



**HAL**  
open science

**Jalons pour une étude géotoponymique de quelques  
distributions macrotoponymiques dans l'espace  
gallo-roman**

Francis Manzano

► **To cite this version:**

Francis Manzano. Jalons pour une étude géotoponymique de quelques distributions macrotoponymiques dans l'espace gallo-roman. L'onomastique gallo-romane alpine : actes de la conférence annuelle sur l'activité scientifique du Centre d'études francoprovençales, Centre d'études francoprovençales René Willien, pp.231-255, 2008. hal-00369371

**HAL Id: hal-00369371**

**<https://univ-lyon3.hal.science/hal-00369371v1>**

Submitted on 19 Mar 2009

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Jalons pour une étude géotoponymique de quelques distributions macrotoponymiques dans l'espace gallo-roman

*Francis Manzano*



## Préambule\*

La communication s'attache à observer quelques distributions géotoponymiques à travers le domaine gallo-roman, ici représenté par le territoire politique de la France métropolitaine.

On se base notamment sur trois types de strats bien identifiés dans la toponymie gallo-romane et française : strat prélatin (ou plus justement pré-indo-européen), toponymie celtique et gauloise, toponymie latine. La toponymie d'origine germanique sera également évoquée rapidement parce qu'il faut, d'une manière ou

d'une autre, la mettre en corrélation avec les couches précédentes pour que le système d'ensemble s'éclaire valablement.

Ces trois (ou quatre) grands types de strats sont envisagés dans leurs distributions géographiques objectives, à partir de données désormais passées dans la vulgate toponymique, ce qui constitue la première étape de la démarche cartographique. Pour autant, il ne s'agit certes pas de données exhaustives mais bien d'une sorte d'échantillonnage qui permet d'imaginer ce que pourrait apporter une cartographie systématique de la toponymie française.

Les strats sont ensuite travaillés à partir d'une méthodologie statistique relativement simple (études de distributions moyennes et définitions de bornes d'écartement). Cette approche des données, en permettant de définir des seuils de pertinence des données observées, donne l'occasion de porter un regard relativement neuf sur la construction toponymique historique et la dynamique sociolinguistique ancienne du domaine gallo-roman.

On voit ainsi se confirmer certaines conclusions classiques, mais aussi quelques surprises apparaître : concentration du fonds pré-roman sur un axe pyrénéo-alpin, faiblesse ou absence du strat celtique dans la zone armoricaine,

au sud-ouest ou dans la zone alpine, scores très élevés de la toponymie eu-latine dans le nord de la France ou, à l'opposé, relativement faibles dans certaines régions méditerranéennes où on les attendrait pourtant.

## 1. La méthode et les données

Volontairement, la cartographie n'est pas basée sur des données exhaustives. Prétendre à des résultats exhaustifs supposerait un travail plus long et méticuleux, systématisé, à cette réserve d'ailleurs que l'exhaustivité n'est guère possible dans ce genre de domaine, où pour certaines composantes (notamment plus on remonte dans le temps) deux solutions diachroniques, ou davantage, sont souvent possibles.

Malgré cet éclairage et ces réserves on constatera que la cartographie de telles données révèle bien que les distributions géotoponymiques permettent de cerner des aires anthropologiques diachroniques fortement significatives, les incertitudes étymologiques qu'on peut avoir par rapport à tel ou tel cas passant au second plan et se noyant en fait dans la logique statistique que les données sûres définissent elles-mêmes.

À partir de quelques types relativement bien connus pour chaque strat, on a effectué préalablement un décompte des toponymes département par département et région par région (régions administratives s'entend). Ces types ont été recueillis et vérifiés, il y a quelques années déjà, à partir des corpus d'Ernest Nègre, des données du *Dictionnaire des noms de lieux de la France*, de celles de Paul Fabre plus tard, d'autres encore.

Prenons un exemple, il y a 23 macrotoponymes foncièrement latins d'origine en Languedoc-Roussillon correspondant aux étymons définis dans la vignette (type : *Luc sur Orbieu, Félines-Termenès, Laroque de Fa* < LUCU, FIGULINAS, FANU, département de l'Aude) et 11 pour le Nord-Pas-de-Calais (type : *Aix en Ergny, Fins, Flines* < AQUIS, FINIS, FIGULINAS, Pas-de-Calais, Nord).

Observons au passage que la densité, déjà remarquable dans l'un et l'autre cas, ne ferait que s'accroître si l'on cartographiait *tous* les types d'origine latine.

Mais telles quelles, dans l'absolu, ces données ne donnent pas grand chose, elles pourraient même induire en erreur, car évidemment la plupart d'entre nous pensent immédiatement que le strat latin pèse forcément plus à Narbonne qu'à Arras !

C'est pourtant une idée reçue. Si l'on élabore en effet un indice de densité (nombre de toponymes du type [multiplié par] 10 000 et [divisé par] la densité de la région en kilomètres carrés) on arrive à 8,37 pour le Languedoc-Roussillon et 8,87 pour le Nord : deux régions vedettes certes, mais pas dans l'ordre attendu.

## **2. Traitement statistique**

En procédant ainsi, on obtient donc des listes de densité qui servent de base à nos calculs. Sans donner l'ensemble des tableaux et des séries statistiques observées, ce qui serait fort long et fastidieux, je me bornerai à préciser rapidement et efficacement les grandes lignes de ce traitement.

Par la suite, une série statistique résultera toujours du croisement entre un ou plusieurs types toponymiques de référence<sup>1</sup> et les individus que constituent les 21 régions administratives de la France<sup>2</sup>.

Parmi les paramètres statistiques, la moyenne arithmétique est le plus accessible. Parfois discutée dans la littérature statistique (pour des raisons techniques), la moyenne arithmétique présente pour nous l'avantage de fixer facilement et au premier coup d'œil des zones de productivité (inférieures ou supérieures à la moyenne de la série), comme on le verra plusieurs fois par la suite. Dans un deuxième temps on calcule l'écart type corrigé de la série. Celui-ci, ajouté ou retranché à la moyenne, permet de définir une borne supérieure (BS) et une borne inférieure (BI). Dans l'exemple cité plus haut, BS est à 5,9, et BI est à 0,58. Cette méthode permet donc d'écarter les données en vue d'une interprétation ultérieure.

Ce qui précède est possible pour les distributions relativement "normales" (ou gaussiennes), ce qui correspond pour notre domaine à des strats relativement diffus dans l'espace. Mais dès que la distribution s'écarte du schéma moyen, il peut ne plus exister de BS ou de BI. Tel est le cas des apports germaniques, tellement concentrés dans le Nord et l'Est que le reste du domaine gallo-roman semble (à tort peut-être) un désert. Mais, évidemment, même cela est intéressant, car dans le cas d'espèce cela révèle aussi bien une frontière historique et culturelle d'envergure entre Romania et Germania que les solutions de continuité trouvées par la Romania pour intégrer les apports germaniques.

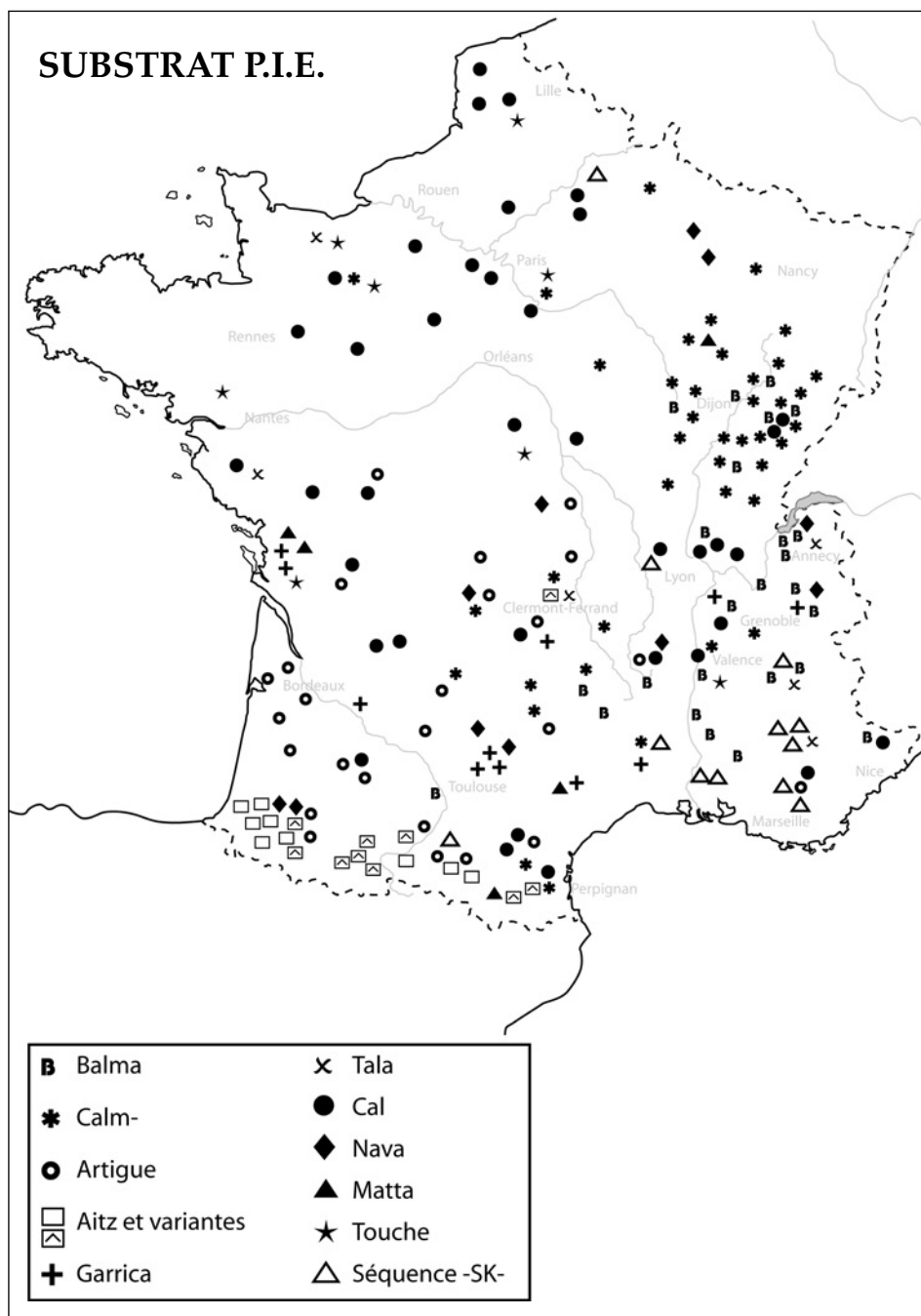
## **3. Les cartes de distribution objective, ou cartes de base**

