

Jean Léo Léonard, UPVM3, Dipralang (EA 739)
(en collaboration avec Guylaine Brun-Trigaud, Université Côte d'Azur,
THESOC) :

**Divisions dialectales, frontières et transitions en
domaine occitan : validation et vicariance
dialectométriques (THESOC)**

Journée d'études Praxiling, 20-21 novembre 2024

Plan

- 1. Problématisation : falsification popperienne des ontologies dialectales (étude de cas : le diasystème occitan)
- 2. THESOC et Gabmap : un duo heuristique
- 3. Horizon de falsification : validation des clusters
 - 3.1. Phonologie (distance d'édition)
 - 3.2. Lexique (binarisme)
- 4. Perspectives heuristiques.

1. Problématisation

Une extension d'un article sous presse pour *Frontiers in Complex Systems*

=> approche de Théorie de la Complexité (TC) implicite dans l'ensemble de la présente communication.



OPEN ACCESS

EDITED BY
Marco Patriarca,
National Institute of Chemical Physics and
Biophysics, Estonia

REVIEWED BY
Gianmario Raimondi,
Università della Valle d'Aosta, Italy
Mikhail Tamm,
Tallinn University, Estonia
Jacques François,
Université de Caen Normandie, France

*CORRESPONDENCE
Jean Léo Léonard,
✉ leonardjeanleo@gmail.com

RECEIVED 07 May 2024
ACCEPTED 15 August 2024
PUBLISHED XX XX 2024

CITATION
Léonard JL (2024) Revisiting Southern Gallo-Romance from a complexity theory standpoint: Occitan.
Front. Complex Syst. 2:1429114.
doi: 10.3389/fcpxs.2024.1429114

COPYRIGHT
© 2024 Léonard. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Revisiting Southern Gallo-Romance from a complexity theory standpoint: Occitan

Jean Léo Léonard*

Laboratoire Dipralang (EA 739), Université Paul-Valéry Montpellier 3, Montpellier, France

In this paper, the inner structure of the Occitan dialect network is revisited in the light of a range of cumulative (Ward's method) vs. reductive (Complete linkage, Groupe Average, Weighted Average) hierarchical algorithms provided by Gabmap, an Online dialectometric application for calculating distance/similarities by edit distance (Levenshtein algorithm). Reticularity of the Occitan geolinguistic space is addressed through connectograms using Gephi, and Multidimensional Scaling is also used to some extent. After sketching the canonical classifications of the Occitan geolinguistic space (Bec, Ronjat), providing the "eponymous dialects", we explore the deep patterns of this diasystem, bringing to light a hierarchy of systemic entities constituting an array of "invisible dialects", corresponding to entities of various size and functions (*macrodialects, dialects, subdialects, varieties, hubs, small worlds, buffer zones, default dialects*). The approach is based on concrete linguistic data from the THESOC database (Université de Nice/CNRS), contrasting the major isoglosses (macrodialectal features) with the "intricate variables", i.e., segmental *nexi*, extracting data relating to strategic points in the complex dialectal network from reductionist algorithms.

KEYWORDS

Occitan, dialect, dialectometry, taxonomy, edit distance, diasystem, external factors, language dynamics

1. Introduction

1. Problématisation

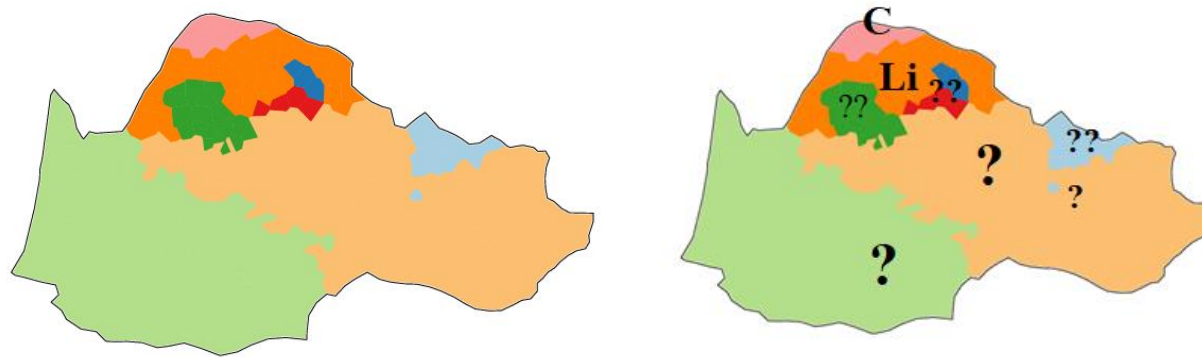
Dialecte *éponyme*
vs dialecte *invisible*

John Nerbonne
et William Kretzschmar
(2003)



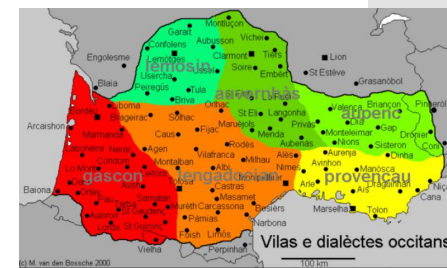
Introducing Computational Techniques in Dialectometry
 Author(s): John Nerbonne and William Kretzschmar
 Source: *Computers and the Humanities*, Vol. 37, No. 3, Computational Methods in Dialectometry (Aug., 2003), pp. 245-255
 Published by: Springer
 Stable URL: <http://www.jstor.org/stable/30204900>
 Accessed: 17-07-2016 14:22 UTC

Il s’agit là, comme (presque) toujours, de « dialectes éponymes », que des approches diasystémiques, comme celle de Bec (en haut, à droite) ou de la dialectométrie dans ses outils réductionnistes (en haut, centre : CL), faisant émerger des « dialectes invisibles » mettent sévèrement à l’épreuve : pour ces entités, des concepts comme les « hubs » et « small worlds » de Duncan Watts seraient heuristiques... (cf. aires bleu et rouge foncés dans carte CL ci-dessous)



Bec (1973)

BLIND DIALECTS. Sometimes we refer e.g. to “the dialect Smith speaks” or the “dialect of South Boston” without reflecting on whether it is distinctive in any way. In such a case we “attribute” a dialect to a location by noting the linguistic features in use there. An attributive dialect is simply the linguistic variety used in a particular place. Note that a field linguist will generally succeed in the task of cataloguing the linguistic features in a given place, i.e. in specifying the attributive dialect, but this does not guarantee success in determining how the local (attributive) dialect compares to the speech of other places. It does not guarantee that the field linguist has noted anything linguistically distinctive about the variety. Dialectologists working in the French tradition most often focus on the careful cataloguing of individual linguistic features (as did Gilliéron himself, especially on their etymologies) instead of the question of what is distinctive to some groups of varieties.



Vilas e dialectes occitanes

1. Problématisation

Et Bec vint...
(*La langue occitane*,
Que-Sais-Je?, 1963)

=> Diasystème et
macrodialectes

Une révolution...

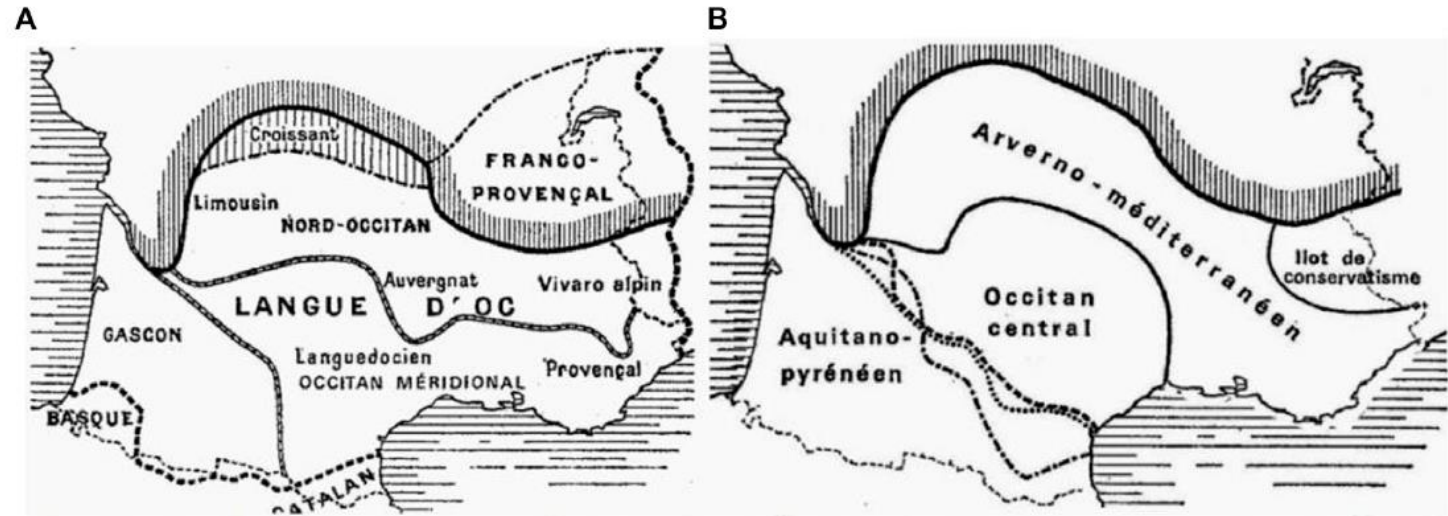
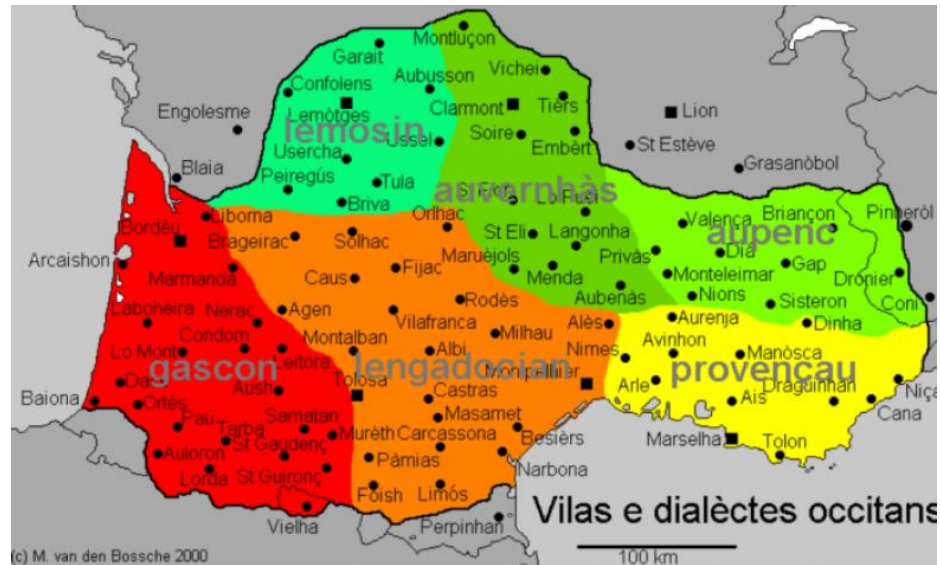


FIGURE 2
Pierre Bec's maps of vicarious hierarchies of Occitan dialects (Bec, 1963). (A) Bec's North vs South Great divide. (B) Bec's NW-SE to central and SW oblique divide.



