



**HAL**  
open science

# MÉMOIRES DU TERRAIN : Enquêtes, matériaux, traitement des données.

Francis Manzano

► **To cite this version:**

Francis Manzano (Dir.). MÉMOIRES DU TERRAIN : Enquêtes, matériaux, traitement des données.. 2011-3, 2011, Nouvelle série, Francis Manzano, 978-2-36442-025-0. hal-02556909

**HAL Id: hal-02556909**

**<https://hal.science/hal-02556909v1>**

Submitted on 28 Apr 2020

**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.

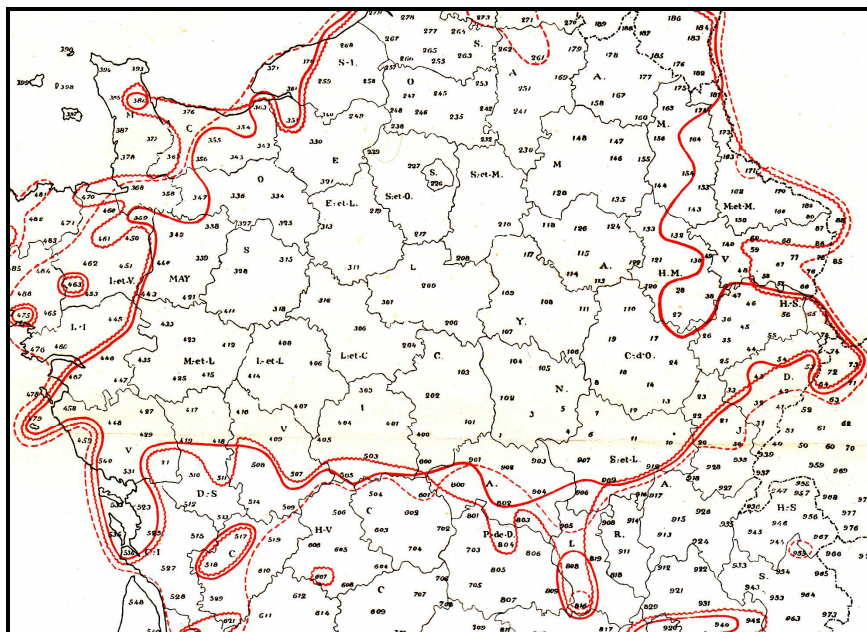
L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# MÉMOIRES DU TERRAIN

*Enquêtes, matériaux, traitement des données*



*Sous la direction de Francis MANZANO*  
*Ouvrage publié avec le concours de la Région Rhône-Alpes*

Rhône-Alpes Région

Publications du Centre d'Etudes Linguistiques  
[Linguistique, dialectologie]

*Nouvelle série 2011-3*

# MÉMOIRES DU TERRAIN

*Enquêtes, matériaux, traitement des données*

Actes du colloque de Lyon  
Centre d'Études Linguistiques  
12 et 13 mars 2009

*Publiés avec le concours de la Région Rhône-Alpes*



*La réalité des choses était autour de nous. Nous avons nous-mêmes parcouru les pays romans ; nous avons recueilli sur le terrain des observations ethnographiques et folkloriques. Comment en rester aux mots ? Tout en sauvegardant le caractère essentiellement linguistique de notre ouvrage, nous croyions devoir fournir à l'historien des mots les données nécessaires pour se faire une idée des choses, afin qu'il ne bâtisse pas dans le vide.*

*Karl Jaberg (1936)*

© Tous droits réservés Francis Manzano

*Les données et positions intellectuelles des chercheurs appartiennent à ces chercheurs comme au terrain dont proviennent ces données. Vous pouvez donc utiliser et citer librement ce document mais en ayant la courtoisie d'en citer les sources exactes.*

[francis.manzano@univ-lyon3.fr](mailto:francis.manzano@univ-lyon3.fr)

**MÉMOIRES DU TERRAIN**  
***Enquêtes, matériaux, traitement des données***  
(CEL, 2009-2011)

Table des matières

Jean Le Dù	
<i>Après la collecte, l'interprétation. Que faire des atlas linguistiques ?</i>	
	[pages 07 à 26]
Jean-Philippe Dalbera	
<i>La dialectologie. Objet, acquis, atouts et lignes de force de son devenir.</i>	
	[pages 27 à 42]
Robert Chaudenson	
<i>La dialectologie française ou « De la servitude volontaire »</i>	
	[pages 43 à 55]
Gérard Bodé	
<i>L'enquête Fortoul de 1856 : état des « idiomes » et « patois » locaux.</i>	
	[pages 57 à 66]
Jean-Paul Chauveau	
<i>Le français d'un rural manceau d'Ancien Régime</i>	
	[pages 67 à 95]
Fernand Carton	
<i>L'exploitation et la valorisation des données de terrain dans le domaine linguistique picard</i>	
	[pages 97 à 109]
Fabrice Jecic	
<i>Du texte dialectal aux cartes d'atlas : mémoires d'un même terrain beauceron. D'après l'Heritage perdu d'André Gilbert (1920-2003).</i>	
	[pages 111 à 127]
Fernande Krier	
<i>Les corpus de l'oral spontané</i>	
	[pages 129 à 137]
Guylaine Brun-Trigaud	
<i>Les chaînes de clitiques : traitements cartographiques</i>	
	[pages 139 à 155]
Aurélien Georges	
<i>Les chaînes de clitiques : l'outil informatique au service de l'analyse comparative</i>	
	[pages 157 à 177]

- Jean-Claude Bouvier  
*Le discours des carnets d'enquête des atlas linguistiques et ethnographiques : l'exemple du maïs et de la truffe dans la Drôme*  
 [pages 179 à 187]
- Jean-Louis Fossat  
*Des données de terrain à leur traitement : 50 ans de recherche*  
 [pages 189 à 239]
- Xavier Ravier  
*Au sujet de la carte ALG IV, 1115, Rouleau de foin*  
 [pages 241 à 254]
- Jean-Pierre Lai & Michel Contini  
*Nouvelles stratégies de recueil de données sur le terrain, exemples à travers le projet AMPER (Atlas Multimédia Prosodique de l'Espace Roman)*  
 [pages 255 à 268]
- Sabina Canobbio  
*Le Piémont des paysans*  
 [pages 269 à 286]
- Matteo Rivoira  
*L'Atlante Linguistico Italiano (ALI) et l'Atlante Toponomastico del Piemonte Montano (ATPM) : histoire, méthodes de récolte et de présentation des données, et perspectives de développement de deux entreprises géolinguistiques turinoises*  
 [pages 287 à 313]
- Melissa Barkat-Defradas & Élisabeth Motte-Florac  
*Dialectologie arabe et atlas linguistique. Les apports du multimédia.*  
 [pages 315 à 331]
- Gérard Taverdet  
*La fin de la dialectologie ?*  
 [pages 333 à 344]
- Francis Manzano  
*Du magasin dialectal de la France : une approche géodiachronique et sociolinguistique des terrains. Du français classique à la dialectologie.*  
 [pages 345 à 369]
- Francis Manzano  
*Postface*  
 [pages 371 à 380]

Page de couverture :  
 Karl Jaberg - *De la diphtongaison française en langue d'oïl*  
 (*Sprachgeographie*, 1908)

*Jean Le Dû*  
*Professeur émérite, Université de Brest*  
*jeanledu@gmail.com*

## **Après la collecte, l'interprétation. Que faire des atlas linguistiques ?**

Il sera surtout question dans mon propos de la géolinguistique en France. La majorité des auteurs d'atlas est maintenant âgée. La relève, si mes informations sont bonnes, est problématique. L'orientation théorique de la linguistique depuis des décennies n'est guère favorable à nos travaux. Mais les avons-nous faits assez connaître ? Existe-t-il même un ouvrage faisant le point sur la question ? Le terrain, naguère fécond, s'est transformé au point d'être méconnaissable : quand je suis entré à l'école en 1944, une grande partie de mes condisciples ne parlait que le breton. Aujourd'hui, en 2009, cette langue est inconnue des jeunes générations, qui ne conservent dans leur français aucune trace de la phonologie du parler de leurs ancêtres.

Problèmes scientifiques, effondrement du terrain : une mise à plat des enjeux est indispensable à ce stade de notre histoire.

### **1. Une discipline marginale**

La géolinguistique n'est plus guère enseignée dans les universités françaises, et n'y a jamais été vraiment très populaire. Combien d'étudiants en Lettres, quelle que soit la langue qu'ils étudient, ont entendu parler des atlas ? Déjà en 1922 Dauzat écrivait, à propos de l'*Atlas Linguistique de la France* :

« En France, en dehors d'un petit cercle de spécialistes, M. Gilliéron fut longtemps ignoré ; il dut faire éditer à ses frais plusieurs de ses ouvrages, dont la première édition porte la forme d'une infime librairie d'une petite ville suisse où ils figuraient sur les rayons entre une bible et un manuel de jardinage ».

La géolinguistique française, il faut le souligner, s'est toujours développée à contre-courant des tendances historiques. L'ALF a été conçu dans le cadre de la revanche. Le premier souci des linguistes qui l'ont inspiré – Gaston Paris et Paul Meyer – était clairement d'ordre idéologique. Niant la coupure entre France d'oïl et France d'oc, Gaston Paris affirmait, s'opposant en cela à la science allemande « qu'aucune limite réelle ne sépare les Français du nord de ceux du Midi et que d'un bout à l'autre du sol national nos parlers populaires étendent une vaste tapisserie dont les couleurs variées se fondent sur tous les points en nuances insensiblement dégradées »... Notons que son souci étant la France romane, il ne mentionnait même pas les alloglottes : Bretons, Basques, Flamands, Alsaciens, Lorrains...

Cet aveuglement volontaire devant la réelle diversité linguistique de la France cadrait avec la vaste campagne de construction d'une France unie tendue vers la reconquête de l'Alsace et de la Lorraine. Chaque enfant apprenait « l'histoire de France » et la langue nationale unique : les temps ne pouvaient être plus défavorables à l'exaltation — ni même à la reconnaissance — de la variété linguistique du pays.

Dauzat avait annoncé dès 1939 son projet atlantographique dans deux articles du *Français moderne* intitulés successivement « Un nouvel atlas linguistique de la France », puis « Le nouvel Atlas linguistique de la France ». En 1941, il parlait de « L'Atlas linguistique régional de France » pour enfin adopter, en 1942, le titre définitif « Le nouvel Atlas linguistique de la France par régions ». Le gouvernement de Vichy, avec sa politique de retour à la terre dont Dauzat était un ferme partisan, encourageait les études régionales (qui ont abouti, entre autres, aux publications issues du chantier 1425 sur l'habitat rural). Mais c'est après la Libération, dans le contexte à nouveau centralisateur de la reconstruction nationale, que l'entreprise a vraiment commencé à décoller. Au prix de combien d'efforts ? Espérons que l'histoire en sera écrite un jour, mais on connaît le rôle majeur de fortes personnalités comme Mgr Gardette et Jean Séguy.

Plus près de nous, on aurait logiquement pu penser que le retour à la régionalisation, avec l'élection du Conseil régional au suffrage universel direct en 1986, aurait favorisé les atlas linguistiques. Tout porte à penser que si le CNRS a dissous le groupe de recherche des atlas en 1996, c'est en prévision des multiples transferts de l'État aux Régions qui se sont multipliés. Certaines régions ont repris le flambeau, comme l'assemblée territoriale corse, qui a soutenu son

atlas dès 1986 ou encore la région Bretagne, qui a permis la publication de mon *Nouvel Atlas Linguistique de la Basse-Bretagne* (Le Dû, 2001). Il apparaît cependant que la plupart du temps les autorités régionales, en recherche de légitimité, ont préféré mettre l'accent sur la « sauvegarde », voire la « renaissance » des « langues régionales ». Comme dans les pays voisins, par exemple en Espagne, chacune de ces nouvelles entités politiques a cherché à accentuer son particularisme : érection de panneaux routiers bilingues, soutien aux écoles en langues régionales comme les Ikastolak basques, les écoles Diwan bretonnes ou les Calandretas occitanes. On parle maintenant, en langage politiquement correct, des langues picarde, normande, bourguignonne etc., plaçant ainsi implicitement ce qu'on appelait avec condescendance des « patois » sur un pied d'égalité avec ce que le langage courant appelle « langues », c'est-à-dire des entités normées et non plus des nuages de parlers apparentés.

En dépit de cette histoire tortueuse, nous nous trouvons aujourd'hui à la tête d'un corpus immense, allant de l'ALF à une collection de plus de 70 atlas régionaux, auxquels vont s'ajouter certains autres encore inédits. D'autres ont été publiés en dehors de la collection : Roussillon (Guiter 1966), Basse-Bretagne (Rolland 1994) etc., tandis que des projets inachevés, comme les atlas des côtes de l'Atlantique et de la Méditerranée, pourraient un jour être relancés.

## 2. Définir la discipline

En France, traditionnellement, on considère la géographie linguistique comme une sous-branche de la dialectologie, elle-même regardée de haut par les linguistes théoriciens. Aux États-Unis, tout ce qui concerne la variation spatiale du langage est classé sous l'étiquette de « sociolinguistique » (il suffit de consulter le site <http://linguistlist.org/> pour le vérifier). Comme tout ce qui vient d'Amérique est parole d'évangile pour nos autorités académiques, nos intérêts paraissent, encore une fois, en porte à faux avec la mode du moment. Pourtant, en y regardant de plus près, on constate que les États-Unis ont de nombreux chantiers atlantographiques en cours (<http://us.english.uga.edu/>) : l'impressionnant *Atlas of North American English* (Labov 2005) devrait suffire à démontrer la modernité de notre discipline !



Dialectologie et géolinguistique fonctionnent très bien dans plusieurs pays d'Europe en tant que disciplines autonomes. Elles sont complémentaires, mais parfaitement distinctes, comme l'indique le nom de la « Société Internationale de Dialectologie et Géolinguistique » et celui de sa revue *Dialectologia et Geolinguistica*, reconnue dans le cadre de ERIH (European Reference Index for the Humanities).

### *Un simple recueil de données synchroniques ?*

On a beaucoup reproché aux géolinguistes d'en être restés au stade de la simple collecte. William Labov écrivait en 1976 :

« Dans la seconde moitié de XIX<sup>ème</sup> siècle, les diachronistes se montraient ouverts aux apports de la dialectologie... Mais, au cours de ce siècle, la dialectologie en tant que discipline paraît avoir perdu tout contact avec la linguistique théorique, et, désormais, les tenants de la géographie linguistique se contentent le plus souvent de recueillir leurs matériaux et de les publier ». (Labov, 1976, p. 362).

Et pourtant Gilliéron lui-même dès 1902 envisageait de :

« ...détailler les aires lexicologiques, phonétiques, morphologiques et syntactiques, au point que le linguiste vienne avec confiance interroger leurs contours et le procès-verbal des conditions où ces aires meurent ou naissent, se dilatent ou se contractent... ». (Gilliéron, 1902, p. 3).

Les pays de langue allemande ont très tôt commencé à interpréter les atlas, mais, en France, une partie de la communauté des linguistes s'est toujours montrée hostile à une utilisation dépassant la description synchronique ou la sauvegarde d'un patrimoine en péril. Dans le domaine breton, François Falc'hun (1963) a été amené, par la seule analyse des 400 premières cartes de l'*Atlas Linguistique de la Basse-Bretagne* de Pierre Le Roux (1924-1963), à soupçonner la prédominance de la ville de Carhaix (l'ancienne Vorgium) sur le territoire de l'actuelle Basse-Bretagne dans l'Antiquité et au Haut Moyen Âge, ce que les archéologues ont démontré depuis. Il a même osé – crime majeur au regard des celtisants et de l'opinion régnante – émettre l'hypothèse d'une probable survivance du gaulois au moment de l'immigration bretonne en Armorique à partir du IV<sup>ème</sup> siècle de notre ère (Loth 1883). Dans son monumental *A Historical Phonology*

of Breton publié à Dublin en 1967, le grand celtisant anglais Kenneth Jackson, après avoir dédié son ouvrage « *to Canon F. Falc'hun whose brilliant researches on Breton phonology have contributed so much to our understanding of the language* », consacre beaucoup de pages à la réfutation de sa méthode.