Elles sont représentées avec le réseau hydrographique, qui n'est pas sans importance et donne des repères, notamment en termes de circulation et de pénétration géographiques. À terme, quelques courbes hypsométriques simplifiées pourraient être intégrées. On a ajouté les situations approximatives de quelques grands centres urbains, afin de faciliter le repérage.

### *Substrat pré indo-européen*

On fait figurer ici quelques types très connus et pour la plupart peu discutés. Ils fournissent différents successeurs du mot de départ et de très nombreux

dérivés ou composés. Observons qu'un bon nombre de ces formations survivent usuellement dans les microtoponymies et les lexiques topographiques régionaux.

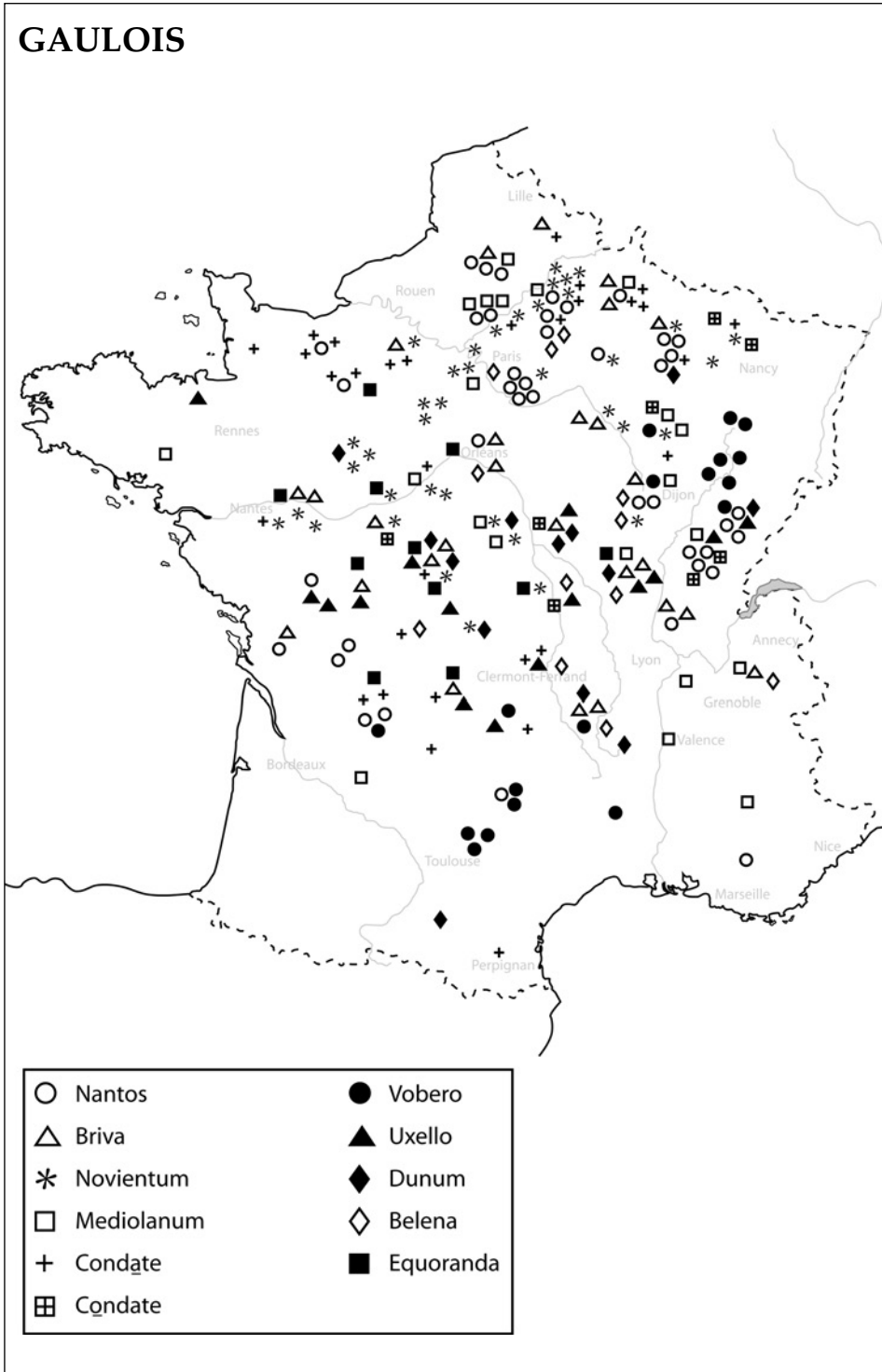


- BALMA** "grotte, cavité", types toponymiques actuels *Balma*, *Balmette* (Oc, Fpr), *Beaume*, *Beaumette* (Oïl) etc. Le type originel est pré-latin, on le dit parfois celtique ou gaulois, mais sa distribution géographique s'accorde en réalité très mal avec la distribution des types eu-celtiques (voir ci-après).
- CALM-**, **CALMA** "hauteur, plat dénudé", types *Lacalm*, *Chaulmes*, *Charmes* (Oc, N-Oc, Fpr), *Chaume*, *Chaumes* (Oïl) etc.
- ARTICA** "friche", types *Artigues*, *Lartigue*, *Artiguelongue* etc. Formations toponymiques principalement méridionales avec haute fréquence dans la zone aquitaine et le sud-ouest<sup>3</sup>.
- AITZ** (basque)<sup>4</sup> "pointe rocheuse", var. *AST*, se retrouve notamment le long de la chaîne pyrénéenne et semble-t-il jusqu'aux régions méridionales du Massif Central. Types *Assas*, *Astis*, *Asté*, *Aston* etc.
- GARRIC**, **GARRICA** "terre calcaire pauvre, pierreuse", probablement en rapport avec la base originelle PIE \*CAR-, \*GAR- "pierre", a désigné en particulier la formation végétale sur sols calcaires des régions eu-méditerranéennes (oc. *garriga*, fr. *garrigue*) et au-delà semble-t-il. Types *Garrigues*, *Lagarrigue* (Oc), *La Jarrie* (Oïl) etc.
- TALA** "argile", types *Tallard*, *Talvon*, *Tallore* etc. Dans divers cas, des confusions ne sont pas exclues avec le germanique (*tal* "vallée") ou même avec le latin (différents anthroponymes à initiale *Tal*-).
- CALA** "pierre" est probablement l'une des bases PIE les plus connues (souvent présentée sous la forme \*KAR-). Elle fournit les très nombreux types *Chelles* (Oïl), *Challes* (Fpr), *Callas* (Oc), et beaucoup d'autres en construction.
- NAVA** "plaine, vallée"<sup>5</sup>, types *Nabas* (Oc), *Naves*, *Naives* (Fpr, Oïl) etc.
- MATTA** "forêt, bois, buisson", apparaît dans les types *Les Mathes* (Oïl), *Matemale*, *Les Matelles* (Cat, Oc) etc.
- TOUCHE**, très commun dans le centre-ouest de la France (microtoponymie) où il désigne des bosquets notamment intercalés entre les cultures, se retrouve dans diverses régions. On le donne généralement comme d'origine pré-celtique. Types *La Touche*, *Les Touches*, *Le Touquet* etc.
- La séquence [sk]**, est présente dans différents toponymes comme *Tarascon*, *Manosque*, *Vénasque* etc. Comme on peut le constater, c'est une distribution plutôt méridionale et même méditerranéenne, bien que des écarts soient possibles. Aussi a-t-on traditionnellement mis ce champ toponymique en relation avec un substrat ligure.