NB : dans les deux cas, on rompt avec l'hypothèse nulle de Meyer et Paris (reprise par Saussure), sans nier son caractère heuristique sur le plan de l'histoire des idées

1. Problématisation

Fil rouge :

Le dialecte
« éponyme » vs le
dialecte « invisible »



**Dialecte « attribué » ou
« éponyme »**

- **Conjoncturel** : corrélats géohistoriques et identitaires
- **Délimité, tracé, stable**
- « **légitime** », socialement déterminé
- **Narré** : une narration existe et s'impose
- **Prédéterminé**, construit interne à la société (psychosocial)

**Dialecte « invisible »
(lit. blind : 'aveugle')
=> de fait, vicariant**

- **Emergent**, par méthode, par procédure (mathématique)
- Libre, instable, voire **labile**
- Invisible, en-dessous du niveau de « conscience linguistique »
- **Opaque** : peu, voire aucune intuition de son existence/ontologie
- **Donné par artefact, construit externe émergent**

=> John Nerbonne and William Kretzschmar (2003), remanié et augmenté



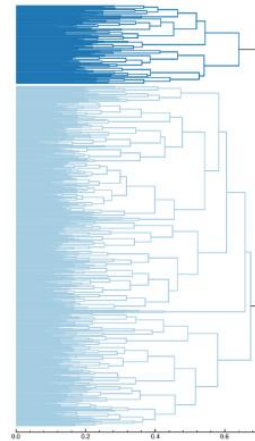
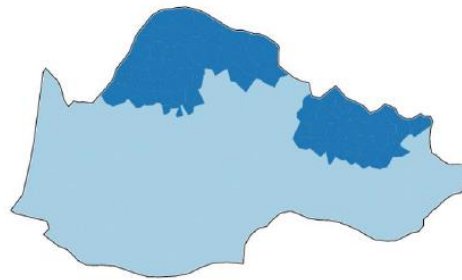
Jeux d'aires, selon une approche réductionniste quantitative :
 Complete Link ou « méthode du voisin le plus éloigné » (à gauche et au centre)
 Vs Méthode des aires par isoglosses (à droite)
 => **LEXIQUE (THESOC)**

1. Problématisation

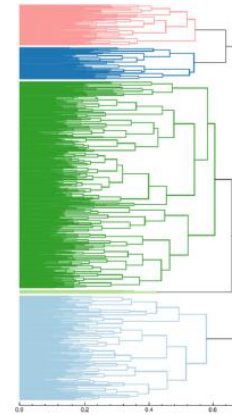
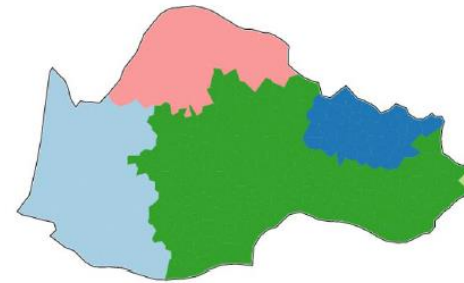
La polarité « dialecte attribué
 " vs " invisible » selon John
 Nerbonne et William
 Kretzschmar (2003) s'avère
 un concept heuristique, qui
 convient bien à la
 dialectométrie

=> à la recherche de
 structures ou motifs ou
patterns cachés dans les
 diasystèmes
 => paramètre du lexique

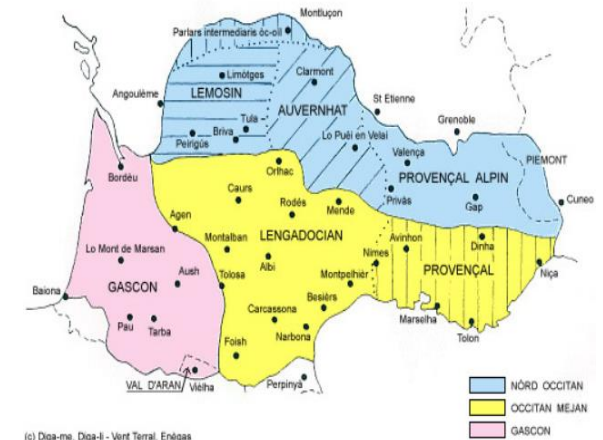
Complete link, 2 classes



Complete link, 5 classes



- «Dialectes et aires
 « éponymes »



(c) Diga-me, Diga-l - Vert Terral, Enègas

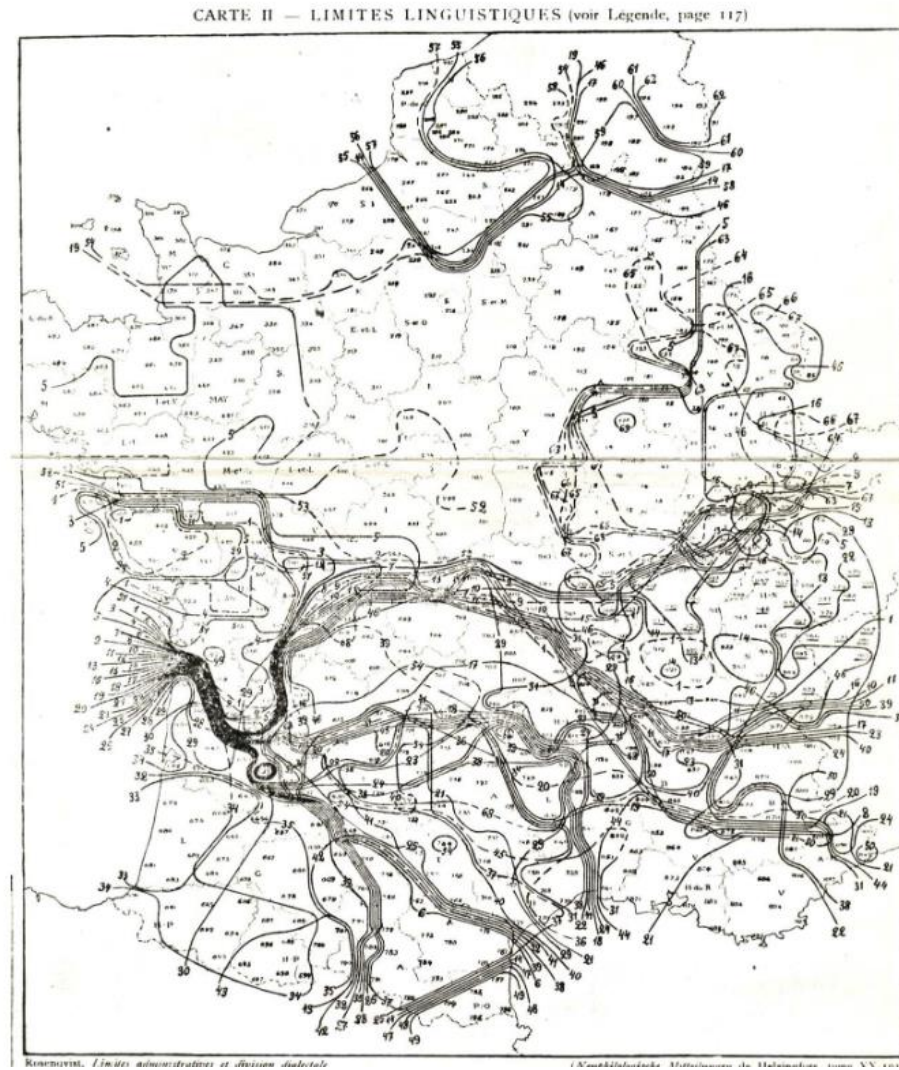
NB : très beaux effets
 de *congruence*
 (plus aisés avec le **lexique**,
 en raison de sa *catégorialité inhérente* ?

Réfutation de l'Hypothèse du Continuum Dialectal Lisse (HCDL)

1. Problématisation

Méthode des isoglosses

(carte de
Rosenquist, à partir
des données de
l'ALF)

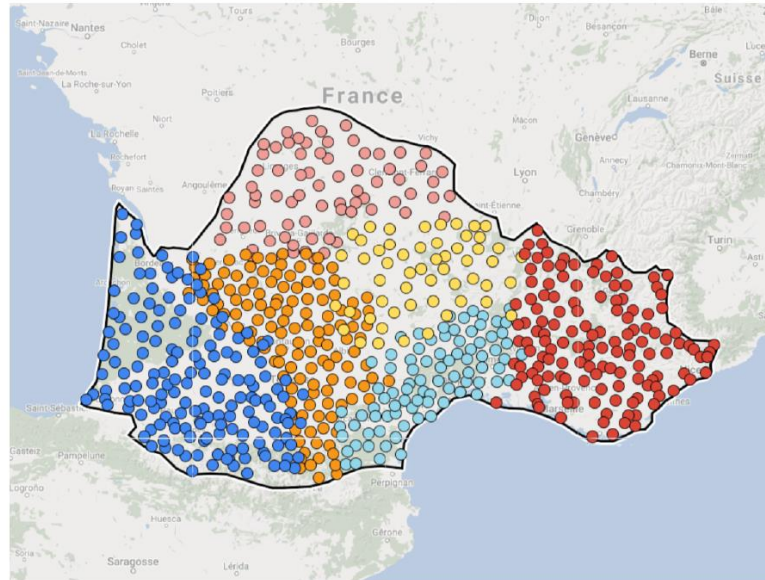


Carte de synthèse :
Rosenquist
(1919)
ALF

2. THESOC et Gabmap : un duo heuristique

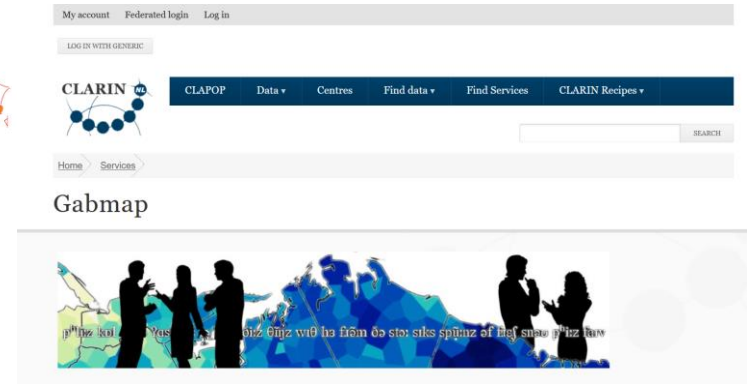
Le réseau dialectal occitan à travers la base de données THESOC, réunie par Guylaine Brun-Trigaud (Université Côte d'Azur, Nice), pour la phonologie et le lexique

Source des données



Procédure (triviale) de l'algorithme de Levenshtein (distance d'édition, calcul sur chaînes de caractères) :

Ajout, suppression, modification (phonologie)



Method: String Edit Distance — Tokenized

Item: abeille

Place: ALAL 23

Show alignments

abeille

Page: « 1 2 3 4 5 6 7 8 9 10 11 »

a	b	'	œ	j	
a	v	'	ε	ʎ	ə
	1		1	1	1

ALAL 23 — ALP 2

a	b	'	œ	j	
a	b	'	e	ʎ	ə
			1	1	1

Place: ALAL 23

Show alignments

cheval

Page: « 1 2 3 4 5 6 7 8 9 10 11 »