« À notre connaissance, la source la plus importante sur les dialectes bretons est évidemment le grand *Atlas Linguistique de la Basse-Bretagne* (...) de P. Le Roux, (...). Il se compose de six cents cartes de la Basse-Bretagne sur lesquelles sont transcrites en alphabet phonétique les différentes prononciations d'un mot ou d'une expression telles qu'elles ont été recueillies par M. Le Roux lui-même des lèvres de ses informateurs dans 77 localités de toute l'aire en question. Cette production splendide est une source de valeur inestimable pour l'étude des dialectes actuels, mais quant aux spéculations concernant leur histoire lointaine – ce qui n'a jamais été son objectif – on devrait l'utiliser avec parcimonie et avec la plus grande précaution ». (Jackson, 1967, p. 5).

Et plus bas :

« 'La géographie linguistique' peut nous dire ce qu'était l'état d'une langue au moment où les sources qu'elle utilise ont été rassemblées (ou plutôt ce qu'était l'état de la langue tel que représenté par certains informateurs), et, jusqu'à un certain point, elle peut bien suggérer la direction possible et l'étendue de changements linguistiques, mais elle ne peut rien nous dire directement sur les dates réelles de production de ces changements ». (Jackson, 1967, p. 6)<sup>1</sup>.

Falc'hun n'a jamais pensé pouvoir atteindre à la datation absolue à partir de la seule étude de cartes d'atlas, mais Jackson, ébranlé dans

---

<sup>1</sup> The most important source for what we know about the Breton dialects is of course the great *Atlas Linguistique de la Basse-Bretagne* (abbreviated ALBB) by P. Le Roux, six parts (Rennes, 1924, 1927, 1937, 1943, 1953, 1963). It consists of six hundred maps of Breton-speaking Brittany, on each of which the pronunciation of some word or phrase is entered in phonetic script as collected by M. Le Roux himself from the lips of his informants for 77 places throughout the area in question. This splendid production is a source of inestimable value for the study of the present-day dialects, but a basis for speculation about their long-past history – for which purpose it was never intended – it should be used sparingly and with the greatest caution.  
[.] 'Linguistic geography' can tell us what the state of a language was at the time when the sources it uses were compiled (or rather, what the state of the language was as represented by certain informants), and up to a point it may well be able to hint at the possible direction and spread of linguistic changes; but it can tell us nothing at all directly about the actual dates when these changes arose.

ses certitudes par les conclusions de son collègue, lui attribue cette prétention en toute mauvaise foi.

### *Le piège d'une vision organiciste de la langue*

Nos manuels nous présentent l'histoire des langues comme linéaire. Comme Benjamin Button, elles naissent très âgées, puis rajeunissent progressivement en franchissant diverses étapes temporelles. Le vieil anglais devient le moyen anglais, puis l'anglais moderne et enfin l'anglais contemporain. On connaît la célèbre citation de Schleicher :

« Les langues sont des organismes naturels qui, en dehors de la volonté humaine et suivant des lois déterminées, naissent, croissent, se développent, vieillissent et meurent ; elles manifestent donc, elles aussi, cette série de phénomènes qu'on comprend habituellement sous le nom de vie. La glottique ou science du langage est par suite une science naturelle ; sa méthode est d'une manière générale la même que celle des autres sciences naturelles ». (Schleicher, 1868, p. 3.)

Bien que cette idée eût très vite été rejetée par la plupart des spécialistes, elle reste dominante dans le grand public, et certains linguistes n'hésitent pas aujourd'hui encore à recourir à cette métaphore en parlant de « mort des langues ».

Pourtant, comme l'écrit Alinei (1996, p. 161) dans son *Origini delle lingue d'Europa* : « Il est évident que les langues n'ont pas de sexe, qu'elles ne s'accouplent pas et n'accouchent pas d'autres langues ». Comme il ne s'agit pas d'être vivants « la tendance des langues, dans des circonstances normales, est la conservation, non pas le changement. Il n'existe pas de 'seuil' chronologique au-delà duquel l'horloge moléculaire se met à faire tic-tac, et à provoquer le changement linguistique par une loi inexorable de type évolutif-organique. » ... Il ajoute « Les changements ... sont dus à des causes extralinguistiques, non organiques »<sup>2</sup>.

---

<sup>2</sup> « la tendenza delle lingue, in circostanze normali, è di conservarsi, non di mutare. Non vi è neanche una 'soglia' cronologica, al di là della quale l'orologio molecolare comincia a ticchettare, e a far mutare una lingua per un' inesorabile legge di tipo evolutivo-organico. » ... « I mutamenti... sono dovuti a cause extralinguistiche, non organiche ».

Ce n'est pas le résultat de lois internes, mais la conquête de l'Angleterre par Guillaume le Conquérant qui explique le passage du vieil au moyen anglais. Si Chaucer utilisait une foule de mots français, en était-il de même pour les paysans de son époque ? On peut en douter, quand on entend, aujourd'hui encore, des anglophones exiger de leur interlocuteur qu'il parle clairement en employant des « one syllable words », des mots d'une syllabe, c'est-à-dire les vieux mots saxons... La langue de Chaucer est une étape dans l'évolution de l'anglais écrit, certainement pas de la langue parlée par les masses paysannes.

### *Les atlas comme instrument d'investigation du passé*

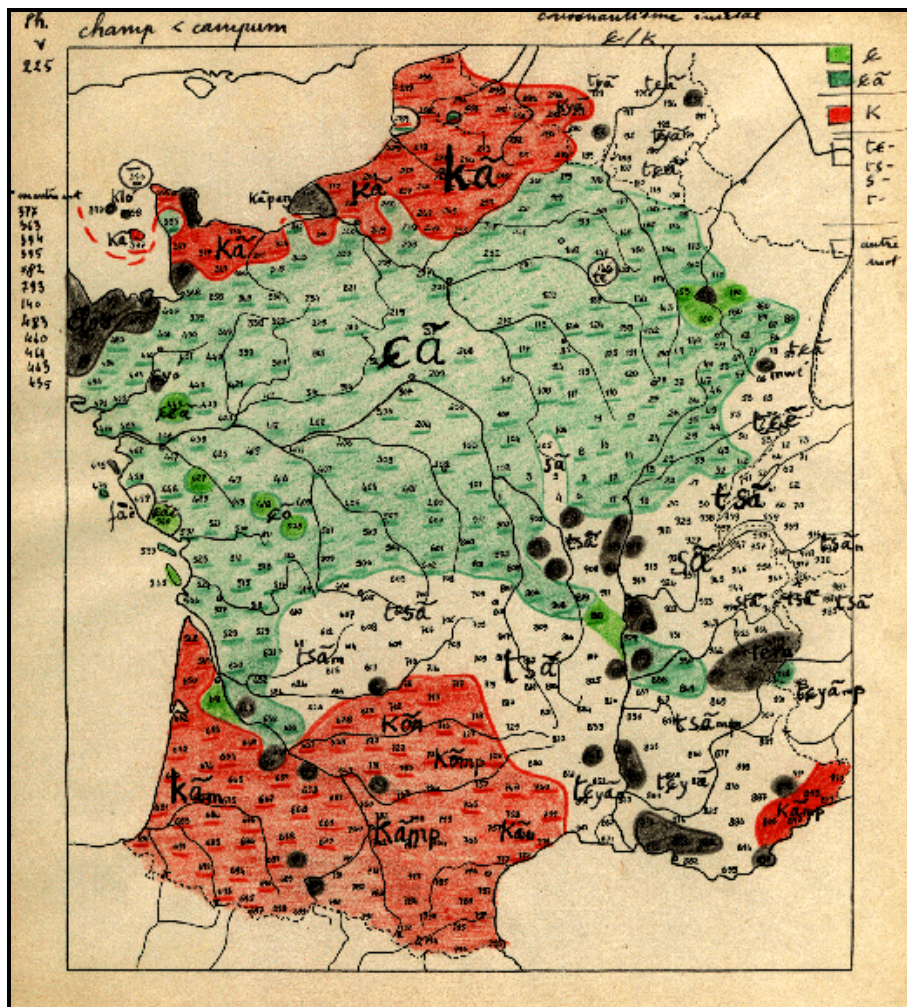


Figure 1 : carte champ de F. Falc'hun, originale et inédite, d'après l'Atlas Linguistique de la France.

On nous présente l'évolution du latin **campum** au français **champ** par un ensemble de stades successifs, un peu comme ceux qui mènent de la larve à la chenille pour aboutir à l'insecte.

Latin **campum** > **camp** > **champ** > **ʃã**

L'examen de cette carte nous apprend que divers stades considérés comme se succédant dans le temps sont en fait contemporains ! Conservation du **k-** initial dans le sud-ouest, le sud-est et le nord. Transformation de ce **k-** en affriquée **tʃ-** dans le sud et le nord-est, en un simple **ʃ-** au centre, ainsi que la présence d'une zone transversale comportant une prononciation **ts-** réduite par endroits à **s-**. Conservation du groupe **-mp**, dans le sud, réduit à **-m** par endroits, et même chute de ce **-m** avec nasalisation de la voyelle précédente.

Si le français normatif dit **ʃã**, c'est que la langue nationale s'est formée au sein d'une région aux habitudes articulatoires très différentes de celles des autres territoires de langues romanes. La succession des faits est à la fois géographique et temporelle. On ne voit pas pourquoi une « loi » phonétique comme la transformation de **k-** en **ʃ-** serait inscrite dans les « gènes » de la langue ! Ce qu'on observe, c'est une succession de phénomènes qu'Alinei appelle hybridation, qui fait que, par exemple, le mot anglais *thriller* devient instantanément *srilleure* en français. L'examen des cartes nous permet de reconstituer les aires dialectales forgées sur la très longue durée par la géographie : opposition plaines — montagnes, aires isolées — points de passage, villes — campagnes etc. et de retracer ainsi l'évolution millénaire des parlers. Le passage de *campum* à *champ* est l'aboutissement d'un long voyage dont seule la géolinguistique permet de reconstituer le trajet et les modalités grâce aux connaissances acquises par la linguistique depuis plus d'un siècle, qui nous permettent d'affirmer que **tʃ-** succède à **k-** et non l'inverse. Loin d'être une annexe peu fiable à l'étude des textes, « La linguistique historique débouche donc sur la géolinguistique, et trouve en elle, non seulement un auxiliaire indispensable et une nouvelle méthodologie, mais aussi la forme la plus accomplie de ses propres explications... nous pouvons dire... que l'interprétation historique de la carte géolinguistique représente l'équivalent moderne et concret tant de la

figure de l'arbre que de celle de l'onde » conclut Alinei (1996, p. 199-200)<sup>3</sup>. Les langues nationales, bien que prestigieuses en regard des humbles patois, sont récentes au regard de l'histoire. Liées à l'écrit et aux classes sociales dominantes, leur étude appartient au domaine de la sociolinguistique.

### *Des parlers ruraux*

Ce sont en effet les parlers ruraux, que la tradition française nomme « patois » et qu'en Basse-Bretagne nous appelons « badumes » 'langues de chez nous', qui constituent la matière première de la géolinguistique. Terrain immense, puisqu'il couvre l'ensemble des campagnes. Son déclin a commencé au moment de l'irruption en son sein de l'économie de marché qui a provoqué ce que Weber appelle « La fin des terroirs » et des parlers qui en étaient l'expression.

Qu'est-ce qui en constitue l'originalité ? Selon Gaston Tuaille (1992) :

« Un dialectophone est quelqu'un qui a appris une langue unitaire, le patois de son village »... Et plus loin : « Le monde dialectal est fait d'une multitude de langues unitaires, qui fonctionnent comme telles, au sein de chacune de ces communautés. »

Ce qui caractérise ces sociétés rurales, c'est qu'elles semblent immuables tant les changements sociaux et économiques qui se produisent en leur sein sont lents. Le conservatisme des patois est tel qu'on doit pouvoir sans vergogne comparer des données collectées à des dates différentes. Nous avons probablement tous relevé au cours de nos enquêtes des archaïsmes qui avaient échappé à nos prédécesseurs. Je ne crois pas qu'une distance d'un demi-siècle change grand-chose aux parlers populaires, dont l'évolution ne devient perceptible qu'au moment de leur déclin. Mon expérience bretonne me dit que le seul changement observable est l'abandon

---

<sup>3</sup> La linguistica storica sbocca dunque nella geolinguistica, e trova in questa non solo un ausilio indispensabile e una nuova metodologia, ma anche la forma più compiuta delle proprie spiegazioni. Anche se dal quadro manca ancora la preistoria, possiamo dire fin da ora che l'interpretazione storica della mappa geolinguistica rappresenta l'equivalente moderno e concreto sia della figura dell'albero che di quella dell'onda ».

brutal, pur et simple de la langue, les « terminal speakers » n'étant, en réalité, que des locuteurs occasionnels.

## 2. Que nous enseigne la géolinguistique ?

Nous avons évoqué l'importance capitale de la géolinguistique pour la linguistique historique. La linguistique générale peut aussi en tirer des enseignements. La comparaison entre les parlers d'aires contiguës montre par exemple la justesse de l'analyse d'un phonème en traits pertinents, puisqu'une variation peut n'affecter qu'un seul d'entre eux tout en préservant les autres. Dans la lénition, par exemple, des consonnes franchissent une étape entre occlusion et spirantisation (**b** > **v**) ou entre sourdité et sonorité (**k** > **g**) en passant de l'occitan aux parlers d'oïl.

### *La définition des langues*

La réflexion sur les frontières dialectales entre directement dans le champ de la géolinguistique. Les mots voyagent sans tenir compte des limites linguistiques supposées. Où s'arrête le territoire de ce qu'on appelle une langue, où commence l'autre ? Si nous ne définissons pas clairement cette notion, nous ne pouvons que déboucher sur une aporie.

Les termes « breton » ou « occitan » aujourd'hui sont des notions liées à l'écriture, donc à la politique. Les bretonnants naïfs admirent ces gens de la radio qui savent le « vrai » breton, tout en avouant ne pas bien comprendre ce qu'ils disent. Les communes belges de Eupen et de Fourons ont comme parler hérité le francique carolingien dit Plattdüsch. Pour des raisons historiques — l'Église avait adopté l'allemand dans l'une, le néerlandais dans l'autre —, on l'a considéré comme une variété d'allemand à Eupen et de néerlandais à Fourons, avec pour résultat concret... l'extinction de la pratique du parler ancestral ! On se heurte sans cesse à la fluctuation des définitions de la « langue », et notre discipline peut aider à approfondir la réflexion.

## ***Ethnologie***

Les atlas qui s'intitulent « Atlas Linguistique et ethnographique » comportent des dessins et des descriptions d'objets ou de pratiques désuètes. Les lents mouvements des isoglosses sont un écho des transformations techniques, des déplacements des centres de pouvoir. Mais les données purement linguistiques permettent à elles seules de reconstituer un passé oublié : jusqu'à l'introduction de l'euro, la langue bretonne n'avait pas encore adopté le franc ! Elle comptait pêle-mêle en sous, réaux, écus, ou livres. Quand des populations abandonnent leur langue au profit d'une autre, des traits anciens peuvent refaire surface dans le nouvel environnement, comme en français le retour du calcul à base vingt (dans quatre-vingt). On pourrait expliquer ainsi nombre de particularités du français teinté de gaulois ou de l'anglais marqué par le celtique insulaire sous-jacent.

## **4. Conclusion : que faire des atlas ?**

La géolinguistique, nous l'avons dit, n'a pas disparu avec la fin du groupe des atlas au sein du CNRS. Les projets internationaux se poursuivent, comme l'*Atlas Linguarum Europae*, et l'*Atlas linguistique Roman* dans lequel Grenoble joue un rôle de premier plan. Le *Thésaurus Occitan* rassemble à Nice les ressources écrites et sonores, toutes d'origine orale, et y met au point des outils d'analyse qui permettront à terme des représentations cartographiques et des instruments d'analyse variés.

\*

Je n'ai évidemment pas la prétention de faire ici un inventaire global de toutes les entreprises.

De nombreux travaux, restés inédits, mériteraient d'être repris et utilisés. Je donnerai comme premier exemple la thèse d'Alain Le Berre, *Ichtyonymie bretonne*, qu'un groupe de chercheurs a cartographiée et publiée sous forme d'un CD-Rom. Il paraît difficile d'exploiter des données figurant sous forme de listes, tandis que leur mise en cartes ouvre la voie aux interprétations.