Chaque type ou série pourrait prêter à un commentaire particulier. Les zonalités peuvent en effet être très différentes : plutôt alpine et jurassienne pour *BALMA*, évidemment pyrénéenne pour *AITZ* etc.

Dans l'ensemble, le phénomène le plus remarquable est le déport du "nuage" au sud / sud-est d'un axe de symétrie que l'on peut tracer facilement et qui irait de l'Alsace au Pays Basque. La grande masse des occurrences tend donc à libérer des espaces pour les couches suivantes.

# GAULOIS



### Gaulois

Les types retenus font également partie du stock généralement admis de la toponymie d'origine gauloise. Certains, comme **CONDATE**, pourraient avoir été repris par les Gaulois à des peuples antérieurs.

**NANTOS** "vallée", types *Nant, Nantua, Nampt* et diverses constructions possibles.

**BRIVA** "pont", également simple ou en construction, types *Brive* (Oc), *Brèves, Brissarthe, Brienne* (Oïl) etc.

**NOVIENTUM** "ville neuve", types *Nogent, Nohant, Nepvant* etc. Très grande représentation du type en langue d'oïl.

**MEDIOLANUM** "point" [ou] "sanctuaire central", types *Meillant, Mâlain, Meulin* (Oïl), *Miolan, Meylan* (Fpr), *Méolans* (N-Oc) etc<sup>6</sup>.

**CONDATE**, accentué sur la pénultième (**CONDATE**) il fournit des toponymes du type *Candé, Condé* (Oïl), *Condat* (Oc), *Conat* (Cat), ainsi que différentes constructions. Accentué sur l'antépénultième (plus rare : **CONDATE**), il aboutit à des types *Condes, Candes*.

**VOBERO** "ruisseau", types simples ou construits *Woivre, Vesvres, Vouvray* (Oïl), *Vorey, Veurey* (Fpr), *Vabre, Lavaur* (Oc) etc.

**UXELLO** "lieu élevé", types *Usseau, Uxeau* (Oïl), *Ussel* (Oc) etc.

**DUNUM** "enceinte fortifiée", simple ou en construction, types *Dun, Dom le Mesnil, Dunet* etc.

**BELENOS**, nom de divinité, sous sa forme féminine \*BELENA, est à la source de différents *Baulne, Beaune* et autres formes construites.

**EQUORANDA**, formé sur le gaulois **RANDA** "limite" (d'eau, de cités, de tribus<sup>7</sup>), types *Eygurande, Iguerande, Ingrande* etc.

Le point le plus remarquable sans doute est la concentration relative "centro-septentrionale" de ces toponymes, qui s'accorde bien avec les données de l'Histoire et de l'archéologie. On doit également remarquer le relâchement du réseau à l'ouest et dans le grand sud, ainsi que dans les régions de forte influence germanique (Alsace, Nord-Pas-de-Calais).

### Celtique : clairières, végétaux

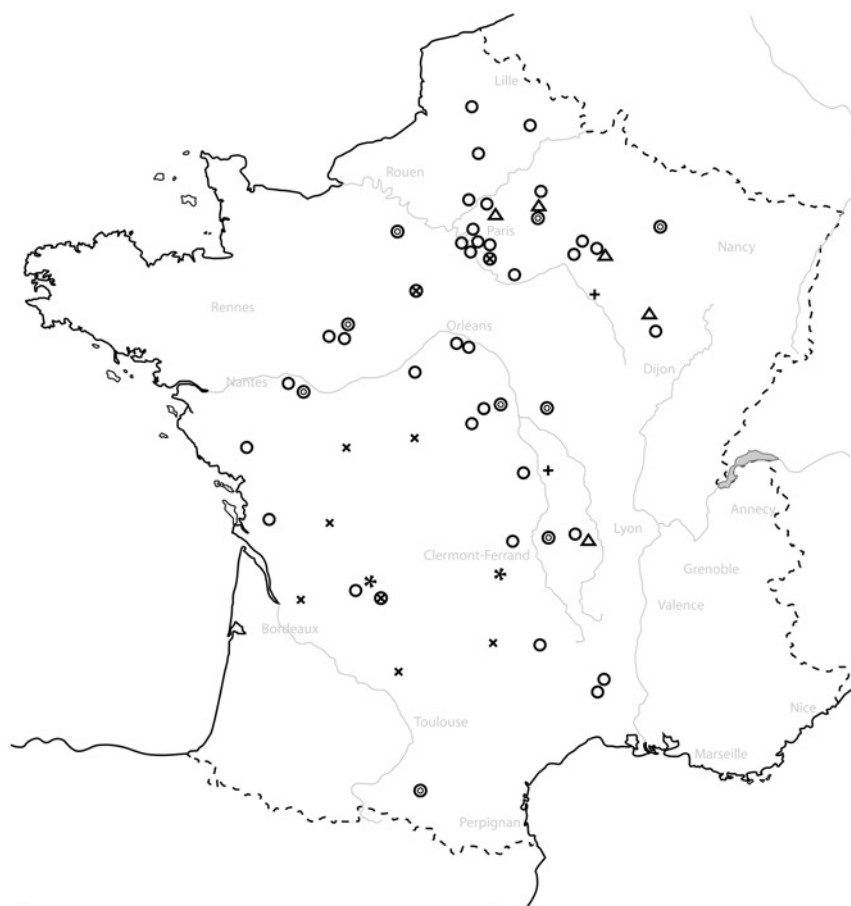
On a volontairement séparé cette carte de la précédente. En effet, bien que les distributions livrent des ressemblances globales normales, cette carte suggère des progressions le long des cours d'eau et fournit donc un cadre interprétatif intéressant pour des périodes où les défrichements se sont produits vraisemblablement de proche en proche, à partir de zones d'accès déjà aménagées et de vallées accessibles.

Les séries observées révèlent des compositions typiques, à partir notamment d'un matériel usuel du celtique, dont notamment le "suffixe" gaulois -IALO,



qui désigne une clairière<sup>8</sup>. Il est intéressant du coup d'observer que le premier élément livre des compléments écologiques d'intérêt quant au couvert végétal et à l'aménagement de cette époque. Certaines de ces compositions apparaissent comme fréquentes (la première et la dernière), les autres sont plus rares.

## CELTIQUE : clairières, végétaux



- Maro (grand) + ialo
- △ Vindos (blanc) + ialo
- \* Aballo (pomme) + ialo
- × Cassanos (chêne) + ialo
- + Ebuo (if) + ialo
- ⊗ Lemo (orme) + ialo
- ◎ Verno (aulne) + ialo

- MARO** "grand" + **IALO**, types *Mareil, Mareuil* etc.  
**VINDO** "blanc" + **IALO**, types *Vandeuil, Vendeuil* etc.  
**ABALLO** "pomme" + **IALO**, types *Valeuil, Valuéjols* (Oc).  
**CASSANO** "chêne" + **IALO**, types *Casseneuil, Casseuil* (Oc),  
*Chasseneuil* (Oïl).  
**EBURO** "if" + **IALO**, types *Avreuil, Ébreuil*.  
**LEMO** "orme" + **IALO**, types *Lumeau* (Oïl), *Limeuil* (Oc).  
**VERNO** "aulne" + **IALO**, type *Verneuil* très courant. Le formant **VERNO** est par ailleurs extrêmement productif dans la toponymie de la France, surtout dans la moitié septentrionale (types *Vern, Vernais* [Oïl], *Vernaz* [Fpr], *Verniolle* [Oc], *Vernet* [Cat] etc.).

Avec cette carte on a également la confirmation du fait que certaines régions semblent avoir échappé à cette progression celtique ou, du moins, semblent en avoir limité les effets. C'est le cas du grand Sud-Ouest et notamment de la Gascogne (rive gauche de la Garonne), zone traditionnelle des Vascons. De l'Alsace aux côtes provençales, c'est l'ensemble de la rive gauche de la Saône et du Rhône qui fait office de repoussoir. Quand à l'Ouest lato-sensu, on voit bien que se confirme son caractère très peu accueillant pour la toponymie d'origine gauloise.

