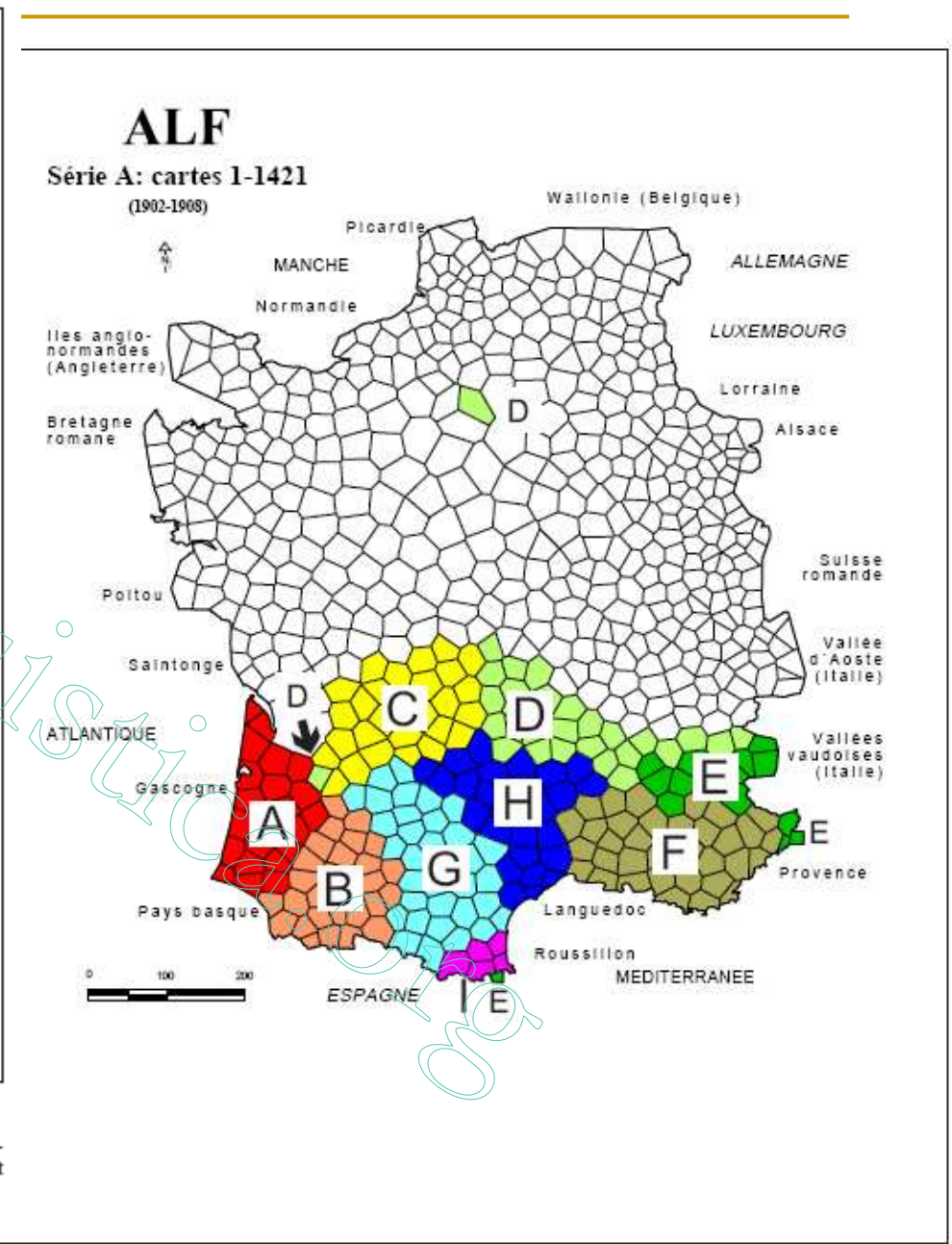
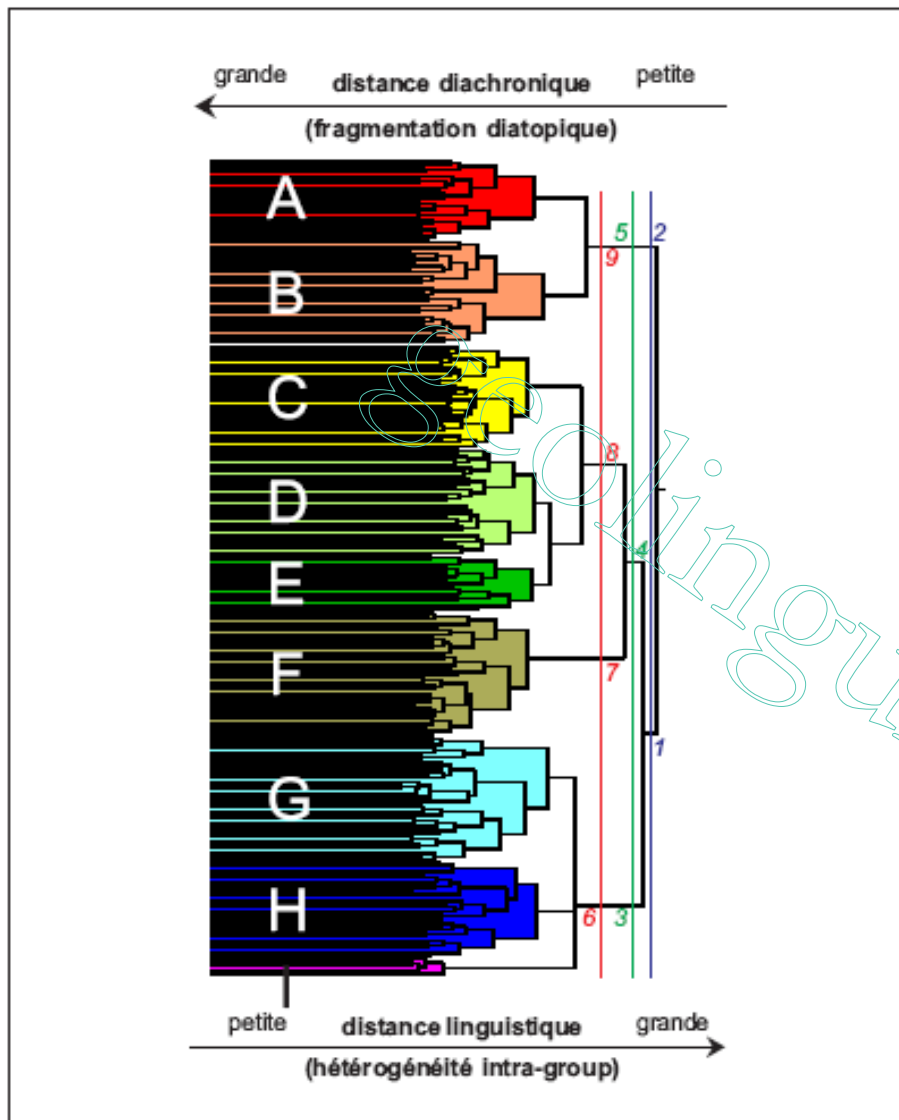
Wortatlas der Region Toskana - Atlante lessicale toscano (ALT)

Relativer Identitätswert (RIW)



Elaboração de Goebel sobre o *Atlante Lessicale Toscano*

<http://www.dialectometry.com>



CARTE 23: Classification dendrographique des vecteurs d'attributs de 221 points-ALF. Algorithme agglomératif: méthode de Ward. Corpus: 1687 cartes de travail; moitié sud du réseau-ALF. Indice de similarité: $IPI(1)_k$. Nombre des dendrogrammes coloriés: 9. Voir aussi le paragraphe 8.5. et la carte 24 (spatialisation de cet arbre).

CARTE 24: Spatialisation des données de la carte 23. Nombre des chorèmes coloriés: 9. Voir aussi le paragraphe 8.5. et la carte 23 (arbre original relatif à cette spatialisation).

Elaborações de Hans Goebel sobre o

<http://www.dialectometry.com>