XII. VERF. AVES	Numerius arquata suite		R 492 sq.
K 2	§eo:leg-'mour	ms §eole:ged-'mour p	CHEOLEG-MOR, CHEOLECED-
K 4	§eo:leg-'mu:r	ms - 'mu:r	-id-
K 6	tro:her-'liP	ms tro:her-'lipa p	TROHER-LIP, TROHER-LIPOU
K 8	er gy'li:en	fs ky'li:et	p KULIENN, KULIED
K 10,16,39	or hefe:leg 'ru:S	ms kefele:ged'ru:S	KEFELEG-ROUZ, -ED R.
K 12,14,41	er hefe:leg-'mu:r	ms kefele:ged-'mu:r	KEFELEG-MOR, -ED-M.
K 17 à 20	er heve:leg'ru:S	ms kevele:ged'ru:S	KEVELEG ROUZ, -ED R.
K 20-B,21,22	-id-		-id-
K 23 à 29,31,32	-id- -o:T	ms	KEVELEG-AOD, -ED-AOD
K 30	keve:loK	ms kevele:geT	p KEVELEG, KEVELEUGEUD
K 33,34,37	kefe:leg-'mu:r	ms kefele:ged-'mu:r	KEFELEG-MOR, -ED-M.
K 34 al.35,39,40,46,46-B	kefele:eg-'o:T	-s:ged-	KEFELEG -AOD, -ED-AOD
K 47,48,50	kefe:leg-o:T	ms kefele:ged-'o:T	p KEVELEG-AOD, -ED-AOD
K 47-B	kefe:leg-'mo:r	ms kefele:ged-'mo:r	p KEVELEG-MOR, -ED-MOR
K 49	er keveleg'o:t	ms -le:ged-'o:T	p KEVELEG-AOD, -ED-AOD
L 3	kefe:loK	ms kefele:jen	p KEFELOG, KEFELEYEN
L 4	keveleg-'soT	ms kevele:ged-'soT	p KEVELEG-AOD, -ED-AOD
L 6 à 14,22,25,26	keve:leg-'mo:r	ms -le:ged-	KEVELEG-MOR, -ED-M.
L 18,19	keve:log-'soT	ms kevele:jen-'soT	p KEVELEG-AOD, -LEYEN-AOD
L 20,21,23,24	-id-	kevele:ged'soT	p - , -LEGED-AOD
L 27,28,28-B	kefe:leg-'mo:r	ms kefele:ged-m.	KEVELEG-MOR, -ED-M.
L 29	kefe:leg-soT	ms kevele:ged-soT	p KEVELEGED-AOD, -ED-AOD
L 31	kefe:leg-soT	ms kefele:ged-soT	p KEFELEGED-AOD, -ED-AOD
L 33	kefe:leK	ms kefele:geT	p KEVELEG, -ED
			("MOR, pas la peine de le dire, on le sait")
L 35	keve:leK	ms kevele:geT	p KEVELEG, KEVELEGED
T 1,2,13	kefe:leg-'mo:r	ms kefele:ged-'mo:r	KEFELEG-MOR, -ED-M.
T 8,9,15,16,17	er heve:leK	ms kevele:ged	KEVELEG, KEVELEGED
T 19	er hevalleK	ms 'ke:vallegeT	p KEVELEUG, -EUD
-1464 (QUEFFELC begace)	-1659 QUEFFELC-MOR	-1826 CORRIGRAU	Houat
-1871 KE(VE)LOK	-1931 KEFELLEKA		
-p.o. on a KEVELEG-KAOD, KEVELEG-LANN, "Bécasse des bois "			

Fig. 2 Liste des termes désignant en breton *Numeria arquata* « le courlis cendré ».

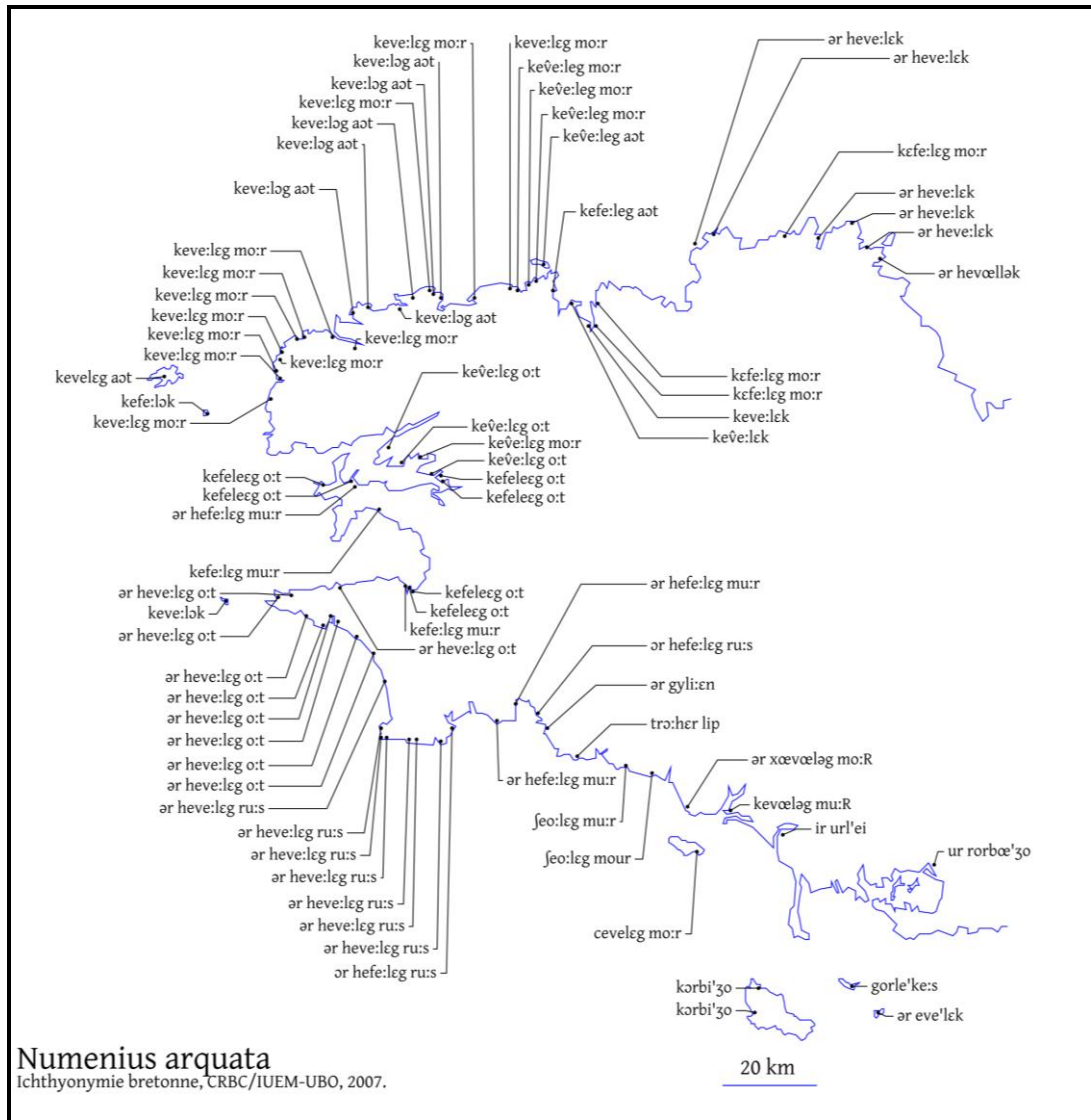


Fig. 3 Carte réalisée à partir des données de la figure 2.

Certains des atlas régionaux de la France ont été publiés sous la forme d'atlas de deuxième génération comportant des cartes interprétatives à aires dégagées. Les progrès techniques nous permettent désormais de conserver les cartes traditionnelles sur lesquelles figurent les données brutes tout en y ajoutant une ou plusieurs interprétations.

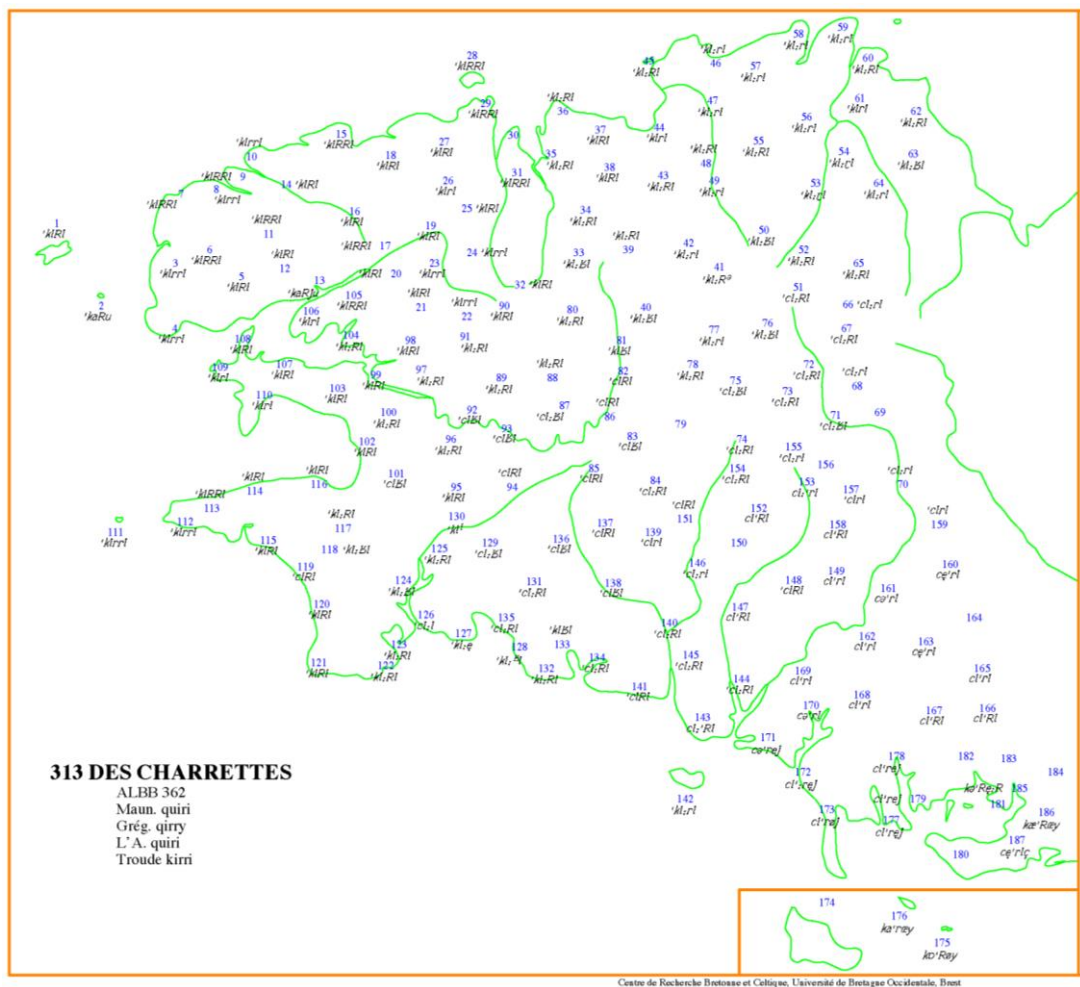
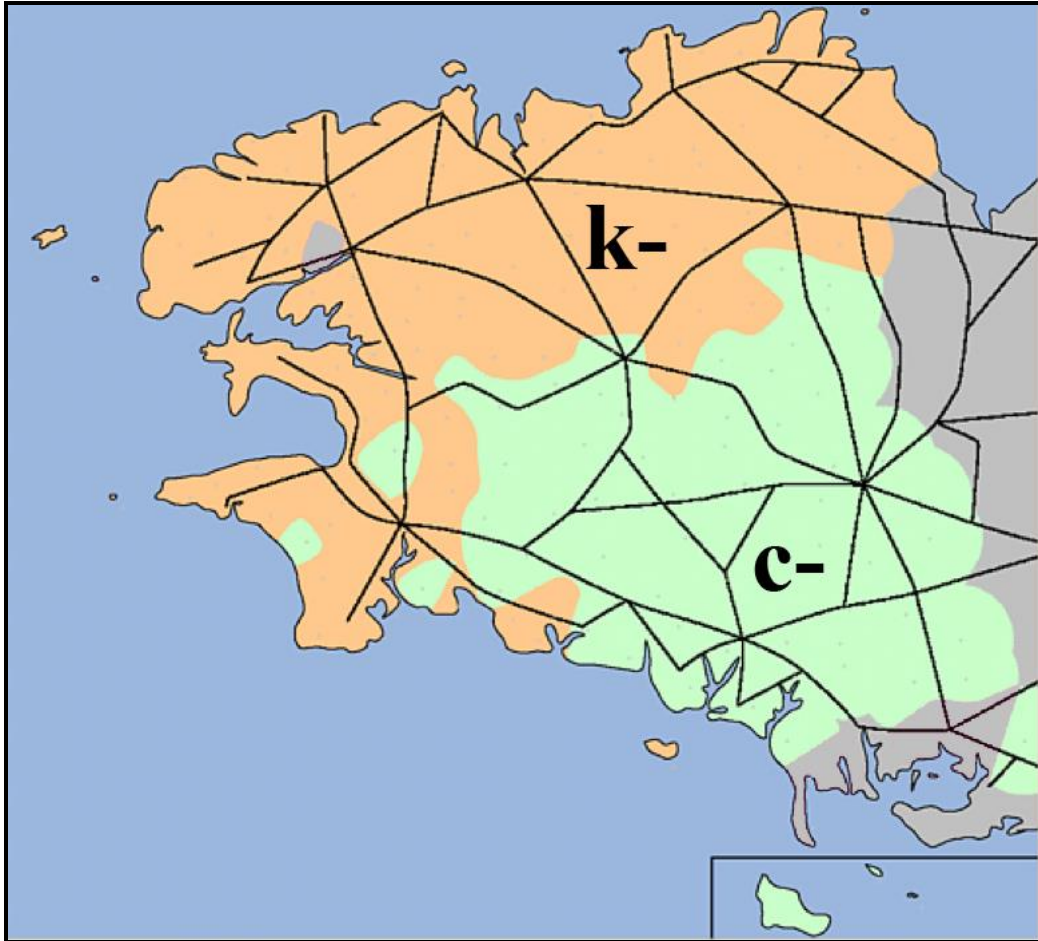


Fig. 4. Carte du Nouvel Atlas Linguistique de la Basse-Bretagne : « des charrettes ».



*Fig. 5. Interprétation assistée par l'informatique d'un trait de la carte « des charrettes » : l'initiale de kirri est vélaire au nord, palatale ou affriquée au sud.*

Cette répartition pourrait nous ramener à la bipartition originelle de la langue bretonne, marquée par l'influence insulaire au nord, tandis que le sud conserve la mouillure qui pourrait être similaire au substrat à l'origine de la palatisation du français d'oïl (le réseau tracé sur la carte représente les routes).

L'Atlas Linguistique des Petites Antilles (ALPA) en voie de réalisation est composé grâce à un logiciel qui permet à la fois de tracer les cartes et de les interpréter de diverses façons.

La carte qui suit présente le démonstratif postposé créole sous ses différentes formes. Son analyse est facilitée par une représentation utilisant des symboles de couleur. Un logiciel ne remplace pas la réflexion, mais il la facilite, en permettant une vérification rapide d'hypothèses, vérification qui aurait naguère demandé une somme de travail énorme, ou n'aurait même pas été envisageable.