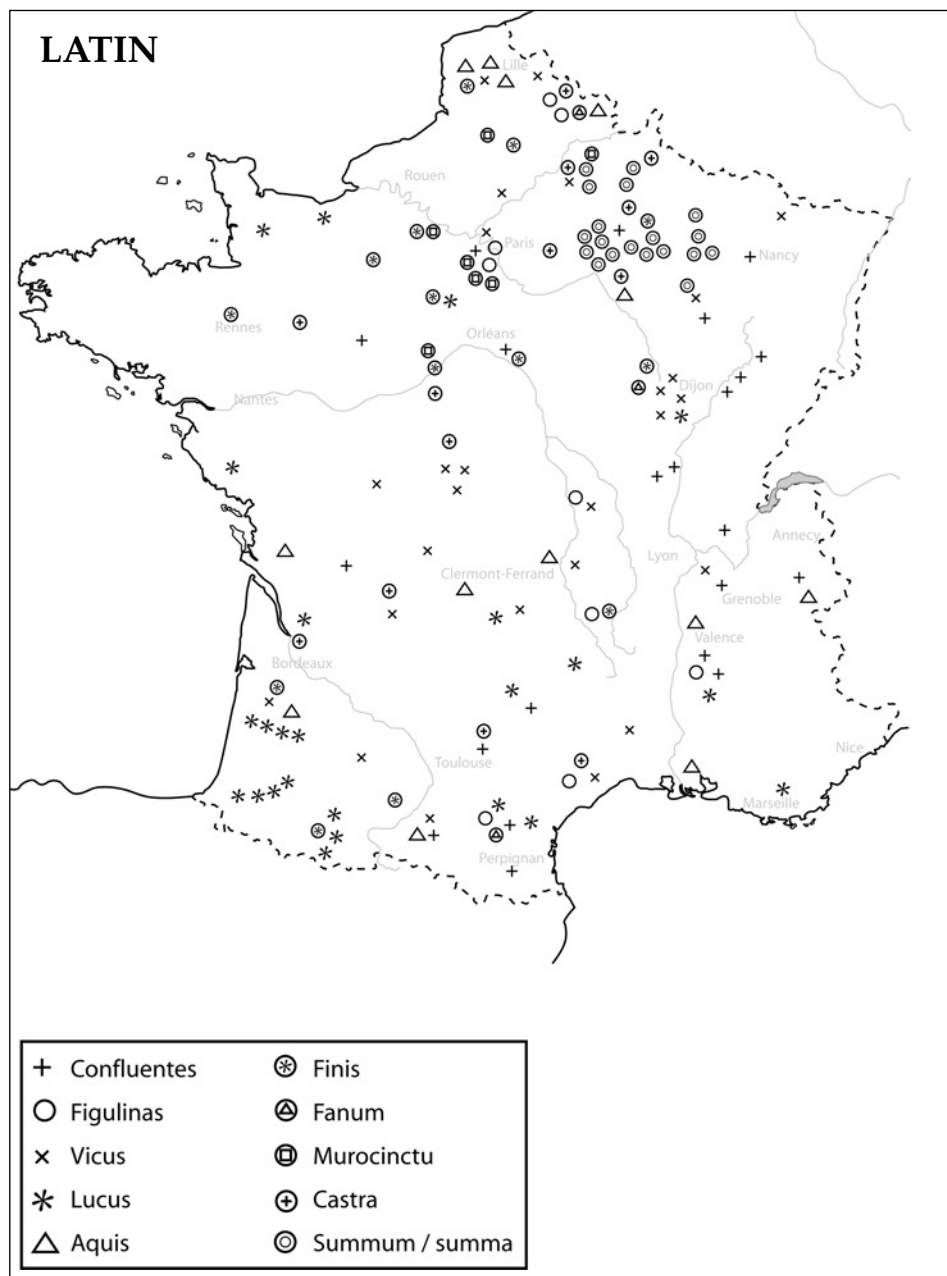
### Latin

Parmi les types retenus, certains sont anciens ou rares. Par là ils sont significatifs d'un strat eu-latin relativement ancien. On remarquera que certains, comme **CONFLUENTEM** ou **FINIS**, pouvaient reprendre la substance de précédents gaulois (**CONDATÉ**, **EQUORANDA**), ce qui probablement explique leur succès.

- CONFLUENTES, CONFLUENTEM** "confluent", types *Conflans* (Oïl), *Couflens, Couffoulens* (Oc), *Conflent* (Cat).  
**FIGULINAS** "atelier de potiers", types *Flines, Félines* etc.  
**VICUS**<sup>9</sup> "village rural"<sup>10</sup>, très productif, types *Vic, Vicq, Vix*. Très fréquent en construction.  
**LUCUS** "bois sacré", comme les précédents, fait partie du fonds eu-latin de la romanisation. On le trouve bien entendu dans le sud, mais aussi plus haut dans le domaine gallo-roman, types *Luc, Lux, Monluc* etc.  
**FANUM** "temple", types *Fa* au sud<sup>11</sup>, *Fain* en langue d'oïl<sup>12</sup>.  
**AQUIS, AQUÆ** "eaux thermales" dont on sait que les romains et gallo-romains étaient grands consommateurs. Types, *Aix, Ax, Dax*.  
**FINIS** "limite, terme" et divers cas (flexion), types *Feins, Fains* (Oïl), *His, Hinx* (Gasc).  
**SUMMUM, SUMMA** "source" (= le point le plus élevé), types *Somme* et différents dérivés et composés : *Sommeilles, Somme-Tourbe* etc.

**MUROCINCTU** "lieu, village fortifié", distribution septentrionale, types *Morsain, Morchain* (Oïl).

**CASTRÀ** "camp fortifié", types *Castres, Castries* (Oc), *Castres, Caëstre* (Oïl), là où la palatalisation de la consonne initiale ne s'est pas produite, ailleurs types *Châtres, Chestres* etc.



La carte révèle une mauvaise pénétration de l'est (Alpes, Jura, Vosges, Alsace), le Rhône-Alpes faisant un peu exception. La Bretagne et l'ouest également sont nettement dépourvus. Le plus surprenant est peut-être que la Provence (dont le nom provient pourtant de la *Provincia [Romana]*) fait pâle figure de ce point de vue. Sous réserve de vérifications plus amples<sup>13</sup>, semblent donc se détacher deux zones de préférence à première vue : l'extrême nord et l'extrême sud. On ne peut manquer de faire également un rapprochement avec la distribution des types gaulois. Émergerait donc le thème d'une forme de concurrence géodiachronique et sociolinguistique entre les deux strats, révélateur probable d'une volonté ou nécessité politique des conquérants d'aller vers l'occupation et le contrôle des zones précédemment contrôlées (au moins en partie) par les Gaulois. On constate toutefois, par simple superposition des cartes brutes, quelques déplacements des aires, notamment vers le sud-ouest.

Deux types n'ont pas été directement cartographiés ici. Ils sont bien connus dans leurs formes et dans leurs distributions depuis un bon siècle, aussi ne fait-on ici que les évoquer tout en les comptabilisant et en les exploitant dans les cartes statistiques qui suivent.

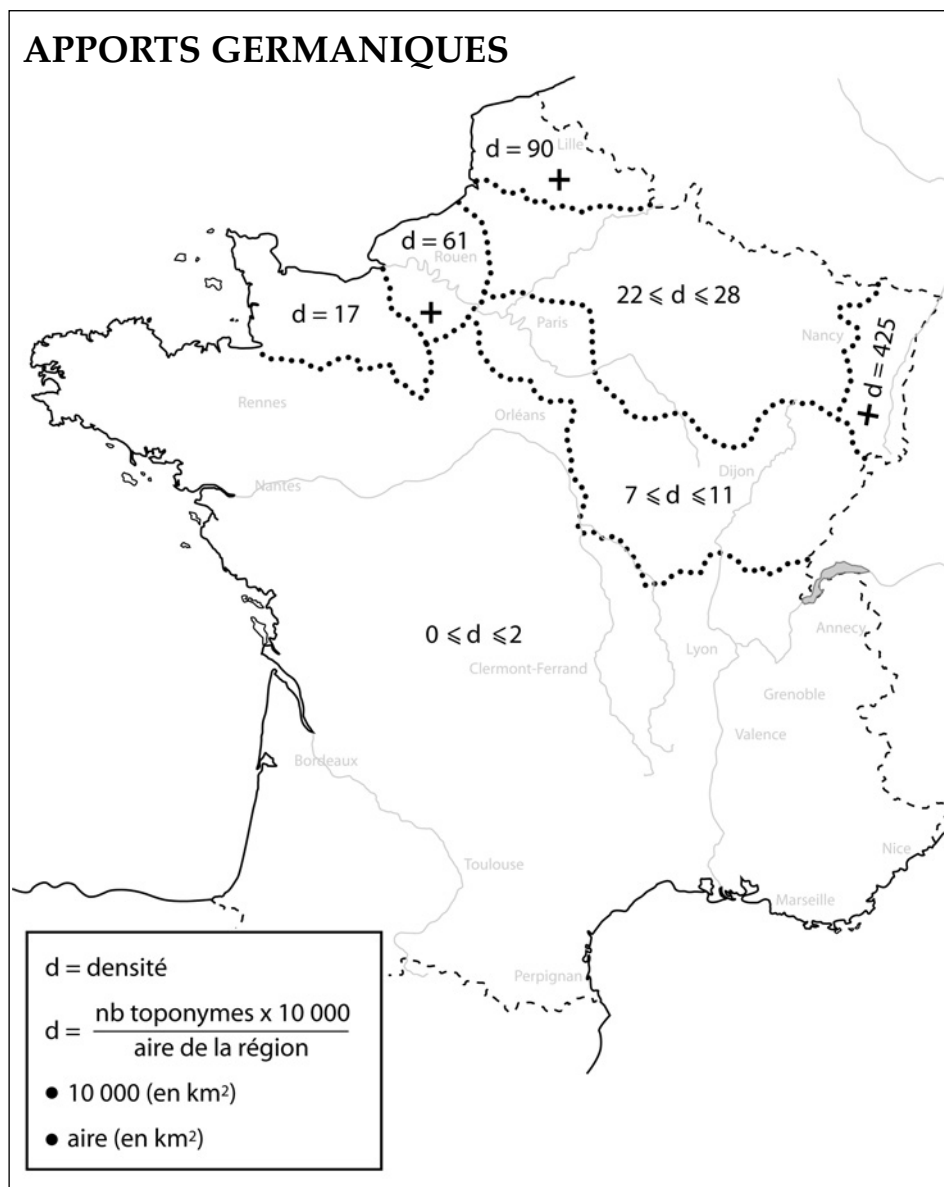
Il s'agit des formations comportant les suffixes –ACU(M) et –ANU(M).