ALAL 23 — ALP 1

ʃ	œ	v	'	o
ʃ	i	v	'	a
	1			1

ALAL 23 — ALP 2

ʃ	œ	v	'	o
ʃ	y	v	'	a
	1			1

ALAL 23 — ALP 3

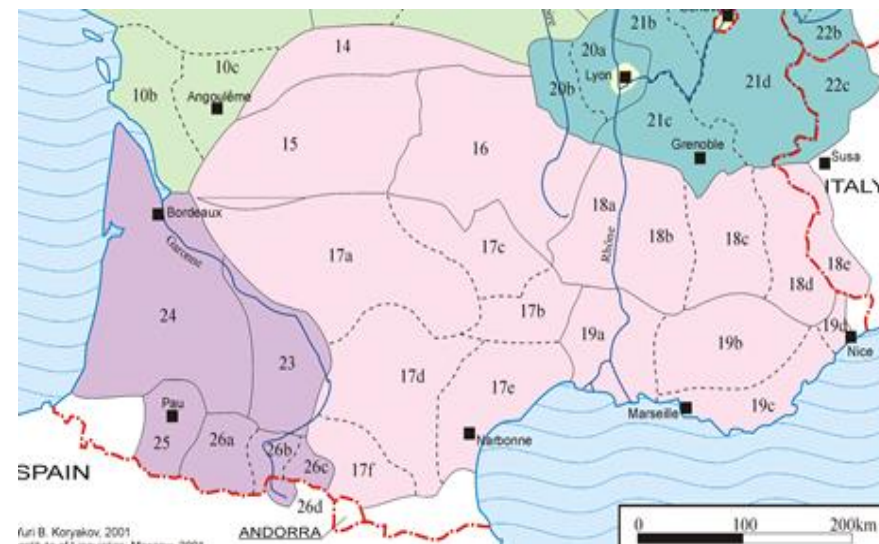
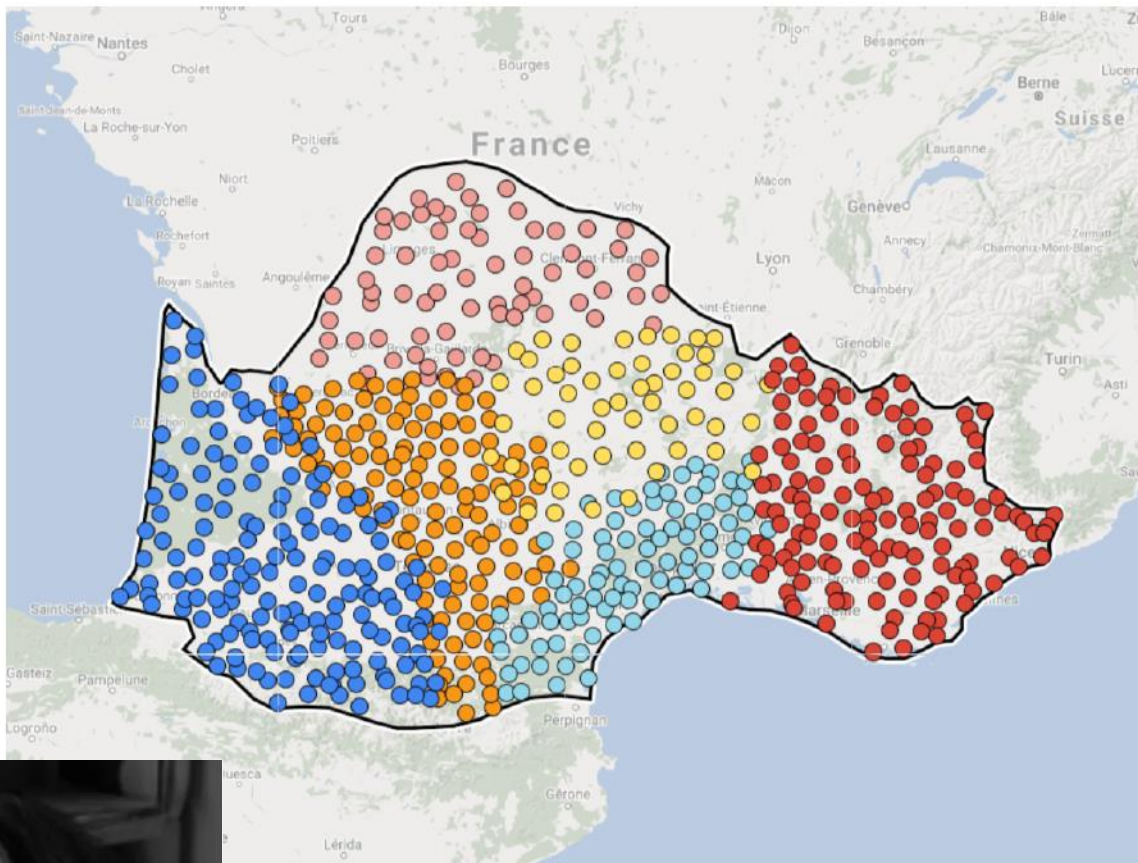
ʃ	œ	v	'	o
ʃ	i	v	'	a
	1			1

ALAL 23 — ALP 6

ʃ	œ	v	'	o
ts	i	v	'	a
1	1			1

THESOC : <http://thesaurus.unice.fr>

Source des données



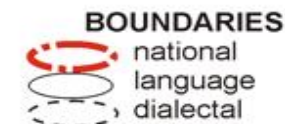
Thésaurus des données du NALF
Initié par A. Dauzat
(esquissé dès 1942 à
Luçon)

En regard :
hyper-éponymisation
Par Y. Koryakov (2001)

- OCCITAN**
- 14. Croissant's** (transitional)
- 15. Limousin**
- 16. Auvergnat**
- 17. Languedocien**
 - a. guyennese
 - b. cevennol
 - c. languedocien-N.
 - d. languedocien-C.
 - e. languedocien-E.
 - f. ariegese
- 18. Vivaro-alpin**
 - a. vivarese
 - b. dromese
 - c. pre-alpine gavot
 - d. alpine gavot
 - e. cisalpine oc
- 19. Provençal**
 - a. rhodanien
 - b. central
 - c. maritime
 - d. niçard

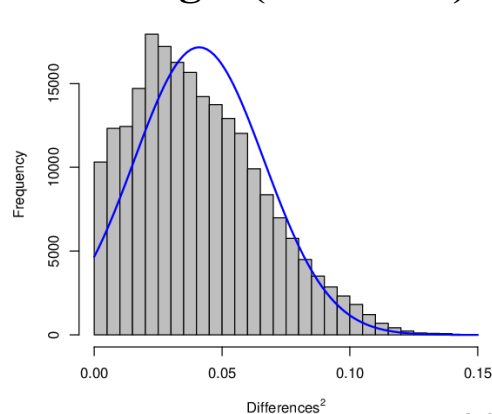


Source: Koryakov Yu.B.
Atlas of the Languages of the World.
Romance languages. Moscow, 2001.



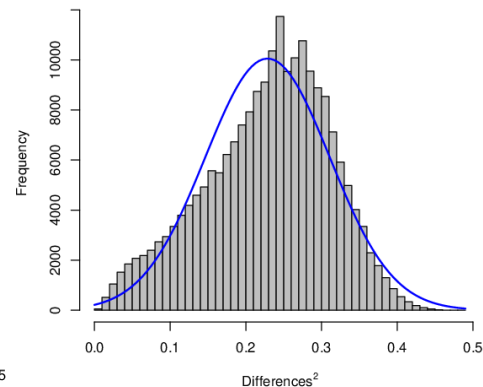
Les volumes du NALF retenus pour constituer la base de données THESOC

Phonologie (dist. édit.)

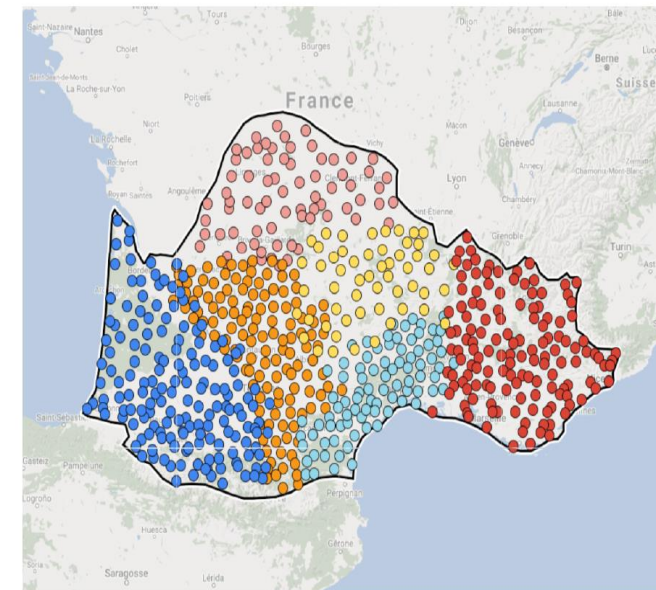


Places:	662
Items:	71
Instances:	44748
Characters:	245029
Unique characters:	99
Tokens:	240562
Unique tokens:	203

Lexique (binarisme) Source des données



Places:	645
Items:	232
Instances:	144025



- ALAL = Potte, Jean-Claude (1975-1992), *Atlas linguistique et ethnographique de l'Auvergne et du Limousin*, Paris, Editions du CNRS (3 vol.).
- ALCe = Dubuisson, Pierrette (1971-1982), *Atlas linguistique et ethnographique du Centre*, Paris, Editions du CNRS (3 vol.).
- ALG = Séguy, Jean (1954-1973), *Atlas linguistique et ethnographique de la Gascogne*, Paris, Editions du CNRS (6 vol.).
- ALLOc = Ravier, Xavier (1978-1993), *Atlas linguistique et ethnographique du Languedoc Occidental*, Paris, Editions du CNRS (4 vol.).
- ALLOr = Boisgontier, Jacques (1981-1986), *Atlas linguistique et ethnographique du Languedoc Oriental*, Paris, Editions du CNRS (3 vol.).

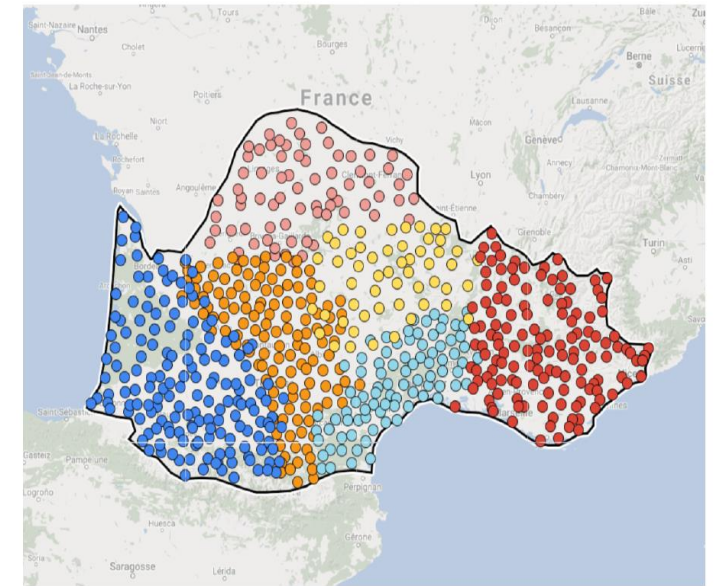
Suite : Les volumes du NALF retenus pour constituer la base de données THESOC

ALLy = Gardette, Pierre (1967-1976), *Atlas linguistique et ethnographique du Lyonnais*, Paris, Editions du CNRS (5 vol.).

ALMC = Nauton, Pierre (1957-1961), *Atlas linguistique et ethnographique du Massif Central*, Paris, Editions du CNRS (4 vol.).

ALO = Massignon, Geneviève/Horiot, Brigitte (1971-1983), *Atlas linguistique et ethnographique de l'Ouest*, Paris, Editions du CNRS (3 vol.).