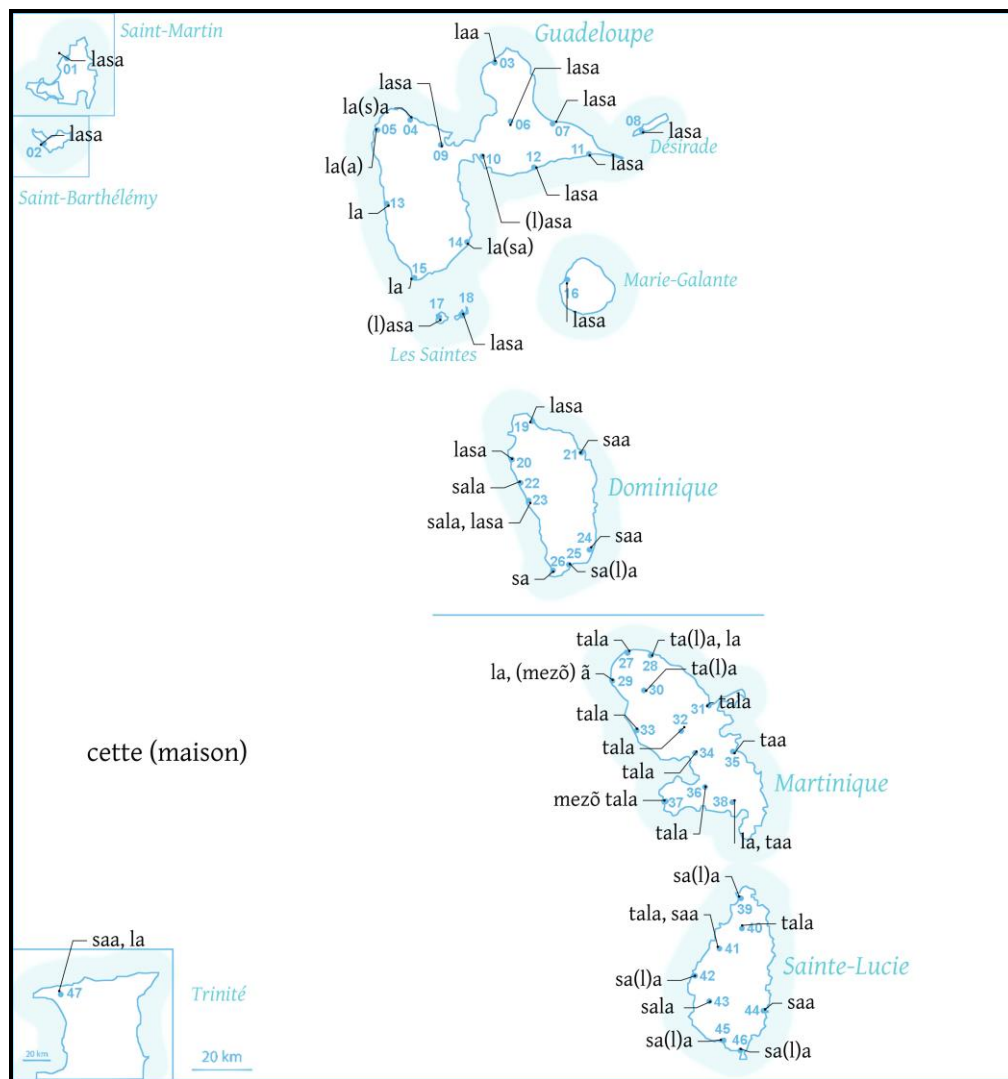


Fig. 6. Carte « cette (maison) » de l'Atlas Linguistique des Petites Antilles.

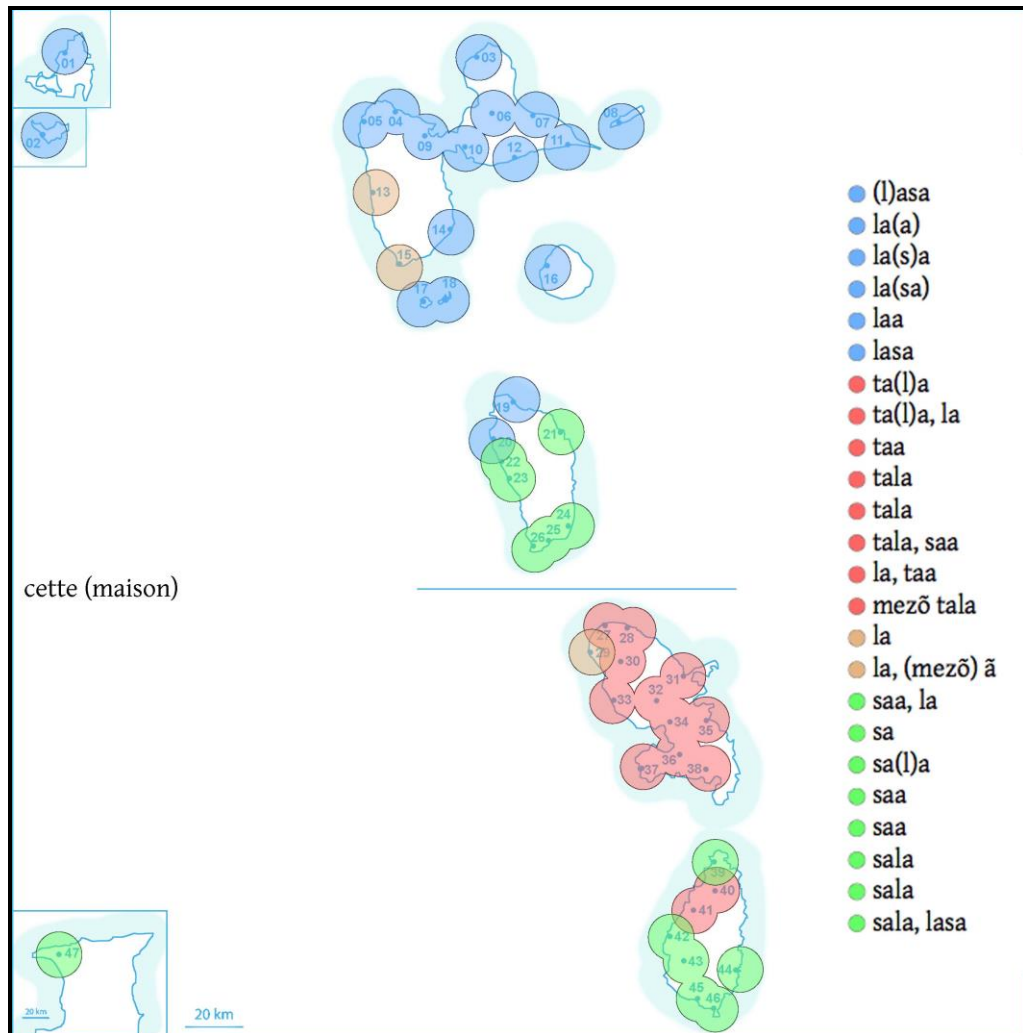


Fig. 7. Interprétation de la carte « cette (maison) » de l'ALPA : on constate une opposition nord / sud nette, et une parenté entre les données des îles ex-anglaises... qui reste à expliquer.

Ce qui m'a toujours surpris et révolté au cours de ma carrière universitaire, c'est de constater le bricolage qui règne dans les sciences humaines. On arrive péniblement à rassembler quelques sous pour un projet de recherche, mais on doit ensuite se débrouiller pour trouver celui ou celle (cartographe, informaticien...) qui va assurer la partie technique. Cette personne agira parfois, on pourrait presque dire par un heureux hasard, dans le cadre de son activité professionnelle. Bien souvent, hélas, il faudra trouver la perle, l'étudiant ou le chômeur génial, qui sera recruté et rémunéré sur de maigres crédits... dont l'utilisation sera soumise à des exigences incompréhensibles, tandis que des réunions de préparation déplaceront des gens de toute la France vers Paris grâce à un budget dépassant largement la somme qui pourrait servir à la réalisation effective du projet.

De nombreuses personnes innovent, créent à droite et à gauche de nouvelles applications en cartographie et en traitement des données, toutes plus ingénieuses les unes que les autres. Ne serait-il pas urgent de mettre en commun ces ressources pour ne pas avoir sans cesse à réinventer la roue ? Une revue est trop coûteuse, et sa publication trop épisodique. L'idéal serait la création d'un site web : mais on va se heurter à nouveau à la difficulté de recruter ou d'utiliser les compétences d'une personne officiellement chargée de le prendre en charge au sein d'un centre de recherche pérenne...

La tendance naturelle des chercheurs est de se regrouper par spécialistes de familles de langues. Pourtant, des contacts entre ces divers chercheurs existent et mériteraient d'être pérennisés, pourquoi pas au sein d'une société française de dialectologie et de géolinguistique ?

Ne serait-il pas possible d'étendre l'initiative occitane du THESOC à l'ensemble du pays ? Il ne fait pas de doute que des études au niveau national dépassant les limites de langues seraient riches d'enseignements.

Une chose est certaine : l'important aujourd'hui est la communication entre chercheurs français et étrangers partageant les mêmes préoccupations. Cela fonctionne à merveille au sein de l'ALE et de l'AliR. Nous avons grâce à internet la possibilité de nous regrouper, alors, pourquoi attendre, faisons-le !

## Références

- Alinei, M. (1996), *Origini delle lingue d'Europa*, volume 1, *La Teoria della continuità*, il Mulino, Bologna.
- Brun-Trigaud G., Le Berre Y., Le Dû, J. (2005), *Lectures de l'atlas linguistique de la France de Gilliéron et Edmont - Du temps dans l'espace*, CTHS, Paris.
- Dauzat A. (1922), *La Géographie linguistique*, Paris, Flammarion.
- Dialectologia et Geolinguistica (revue)*, Mouton de Gruyter.
- Falc'hun, F. (1963), *Histoire de la Langue bretonne d'après la géographie linguistique*, PUF, Paris.
- Géolinguistique (revue)*, centre de dialectologie, Grenoble.
- Gilliéron, J. (1902), *ALF, Notice servant à l'intelligence des cartes*, Paris, p. 3.
- Guitier, H. (1966), *Atlas linguistique des Pyrénées Orientales*, CNRS, Paris.
- Jackson K. (1967), *A Historical Phonology of Breton*, Dublin.
- Labov, W. (1976), *Sociolinguistique*, Paris, Éditions de Minuit.
- Labov, W., Ash, S., Boberg, C., (2006), *Atlas of North American English: Phonetics, Phonology and Sound Change*.
- Le Berre, A. (1973), *Ichthyonymie bretonne*, Université de Bretagne Occidentale, Brest.
- Le Berre, A., Le Berre, I., Le Dû, J. et alia (2009), *Ichthyonymie bretonne, Atlas linguistique de la faune marine de Bretagne*, Centre de Recherche Bretonne et Celtique, Université de Bretagne Occidentale, Brest (CD-Rom).
- Le Dû, J. (2001), *Nouvel Atlas Linguistique de la Basse-Bretagne*, CRBC, Brest.
- Le Roux, P. (1924-1963), *Atlas Linguistique de la Basse-Bretagne*, Plihon, Rennes.
- Loth, J. (1883) *L'Émigration bretonne en Armorique du V<sup>e</sup> au VII<sup>e</sup> siècle de notre ère*, Rennes. [réédition 1980, Slatkine Reprints, Paris-Genève-Gex].
- Rolland, P. (1994), *Atlas linguistique de la région de Pontivy, Morbihan – Côtes d'Armor*, Brest, Ar Skol Vrezoneg.
- Schleicher, A. (1868), *La théorie de Darwin et la science du langage. De l'importance du langage pour l'histoire naturelle de l'homme*, Paris, A. Franck.



Tuailon G. (1992), "Diatopie dans un espace dialectal et diatopie dans les usages d'une langue unitaire", in Rika van Deyck, *Diatopie, diachronie, diastratie. Approches des variations linguistiques*, Gent.

*Jean-Philippe Dalbera*  
*Université de Nice, UMR 6039*  
*dalbera@unice.fr*

## **La dialectologie**

*Objet, acquis, atouts et lignes de force de son devenir*

### **1. L'image de la discipline : les atlas, la mémoire vivante, le patrimoine**

L'image de la dialectologie que nous renvoient les miroirs disponibles (indicateurs d'activité, échanges disciplinaires, publications...) n'est guère brillante. Noyée au milieu des approches multiples de la variation, la dialectologie a beaucoup de mal à se faire entendre. Sa présence dans les conférences scientifiques touchant au langage se fait problématique ; c'est ainsi que dans le cadre du récent Congrès Mondial de Linguistique Française (CMLF 2008) qui s'est tenu à Paris, elle n'a figuré que sous la bannière de l'Écologie des langues<sup>1</sup>. Certes, pour ce qui est de la dialectologie française, le chantier des atlas jouit d'un prestige certain mais toutes sortes de réserves (formulées ou implicites) sont présentes en permanence dans l'esprit du grand public comme des spécialistes des autres disciplines linguistiques. De quoi est faite cette image et à quoi cela tient-il ? Ce que la plupart des gens voient dans la dialectologie, c'est le travail de terrain, la phase de recueil de données linguistiques ; c'est l'enquête, si possible en situation, auprès de locuteurs-témoins. Ce cadre, en France, prédéfinit la suite : ces témoins sont âgés, du fait de la perte de vitalité des parlers locaux, la thématique des entretiens est pour l'essentiel axée sur des activités du passé car c'est là effectivement

---

<sup>1</sup> Il se trouve qu'en ce même lieu en 2006, à l'invitation de Brigitte Horiot, j'ai été amené à présenter les recherches en dialectologie en cours à Nice (UMR 6039, Bases, corpus et langage, CNRS/Université de Nice-Sophia Antipolis), et à esquisser un avenir possible de la discipline. Mon point de vue n'ayant pas radicalement changé depuis lors, des redites sont inévitables. Vous voudrez bien ne pas m'en tenir rigueur.

que la conservation de la langue est le plus manifeste. La dimension patrimoniale est donc au premier plan : les mots et les choses de nos grands-parents et arrière-grands-parents font l'objet d'un sauvetage concerté et organisé et d'un archivage spécifique à un moment où — nul ne l'ignore — les techniques, les pratiques, les modes de pensée des acteurs de cette période sont en passe de s'effacer avec leur langue. Ce qui est retenu donc est l'attachement au passé, le souci de ne pas perdre le fil qui nous relie aux anciens, de garder et de vivre une certaine identité. La dialectologie réceptacle de la mémoire vivante, telle est la face sympathique, positive de la discipline. En contrepartie le caractère scientifique du travail dialectologique en souffre. On est dans l'humain, dans l'intime ; l'intérêt proprement scientifique n'apparaît guère. Il y a même une sorte de dérision qui s'attache à l'objet et qui finit par s'étendre à celui qui l'étudie. L'étiquette de patois n'est plus guère d'usage mais la représentation demeure : ces parlers en valent-ils la peine ? N'y a-t-il pas des langues plus importantes à soumettre à l'étude ? À quoi bon se préoccuper de parlers dont le nombre de locuteurs actifs n'excède pas dans certains cas le nombre des doigts des deux mains ?

Au-delà de ces réserves se pose avec acuité la question de savoir en quoi consiste l'étude de ces parlers et quelle est sa finalité. Que faire de cette masse énorme de données qui manifestent une variabilité quasi infinie ? Qu'en attendre ? La sauvegarde d'un moment de notre culture a été réalisée avec l'insertion de témoignages dans l'écrin des atlas linguistiques et ethnographiques. Et après ?

Bref l'idée qu'on se fait de la dialectologie tient essentiellement à deux choses : le passéisme et l'absence de propos scientifique clair. Qu'en est-il en réalité, ou du moins vu de l'intérieur ?

## **2. Genèse de la discipline : fausses questions et vraies réponses**

Un bref retour en arrière sur les conditions de l'émergence de la dialectologie n'est sans doute pas superflu avant d'esquisser les projets d'avenir. Pourquoi a-t-on décidé de considérer ces usages locaux ? De quoi est née la dialectologie ? La réponse figure dans la plupart des manuels : du besoin de répondre à la question — épistémologiquement fondamentale — de savoir si les 'lois

phonétiques’ sont de ‘vraies lois’, comparables, dans d’autres domaines aux lois formulées pour la pesanteur ou les gaz parfaits. Sachant que les grandes langues sont rétives à l’application inéluctable des lois phonétiques et attestent de multiples cas de dérogation<sup>2</sup>, est-il possible de mettre cela sur le compte de leur caractère composite, construit, enrichi, normé... ? Un bon moyen de s’en assurer consiste à tester les effets des lois phonétiques dans des langues plus « naturelles », présumées moins soumises à des pressions académiques ou des normes de hiérarchie sociale et parlées dans des communautés non pas repliées sur elles-mêmes mais moins ouvertes à l’échange et ayant connu un moindre brassage. C’est pour cela que l’on s’est tourné vers les idiomes locaux, ces patois que l’on disait jusque là rustres et grossiers. Et c’est pour cela qu’on a songé à prendre en compte les usages autres que ceux de la ‘bonne société’ et à forger des atlas linguistiques. Le résultat a été important mais n’a pas satisfait les attentes.

Les données établies par les dialectologues ne confirment pas vraiment les lois phonétiques : des exceptions de types multiples demeurent ; les isoglosses se refusent à livrer une image lisible d’entités aréales ; elles évoquent plutôt un écheveau de fils embrouillés ; le dialecte conçu comme un ensemble de parlers apparentés entre eux de manière plus étroite que chacun d’entre eux avec des entités voisines ne se laisse pas délimiter nettement ; les différenciations sont progressives. Sur le plan de la conception du changement linguistique, c’est un peu le désenchantement. La dialectologie est-elle mort-née ? *Lois phonétiques, dialectes, isoglosses et frontières dialectales* se délitent ; on se rabat sur l’histoire des mots : « chaque mot a son histoire »<sup>3</sup>.