Du premier type sont les nombreux *Neully*, *Neuillay*, *Neuillé* de la langue d'oïl, *Noaillac*, *Noalhac* de la langue d'oc. C'est un paradigme extrêmement productif qui implique des centaines de toponymes du genre.

Du second type, moins fréquent et diffus, sont les *Draguignan*, *Frontignan*, *Lézignan* ou *Perpignan* (pron. *Perpignà*) de la région méditerranéenne. Sans compter les formes résultant d'un déplacement d'accent (*Prouille*, *Bize*). À ce corpus méditerranéen correspondent plus au nord les types du Fpr *Poncin*, *Tullins* etc. Plus au nord encore le type proprement dit (–ANU) disparaît il est vrai, mais un foyer de formes de féminin pluriel (–ANAS) existe dans le Nord (*Valenciennes*, *Marchiennes*)<sup>14</sup>, ajusté à la région de forte romanisation qui a été déjà évoquée.

Pour résumer très brièvement le dossier, le premier type fut prolifique à époque gallo-romaine, mais le suffixe était d'origine gauloise (–ACOS). Quant aux formations en –ANU(M), elles semblent liées à des structures d'aménagement agropastoral d'envergure comme les *villæ*, largement significatives d'un aménagement paysager de type "romain".

Sur la base de cette opposition binaire, certains toponymistes ont tendu presque naturellement à faire du premier type un révélateur "celto-septentrional" et du second un révélateur "latino-méditerranéen". Si tout cela, globalement, n'est pas faux (les cartes suivantes le montrent), dans cette approche c'est la simplification, la volonté de trancher (assez significative d'une époque des études romanes) qui pose problème.



### *Regard sur les apports germaniques*

Pour les types d'origine germanique, on observera la carte simplifiée des densités, établie à partir de quelques types significatifs. Les distributions révélées par cette carte sont très explicites en terme de zonalité, mais quelques précisions seront données plus loin. Il s'agit comme on peut le voir, de distributions massivement septentrionales, bien que certains types puissent se retrouver plus au sud, comme par exemple le type FARA.

Ces types renvoient à des composantes différentes (diachroniques, ethniques etc.) du superstrat germanique d'ensemble comme : FARA "famille" (*Lafare, La Fère*), TAL ou THAL "vallée" (*Dieffenthal, Dieppedale* ou *Dieupentale* suivant les régions), BACH "ruisseau" (*Murbach, Forbach*), HEIM "village" (*Dieholsheim*), BERG "montagne" (*Berck, Berg, Bergheim* etc.). Ils sont généralement productifs du Pas-de-Calais à l'Alsace.

Les côtes de la Manche et de la Normandie présentent de leur côté un faisceau fort caractéristique, d'origine le plus souvent scandinave. HAM, "village" doit être rapproché de HEIM qui précède (*Ham, Hem, Ouistreham*). On citera aussi les noms en *-beuf* (*Elbeuf, Criquebeuf*), en *-fleur* (*Honfleur, Barfleur*), en *-bec* (*Bolbec, Varenguebec*), en *-lon* (*Bouquelon, Yquelon*), en *-tot* (*Criquetot, Yvetot*) etc.

#### **4. Les cartes géostatistiques**

Les cartes abordées précédemment, tout en nous mettant sur la voie, nous amènent à faire des constats le plus souvent intuitifs et qualitatifs. Le petit traitement statistique que nous leur appliquons va permettre une meilleure saisie des problèmes.

##### *Substrat pré indo-européen*

Sur la base de la moyenne, on sépare très facilement le sud et le nord du domaine gallo-roman. Se confirme ainsi l'idée ancienne que le substrat pèse davantage sur les paysages linguistiques du sud. Il est intéressant de constater que la région parisienne possède un statut bien particulier au sein des terres d'oïl. La carte de synthèse, plus loin, le confirmera.

BS et BI permettent de détacher très clairement la zone la plus favorable au substrat qui se déroule des Pyrénées méditerranéennes au Jura (> BS), croisant dimension méridionale et dimension orientale. Les régions < BI sont probablement à mettre en relation avec d'autres phénomènes (apport breton, apport germanique notamment, mais pas seulement sans doute).

Les rangs extrêmes confirment cela. Mais surtout ils révèlent un contraste saisissant, une vraie rupture géolinguistique entre la Franche-Comté et l'Alsace, le 1<sup>er</sup> et le dernier de la série se confrontant directement dans l'espace. Ce phénomène se retrouvera plus loin et semble révéler une vraie frontière diachronique notamment fondée sur la germanisation extrême de l'Alsace.

La Bretagne, comme cela apparaît plusieurs fois dans nos cartes par ailleurs, rejette nettement ce strat, du moins pour les types récurrents observés. Faut-il en conclure que ces régions étaient peu habitées (et donc peu marquées toponymiquement) avant l'arrivée des migrants indo-européens ?

*Gaulois*

Les scores supérieurs à la moyenne révèlent bien la difficulté croissante de la couche celtique à s'établir vers l'ouest et le sud notamment. Que l'on raisonne à partir des moyennes ou des bornes (BS, BI), vers le sud l'ensemble du champ bute visiblement sur le substrat et ses composantes anthropologiques, ainsi que sur les apports latins qui l'ont semble-t-il contrecarré.

Les cartes de rangs et de bornage suggèrent aussi une sorte de pont gaulois-celtique, de la Franche-Comté à la Picardie, ce qui correspond assez bien à nos connaissances de la distribution anthropologique des premiers peuples celtes du futur territoire gallo-roman (voir plus bas).

Comme de juste, l'Alsace s'extrait du système<sup>15</sup>, la Bretagne aussi, mais c'est aussi le cas de la région PACA<sup>16</sup>.

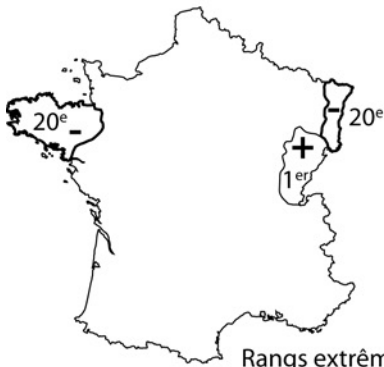
### SUBSTRAT P.I.E.



scores  $> \bar{x}$



+ Régions  $>$  Borne Sup.  
- Régions  $<$  Borne Inf.



Rangs extrêmes

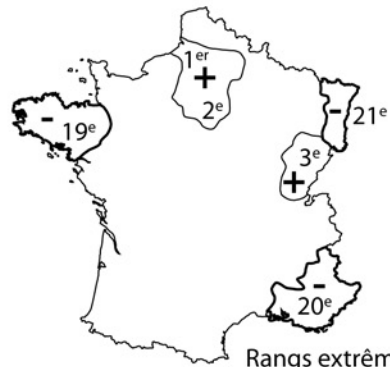
### GAULOIS



scores  $> \bar{x}$



+ Régions  $>$  Borne Sup.  
- Régions  $<$  Borne Inf.



Rangs extrêmes



## CELTIQUE GLOBAL



Régions >  $\bar{x}$

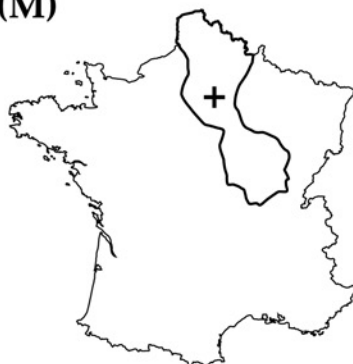


Régions  $\geq$  Borne Sup.

### -ACU(M)



Régions >  $\bar{x}$

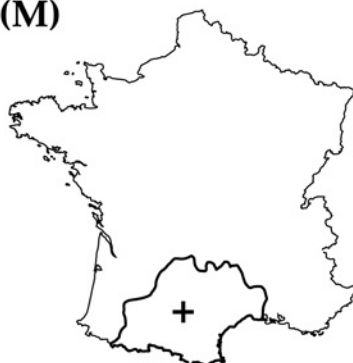


Régions  $\geq$  Borne Sup.

### -ANU(M)



Régions >  $\bar{x}$



Régions  $\geq$  Borne Sup.