ALP = Bouvier, Jean-Claude/Martel, Claude (1975-1986), *Atlas linguistique et ethnographique de la Provence*, Paris, Editions du CNRS (3 vol.) et Bouvier, Jean-Claude/Martel, Claude/Brun-Trigaud, Guylaine (2016), *Forcalquier, Alpes de Lumière* (1 vol.).



3.1. Paramètre 1 : la phonologie par distance d'édition (AL)

	abeille	agassa	agneau	agneler	aiguille	âne	arbre	blé	boeuf	caille	chapeau
ALAL 55	b'æj	ʒ'as	ɑŋ'e	ɑŋəl'a	g'yj	'an	'æRBR ^a	bl'a	bj'a ^o	k'aj	ʃap'e
ALG 680	aθ'æʎə	ɑɣ'asə	ɑŋ'et	ɑŋər'a	ɑɣ'yʎə	'azu	'awbrə		b'yw	k'aʎə	ʃap'ew
ALP 76	ab'ejɔ		ɑŋ'el	ɑnel'a	g'yja	'g ⁿ zu	'arbu				kap'εε
ALP 149	ab'ijɔ	ɑg'asœ	ɑnj'ew	ɑnel'a	ægw'yjɔ	'aj	'awbre	bl'a	b'yw	k'ajɔ	kap'ew
ALLOc 81.20	aθ'eʎɔ	ɑɣ'asɔ	ɑŋ'el	ɑnel'a	ɑɣ'yʎɔ	'aze	'arθre	bl'at	bj'ɔw	k'al:o	kap'el
ALLOr 07.02	ab'ejɔ	adʒ'aso	ɑŋ'ew	ɑnel'a	ɑg'yjɔ	'aze	'awbre	bl'a	bj'ɔw	k'ajɔ	tʃap'ε ^w
ALLOr 07.05	ab'ejɔ	ɑg'aso	ɑŋ'ew	ɑnel'a	ɑg'yjɔ	'aze	'awbre	bl'a	bj'ɔw	k'ajɔ	kap'ε ^w

MAINTIEN : [- continu, dorsal]
g → ^{g^{ve}} / V V
distmap AGASSA Bloc Vélaire -
cont aGasa



FRICATIVE : [+ continu, dorsal]
g → ʒ / V V
distmap AGASSA Bloc Vélaire +
cont aGasa



2. THESOC et Gabmap : un duo heuristique



Echantillon de données du corpus retenu pour la phonologie, pour quelques points d'enquêtes du NALF

	abeille	agassa	agneau	agner	aiguille	âne	arbre	blé	boeuf	caille	chapeau
ALAL 55	<i>b'æj</i>	<i>ʒ'as</i>	<i>ɑn'e</i>	<i>ɑnəl'a</i>	<i>g'yj</i>	<i>'an</i>	<i>'ærbR^ø</i>	<i>bl'a</i>	<i>bj'a^o</i>	<i>k'aj</i>	<i>ʃap'e</i>
ALG 680	<i>ab'ælə</i>	<i>ɑγ'asə</i>	<i>ɑn'et</i>	<i>ɑnər'a</i>	<i>ɑγ'ylə</i>	<i>'azu</i>	<i>'awbrə</i>		<i>b'ɣw</i>	<i>k'alə</i>	<i>ʃap'ew</i>
ALP 76	<i>ab'ejɔ</i>		<i>ɑn'el</i>	<i>ɑnel'a</i>	<i>g'yja</i>	<i>'gⁿzu</i>	<i>'arbu</i>				<i>kap'εε</i>
ALP 149	<i>ab'ijɔ</i>	<i>ɑg'asœ</i>	<i>ɑnj'ew</i>	<i>ɑnel'a</i>	<i>ægw'yjɔ</i>	<i>'aj</i>	<i>'awbre</i>	<i>bl'a</i>	<i>b'ɣw</i>	<i>k'ajɔ</i>	<i>kap'ew</i>
ALLOc 81.20	<i>ab'elɔ</i>	<i>ɑγ'asɔ</i>	<i>ɑn'el</i>	<i>ɑnel'a</i>	<i>ɑγ'ylɔ</i>	<i>'aze</i>	<i>'arθre</i>	<i>bl'at</i>	<i>bj'ɔw</i>	<i>k'al:o</i>	<i>kap'el</i>
ALLOr 07.02	<i>ab'ejɔ</i>	<i>adʒ'aso</i>	<i>ɑn'ew</i>	<i>ɑnel'a</i>	<i>ɑg'yjɔ</i>	<i>'aze</i>	<i>'awbre</i>	<i>bl'a</i>	<i>bj'ɔw</i>	<i>k'ajɔ</i>	<i>tʃap'ε^w</i>
ALLOr 07.05	<i>ab'ejɔ</i>	<i>ɑg'aso</i>	<i>ɑn'ew</i>	<i>ɑnel'a</i>	<i>ɑg'yjɔ</i>	<i>'aze</i>	<i>'awbre</i>	<i>bl'a</i>	<i>bj'ɔw</i>	<i>k'ajɔ</i>	<i>kap'ε^w</i>

Place: ALAL 23 ▾

Show alignments

cheval

Page: « 1 2 3 4 5 6 7 8 9 10 11 »

ALAL 23 — ALP 1

ʃ	æ	v	'	o
ʃ	i	v	'	a
	1			1 2

ALAL 23 — ALP 2

ʃ	æ	v	'	o
ʃ	y	v	'	a
	1			1 2

ALAL 23 — ALP 3

ʃ	æ	v	'	o
ʃ	i	v	'	a
	1			1 2

ALAL 23 — ALP 6

ʃ	æ	v	'	o
ʃ	i	v	'	a
1	1			1 3

Jeux d'aires, selon une approche réductionniste quantitative :
 Complete Link ou « méthode du voisin le plus éloigné » (à gauche et au centre)
 Vs Méthode des aires par isoglosses (à droite)
 => **LEXIQUE** (THESOC)

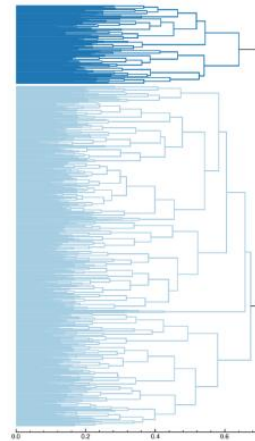
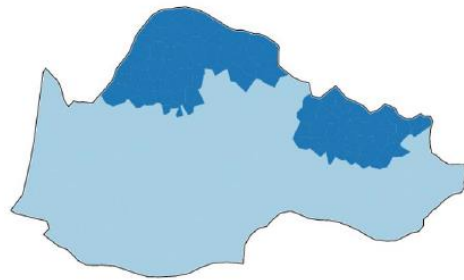
2. THESOC et Gabmap : un duo heuristique

=> La polarité « dialecte attribué " vs " invisible » selon John Nerbonne et William Kretzschmar (2003) s'avère un concept heuristique, qui convient bien à la dialectométrie

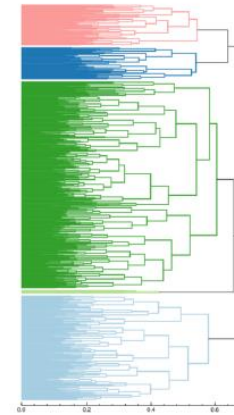
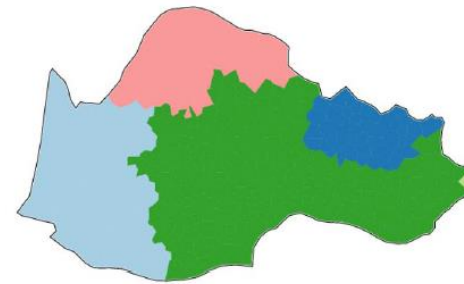
=> à la recherche de structures ou motifs ou *patterns* cachés dans les diasystèmes

=> paramètre du lexique

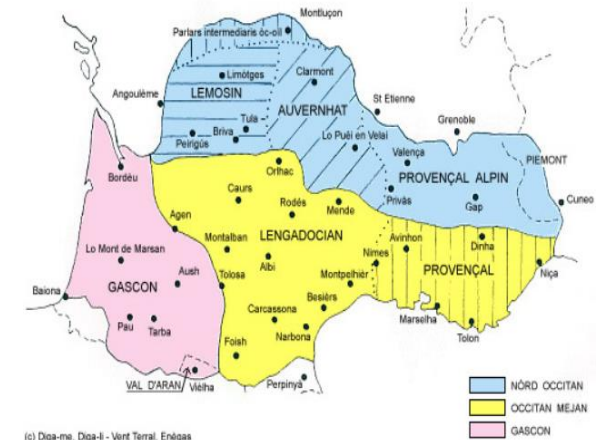
Complete link, 2 classes



Complete link, 5 classes



- «Dialectes et aires
« éponymes »



(c) Diga-me, Diga-l - Vert Terral, Enégas

NB : très beaux effets
de *congruence*
(plus aisés avec le **lexique**,
en raison de sa *catégorialité inhérente* ?

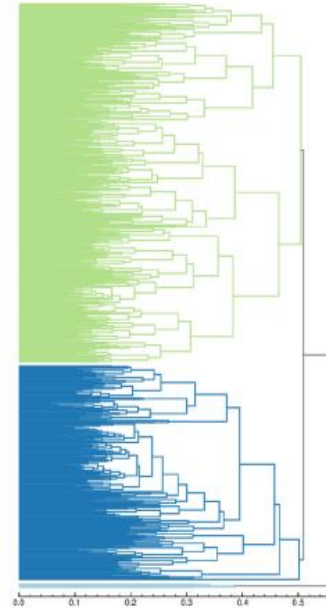
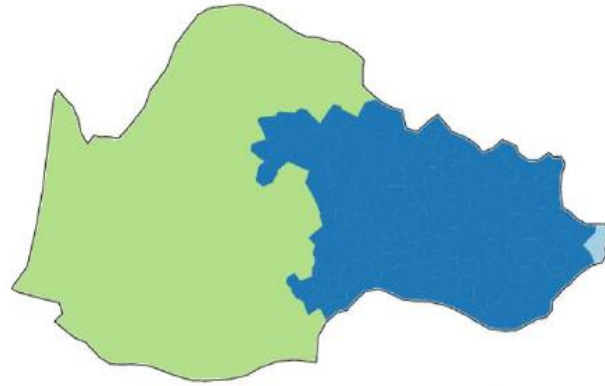
2. THESOC et Gabmap : un duo heuristique

Moyenne de groupes,
pour le paramètre du
lexique...

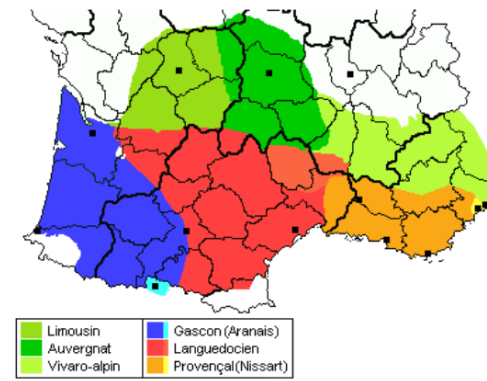
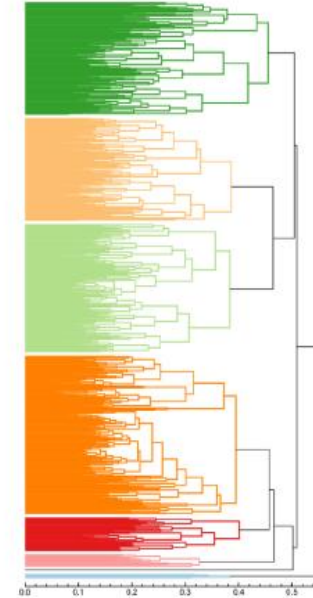
=> fonction binariste
(‘categorical’) dans
Gabmap

NB : \neq distance d’édition
(ne pas confondre...)