De fait, la situation est loin d’être aussi noire qu’on a bien voulu le dire. Les lois phonétiques ne sont pas des lois comparables aux lois de la physique : —soit. Les isoglosses ne se superposent pas et ne livrent pas des frontières dialectales : —il faut en convenir. Les dialectes n’existent pas : —c’est aller un peu vite en besogne. Car les lois phonétiques ont tout de même une validité évidente ; les frontières

---

<sup>2</sup> Paveau-Sarfati (2003) mentionnent notamment l’inflexion que connaît la doctrine néogrammairienne chez H. Paul (1880) : « La loi phonétique n’énonce pas ce qui, étant données certaines conditions générales, doit nécessairement et perpétuellement se produire, elle constate seulement la congruence existant au sein d’un groupe de phénomènes historiques donnés ».

<sup>3</sup> Cf. Gilliéron-Mongin (1905), 27 : « Force nous est donc de repousser le patois comme base d’opération scientifique [...] : à l’étude du patois nous opposerons l’étude du mot ».

dialectales sont quelquefois nettement indiquées par un épais faisceau d'isoglosses...<sup>4</sup> Alors où est le vice de raisonnement ? La première évidence est que l'on a peut-être postulé des propositions intenable. Qu'implique, en effet, la notion de loi phonétique ? —que le niveau phonétique est seul responsable de l'évolution des langues. Dite ainsi, la proposition apparaît difficile à soutenir ; pour quelles étranges raisons les sons seraient-ils garantis contre toute influence, contre toute pression interne ou externe ? On sait pertinemment par exemple que l'organisation morphologique va parfois à l'encontre des exigences du phonétisme ; il faut donc envisager que les lois phonétiques constituent une tendance, une sorte de pente sur laquelle roulent les sons mais que d'autres facteurs linguistiques pèsent également, interfèrent et font que certains aboutissements se révèlent déviants par rapport à ce qui serait mécaniquement attendu. Par ailleurs on n'ignore pas que les changements linguistiques s'opèrent dans le temps, ont une durée d'effet limitée et se diffusent plus ou moins vite. Quelle difficulté y a-t-il alors à admettre qu'une loi phonétique puisse se voir appliquée de manière quasi complète, hors interférences, sur une aire donnée et incomplètement dans une autre aire ? Troisièmement, que penser de l'exigence de superposabilité imposée aux isoglosses ? Par quel miracle deux changements phonétiques survenus à trois ou quatre siècles d'intervalle devraient-ils connaître la même diffusion ou les mêmes limites d'effet dans l'espace géographique ? Chacun sait que les délimitations des états, les divisions religieuses, les circonscriptions administratives ont changé plusieurs fois au cours de l'histoire, définissant de nouveaux contours de communautés, voire de nouvelles communautés. Pourquoi n'en serait-il pas de même pour le linguistique ? On prend conscience en touchant du doigt cela qu'un certain nombre de non-dits ont faussé le regard sur le changement diachronique. L'établissement de correspondances métachroniques est évidemment un point crucial dans l'appréhension de la parenté ; mais la formalisation la plus courante de ces correspondances induit subrepticement l'idée que traiter de l'évolution diachronique revient à prendre en compte, terme à terme, les éléments qui se correspondent (l'unité source et le point d'aboutissement) et à vérifier la vraisemblance des changements intervenus. En d'autres termes, tout se passe comme si la

---

<sup>4</sup> La frontière dans l'espace entre occitan et ligurien, par exemple, est très nettement marquée dans les Alpes-Maritimes par un bourrelet d'isoglosses ; cf. Dalbera (1994) pour une analyse détaillée.

correspondance était directe entre deux états de système (origine-arrivée) alors que nous savons bien qu'il existe de multiples états intermédiaires. Conséquence évidente : les arabesques apparemment fantaisistes ou incohérentes que dessinent les isoglosses sont largement le fait de l'à-plat opéré, c'est-à-dire de l'écrasement des systèmes intermédiaires. Si l'on présente les classes de changements importants survenus (lénitions, palatalisations, diphtongaisons...) chacune sur un calque propre, on obtient des images lisibles. Cela ne change rien au fait que, une fois la pile de calques aplatie, on retrouve la complexité aréale ; mais on la comprend beaucoup mieux.

Conclusion : les notions de *loi phonétique*, *isoglosse*, *dialecte* peuvent rester opératoires pour peu que l'on réajuste leur définition. Et il est clair que, en pratique, c'est bien ainsi que l'on a travaillé jusqu'ici ; ces lois sont un pilier majeur dans toute étude des changements diachroniques mais il faut composer la pression qu'elles exercent avec d'autres forces. Les lois phonétiques sont des mirages ; —certes ! Mais cela n'empêche pas qu'on les utilise, par exemple pour faire de l'étymologie où elles demeurent une régularité majeure. Ce que l'on ajoute ici c'est l'idée qu'elles ne sont pas les seules à produire de l'évolution. Elles n'ont pas l'exclusivité du changement diachronique. Et il nous appartient par conséquent de déterminer ce que sont les autres facteurs et quel est leur poids respectif dans l'exercice des mutations.

Ce qui demeure, au fond, de cette démarche dialectologique, au delà de son échec apparent, c'est que dans le tissu dialectal a été découverte une richesse insoupçonnée : cette variété que l'on y observe n'est pas le témoignage d'un écart ou d'une série d'écarts par rapport à la langue-étalon (la vraie) mais d'une dynamique d'évolution démultipliée ; les matériaux dialectaux objectivent les voies potentielles d'évolution de la langue. Du reste, des idées jaillissent. Mentionnons en vrac : l'espace ne livrerait-il pas une représentation du temps ? l'oralité ne primerait-elle pas sur l'écrit, jusque-là sacralisé ? la diffusion du changement linguistique ne gagnerait-elle pas à être modélisée en termes de centres et de périphéries ? Comment utiliser au mieux la cartographie et le concept d'isoglosse ?

### 3. Les premiers acquis et les prémisses d'une redéfinition de l'objet de la dialectologie

Ce sont là les premiers acquis d'une dialectologie qui s'annonce comme un grand courant novateur qui traverse la linguistique. Il n'en reste pas moins qu'au départ ce qui caractérise la dialectologie et la sépare de la philologie et plus globalement de la linguistique, c'est le fait de travailler sur l'oralité ; en liaison avec cet aspect une seconde caractéristique intervient, qui touche à la préparation de l'enquête, au mode de recueil des faits, au mode de transcription et de consignation des données linguistiques, au mode de choix des témoins consultés, bref au travail dit de terrain et à la technique d'enquête. En troisième lieu figure la prise en compte de la dimension spatiale, la variabilité de proche en proche dans le continuum géographique.

Mais ni l'étude des données orales, ni l'enquête de terrain, ni la variation dans l'espace ne sont plus à même de caractériser la dialectologie : l'oralité est désormais incontournable pour la linguistique dans son entier, les sociolinguistes ont approfondi la question de l'enquête et de son impact sur les faits recueillis, les géolinguistes se sont aussi approprié l'espace géographique... Que reste-t-il au dialectologue ? En outre, à l'enthousiasme des débuts succède une période difficile. Et un reflux ne tarde pas à se manifester. En effet, l'élaboration des atlas nécessite et consomme beaucoup d'énergie de telle sorte que l'exploitation des données ne suit pas comme on l'aurait espéré : elle semble devoir être renvoyée à plus tard.<sup>5</sup> *Sine die* ?

Lorsque les enquêtes révèlent l'étendue de la variation dans l'espace, la question se pose de savoir si celle-ci est structurée. Ce que visaient, une fois prise la mesure de l'étendue de la variation, les pionniers de la dialectologie, c'était le mode de structuration de l'espace. Notamment la délimitation dans l'espace géographique d'aires témoignant d'une certaine homogénéité, susceptibles de définir les contours d'un dialecte ou d'une langue ; en d'autres termes, ceux-ci cherchaient à mettre en évidence des entités localisables sur la base

---

<sup>5</sup> Pour s'en tenir à quelques traits. Mais ceux-ci ne sont pas à prendre comme une esquisse d'histoire des idées dialectologiques, loin s'en faut. La chronologie des concepts, leurs auteurs, leur genèse excédaient manifestement le cadre de cet exposé.

du critère d'écart. Après quoi il devenait loisible de s'interroger sur la signification possible des configurations révélées et de tenter de les mettre en rapport avec les données de l'histoire. La tâche de la dialectologie se dessine : mettre en œuvre une analyse comparative en vue d'établir le degré de parenté entre les parlers d'une aire donnée.

#### **4. Un champ d'investigation passionnant**

Un champ formidable est désormais ouvert. Et dans ce champ l'objet propre de la dialectologie se dégage peu à peu. C'est la notion de *reconstruction* qui va permettre d'opérer la synthèse de tout ce qui vient d'être dit. Le dialectologue se présente désormais comme un comparatiste qui opère en milieu balisé puisque la parenté génétique des systèmes sur lesquels il se penche ne fait pas de doute. L'objet de son étude consiste à rendre compte de la genèse d'une aire linguistique, à reconstruire les états successifs constituant les étapes de l'évolution entre deux points du temps ; par exemple, en matière de langue d'oc, du latin à la langue des troubadours ou de la langue des troubadours jusqu'à l'époque contemporaine.

Les données du problème ne sont certes pas les mêmes dans les deux cas, les techniques non plus : la deuxième reconstruction est spécifiquement établie à partir de l'analyse comparative de données de source orale représentatives des parlers modernes tandis que la première s'appuie, par nécessité, sur des textes, qui donnent moins de prise à la variation. Mais si l'on admet que l'état de langue attesté sous le nom de langue des troubadours représente une phase intermédiaire de développement, comportant son lot de variations, on se rend compte que le propos de reconstruction fédère les deux analyses, les fait apparaître en complémentarité. Il n'est pas toujours possible à partir de sources écrites de dégager un système phonologique complet et sûr ni de gérer les variantes ; il n'est pas toujours possible à partir de la comparaison des entités observables à l'époque moderne de déterminer à coup sûr l'enchaînement de certains changements ; c'est bien souvent la convergence des deux regards qui permet d'aboutir. (Re)lire les textes 'anciens' à la lumière des données dialectales modernes, assurer une chronologie des mutations en y intégrant les relais (datés) des textes, ce sont là des



démarches de bon sens que les chercheurs ont spontanément engagées sans attendre qu'elles soient théorisées.

Notons que tout cela justifie, à notre sens, la consubstantialité ou l'affinité des analyses dialectologiques, diachroniques et génétiques. Gardons-nous de confondre ces trois perspectives mais gardons à l'esprit les implications qui existent entre elles. Remarquons en outre que la reconstruction qui opère en amont encore, c'est-à-dire celle qui du latin conduit à l'indo-européen, bien que fondée sur des données matériellement encore différentes (cette fois les données tangibles sont des documents exclusivement écrits [latin, grec, sanskrit etc...]) et le système reconstruit une abstraction élaborée sans filet, sans trace directe du système d'origine) relève de la même problématique. Et nous avons eu l'occasion de signaler que dans quelques cas — privilégiés, il va sans dire — des traces dans les parlers actuels de strates très anciennes (IE) sont encore repérables dans les parlers modernes, alors même que les niveaux intermédiaires, latin ou langue des troubadours semblent les avoir éliminées.

Pour effectuer la mesure du degré de parenté génétique entre les parlers d'une aire donnée, le dialectologue est conduit à forger ses propres concepts et à constituer un corps d'outils propres à la discipline. Tels sont, par exemple les concepts de changements graduels, changements en rupture et d'innovation partagée ; les premiers (changements graduels) s'inscrivent dans le cadre d'un héritage partagé et dans une évolution linéaire commune, et témoignent du fait que l'écart entre les parlers soumis à ce type de changement relève simplement d'un décalage dans le temps et, par suite, d'un degré d'accomplissement différent ; les seconds (changements en rupture), sont généralement des réactions du système à un désordre induit par une mutation ; un changement de ce dernier type révèle alors souvent une dynamique de convergence, une innovation partagée.

Le dialectologue est dès lors, à notre sens, dans son rôle majeur, celui d'un comparatiste, peut-être même d'un comparatiste privilégié. De fait, entre la démarche du comparatiste-dialectologue et celle du comparatiste indo-européaniste, seule ou presque, l'échelle change. Les problèmes que l'un et l'autre ont à traiter sont de même nature. Dans tous les cas, les systèmes reconstruits ne sont que des systèmes

intermédiaires et de pures abstractions. Mais il est vrai que les conditions de travail ne sont pas les mêmes : le recul (ou l'absence de recul) par rapport aux données, des matériaux de type différent, des univers de référence différents...

Le bilan du travail dialectologique n'est finalement pas maigre, tant s'en faut. Mais il reste un aspect à souligner. Au delà des descriptions qu'il donne des parlers, au delà de la parenté génétique qu'il mesure, au delà de la représentation de la langue dans l'espace qu'il propose, au delà des strates diachroniques qu'il révèle, le dialectologue se fait le témoin —admiratif— de la façon dont le langage se déploie dans son évolution. La reconstruction d'un ensemble dialectal se fait souvent avec des données qui ne changent pas spectaculairement d'un parler à un voisin immédiat mais qui peuvent, quand on confronte un bout de la chaîne à l'autre, manifester un écart considérable. C'est alors toute une architecture sous-jacente qui se révèle, susceptible d'inspirer ou de nourrir les modélisations du changement linguistique. C'est là qu'on discerne ce qui est possible en matière de langage, qu'on touche du doigt comment des systèmes cousins que peu d'écarts séparent, confrontés à une même difficulté sont susceptibles de s'engager dans des voies diverses pour y porter remède. C'est là qu'on appréhende la dynamique des systèmes.

Il est certain que, posée ainsi, comme reconstruction d'un espace dialectal, la dialectologie reprend fière allure. Elle (re)devient le laboratoire du changement linguistique entrevu par les premiers dialectologues. Elle offre *in vivo* un modèle de la variabilité de l'évolution, renseigne sur le degré d'autonomie des divers sous-systèmes constitutifs de la langue (ou des différents modules de la grammaire) et légitime, en illustrant les différents aboutissements possibles d'une configuration donnée à un moment donné, à l'intérieur d'un domaine homogène, c'est-à-dire presque toutes choses égales par ailleurs, le couplage *absence de déterminisme - contraintes fortes* qui caractérise le langage. On a ainsi une vision du jeu du hasard et de la nécessité, façon langage.

On notera par ailleurs, au plan méthodologique, que le problème de la mesure de l'écart qui sépare les parlers d'une aire donnée peut s'envisager de deux manières : si on s'intéresse aux degrés de parenté génétique, la mesure de l'écart s'opère par rapport au système source : plus le nombre de changements est élevé, plus la distance est grande ;

si l'on mesure les écarts des parlers entre eux à un moment de leur développement, c'est une parenté typologique que l'on estime.

## 5. Les nouveaux champs de la reconstruction

Cela souligne la contribution de la dialectologie à la linguistique générale. Nous avons rappelé<sup>6</sup>, en reprenant l'idée que *l'espace est une projection du temps*, que la reconstruction d'un espace dialectal, qui consiste —schématiquement— à construire de manière cyclique un couple *système source-règles de changement* pouvait contribuer à une modélisation intéressante du changement diachronique en général. Notamment en montrant la cohérence mais en même temps l'absence de déterminisme dans les métamorphoses d'un système, et en attestant que, en réponse à un désordre induit par une mutation, le rééquilibrage n'est pas aléatoire mais reste différent d'un dialecte à l'autre. Ce genre de constat met vigoureusement en garde contre la tentation de généraliser, à partir de quelques langues, fussent-elles dominantes, des principes ou des paramètres à vocation d'universaux.

Mais tout cela n'est pas vraiment neuf. Pourtant le renouveau devient palpable si l'on ajoute que cette reconstruction d'un espace dialectal n'est guère pratiquée jusqu'ici qu'en matière de phonétisme et de morphologie et mériterait d'être étendue au reste. C'est dans les deux domaines précités que la reconstruction comparative a fait ses preuves. Il semble inutile de s'attarder sur ce point. En revanche la question peut être intéressante de savoir si d'autres reconstructions sont envisageables ; on pense à la syntaxe et surtout au lexique dans la mesure où l'immense majorité des matériaux issus des atlas sont de l'ordre du lexique. Et on se demande, pour le cas où la réponse serait positive, à quoi ces reconstructions pourraient mener<sup>7</sup>.