*Celtique global*

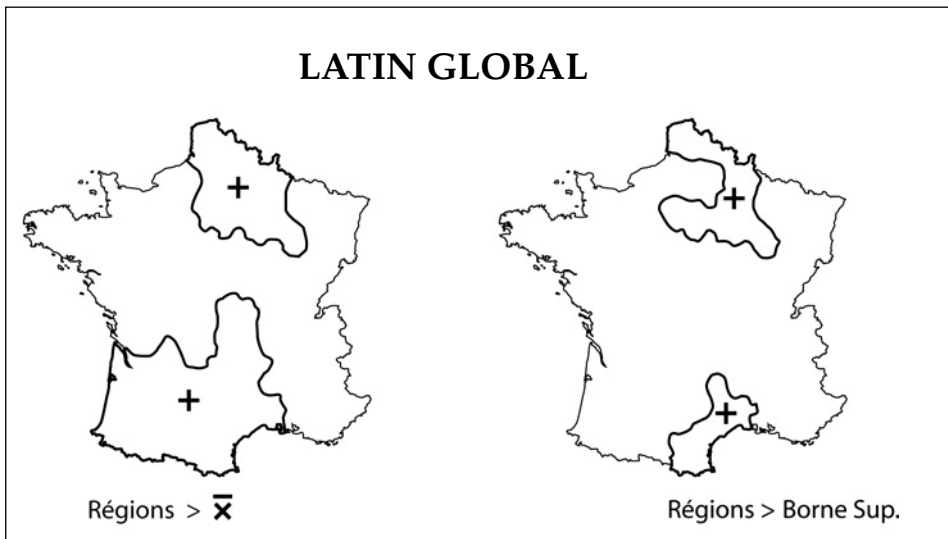
Quand on cumule les 2 cartes relatives au gaulois-celtique (Celtique global), on n'est pas très éloigné des cartes précédentes. Mais une brève superposition montre le gain vers l'ouest et le sud, signe probable d'une descente historique des défrichements et aménagements gaulois.

**-ACU** Comme nous l'avons rappelé plus haut, ce suffixe exprime bien la rencontre entre fonds gaulois et superposition romane. Son extension permet de retrouver plus nettement encore le mouvement de progression vers le sud et l'ouest déjà posé par la carte précédente. Les régions > BS sont encore toutefois à l'arrière du système, assez proches du "pont" évoqué plus haut.

Parmi les régions où le type réussit mal, on ne s'étonnera pas de retrouver l'Alsace et l'Est, la Bretagne, et, globalement le grand "Sud" (notamment méditerranéen). Certes, on trouvera toujours dans ces régions des toponymes en -AC, mais visiblement, le type rencontre ici le suivant, à ses dépens.

**-ANU** La distribution est véritablement méridionale, que l'on se base sur les moyennes ou sur le bornage positif.

On remarquera que la région PACA se révèle relativement fragile face à un Languedoc qui, visiblement, a concentré l'aménagement de type latin.

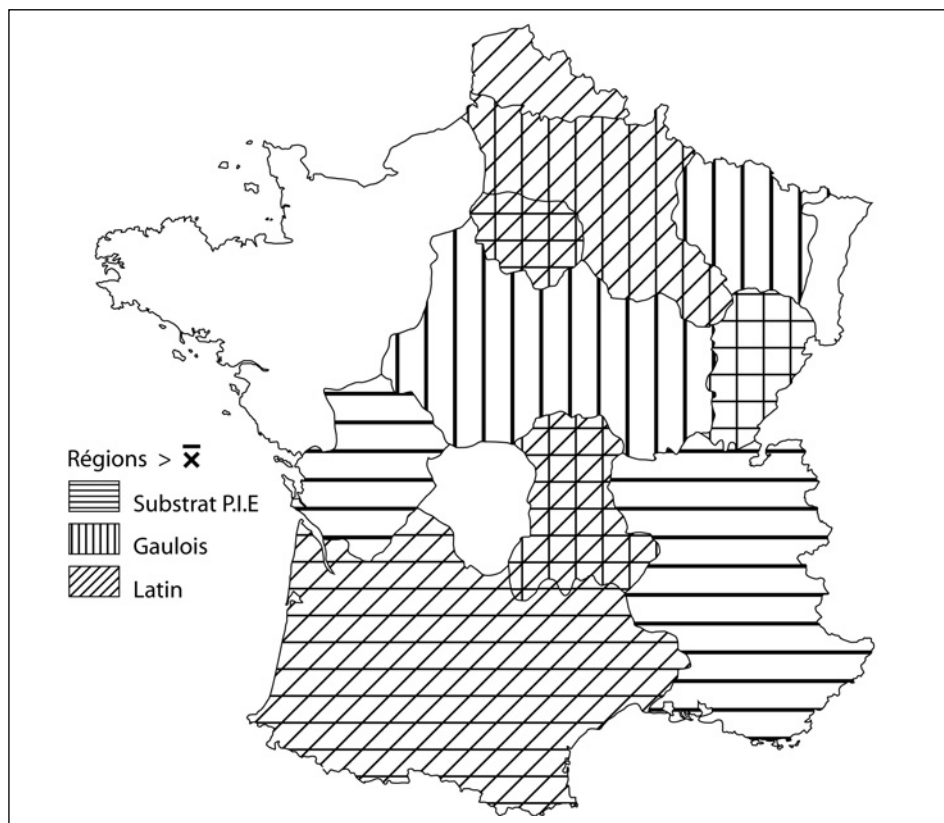


*Latin global*

Quand on cumule les différents types eu-latins exposés et le type –ANU(M), on découvre ou l'on redécouvre ce très beau phénomène de fractionnement du strat, entre extrême nord et extrême sud, tant dans la distribution par la moyenne que dans le bornage.

Les archéologues savent mieux que les linguistes à quel point le nord du pays a été une zone de colonisation intensive et de romanisation. Il y a une trentaine d'années déjà, les travaux de R. Agache, notamment fondés sur la photographie aérienne, avaient ainsi révélé un intense aménagement romain de la Picardie.

On observera à nouveau le comportement divergent de la Provence, bien moins "latine" que son voisin occidental sur le plan macrotoponymique. Mais si la Provence tire son nom de la "Provincia" (comme je l'ai rappelé plus haut), rappelons aussi que le Languedoc-Roussillon fut le siège historique de la Narbonnaise.

*Carte cumulative*

*Carte cumulative*

Cette dernière carte, établie à partir des moyennes arithmétiques, nous donne en quelque sorte le profil global de chaque région quant à sa position par rapport aux strats observés. Il faut comprendre qu'elle présente, ou combine plusieurs possibilités, ce qui va nous aider pour les lignes d'interprétation :

- aucune couche ne s'établit de manière significative, il n'y a aucune dominance statistiquement définie (blanc : Alsace, Bretagne, Normandie, Pays de la Loire, Limousin)
- le substrat PIE l'emporte et minimise les autres couches (Rhône-Alpes, PACA, Poitou-Charentes)
- dominance du gaulois (Bourgogne et Centre)
- dominance du latin (Nord-Pas-de-Calais)
- bi-dominance substrat PIE + gaulois (Franche-Comté)
- bi-dominance substrat PIE + latin (Languedoc-Roussillon, Midi-Pyrénées, Aquitaine)
- bi-dominance gaulois + latin (Champagne-Ardenne, Picardie)
- tri-dominance substrat PIE + gaulois + latin (Région parisienne, Auvergne)

On ne sera pas surpris de retrouver le grand Ouest sous forme de zone blanche, ainsi que l'Alsace. La présence de toponymies non romanes (superstrats breton, scandinave, germanique) peut permettre de comprendre une partie du phénomène. Mais l'argument est moins justifié de la Basse-Normandie aux pays de la Loire, en passant par la Haute-Bretagne. Pour autant qu'on sache, ces régions ont été celtisées anciennement puis romanisées à des degrés divers, mais apparemment sans que ces phénomènes ne laissent de marquage puissant<sup>17</sup>. L'une des hypothèses serait alors que le marquage de telles régions pourrait être postérieur (toponymie romane plus tardive, outre les apports récents de superstrat).

Des remarques analogues pourraient être faites à propos du Limousin, de statut étrange. Souvent rapproché de l'Auvergne voisine (cette idée vient assez facilement à l'esprit), on mesure en fait que ces deux voisins ont peu de choses à voir de notre point de vue. Notamment et de manière quelque peu surprenante la caractérisation par le substrat PIE n'est pas suffisante, un substrat qui pourtant devrait semble-t-il caractériser l'ensemble des latitudes du centre de la France, comme le suggère bien la carte.

Cette faiblesse du PIE, comme pour la Bretagne, finirait par faire penser que de telles zones auraient pu être en fait de véritables *no man's land* tant dans la pré-que dans la proto-histoire, pour des raisons qu'il conviendrait alors d'explicitier.

De manière claire, la zone de dominance gauloise vient buter sur le système méridional PIE, avec de bonnes incursions vers le Jura et l'Auvergne. Tout cela paraît relativement conforme aux enseignements de l'Histoire et de l'archéologie et à ce que nous avons vu plus haut.

Se vérifie également le dimorphisme polaire important de la latinisation toponymique, opposant clairement un sous-système méridional (Auvergne, Languedoc-Roussillon, Aquitaine) et un sous-système septentrional (Nord, Ardennes, Champagne, Bassin Parisien, Picardie), trace possible d'un aménagement latifundiaire de type colonial.