Group average, 3 classes

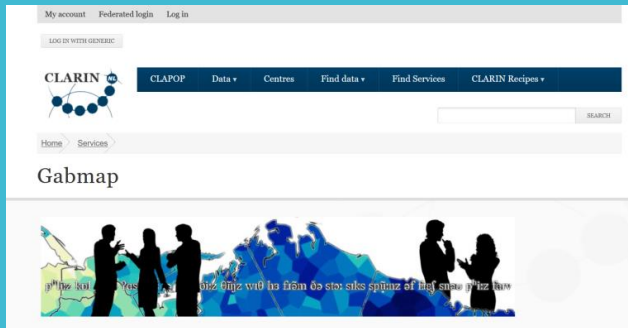


Group average, 8 classes



2. THESOC et Gabmap : un duo heuristique

Implémentation avec Gabmap des "Trois mousquetaires" algorithmiques



Méthode de Ward (ou CAH)

Complete link (ou « Voisin lointain »)

Group average (Moyenne de groupe)

Weighted average (Moyenne pondérée)

Quelques exemples : base de données phonologique (71 entrées) :

Group average 2 classes



Group average 4 classes



Weighed average, 3 classes



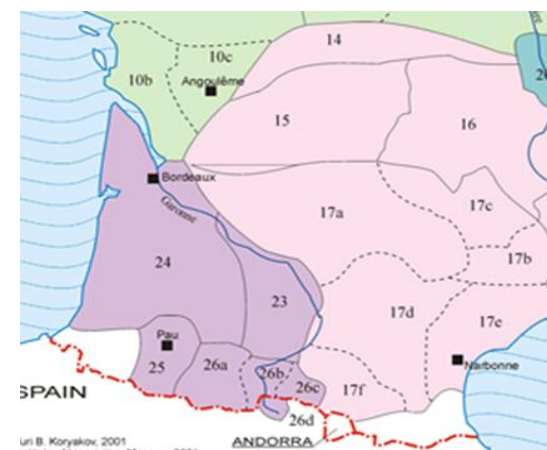
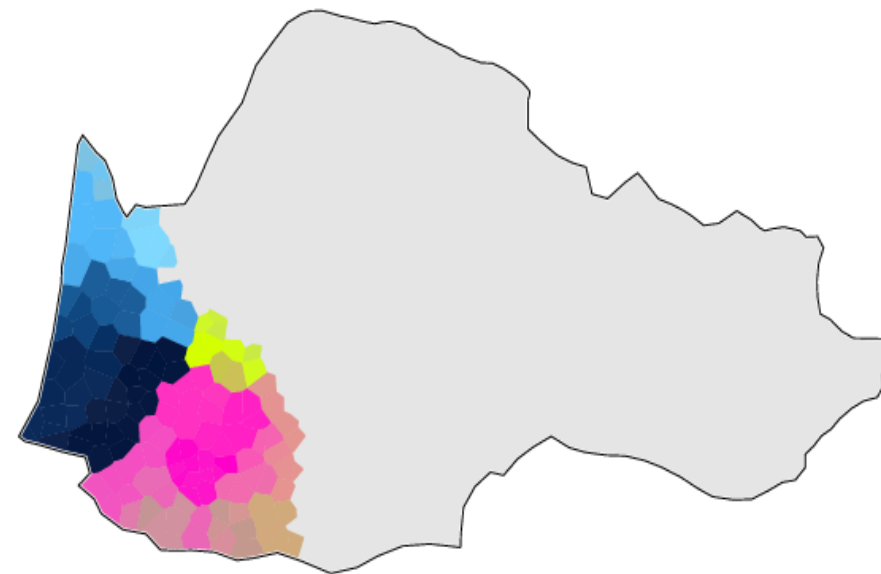
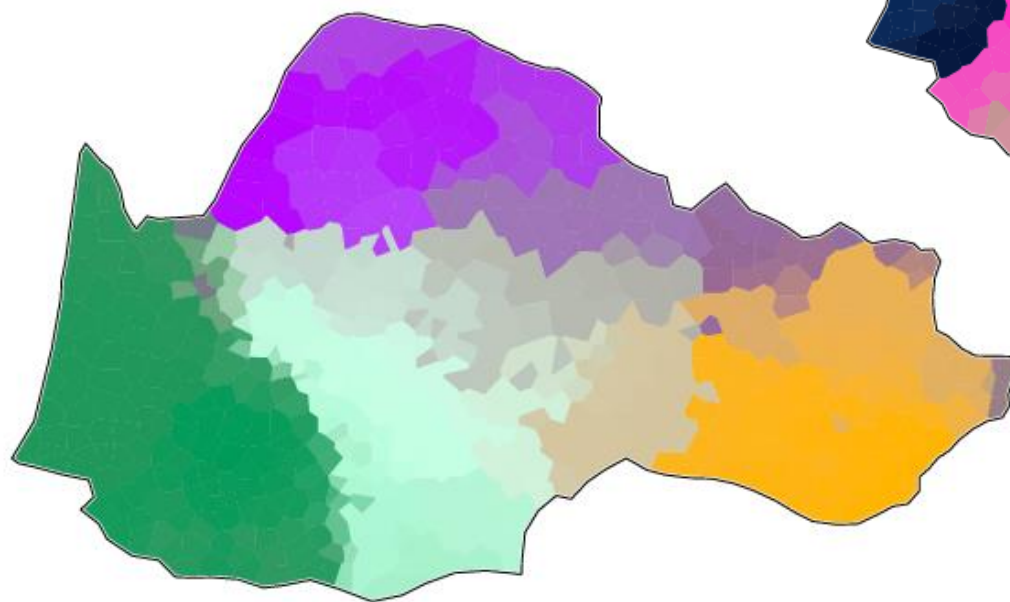
Weighed average, 8 classes



3. Horizon de falsification : validation des clusters

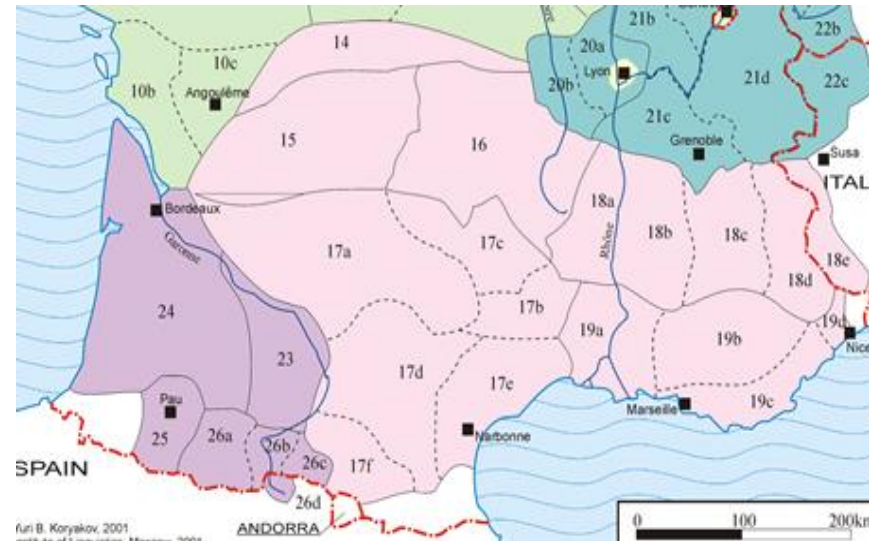
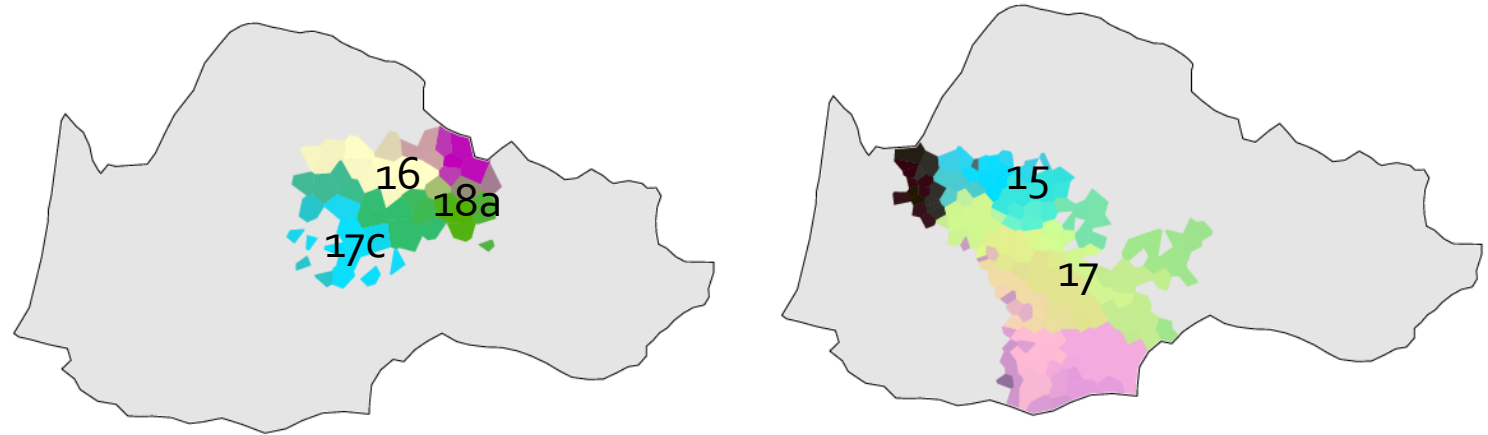
MDS et Aires
bruitées (FCM)

Phonologie, 71
items,
THESOC



3. Horizon de falsification : validation des clusters

FCM ou *aires
bruitées*,
phonologie
(71 items) :
auvergnat
(post)ronjation et
languedocien
occidental vs aires
« éponymes » de
Koryakov



- OCCITAN**
- 14. Croissant's** (transitional)
 - 15. Limousin**
 - 16. Auvergnat**
 - 17. Languedocien**
 - a. guyennese
 - b. cevennol
 - c. languedocien-N.
 - d. languedocien-C.
 - e. languedocien-E.
 - f. ariegese
 - 18. Vivaro-alpin**
 - a. vivarese
 - b. dromese
 - c. pre-alpine gavot
 - d. alpine gavot
 - e. cisalpine oc
 - 19. Provençal**
 - a. rhodanian
 - b. central
 - c. maritime
 - d. niçard