Tous ceux qui depuis longtemps jouent avec les isoglosses connaissent l'embarras dans lequel on se trouve quand on essaie de composer une image de synthèse à partir d'isoglosses ; les isoglosses phonétiques, phonologiques ou morphologiques, bien que distinctes dans le type de propriété qu'elles expriment s'apparient sans que l'on puisse crier au scandale ; cumuler dans une image cartographique

---

<sup>6</sup> §2, dernier alinéa.

<sup>7</sup> Ces questions font l'objet notamment de programmes de recherche à Nice, dans le cadre de l'UMR 6039.

l'isoglosse qui sépare l'aire de palatalisation des consonnes latines C- et G- devant A de celle où ces consonnes demeurent inchangées (phonétique), l'isoglosse qui sépare l'aire de neutralisation de l'opposition /aw/ ~ /o/ de celle qui conserve cette distinction (phonologique), enfin l'isoglosse qui sépare l'aire où, dans le système verbal s'introduit un infixe -g- au subjonctif, de celle qui ne connaît pas ce phénomène, ne fait hurler personne.

En revanche, mettre la ligne de partage entre le type lexical *trida* « grive » et le type lexical *cesèra* sur le même plan que ces trois propriétés, fait hésiter.

Et que dire quand les types lexicaux se multiplient ?<sup>8</sup> On lit souvent que les isoglosses lexicales ne sont pas comme les autres. Pourquoi ? Où est l'anomalie ? Elle procède du fait, semble-t-il, que chaque isolexe ainsi conçu représente non pas une propriété systématique mais un comportement quasi individuel ; il peut y avoir autant d'isoglosses que de mots représentés ; et par suite la signification de l'isoglosse, du point de vue de la partition aréale qu'elle est censée opérer, est réduite à peu de choses. Mais cette voie n'est peut-être un cul de sac qu'en apparence. Peut-être cela tient-il à ce qu'on ne compare pas les bonnes unités.

L'unité lexicale, c'est le mot. Or ce ne sont pas des mots, signes à deux faces, que l'on confronte habituellement mais seulement leur signifiant ; l'impasse est faite sur le signifié. L'analyse s'en tient au couple signifiant-référent : ainsi par exemple, l'oiseau auquel se réfère la dénomination *Turdus viscivorus* des naturalistes est appelé, dans celles de nos régions où il passe une partie de son existence, de plusieurs façons : fr. (grive) *draine*, oc. *cesèra*, *trida*, it. *sassello*, étant entendu que chacun de ces types lexicaux se réalise selon de multiples variantes. Que signifient ces mots ? Comment ont-ils été forgés ? Peut-on reconstruire leur genèse ?

Ce genre de questionnement est au centre des préoccupations de l'*Atlas Linguarum Europae*. L'objectif des chercheurs qui y travaillent est de tirer parti des cartes lexicales des atlas et d'établir s'il existe des constantes dans la motivation des mots que l'on crée et qui entrent dans l'usage. Pour y parvenir, ces chercheurs procèdent à la comparaison de tous les mots susceptibles d'exprimer une notion donnée dans l'ensemble des dialectes et langues de l'Europe et y

---

<sup>8</sup> Comme c'est le cas, par exemple, pour les dénominations occitanes du martinet : *balestrier*, *fauçil*, *trenquet*, *sbiro*, *passovolan*...

cherchent des constantes sémantiques soit directement si ces mots sont transparents, soit en recourant à leur motif s'ils sont opaques. Les résultats sont édifiants et prometteurs<sup>9</sup>. Les dialectologues travaillant dans les langues romanes, de leur côté, se sont demandé à quoi pourrait conduire une analyse de ce genre (comparaison des formes dialectales correspondant à une notion donnée) dans le cas de langues de parenté génétique très étroite (toutes sont issues du latin). Les résultats ont dépassé toutes les espérances : des récurrences massives au niveau de la motivation mises en évidence, des étymologies élucidées, des filiations sémantiques rétablies...

## 6. Que serait une reconstruction lexicale ?

En prenant appui sur ces résultats, on peut revenir à la question de la reconstruction du lexique. Il importe de définir clairement la place de la reconstruction lexicale par rapport à l'étymologie. Celle-ci, comme chacun sait, comporte une démarche de reconstruction ; en témoignent les étymons marqués de l'astérisque dans les dictionnaires. Mais cette reconstruction, conçue sur base comparative, se limite au plan du signifiant ; il n'y a pas de reconstruction comparative du signifié. Or on vient de voir que les recherches conduites dans le cadre des atlas interprétatifs permettent d'accéder dans bon nombre de cas au motif du signe, c'est-à-dire à son signifié premier ou, du moins, au signifié du cycle motivationnel précédent. Il devient dès lors possible, sur la base d'une hypothèse d'homogénéité aréale et en se servant des signes transparents pour décrypter ceux qui paraissent, en première approximation, opaques, de dégager des constantes et de reconstruire, sur le mode comparatif, le sémantisme source.

La démonstration a été ainsi faite qu'en accédant ainsi au signifié, on se met en mesure de résoudre certains problèmes étymologiques délicats, et d'envisager une reconstruction lexicale : celle-ci se présente alors comme une double reconstruction respectant l'autonomie de chaque plan (phonique et sémique).

---

<sup>9</sup> Cf. les cartes publiées et les commentaires afférents de l'ALE et notamment les analyses d'Alinei ; par exemple Alinei (1986) Belette, ALE carte 28 et commentaire pp. 145-224 ; cf. également le chapitre II. Semantic dictionary of proto-Indo-European and reconstruction of Indo-European proto-culture dans Gamkrelidze-Ivanov (1995).

Soit un exemple pour fixer les idées. On peut montrer que les types lexicaux évoqués un peu plus haut utilisés pour désigner la grive procèdent d'une représentation quasi constante : ces turdidés sont perçus comme des oiseaux tachetés ; seule l'expression, la formulation de cette image varie d'un parler à l'autre ; et, de fait, chacun sait que cette notion de tache peut s'exprimer à partir de bases lexicales diverses : le français dira *tacheté*, *moucheté*, *grivelé*, *marqueté*... ou encore *à pois*. C'est bien cette représentation qui revient avec insistance dans les noms dialectaux de la grive ; celle-ci se nomme *gribo* « la criblée, la grivelée » (< CRIBRARE), *trido* « la triée » (< TRITARE), *sassello* « le sassé » (afr. *sasse* < SAETACIU), *cesèro* « la piselée » (< CICER(A) « pois chiches »)...

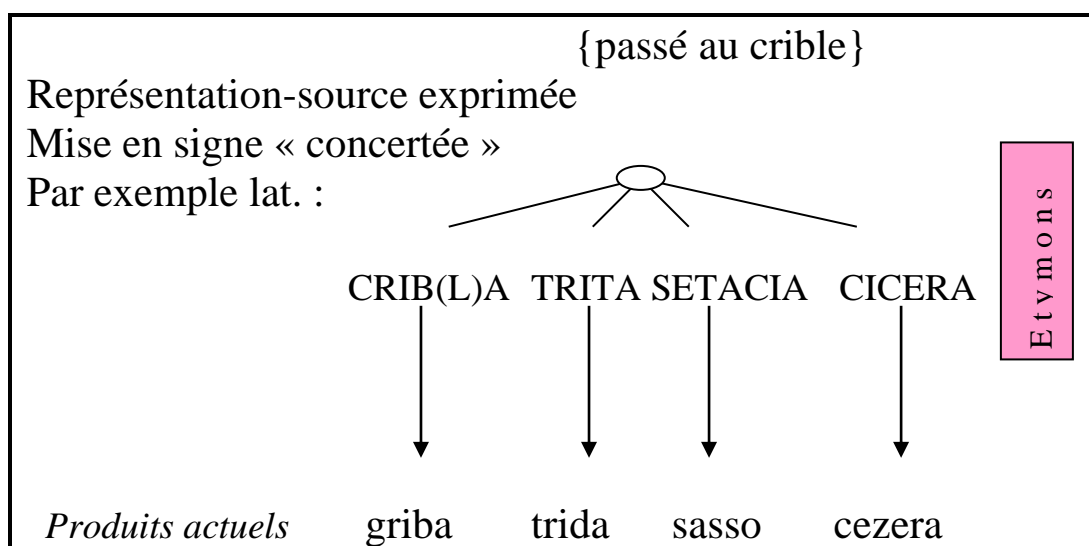
Il ressort de là qu'on est bien en présence de deux reconstructions autonomes, l'une qui reconstitue une formule phonique qui subsume un faisceau de formes dialectales (par ex. [k'ikera] d'où sont issues [s'ezera, s'ezura, s'ejra, se'zero...], l'autre qui reconstitue une représentation source, renvoyant à un ensemble de petites taches régulières du type de celles que génère le passage à travers une surface percée de trous. Le fait que la continuité, la filiation entre mots dialectaux modernes et mots sources est commandée non par une dynamique unique, la continuité du signifiant à travers les modifications phoniques successives, mais par deux dynamiques distinctes, celle du phonétisme certes, mais aussi celle du sémantisme, dans la petite série de mots cités plus haut, est démontré par le nom d'un autre turdidé, le merle ; les formes romanes renvoient toutes à lat. MERULA mais ce mot reste d'étymologie obscure<sup>10</sup>; il est pourtant corrélé aux autres, mais le fil qui le relie ne relève en rien du phonétisme : le merle, c'est le turdidé au plumage uni, immaculé, c'est TURDUS MERUS, -A ou MERULA. La relation est de l'ordre de l'antonymie, d'une antonymie qui ne passe pas par le signifiant ; le merle aurait sans doute pu s'appeler IMMACULATUS mais le sort en a voulu autrement.

Les rapports entre étymologie et reconstruction lexicale deviennent plus précis. Au commencement (du cycle) est une représentation (à exprimer) ; celle-ci donne lieu à plusieurs formulations approximativement équivalentes (synchrones ou successives ? calquées les unes sur les autres ou réinventées ? cela reste à voir). Une fois formulées elles peuvent se fixer et devenir des mots dans une

---

<sup>10</sup> Ernout A, Meillet (1979).

langue particulière. Après quoi le cycle motivation—convention—arbitraire fait son office. Et ces mots, s'ils se transmettent, constituent à nos yeux les étymons. Les types lexicaux apparaissent en définitive comme des variantes d'une proto-unité (= signe) qui associe à une certaine représentation une certaine formulation.



Cela ne peut manquer d'ouvrir une question importante : la parenté génétique se juge-t-elle seulement sur les liens de niveau 2 (continuité du signifiant) ou faut-il prendre en compte des « représentations sous-jacentes communes » ? C'est tout simplement le problème des universaux du langage qui se repose là, vu sous un certain angle.

## 7. Épilogue : du pain sur la planche

En guise d'épilogue, il nous faut insister sur les atouts essentiels que la dialectologie peut mettre au service de la linguistique générale, dans le domaine, notamment, de la reconstruction. Modèle réduit ou prototype quand il s'agit de reconstruction du phonétisme dans un espace dialectal, la dialectologie apparaît comme la clef en grandeur réelle de la sémantique lexicale ; elle seule en effet est de nature à fournir les listes d'expressions équivalentes d'une notion donnée que requiert la démonstration ; la motivation opacifiée d'un mot est inaccessible si elle n'est pas suggérée puis authentifiée par des séries. La variation que connaît le phonétisme peut s'appréhender dans le

cadre d'une (seule) langue ; cette variation est en général conditionnée par le contexte phonique ; de sorte qu'une reconstruction interne est sans doute plus difficile qu'une reconstruction comparative mais qu'elle demeure envisageable. En matière lexicale, la variation dans le cadre d'une langue unique reste très marginale ; la reconstruction ne saurait donc être que comparative ; et encore faut-il que cette variation soit suffisamment ample pour que des récurrences significatives puissent y être décelées. Seule la variété observable dans un espace dialectal remplit les conditions adéquates, ce qui revient à dire que la sémantique lexicale a beaucoup à attendre de la dialectologie. Cela toujours en fonction du même principe : la clef de la reconstruction se trouve dans la variation, ou plus précisément dans les récurrences que révèle la variation. De la même façon que la parenté se juge sur le plan du signifiant non en fonction de ressemblances immédiates mais sur la base de correspondances systématiques, de même la filiation sémantique se jugera à partir de la récurrence des représentations, fussent-elles opacifiées. La reconstruction lexicale telle que nous l'avons conçue (reconstructions autonomes sur chacun des deux plans du signe linguistique) renouvelle profondément la démarche étymologique : disons que l'idée d'utiliser en ce sens l'extraordinaire variété que recèlent les corpus lexicaux dialectaux ouvre des pistes insoupçonnées en matière de lexicologie diachronique, de sémantique lexicale et suggère des passerelles vers l'anthropologie.

De sorte que l'étude des dialectes nous semble sinon le chemin le plus court du moins un chemin praticable et sûr pour accéder au langage.<sup>11</sup>

---

<sup>11</sup> Dalbera (2006).



## Ouvrages cités

- Alinei M. (1986), Belette, *Atlas Linguarum Europae*, carte 28 et commentaire pp. 145-224, Assen, Van Gorcum.
- Dalbera J. Ph. (1994), *Les parlers des Alpes-Maritimes. Etude comparative. Essai de reconstruction*, Londres, AIEO.
- Dalbera J. Ph. (2006), *Des dialectes au langage. Une archéologie du sens*, Paris, Champion.
- Ernout A., Meillet A. (1979), *Dictionnaire étymologique de la langue latine*, Paris, Klincksieck.
- Gamkrelidze T., Ivanov V. (1995), *Indo-European and the Indo-Europeans*, Berlin, New York, Mouton de Gruyter.
- Gilliéron J., Mongin J. (1905), *Scier dans la Gaule romane du Sud et de l'Est*, Paris.
- Paveau M.-A., Sarfati G.E. (2003), *Les grandes théories de la linguistique. De la grammaire comparée à la pragmatique*, Paris, A. Colin.

*Robert Chaudenson*  
*Professeur émérite, Université Aix-Marseille 1*  
*robert.chaudenson@wanadoo.fr*

## **La dialectologie française ou « De la servitude volontaire »**

Il est un peu étrange pour un lyonnais comme moi de venir parler de dialectologie à Lyon sur la rive gauche du Rhône, alors que, dans mon esprit du moins, le temple de cette science, si c'en est une, s'est définitivement situé, sur l'autre rive du fleuve, dans la presque île et, plus précisément, rue du Plat, dans ce qui fut longtemps l'Institut de linguistique romane de la Faculté catholique.

À lire la liste des participants prévus pour ce colloque, je pense que je ne dois pas être le seul à avoir cette impression, puisque, sans même parler de Brigitte Horiot, plusieurs des conférenciers pressentis, de Jean-Claude Bouvier à Xavier Ravier, ont hanté, comme moi, cet Institut.

Cette évocation du passé n'est pas pour moi qu'une façon commode d'introduire mon propos ; elle nous situe d'emblée au coeur de plusieurs des questions que le Centre d'Études Linguistiques de l'université Lyon3 a choisi de placer au centre de ces débats. Le résumé que j'ai envoyé donne une idée de mes positions, ne serait-ce que par son titre « La dialectologie française ou « de la servitude volontaire » » ainsi que par la citation d'Homère qui ouvre le *Traité* de La Boétie auquel je fais référence (« Il n'est pas bon d'avoir plusieurs maîtres ; n'en ayons qu'un seul. Qu'un seul soit le maître, qu'un seul soit le roi. »). Comme je le disais dans ce petit texte, rédigé il y a quelques mois, cet intitulé ne convenait qu'imparfaitement à mon sentiment qu'aurait pu exprimer, de façon plus claire,

l'expression « De l'ilotisme volontaire » que je n'ai pas retenue, dans la crainte qu'elle pût choquer certains participants.

## **Brève histoire personnelle de la dialectologie française**

Pour reprendre les mots d'Homère, la dialectologie française n'eut qu'un « seul maître et un seul roi », Jules Gilliéron, dont l'*Atlas linguistique de France* servit à la fois de fondement et de modèle à la tradition dialectologique française des atlas. Son successeur fut Dauzat qui, à partir de 1939, continua, pour la compléter l'oeuvre de Gilliéron, sous la forme des *Atlas linguistiques de la France par régions*. Après la guerre, dans le cadre du tout nouveau CNRS, le projet reprit avec, à sa tête, Jean Séguy et Pierre Gardette. Le retrait de Séguy fit de Mgr Gardette le nouveau roi de la dialectologie française jusqu'à sa mort en 1973.