On peut ajouter que le grand Sud, qu'il soit par ailleurs plutôt PIE ou plutôt latin et croisé, se présente bien comme une franche zone de rejet du celtique ou du germanique.

Quant à la latinisation venue de la Narbonnaise, de la Provincia ou de la Lyonnaise, à l'évidence elle n'est parvenue qu'à un décapage partiel sans jamais parvenir à reconditionner un axe PIE qui va de la Charente à la Suisse et à l'Italie. Contrairement aux idées reçues, l'ensemble alpin en particulier, du Rhône-Alpes à la Provence est un véritable repoussoir des strats gaulois et latin : la romanisation n'a forcément porté ici que sur des secteurs par définition limités (contrairement au Languedoc ou aux régions septentrionales par exemple).

#### *Quelques remarques particulières pour finir*

La Franche-Comté maximise les composants PIE et gaulois. Elle semble donc se définir comme une plaque tournante, une intersection concrète entre l'aire gauloise et la grande zone alpine du PIE. Comme le révèlent différents travaux historiques et archéologiques, c'est dans ce secteur notamment que semblent s'être produites les infiltrations croissantes des Celtes vers 1200 av. J.-C. et par la suite (Champs d'urnes, Hallstatt, La Tène, en relation probable avec leurs positionnements antérieurs (Allemagne du Sud, Europe Centrale). C'est dans ce secteur encore, au contact de l'Alsace actuelle, que les alliances et affrontements entre tribus celtiques et germaniques (vers le milieu du 1<sup>er</sup> siècle av. J.-C.) provoqueront directement l'intervention de César (prise de Besançon, campagne contre Arioviste).

On remarquera pour finir que deux régions seulement sont interlopes, qui croisent et équilibrent tous les strats : Auvergne et Île de France. C'est suffisamment rare pour être souligné. Cela veut donc dire que ces deux régions fonctionnent de fait comme des "centres statistiques" où les strats observés s'équilibrent.

On peut, pour cette raison même, les approcher comme des centres représentatifs de l'ensemble du domaine gallo-roman. L'un plutôt rural représentant l'intersection anthropologique des éléments ayant préparé la "Gallo-romania", l'autre centré sur un noyau urbain fort, puisque Paris fut dès le Moyen Âge l'une des plus grandes villes d'Europe.

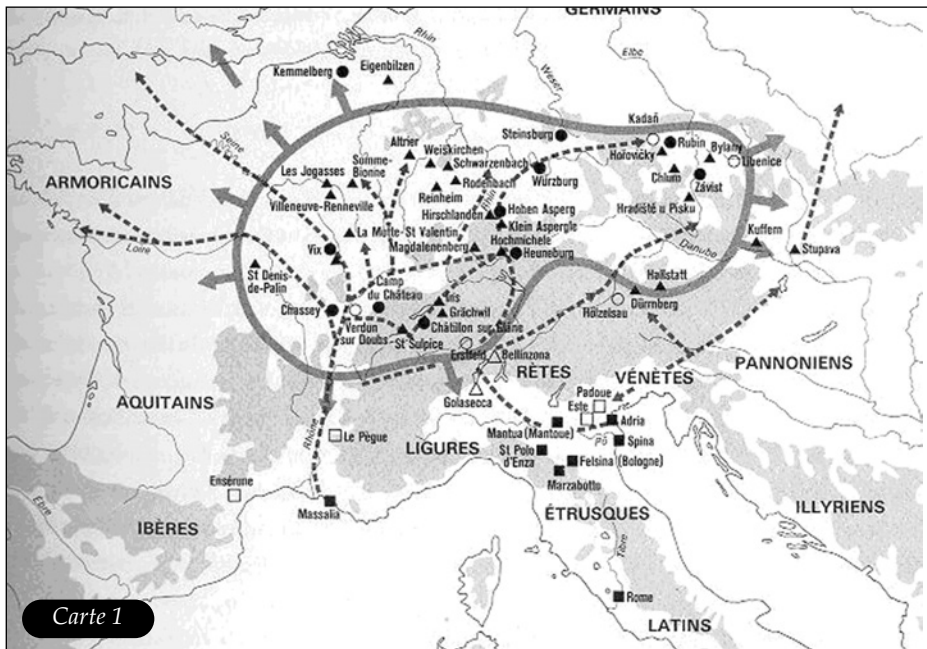
## 5. Données historiques et données toponymiques

Dans cette perspective, je terminerai par un commentaire rapide de trois cartes établies à partir de données historiques et archéologiques relatives à l'expansion des Celtes. Afin d'éviter de multiplier les cartes de ce genre, les trois cartes en question permettent de revenir globalement sur les pages précédentes et les quelques leçons toponymiques qui leur sont liées. On ne portera essentiellement qu'un commentaire sur les limites de l'expansion et l'articulation avec les marquages toponymiques.

### Carte 1 : expansion celtique vers le VI<sup>e</sup> / V<sup>e</sup> siècles av. J.-C.

La carte 1 est celle de la distribution des Celtes généralement admise vers les VI<sup>e</sup>-V<sup>e</sup> siècles av. J.-C.<sup>18</sup> Durant la première partie de l'âge du fer, on remarquera la difficulté déjà évoquée des Celtes à sortir de régions qui correspondent aujourd'hui dans les grandes lignes à l'Autriche, l'Allemagne du sud, la France de l'est. Par exemple, le lien avec notre carte de synthèse est manifeste et concrétise d'une certaine manière l'existence d'une base arrière eu-celtique que les expansions ultérieures ne suffiront pas à modifier fondamentalement.

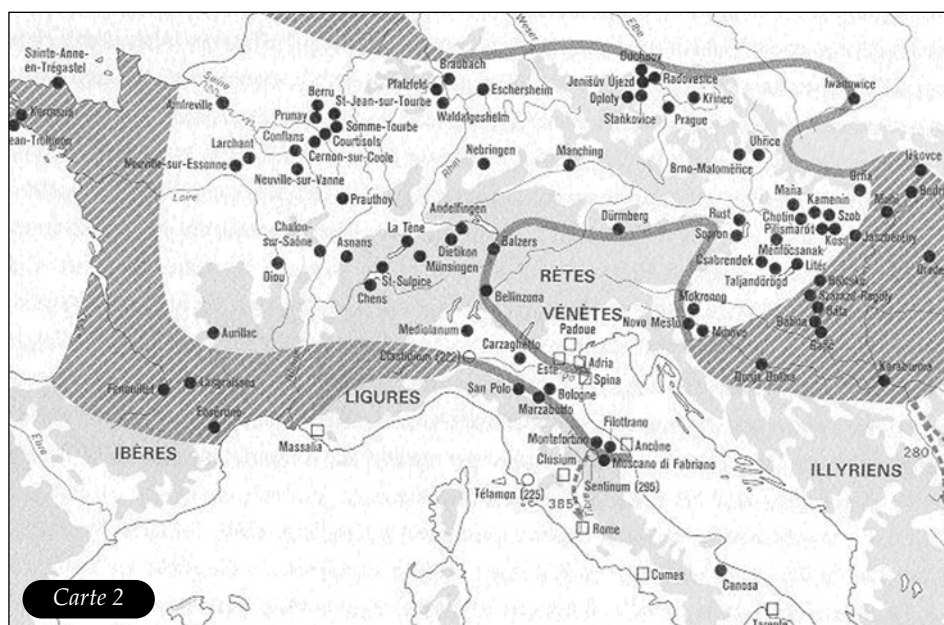
Deux phénomènes majeurs, plus haut récurrents, sont soulignés. C'est tout d'abord la très grande difficulté de progression vers le sud, notamment le massif alpin, qui n'est affecté que dans sa partie septentrionale et occidentale. C'est ensuite une difficulté de même nature vers l'ouest et le sud-ouest.



## Carte 2 : expansion celtique (second âge du fer, La Tène, IV<sup>e</sup>/ III<sup>e</sup> av. J.-C.)

Les données ultérieures (*carte 2*) confirment de telles zones de compression. La progression celtique piétine toujours en direction de l'ouest, de la partie méridionale des Alpes et des régions méditerranéennes. On pourrait dire, certes en simplifiant, qu'une ceinture reliant territoires armoricain, aquitain, ibérique<sup>19</sup>, ligure, rête, vénète, s'oppose constamment à toute pénétration d'envergure. Comme preuve de cela, la pénétration dans l'Italie du nord n'est en réalité qu'une relative trouée confirmant toujours la très haute résistance des Alpes méridionales et le l'Apennin, et glissant donc le long des régions méditerranéennes en direction de l'Adriatique.