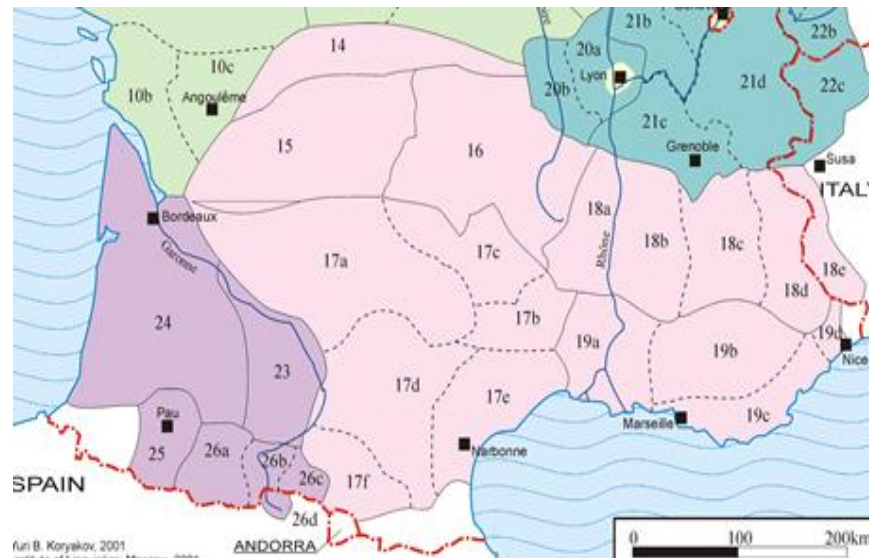
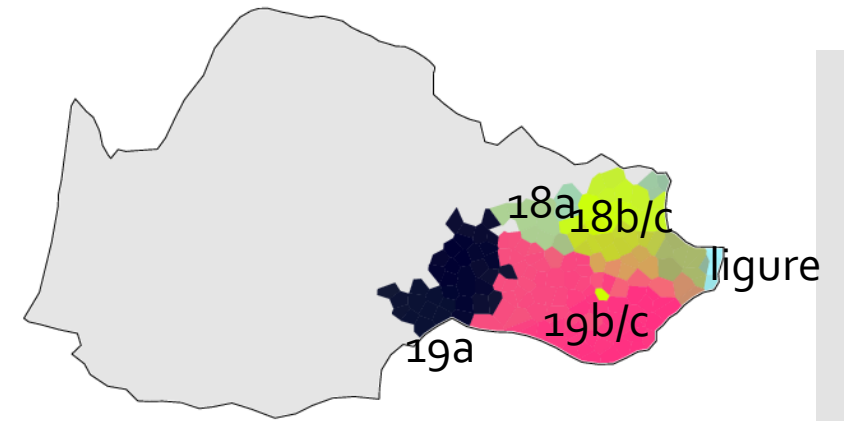
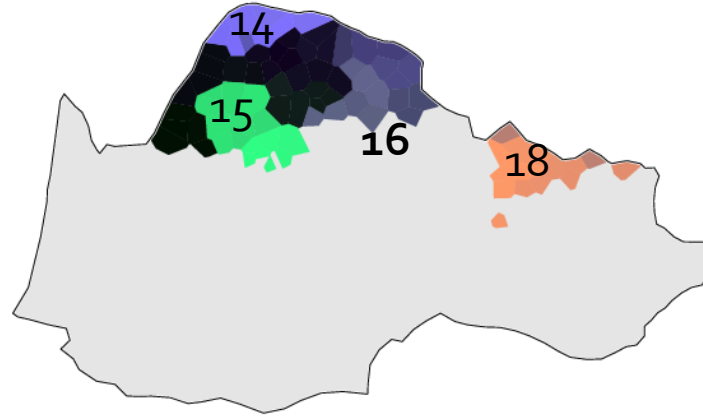
- BOUNDARIES**
- national
 - language
 - dialectal

Koryakov Yu.B.

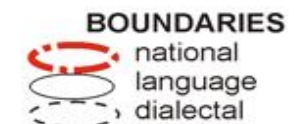
Atlas of the Languages of the World.
Romance languages. Moscow, 2001.

3. Horizon de falsification : validation des clusters

FCM ou aires
floutées/bruitées par
DistEdit et structure interne
du macrodialecte
« arvernoméditerranéen »
de Bec



- OCCITAN**
- 14. Croissant's (transitional)
 - 15. Limousin
 - 16. Auvergnat
 - 17. Languedocian
 - a. guyennese
 - b. cevennol
 - c. languedocian-N.
 - d. languedocian-C.
 - e. languedocian-E.
 - f. ariegese
 - 18. Vivaro-alpin
 - a. vivarese
 - b. dromese
 - c. pre-alpine gavot
 - d. alpine gavot
 - e. cisalpine oc
 - 19. Provençal
 - a. rhodanian
 - b. central
 - c. maritime
 - d. niçard



Koryakov Yu.B.

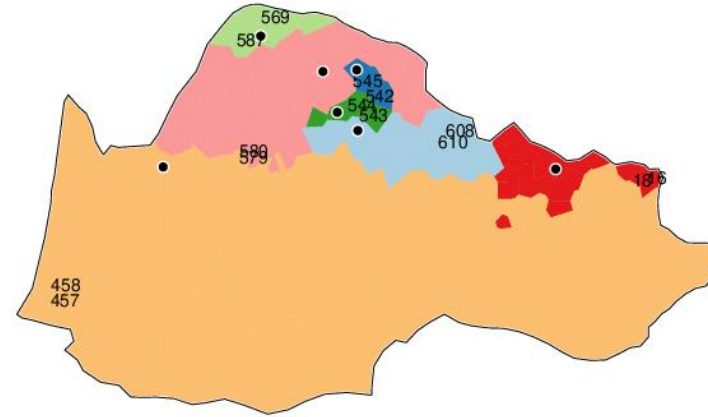
Atlas of the Languages of the World.
Romance languages. Moscow, 2001.

3. Horizon de falsification : validation des clusters

(Validation de
la) Moyenne de
groupe vs
dialectes
éponymes de
Koryakov

NB : Hiérarchie plate en
sud-occitan vs granularité
du nord-occitan

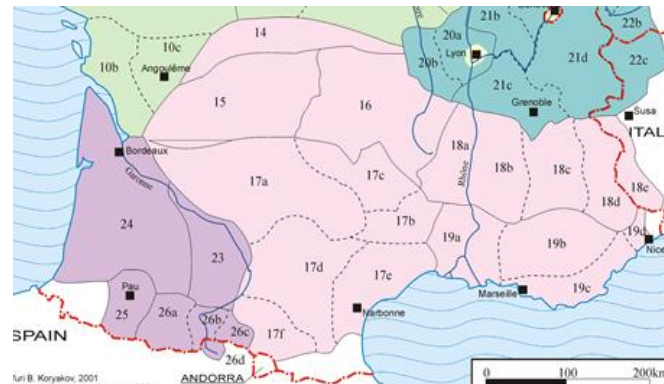
Moyenne de groupe, 7 classes
⇒ Validation de custers
« Chorèmes »



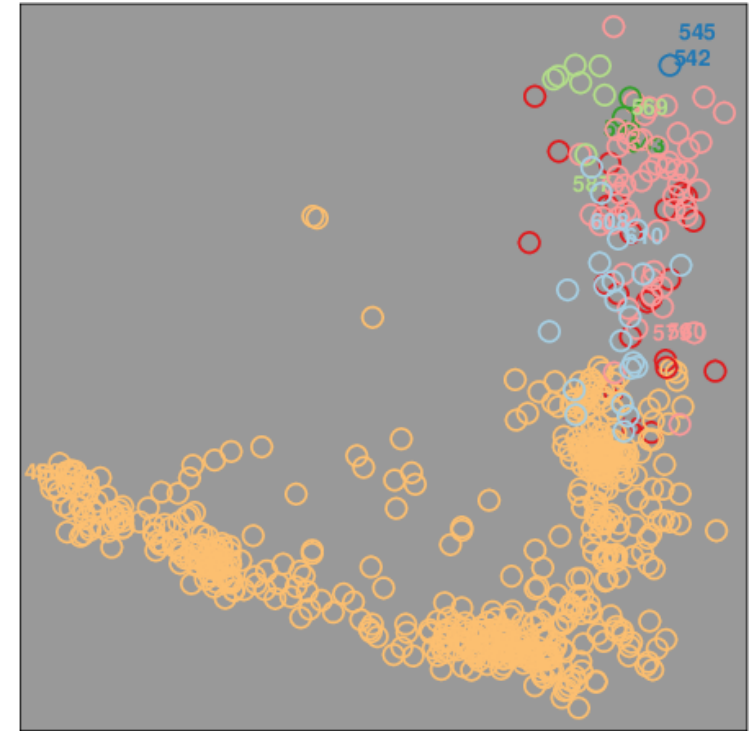
- OCCITAN**
- 14. Croissant's (transitional)
 - 15. Limousin
 - 16. Auvergnat
 - 17. Languedocian
 - a. guyennese
 - b. cévennol
 - c. languedocian-N
 - d. languedocian-C
 - e. languedocian-E
 - f. ariegese
 - 18. Vivaro-alpin
 - a. vivarese
 - b. dromese
 - c. pre-alpine gavot
 - d. alpine gavot
 - e. cisalpine oc
 - 19. Provençal
 - a. rhodanien
 - b. central
 - c. maritime
 - d. niçard

BOUNDARIES

- national
- - - language
- ⋯ dialectal



Moyenne de groupe, 7 classes
⇒ Validation de custers
MDS



3. Horizon de falsification : validation des clusters

3.2. Paramètre 2 : le lexique, par calcul binaire (fonction « catégorielle »)

Item:

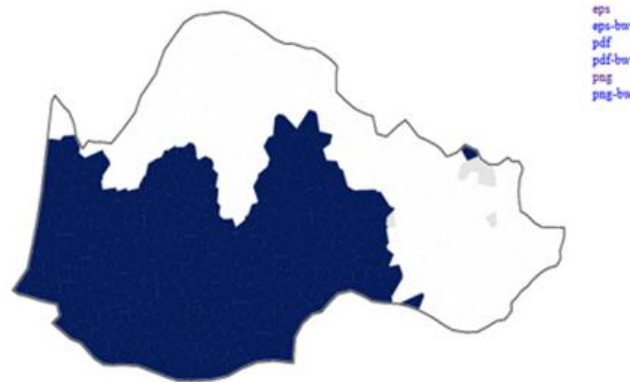
There are 5 variants for *faucher*
Select one or more variants:

- copar (1)
- dalhar (384)
- fauçar (63)
- segar (175)
- seitar (17)

— or —

Regular expression:

Distribution map for "dalhar" in faucher



Item:

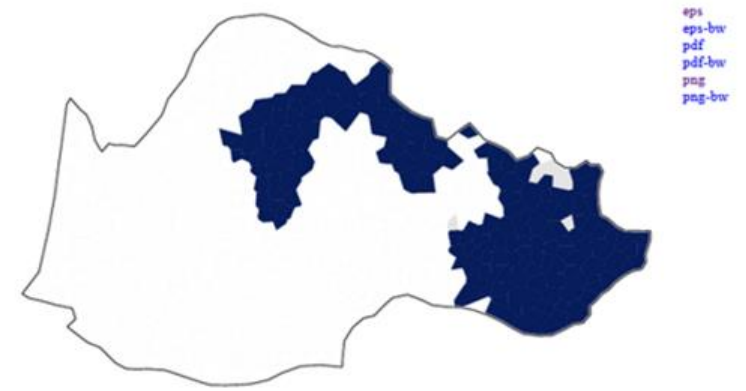
There are 5 variants for *faucher*
Select one or more variants:

- copar (1)
- dalhar (384)
- fauçar (63)
- segar (175)
- seitar (17)