Pour me situer dans le domaine, mes premiers contacts avec la dialectologie se sont établis, un peu par hasard, vers 1968, au moment où, dans la préparation de ma thèse de doctorat sur le lexique du créole de la Réunion, j'ai été amené à travailler, pour les aspects étymologiques, sur les dictionnaires et glossaires des parlers français d'oïl. Mon directeur de thèse était Raymond Arveiller qui, sans être lui-même dialectologue, au sens restreint du terme, avait fait sa thèse complémentaire sur le parler de Monaco.

Je me suis vite rendu compte que les séjours parisiens que je faisais pour travailler à la Bibliothèque Nationale (dans les conditions de travail que nombre d'entre vous ont dû connaître) étaient infiniment moins fructueux que les recherches que je pouvais conduire, à Lyon même, à l'Institut de Linguistique Romane. Monseigneur Gardette, à qui j'avais exposé mon projet de recherche, m'ouvrit les portes de l'Institut. Durant mes séjours à Lyon où je venais fréquemment, je pus donc y consulter, à ma guise, avec des horaires qui n'étaient pas aussi réduits que ceux de la rue de Richelieu et surtout en accès direct, tous les glossaires dialectaux qui m'étaient nécessaires. J'ai pu ainsi conduire toutes mes recherches sans avoir recours au FEW, que je n'ai consulté que dans la phase finale de rédaction de mon travail, ce qui

m'a d'ailleurs permis d'y constater un certain nombre d'omissions, de lacunes ou d'erreurs.

Je devins ainsi, au fil des ans, à la fin des années 60 et au début des années 70, un familier de l'Institut où régnait une ambiance studieuse et amicale ; chacun y travaillant dans son coin : Monseigneur à son grand bureau, près de la fenêtre, face à celui de Brigitte Horiot ; à droite en entrant, Paulette Durdilly et Laure Malapert veillaient, en particulier, sur la préparation et l'élaboration de la *Revue de linguistique romane* ; Jean-Baptiste Martin, fidèle au poste, mais sans affectation de place aussi précise, moi-même et quelques visiteurs, réguliers ou de passage, nous nous installions sur l'une ou l'autre des grandes tables, pour ce qui me concerne plutôt à gauche de l'entrée, tout près des glossaires dialectaux dont je me nourrissais. Vers 16 heures 30, tout ce petit monde gagnait la pièce voisine, pour le thé et les petits biscuits rituels et des bavardages qui étaient hors de question dans le sanctuaire voisin. Certes, il me fut donné de voir des dialectologues de passage, et en particulier Gaston Tuailon qui venait en voisin, mais mes contacts avec la dialectologie n'allèrent guère au-delà jusqu'à la soutenance de ma thèse en juin 1972. Un moment, peut-être historique, me revient en mémoire. Celui où un jeune et blond linguiste batave (de Nimègue me semble-t-il), vint présenter, à l'heure un peu légère du thé quotidien, un projet d'Atlas linguistique de l'Europe qui recueillit un intérêt sympathique mais mesuré de la part de Monseigneur.

Je ne vais naturellement pas reprendre ici l'historique des *Atlas linguistiques de la France par régions* que Jean Le Dû a déjà fait, dans son détail et avec une connaissance des faits bien plus étendue et précise que la mienne ("La disparition du groupe des atlas et l'avenir de la géographie linguistique", *Le français moderne*, LXV, 1, 1997, pp. 6-12). Je ne parlerai pas davantage de la politique du CNRS en matière de sciences du langage, toujours soumise au croisement absurde de réformes administratives sans cesse remises en cause (on veut toujours regrouper pour réduire les coûts, jusqu'à ce qu'on doive casser ces conglomerats ingouvernables et stériles qu'on a ainsi constitués (comme, en leur temps, l'INALF ou le Lacito), ni des changements de cap scientifiques liés aux vagues successives de telles

ou telles sous-disciplines, selon les intérêts ou les ambitions des responsables sectoriels.

Mon témoignage est donc celui d'un modeste chercheur, d'abord extérieur à ce groupe, qui en découvrit peu à peu l'existence et les méthodes, avant de s'y adjoindre, mais dans une position qui resta toujours quelque peu marginale, pour des raisons qui apparaîtront dans la suite de cet article et en constituent l'essentiel.

À la fin des années 60, l'Institut de linguistique romane de Lyon (désormais ILR) est au centre des activités de dialectologie en France. Il est à la fois le lieu de publication des premiers atlas (P. Nauton pour le Massif Central, en fait, la Haute Loire, P. Gardette pour le Lyonnais), le siège de la *Revue de linguistique romane* qui, à cette époque, ne se consacrait pas encore, de façon quasi exclusive, aux études médiévales, et le laboratoire d'implantation de la RCP 160 (Recherche Coopérative sur Programme) consacré aux Atlas linguistiques de la France, qui a pris, institutionnellement, la succession de la « Commission des Atlas », dont le président était certes Michel Lejeune, mais dont la direction effective était entre les mains du secrétaire Pierre Gardette<sup>1</sup>. Monseigneur Gardette voit d'ailleurs sa position officialisée lors de la création de la RCP 160 qu'il dirigera jusqu'à sa mort en 1973. Il est alors remplacé par Gaston Tuaille qui, quoique installé à Grenoble, sera régulièrement présent à l'ILR de 1973 à 1979.

C'est donc Gaston Tuaille qui, en 1977, vivra la nouvelle réforme qui fera de la RCP 160 le GRECO 9 (GRECO = Groupement de recherches coordonnées), qui lui-même, en 1988, deviendra le GdR 9 (= Groupement de recherches)<sup>2</sup>. Sauf erreur de ma part, la dernière réunion du GRECO9/GdR 9 à laquelle j'ai participé s'est tenue en 1988, sous la présidence de Jean Lanher, avant la nomination de Jean Le Dû. Mon propos concernera donc essentiellement la période 1968-1988.

---

<sup>1</sup> Michel Lejeune, éminent helléniste, dont le *Précis d'accentuation grecque* et la *Phonétique historique du grec* ont tant fait souffrir beaucoup d'entre nous, abandonnait sans doute volontiers à Pierre Gardette la conduite de cette commission des Atlas de la France.

<sup>2</sup> On constate, par ce bref résumé de l'histoire institutionnelle de la dialectologie française au CNRS, la manie des changements de sigles due à cette structure de recherche archéo-soviétique qui, faute de pouvoir réellement se réformer, fait sans cesse valser les étiquettes, sans le moindre changement réel, sauf, souvent, en faveur de ceux qui, du fait qu'ils président à ces changements, les opèrent en faveur de leur institution de rattachement et/ou d'eux-mêmes. Le cas de Lyon et des sciences du langage est particulièrement parlant sur ce point précis.

## De ma thèse à l'Atlas linguistique et ethnographique de la Réunion

En ayant travaillé régulièrement à l'ILR entre 1968 et 1972, j'ai été naturellement en contact avec la problématique des Atlas linguistiques de la France par régions, même si je n'ai pas eu recours aux Atlas eux-mêmes dans ma recherche, puisque aucun de ceux qui pouvaient le plus m'intéresser (parlers d'oïl du Nord et de l'Ouest français) n'était encore publié à cette époque où l'on ne disposait que des *Atlas* de P. Nauton (Haute Loire) et de P. Gardette (Lyonnais). Les premiers volumes d'introduction de ces ouvrages constituaient alors la bible de tout chercheur en dialectologie et, *a fortiori*, de tout chercheur français désireux de mettre en chantier un atlas.

Nous étions à cette époque, à l'Université de la Réunion, plusieurs enseignants-chercheurs à nous intéresser aux problèmes du créole local (Michel Carayol, Pierre Cellier et moi-même en particulier) ; comme le terrain était à peu près totalement vierge, il nous est apparu qu'un Atlas linguistique et ethnographique de la Réunion, dans le cadre et sur le modèle des Atlas de la France par régions (puisque après tout, la Réunion était et demeure un département français) pouvait constituer un projet d'envergure pour la jeune équipe associée au CNRS que nous venions d'obtenir. Cette formation, l'ERA 583, dans la suite, devait, elle aussi, changer souvent de dénomination administrative, mais non de programme ni de directeur puisque, contre toutes les règles, je l'ai dirigée de 1972 à 1995 et que j'en ai quitté la direction, à ce moment, de ma propre initiative.

La première personne auprès de laquelle je me suis ouvert de ce projet fut naturellement Monseigneur Gardette, à la fois le plus éminent spécialiste du domaine, mais aussi le responsable de la RCP 160 à laquelle nous souhaitions nous rattacher. Nous n'attendions guère de moyens de la part du CNRS, car je savais que nous n'aurions pas grand mal à obtenir du Conseil Général de la Réunion une aide financière sans doute bien plus immédiate et plus importante que celle du CNRS, dont les maigres crédits n'ont jamais couvert, en gros, que les dépenses de courrier et de téléphone.

À ce stade, ce que je voulais solliciter de Monseigneur Gardette était bien plutôt des conseils que des moyens, car il me semblait que les choses avaient beaucoup évolué depuis les années 50, où avaient été mis en chantier les premiers atlas du CNRS. Si Monseigneur Gardette fut d'emblée intéressé par le projet, dès le début de nos échanges, l'une des premières questions posées, qui nous fit immédiatement diverger, fut celle de l'enregistrement des enquêtes. Nous entendions, en effet, les enregistrer intégralement pour constituer une base documentaire, ce que rendaient aisément envisageable la taille réduite de l'île (2500 km<sup>2</sup>), le nombre limité de points d'enquête prévus (40, ce qui donne toutefois une densité inégale) et le soutien financier local dont nous espérions disposer. Monseigneur Gardette, fidèle au crayon, à la gomme et au carnet d'enquête (que nous ne récusions nullement par ailleurs !) était, quant à lui, délibérément et résolument hostile aux enregistrements<sup>3</sup>.

Cette position tenait sans doute, à ce que, pour lui qui avait commencé à enquêter dans les années 50, l'enregistrement des données, avec les moyens techniques de l'époque, était à peu près impossible pour de multiples raisons, sur les terrains où il devait opérer. L'évolution technologique (miniaturisation et autonomisation des appareils d'enregistrement) n'avait pas été prise en compte, alors que ces améliorations levaient largement tous les obstacles, y compris l'objection quant au fait que l'importance des appareillages entraînait chez les informateurs des gênes qui compromettaient le bon déroulement des enquêtes.

À cela s'ajoutait, dans le cas de Pierre Gardette lui-même, ses responsabilités au sein de la RCP 160 ; elles lui faisaient craindre, en effet, que la systématisation des enregistrements n'entraînent la nécessité d'équiper de magnétophones coûteux et de pourvoir en kilomètres de bandes magnétiques toutes les équipes et tous les chercheurs de la formation, ce que ne permettaient sans doute pas les crédits de cette dernière qui étaient, pour l'essentiel, consacrés aux inévitables frais de terrain des enquêteurs. Dans le cas de l'*Atlas*

---

<sup>3</sup> Quand je lis que le « le Projet Gardette » est consacré à la « numérisation du fonds sonore franco-provençal de l'Institut Pierre Gardette », on comprendra donc que je sois, comme on dit aujourd'hui, MDR (= mort de rire) au simple souvenir des propos de Mgr Gardette sur la question de l'enregistrement des données orales.

*linguistique de la Réunion* (désormais ALRé, pour le distinguer de l'*Atlas linguistique de Rodrigues* – ALRo – qui sera réalisé dans la suite par la même équipe), ce problème ne se posait pas, puisque nous ne demandions rien et que nous pouvions assumer, grâce au Conseil Général de la Réunion, l'équipement en matériel d'enregistrement comme les frais d'enquête, l'enquêteur, Christian Barat étant mis à disposition de l'entreprise par le Vice-Rectorat, que nous avons rendu sensible aux retombées pédagogiques d'une telle entreprise.

Toutefois, on peut imaginer que, bien que ce point n'ait jamais été soulevé, l'enregistrement systématique des données aurait posé un tout autre problème, lié au statut même de l'enquêteur.

## **Le statut professionnel et scientifique de l'enquêteur d'Atlas linguistique**

Comme je l'ai déjà dit, je ne veux nullement faire ici un historique précis de la RCP 160 et du programme des Atlas de la France par régions. Même si l'ALRé a été publié dans la série des Atlas du CNRS, ce dernier organisme n'a pris en charge que la publication elle-même (ce qui est déjà très important), mais n'a financé en rien les enquêtes, ni par des crédits, ni par des postes, puisque notre équipe ne disposait pas alors d'emploi de chercheur ou d'ITA<sup>4</sup>.

Au début des années 70, la RCP 160 comprenait plus de 20 postes de chercheurs et d'ITA.

Le coût réel de cette RCP était donc très élevé, si l'on prend en compte les seuls salaires et charges sociales, sans parler, bien entendu, en termes de coûts consolidés. On a là un problème typique et récurrent du CNRS, où les frais de personnel, par nature incompressibles, absorbent la plus grande partie du budget et empêchent, de ce fait, les chercheurs de conduire les recherches pour lesquelles ils sont pourtant recrutés et payés. Je ne m'y attarde pas davantage, sauf pour faire observer qu'avec le coût d'un seul de cette

---

<sup>4</sup> ITA dans le jargon « cénéersien » signifie « Ingénieurs, techniciens, administratifs », le mot « ingénieur » ayant, au CNRS, un sens très différent de celui que le terme a dans le français ordinaire et sans rapport avec lui. Dans la publication des trois volumes de l'ALRé, tant pour la composition que pour l'édition elle-même, la contribution de Claude Martel et de Jacques Tourrel, l'une et l'autre ITA du CNRS, fut essentielle et décisive.



vingtaine de chercheurs et d'ITA converti en crédits d'équipement et de fonctionnement, on aurait pu accroître dans des proportions énormes les potentialités d'action et de recherche de la formation. Mais le but des syndicats-maison, tout puissants dans l'institution, a toujours été non de faire mieux fonctionner la recherche mais d'avoir toujours plus de postes, donc de clients !

Or, sans faire de comparaison désobligeante pour qui que ce soit, force est de constater que l'ALRé, sans aucun moyen du CNRS, ni en postes, ni en équipement, ni en frais d'enquêtes, a été mené à bien beaucoup plus vite que la plupart, pour ne pas dire la totalité, des autres atlas.

Alors qu'on aurait pu penser que, dans une formation comme la RCP 160 s'établissait une répartition logique des activités, les chercheurs faisant les recherches et les ITA se vouant aux tâches techniques (la principale étant la mise en cartes des données), on observait une confusion permanente des deux fonctions, les chercheurs accomplissant des travaux de techniciens, quand des ITA faisaient des recherches et cela d'autant plus que les nominations sur des postes d'ITA étaient alors plus faciles que sur des postes de chargés de recherche.

Le lecteur, s'il s'en trouve un, commence sans doute à se demander où je veux en venir et où se situe la « servitude volontaire » ou « l'ilotisme » avec lesquels je l'ai appâté ou effrayé plus haut.

La situation que je viens de décrire avait deux conséquences majeures, l'une qu'on pourrait (à tort sans doute) juger positive, l'autre négative.

L'enquêteur d'atlas (chargé de recherche ou ITA, selon les cas) avait pour fonction unique et exclusive de collecter des données, en les notant « au vol », sans les enregistrer, et de les rendre ensuite accessibles sur des cartes dans la forme qu'il avait choisie. Cette tâche de collecte est plus une activité de technicien que de chercheur. On peut l'accomplir en ignorant tout des sciences du langage et cela d'autant plus qu'on peut tout à fait dissocier la collecte de terrain elle-

même, de la transcription et de la restitution ultérieures, ce qui a parfois été le cas.

Dans un tel schéma opérationnel où rien n'est enregistré, l'activité de l'enquêteur est certes modeste, mais, en même temps, elle lui donne aussi un pouvoir considérable, puisqu'il est le seul garant, et de façon absolue, des faits qu'il rapporte. Rien ne lui est « opposable », ce qui n'est évidemment pas le cas si les corpus oraux ont été enregistrés et inclus dans une base de données.