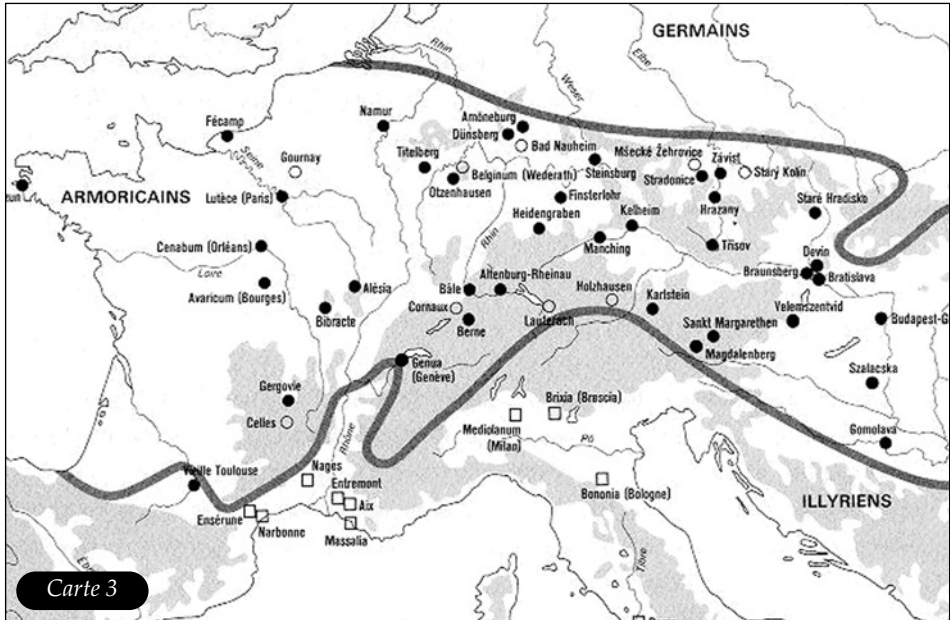
On soulignera à quel point serait significative la superposition entre cette carte et celle du substrat gaulois plus haut proposée.



## Carte 3 : expansion celtique avant l'intervention romaine (vers II<sup>e</sup> av. J.-C.)

Au moment où Rome va intervenir directement, la carte des sites significatifs révèle toujours la limitation potentielle des déplacements vers l'ouest et vers le sud. Pour la région observée, la base arrière du système se trouve toujours au nord des Alpes et dans la moitié orientale/septentrionale de la France actuelle. La France méditerranéenne, les Alpes du sud et l'Italie du nord restent fondamentalement hors d'atteinte.

Les conséquences d'un tel constat sont au moins de deux ordres. D'une part le développement attesté du peuplement celtique ancien vers l'ouest et le sud ne fut jamais assez puissant pour modifier en profondeur le paysage toponymique antérieur, alors même que l'expansion méditerranéenne (latine) allait commencer et donc renouveler le marquage toponymique du sud, comme nous l'avons vu. Il paraît donc finalement normal que la combinatoire macrotoponymique des régions méridionales de France implique prioritairement les substrats et la latinisation et non la couche gauloise.



Dans cette communication, je n'avais pas l'intention d'apporter de nouvelles données toponymiques. Je souhaitais plutôt montrer comment un traitement géographique et statistique de données généralement vues comme des masses historiques relativement compactes et intéressant les chercheurs surtout par leur dimension étymologique, pouvait conduire à une saisie dynamique et à l'éventuel repérage de mouvements anthropologiques derrière les cartes et les étymologies traditionnelles retenues. On prend conscience aussi, chemin faisant, des limites d'approches étymologiques trop ouvertes et imprécises par défaut d'alignement sur les données historiques et archéologiques. En cas de doute et d'étymologies multiples (chose très courante en onomastique ancienne) on comprendra par exemple que faire le choix d'étymologies celtiques pour la Provence ou pré-indo-européennes pour la Bretagne (par exemple) a plus de chances de conduire à de mauvais qu'à de bons choix. Une cartographie statistique d'ensemble, corrélée à une approche historique cohérente pourrait alors jouer le rôle de garde-fou et nous aider à assurer notre perception de mouvements anthropologiques anciens.



Au-delà des discussions possibles sur tel ou tel point on voit l'intérêt que pourrait présenter une cartographie systématique des données toponymiques de la France et du domaine gallo-roman. Dans cette perspective, l'utilité d'un *atlas toponymique* du domaine gallo-roman serait grande et, à terme, celle d'un atlas roman puisque les mouvements anthropologiques ne sont en réalité vraiment visibles que lorsqu'on porte un regard d'ensemble (comme on l'a constaté quelques lignes plus haut). Ce serait là un complément naturel aux études de dialectologie que beaucoup d'entre nous défendent et un miroir intéressant pour les historiens, les archéologues, les analystes du paysage.

## N O T E S

\* Abréviations :

**Oïl**, langue d'oïl

**Oc**, langue d'oc

**N-Oc**, nord-occitan

**Cat**, catalan

**Gasc**, gascon

**Fpr**, francoprovençal

**PIE**, pré indo-européen

**Var.** variante

Les formes actuelles sont données en minuscules, les étymons en majuscules. Les voyelles toniques sont soulignées dans les étymons en cas de besoin.

<sup>1</sup> Comme AQUIS, FIGULINAS etc. vus plus haut. Tous ces types, analysés ou cumulés, se retrouvent dans les légendes des cartes. Pour cet exemple, la carte "LATIN" cumule 9 types, lesquels peuvent évidemment être observés à part une autre fois, en donnant sans doute d'autres renseignements.

<sup>2</sup> Paris (Île-de-France), Champagne-Ardenne, Picardie, Nord-Pas-de-Calais, Lorraine, Alsace, Haute-Normandie, Basse-Normandie, Centre, Bourgogne, Franche-Comté, Pays de Loire, Bretagne, Poitou-Charentes, Aquitaine, Midi-Pyrénées, Limousin, Rhône-Alpes, Auvergne, Languedoc-Roussillon, Provence-Alpes-Côte d'Azur (PACA).

<sup>3</sup> Le type est par exemple très productif dans la microtoponymie pyrénéenne.

<sup>4</sup> La thèse d'un substrat de type basque, ou bascoïde, est relativement ancienne. Elle a été définitivement établie durant la seconde moitié du xx<sup>e</sup> siècle par les travaux de chercheurs catalans comme Joan Coromines et Henri Guiter.

<sup>5</sup> Ce type est parfois considéré comme indo-européen (rapprochement avec lat. NAVIS / NAVEM ou irl. NAU "navire").

<sup>6</sup> Le type est évidemment à la source du *Milan* italien, lequel s'inscrit, comme on le verra sur la fin, dans un couloir celtique de l'Italie du nord.

<sup>7</sup> Ce type reste relativement imprécis sur le plan sémique. Il semble toutefois avoir contribué à marquer des territoires, des termes et des zones de contact entre groupes celtiques. Le premier élément (gaul. EQUA) est régulièrement rapproché du latin ÆQUUS, ÆQUA "égal".

<sup>8</sup> La présentation du type est limitée. IALO peut se combiner naturellement avec des formants de la carte précédente, type *Nanteuil* (NANTO), *Neuil* (NOVIO), voire même avec certains des couches antérieures, type *Naveil* (NAVA).

<sup>9</sup> Comme de juste, on doit partir d'accusatifs dans la plupart des cas ici et par la suite.

<sup>10</sup> Se différencie de la Cité, du Municipie, de la Colonie etc.

<sup>11</sup> Laroque-de-Fa (Aude).

<sup>12</sup> Voir le parallélisme avec MANU > *main* en français.

<sup>13</sup> Par exemple en cartographiant ultérieurement l'ensemble de la toponymie d'origine latine.

<sup>14</sup> Ce sous-type existe aussi dans le sud méditerranéen, avec des formations du type *Belvianes*, *Naussanes*. On trouve par ailleurs très régulièrement au sud des féminins singuliers (*Bassane*, *Pélissane*). Rencontrant en partie les précédents, on rappellera l'existence d'un type -ANICAS (> *-argues*, *Meyrargues*, *Vauvenargues*) en Languedoc et en Provence, aire dont on trouve la trace jusqu'au massif Central.

<sup>15</sup> Révélant toujours probablement un profond décapage par le superstrat germanique.

<sup>16</sup> Provence-Alpes-Côte d'Azur.

<sup>17</sup> En outre, dans la partie la plus occidentale de ce « grand Ouest », l'influence des superstrats (de distributions plutôt côtières) est généralement très limitée, voire nulle.

<sup>18</sup> D'après l'Arbre Celtique ([www.arbre-celtique.com](http://www.arbre-celtique.com)).

<sup>19</sup> On rappellera que durant cette période, les Ibères sont fortement installés dans les zones (notamment littorales) du Languedoc-Roussillon.

#### O U V R A G E S É V O Q U É S

AGACHE, R., BRÉART, B., *Atlas d'Archéologie aérienne de Picardie* (2 vol. in plano).  
Société des Antiquaires de Picardie, 1975.

AGACHE, R., *La Somme préromaine et romaine d'après les prospections aériennes*.  
Société des Antiquaires de Picardie, 1978.

DAUZAT, A., ROSTAING, CH., *Dictionnaire étymologique des noms de lieux en France*.  
Larousse, différentes éditions.

FABRE, P., *Les Noms de lieux et de personnes*. Paris, Nathan, 1982, 278 pages.

NÈGRE, E., *Les noms de lieux en France*. D'Artrey, 1977.

NÈGRE, E., *Toponymie générale de la France*. Droz, 1991 (3 volumes).