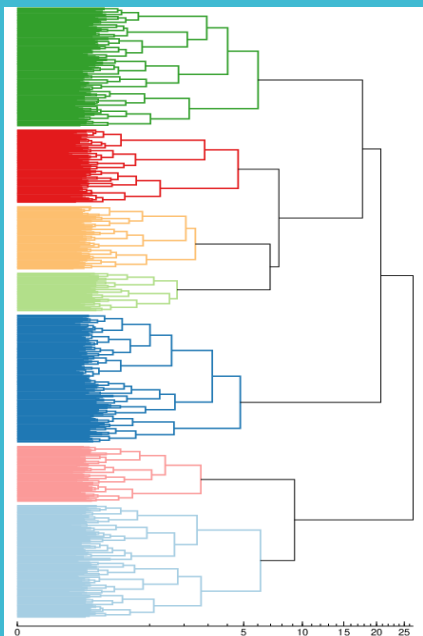
— or —

Regular expression:

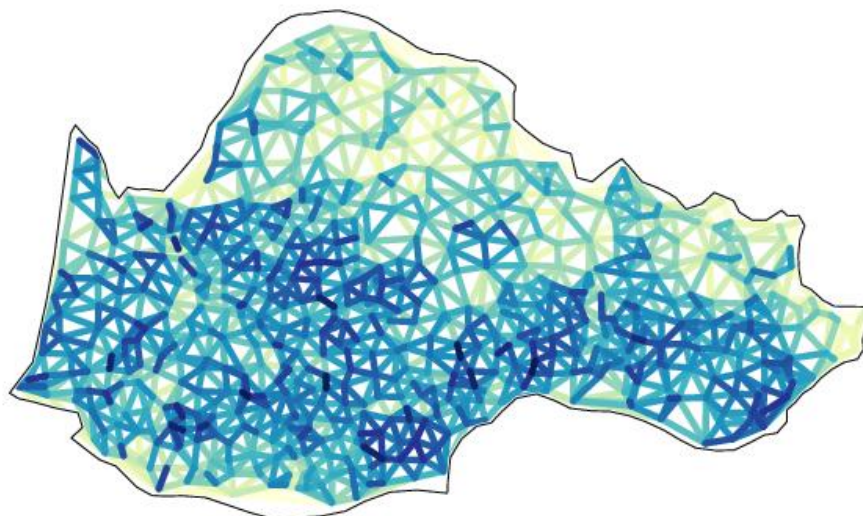
Distribution map for "segar" in faucher



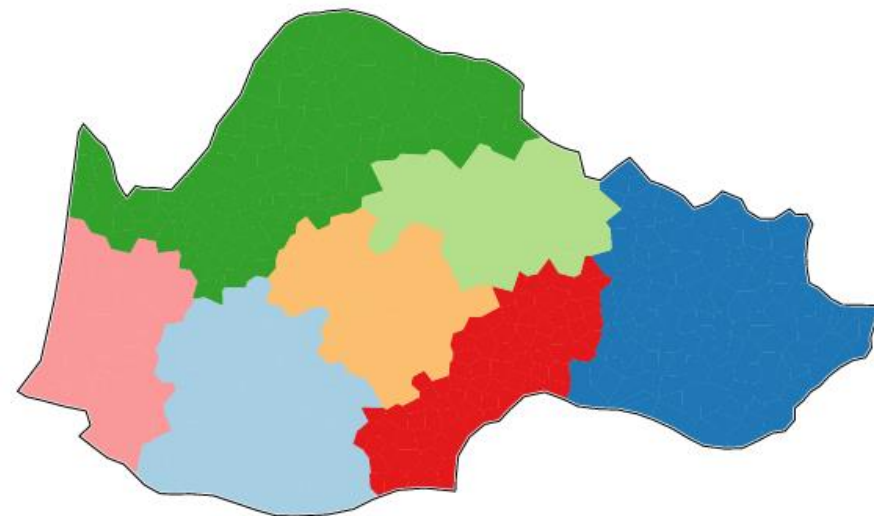
Lexique, THESOC :
regard sur l'algorithme
« canonique »
(Méthode de Joe Ward
Jr.)



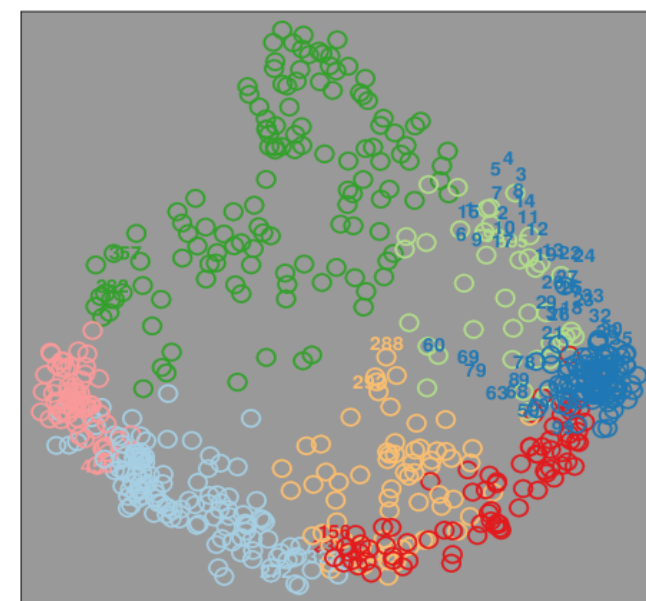
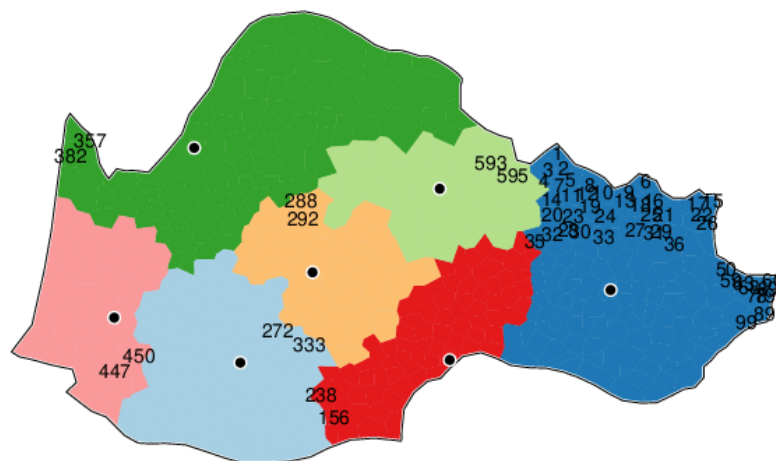
Réticularité des liens lexicaux en domaine d'oc



CAH, 7 classes, lexique



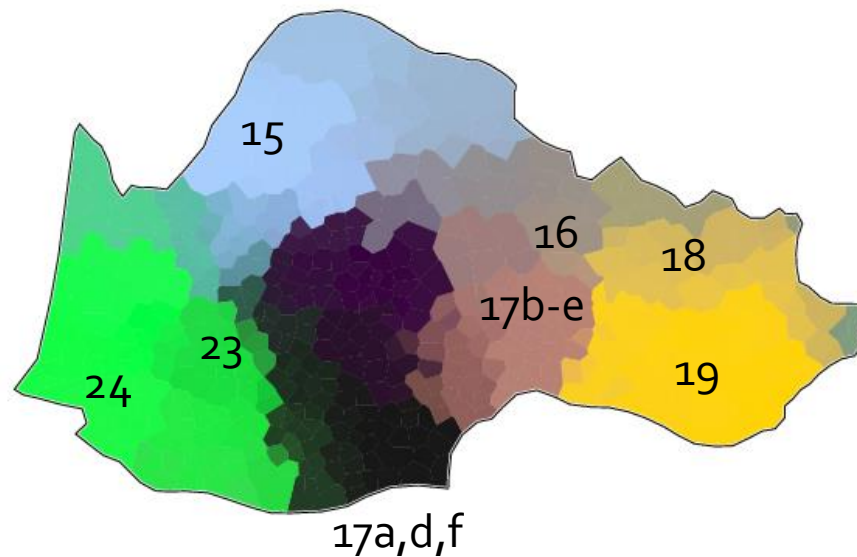
Validation, CAH, 7 classes, lexique



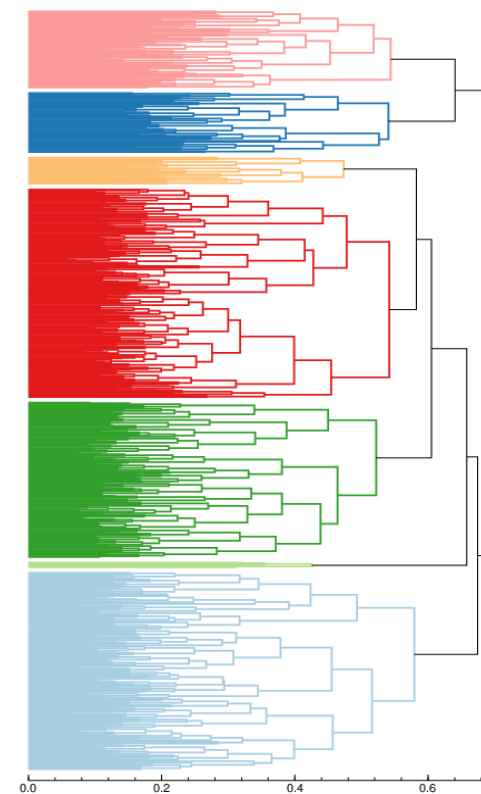
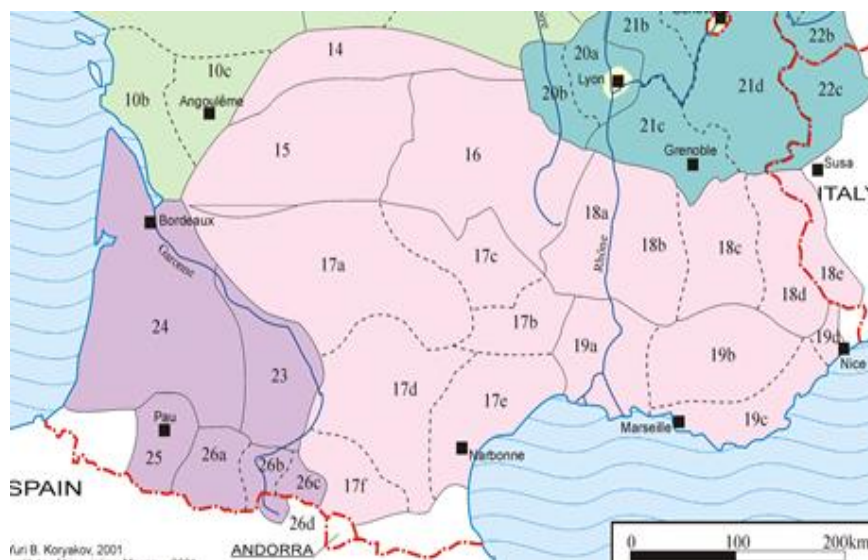
Lexique, THESOC :

Le limousin et le vivaro-alpin (nord-occitan) différent davantage que le gascon du reste du diasystème lexical.
=> Nord-Occ. Et gascon sont **clades structurants externe et interne** respectivement

Positionnement multidimensionnel



Complete Link, 7 classes



Lexique, THESOC :

Moyenne de groupe :

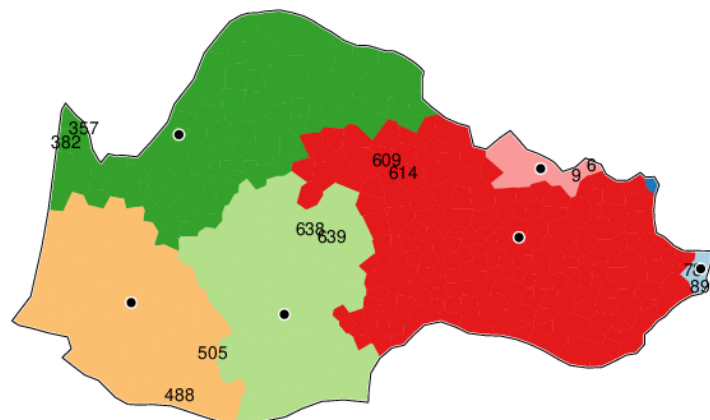
Occitan central & occidental
vs occitan oriental (deux
blocs) : configuration très
robuste

Moyenne pondérée :

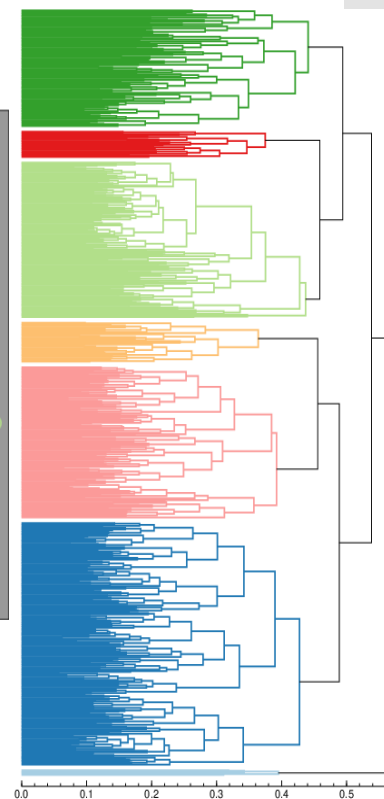
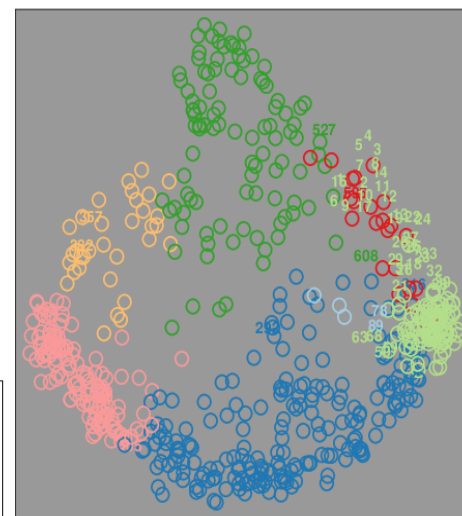
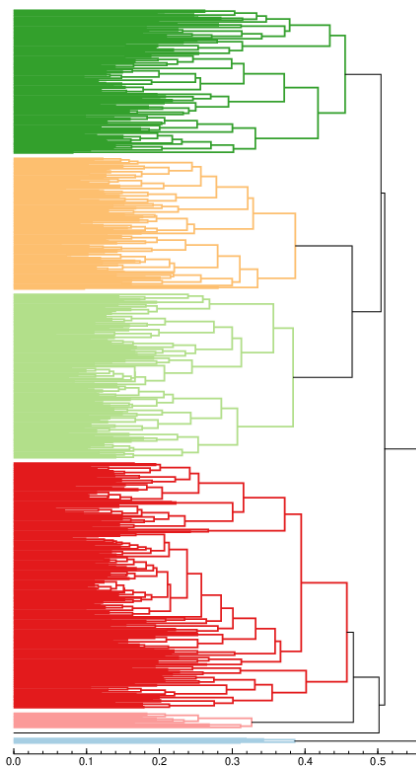
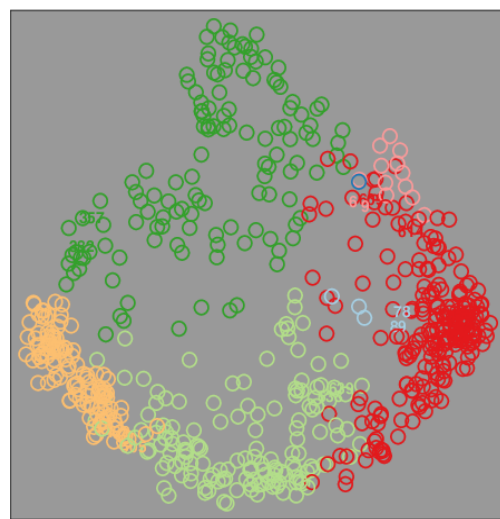
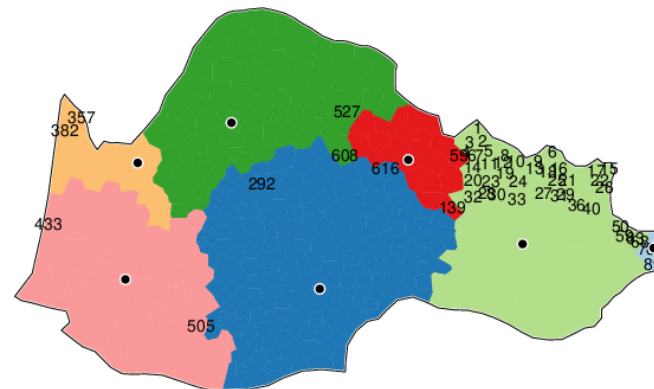
arverno-méditerranéen
labile sur son flanc oriental
vs bloc robuste de l'occitan
centre-occidental

NB : donne détection du
singleton ligure au SE

Moyenne de groupe, 7 classes

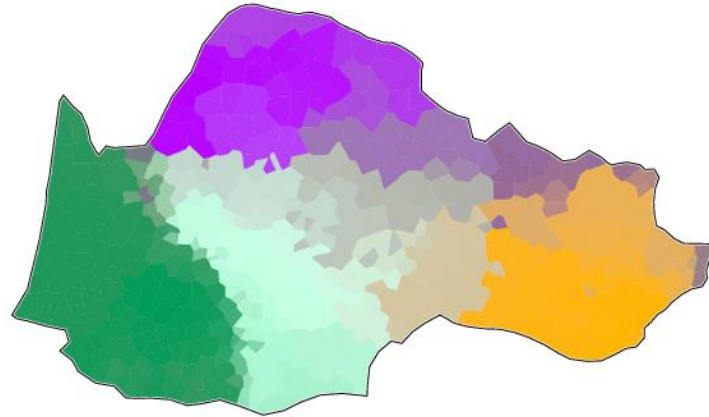


Moyenne pondérée, 7 classes

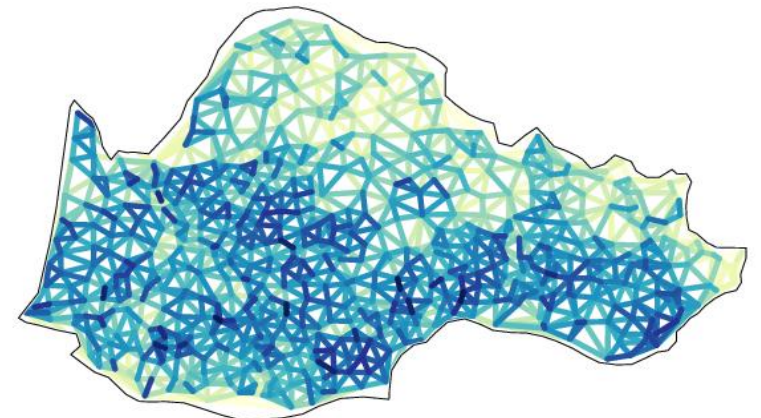
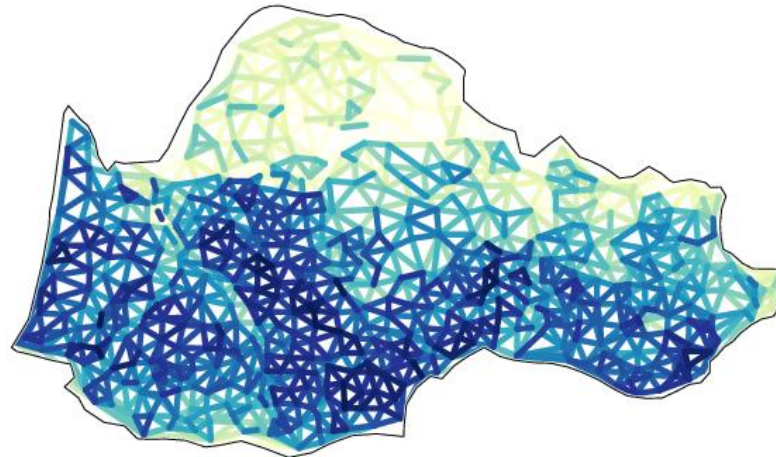
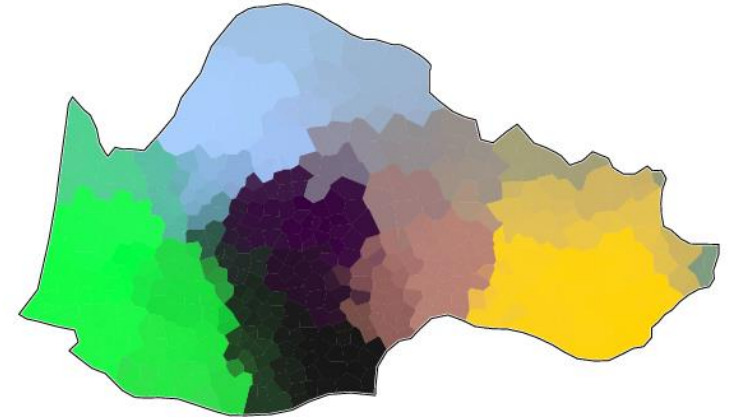


4. Perspectives heuristiques

Phonologie, distance d'édition



Lexique, traitement
« catégoriel » (binaire)



Perspectives heuristiques, 1

- Transcender les classifications canoniques des manuels et explorer des gammes de configurations (aires) et de dispositions (taxinomies) vicariantes permet de reconsidérer en profondeur l'hologramme diasystémique d'un domaine comme l'occitan.
- Les conséquences pour la dialectologie générale sont notables : un dualisme émerge qui oppose un chiasme nord/sud dans la composante phonologique (systèmes fermés) à un chiasme est/ouest dans la composante lexicale (systèmes ouverts).
- **Congruence** avec la « manuélistique » d'une part et **falsification/validation** par bruitage/tuilage des structures internes du diasystème pour ces deux paramètres permettent de tester ce dualisme diasystémique. Sa robustesse est confirmée.

Perspectives heuristiques, 2

- Une telle approche confirme l'hypothèse de l'Ecole de Salzburg (Hans Goebel) selon laquelle la structuration de l'espace géolinguistique reflète l'auto-organisation de structures émergentes d'aménagement linguistique entre ces décideurs multiples que sont les dialectes d'un même domaine (une sphère d'*Abstand* klossien).
- L'une des conséquences est que, paradoxalement, à la fois toutes les ontologies connues (dialectes éponymes) sont à la fois confirmées, mais aussi qualifiées sur le plan hiérarchique et, partant, remises en question.
- Ontologie(s) des « dialectes invisibles » de Nerbonne & Kretzschmar (2003) : pour ces entités, des concepts comme les « hubs » et « small worlds » de Duncan Watts seraient heuristiques, en tenant compte de la critérologie statistique proposée par cet auteur. Voir aussi les ontologies de Geoffrey West, en TC.