Cette situation et cette position ont joué selon moi, un rôle déterminant dans le fait que les enquêteurs et les rédacteurs d'atlas linguistiques de la France par régions, à de rares exceptions près, ont accepté, pour ne pas dire souhaité (« servitude volontaire » ou « ilotisme » comme on voudra), de se limiter à la pure et simple collecte des données et à leur restitution aussi dépouillée de commentaires ou d'analyse que possible sous forme cartographique.

Pour l'anecdote, je rappellerai ici une réunion de la RCP 160, qui s'est tenue à l'ILR, sans doute après 1973, car Monseigneur Gardette n'y était pas. Bon nombre d'atlas touchaient déjà à leur fin, certains étaient terminés (au moins pour les enquêtes) et l'équipe de la RCP 160 s'interrogeait sur son avenir et envisageait de s'orienter vers l'étude des français régionaux<sup>5</sup>. On considérait, en effet, de toute évidence, que le travail serait fini avec la publication du dernier volume du dernier atlas et qu'il fallait passer à autre chose, en laissant aux linguistes le soin de travailler sur les données ainsi recueillies. Comment démontrer avec plus de clarté et de force qu'on entendait se limiter strictement, comme toujours, aux tâches ancillaires de collecte et de restitution ?

Ce n'était naturellement pas mon avis, mais je n'ai guère trouvé de soutien. Je crois que c'est le souvenir de cette réunion, où j'ai essayé de faire valoir que bien d'autres travaux pouvaient prolonger la

---

<sup>5</sup> Ce problème fut aussi, à une échelle bien plus considérable, celui de l'INALF, qui comprenait 130 ou 150 postes quand la rédaction du *Trésor de la langue française* finit par toucher à sa fin, quels qu'aient pu être les efforts pour ralentir au maximum l'achèvement du TLF. On retrouve là les problèmes structurels du CNRS qui l'empêchent d'avoir une politique scientifique cohérente et adaptable, puisque l'institution n'a, pratiquement, aucun pouvoir sur ses personnels, ni pour la localisation, ni pour l'activité.

rédaction et la publication de ces atlas, qui est à l'origine lointaine du présent texte. J'ai senti alors, avec une parfaite netteté, que la plupart des participants entendaient bien ne pas donner de prolongements réellement scientifiques à leurs activités de collecte, quelles que soient les raisons d'une telle attitude. Dans les projets envisagés pour les français régionaux, la majorité des participants concernés entendaient, par exemple, se limiter strictement au lexique, en laissant de côté les faits grammaticaux, déjà largement absents des atlas, de façon à transposer, sans trop de mal, dans le domaine quasi exclusif de la littérature régionaliste, les collectes lexicales qui avaient été celles des atlas eux-mêmes.

Si l'ALRé a été inévitablement réalisé, *grosso modo*, selon le modèle classique, dans la mesure où, publié par le CNRS, il entrerait dans la série des atlas de la France par régions <sup>6</sup>, j'ai fait ou suscité, dans la suite, pour d'autres ouvrages, d'autres choix. Ce fut le cas du premier volume de l'*Atlas Linguistique de Rodrigues* (1992), où les faits sont analysés et interprétés sur divers plans au sein même de l'atlas ; j'ai aussi incité Dominique Fattier à faire de même, quand elle a préparé comme thèse d'Etat sous ma direction, l'*Atlas linguistique d'Haïti* (1998). Même si je vois bien les raisons qui poussaient bon nombre de participants de cette réunion de la RCP 160 à Lyon à s'interdire toute réflexion ou toute analyse scientifique dans le cadre des atlas, je suis en total désaccord avec ce point de vue qui, je le pense, a contribué à la minoration de la position et de la place scientifiques de la dialectologie, au moment même où on assistait à l'essor de la sociolinguistique.

On aurait pourtant tout à fait pu imaginer un développement conjoint et complémentaire de la dialectologie (opérant dans les zones rurales et/ou étendues) et de la sociolinguistique (qu'on a appelé, un moment, à ses débuts, « dialectologie urbaine » et, plus rarement,

---

<sup>6</sup> Il présente tout de même des innovations comme, par exemple, les « données négatives », dont avait commencé à user Jean Séguy. Un terme, qu'on peut attendre, est proposé au témoin, de façon à éviter que la non-mention de cet item puisse être interprétée de façon plus précise. Plusieurs signes notent les réponses obtenues : le signe moins (-) si le terme est refusé par le témoin ; le signe égal (=) quand le terme est connu et admis par le témoin, même s'il ne l'a pas proposé spontanément ; le signe « égal barré » indique que le terme est admis, mais avec un sens différent. D'autre part, dans l'ALRé, sont proposés de nombreux textes, d'étendue variable mais d'une dimension très supérieure à celle de simples exemples. L'aspect ethnographique est aussi très présent (nombreux dessins et photos).

« dialectologie sociale »). Il est d'ailleurs significatif de voir que W. Labov, en préfiguration de sa thèse, fait son MA sur Martha Vineyard's (travail exemplaire de ce que pourrait être une véritable dialectologie, qui ne se bornerait pas à la pure et simple collecte de données) et son Phd sur le parler de New-York. Le second travail est regardé comme fondateur de la sociolinguistique moderne, alors que le premier est, à mon sens, infiniment plus riche et plus intéressant<sup>7</sup>.

On pourrait s'interroger longuement sur ces faits. Ce n'est évidemment pas le lieu de le faire ici. Pour dire les choses en un mot, je crois que la sociolinguistique urbaine a bénéficié, aux Etats-Unis où elle se développe, de l'appui du courant démocrate des années Kennedy et que les sociolinguistes du cru ont été assez habiles pour faire croire aux politiques, dispensateurs de crédits, que leurs travaux allaient aider à régler les problèmes sociaux et éducatifs des ghettos. Ce n'est évidemment pas un hasard si le vernaculaire noir américain a été l'objet de toutes les attentions, même si les résultats attendus n'ont jamais été atteints<sup>8</sup>.

Le succès de la démarche labovienne a reposé aussi sur le caractère quelque peu mécaniste de l'analyse. Avec une irrévérence que je confesse (mais il y a prescription), j'ai souvent comparé, pour mes étudiants, la méthode de Labov à une machine à faire de la saucisse. On introduit, d'un côté, des ingrédients soigneusement choisis, inventoriés, pesés et analysés (les faits linguistiques et les données sociales, ces dernières reprenant d'ailleurs les catégories de la sociologie américaine dominante (ce qui explique la si pertinente remarque de Bickerton), on choisit un mode de mise en forme (le boyau) adéquat et approprié (les outils de la mise en forme statistique et graphique, dont le « camembert » ce qui prolonge la métaphore

---

<sup>7</sup> On peut dire à propos des notations phonétiques de Labov, prises au vol, dans le brouhaha d'un grand magasin new-yorkais, qu'elles bénéficient du même privilège de « non-opposabilité » que j'ai déjà évoqué. Pour avoir fait, en compagnie de Michel Carayol, de multiples expériences de notations phonétiques collectives des mêmes enregistrements, nous pouvons assurer qu'en fait, nous entendons et, de ce fait, nous notons tous des choses différentes, cette diversité croissant encore quand il s'agit de faire en sorte que ces notations cadrent avec une théorie pré-établie les concernant. Après des décennies d'une vogue inexplicable, le soufflé de la sociolinguistique labovienne est aujourd'hui bien retombé. Je ne sais plus, les années ayant passé, si Derek Bickerton m'a dit ou a écrit quelque part que « la grande découverte de Labov est que les Américains de la classe moyenne parlent comme des Américains de la classe moyenne ».

<sup>8</sup> On peut, *mutatis mutandis*, faire une analyse un peu analogue pour les premiers travaux de Chomsky, que l'armée et la marine américaines, ont très largement subventionnés, en se laissant persuader que l'approche générativiste, par ses hypothèses universalistes, allait conduire à des systèmes performants de traduction automatique.

alimentaire) et on met le tout dans la machine, qui livrera, de l'autre côté, le produit attendu. C'est évidemment ce qui explique le succès de cette démarche, dans toute l'Amérique du Nord, et le nombre de travaux et de thèses qui ont été conduits selon ce modèle, de tels produits étant, en outre, aisément vendables auprès des pouvoirs publics dans le registre de la remédiation des malaises sociaux.

### **En guise de conclusion**

Pour en revenir à la RCP 160 et, au-delà, à la dialectologie française, ce sont les divergences de vue sur ces points qui m'ont conduit à m'éloigner définitivement du groupe des dialectologues français dans la suite, en dépit du fait que je continuais à m'intéresser aux atlas linguistiques (Rodrigues, Haïti, Seychelles). Si Dominique Fattier a achevé et publié son monumental *Atlas linguistique d'Haïti*, je n'ai pas eu le temps de publier les deux volumes de l'*Atlas linguistique de Rodrigues*, quoique les données, intégralement enregistrées, soient entièrement dépouillées et que j'ai même commencé, mais depuis bien longtemps, la rédaction du premier volume (le tome publié en 1992 est, en fait, le troisième sur la vie rurale et la pêche).

En 1997, dans l'article que j'ai déjà évoqué, sous une forme bien plus nuancée et surtout infiniment plus prudente que moi, au moment où il est responsable du GdR 9, Le Dû écrit, pour essayer de sauver le projet des atlas au CNRS : « Il ne paraît pas qu'on ait jusqu'à présent pleinement tiré parti du sens des faits que recueillent les atlas linguistiques et ethnographiques ». Il a totalement raison, mais il faudrait ajouter que la majorité des membres de la RCP 160, puis du GRECO 9, portent l'entière responsabilité de cet état de fait par leur refus obstiné d'essayer de sortir de leur rôle de simples collecteurs et transpositeurs de données. La dialectologie française a, pour une bonne part, été tuée par les dialectologues eux-mêmes, même si la structure française de recherche, le CNRS en la circonstance, porte aussi une bonne part de cette responsabilité.

## Bibliographie

- Carayol M., Chaudenson R., Barat Ch., *Atlas linguistique et ethnographique de la Réunion*, 3 volumes : tome 1, 1985 ; tome 2, 1987 ; tome 3 1991, cartes + planches dessins et photos, Paris, CNRS
- Chaudenson R., 1974, *Le lexique du parler créole de la Réunion*, Paris, Champion.
- Chaudenson R, Carayol M, Barat Ch., 1992, *Atlas linguistique de Rodrigues*, tome 3, AIF, Paris.
- Fattier D., 1998. *Contribution à l'étude de la genèse d'un créole : l'Atlas linguistique d'Haïti, cartes et commentaires*. Villeneuve d'Ascq, Presses Universitaires du Septentrion (collection thèse à la carte).
- Labov W., 1963, "The social motivation of a sound change", in *Word*, 19, p. 273-309 [étude sur Martha's Vineyard].
- Labov W. 1966, *The Social Stratification of English in New York City*.
- Le Dû J., 1997, "La disparition du groupe des atlas et l'avenir de la géographie linguistique", in *Le français moderne*, LXV, 1, pp. 6-12.

*Gérard Bodé*  
*École Normale Supérieure de Lyon*  
*Institut Français de l'Éducation<sup>1</sup>*  
*gerard.bode@ens-lyon.fr*

## **L'enquête Fortoul de 1856 : état des « idiomes » et « patois » locaux**

Dans un colloque peuplé d'éminents linguistes et consacré aux aspects méthodologiques de l'enquête linguistique, les organisateurs ont fait appel à un historien, donc à un non-linguiste, habitué à manipuler les sources d'archives mais peu au fait des études de terrain, et qui fait donc un peu figure d'Iroquois dans cette assemblée.

De plus, il traitera d'une enquête qui n'est même pas une enquête linguistique mais une enquête administrative portant sur l'état de l'instruction primaire à la fin de l'année 1855. Néanmoins, j'ose espérer que la visite de l'historien dans le temple du langage puisse finalement être utile au dialogue entre les disciplines en offrant une vision diachronique de la perception du fait linguistique par les Français du Second Empire ainsi qu'en présentant un aperçu du comportement langagier de diverses composantes de la société.

Ce travail étant actuellement en cours, il ne sera cependant pas possible d'en dresser une synthèse mais plutôt d'éclairer certains aspects qui semblent particulièrement révélateurs. Le premier point décrira l'enquête, avant d'en pointer et d'en analyser certains aspects socio-historiques qui entrent autant dans les préoccupations des dialectologues que dans celles des historiens.

---

<sup>1</sup>UMR 5190 LARHRA, Équipe Histoire de l'éducation.

Premier point : quelle est donc cette enquête de 1856 ? Depuis la fin du XVIII<sup>e</sup> siècle au moins, la préoccupation à l'égard de ce que nous appelons aujourd'hui les langues régionales avait donné naissance à diverses enquêtes. La plupart de ces enquêtes sont bien connues et ont fait l'objet d'études savantes : celle de l'abbé Grégoire de 1794 a été longuement disséquée par Dominique Julia, François Revel et Michel de Certeau<sup>2</sup> ; celle de Coquebert de Montbret de 1806 est évoquée par Ferdinand Brunot dans le tome IX de son *Histoire de la langue française*<sup>3</sup> ; celle de Victor Duruy de 1863 est brièvement analysée par Eugène Weber dans son ouvrage sur *La fin des terroirs*<sup>4</sup>. Mais l'enquête ordonnée par Hippolyte Fortoul, ministre de l'Instruction publique, a été complètement oubliée. Dans son étude de référence, *La statistique de l'enseignement primaire, 19<sup>e</sup>-20<sup>e</sup> siècles*, l'historien Jean-Noël Luc<sup>5</sup> note qu'elle n'avait pas abouti « à cause du décès du ministre ». Certes, elle n'a jamais été publiée, mais elle existe bel et bien et repose tranquillement dans 15 cartons des Archives nationales à Paris<sup>6</sup>.

Fortoul avait demandé, par une circulaire du 15 février 1856<sup>7</sup>, aux inspecteurs d'académie de chaque département, ainsi qu'aux inspecteurs de l'enseignement primaire pour chaque arrondissement, de répondre à une liste de 27 questions afin de déterminer l'état de l'enseignement primaire à la date du 31 décembre 1855. Les cinq premières questions servent à dresser un état géographique, moral, religieux, linguistique et économique du département ou de l'arrondissement, et le point 4 s'intitule « Idiomes et patois locaux » et précise « En indiquer autant que possible l'origine ; rechercher les moyens de propager l'usage de la langue française dans les campagnes et signaler les obstacles qui s'opposent à cette propagation ». Pour chaque département, un rapport est établi par arrondissement par l'inspecteur des écoles primaires territorialement compétent et un rapport général est dressé par l'inspecteur d'académie.

---

<sup>2</sup> Certeau, Michel de, Revel, Jacques & Julia, Dominique. *Une politique de la langue : la Révolution française et les patois*. Paris : Gallimard, 1987. Réédité en 2002 dans la collection Folio Histoire.

<sup>3</sup> Brunot, Ferdinand (dir.). *Histoire de la langue française des origines à nos jours. Tome 9 : La Révolution et l'Empire*. Paris : A. Colin,

<sup>4</sup> Weber, Eugen. *La fin des terroirs : la modernisation de la France rurale, 1870-1914*. Paris : Fayard, 1984.

<sup>5</sup> Luc, Jean-Noël. *La statistique de l'enseignement primaire, XIX<sup>e</sup>-XX<sup>e</sup> siècles. Politique et mode d'emploi*. Paris : INRP : Economica, 1985.

<sup>6</sup> Archives nationales F<sup>17</sup> 9321-9335.

<sup>7</sup> *Circulaires et instructions officielles relatives à l'Instruction publique*, Paris, Impr. Impériale, 1867, t. 5, 1856-1863, p. 24-26.



L'enquête concerne les 87 départements de l'époque, donc sans la Savoie, la Haute-Savoie et les Alpes-Maritimes qui ne seront français qu'en 1860, mais avec les trois départements d'Algérie (Alger, Constantine et Oran). Toutes les réponses ne sont pas parvenues. Elles manquent complètement pour deux départements, à savoir le Doubs et les Pyrénées-Orientales. La statistique suivante permet de connaître le taux de réponses conservées :

Nombres de rapports	Réponses	%
87 rapports généraux de l'inspecteur d'académie	70	80,5 %
365 rapports des inspecteurs de l'enseignement primaire	349	95,6 %
Total : 452 rapports	419	92,7 %

Il faut noter que dans trois cas (l'Ain, l'Allier et les Hautes-Alpes), les documents existent bien aux Archives nationales, mais la page sur les patois a été arrachée par un lecteur indélicat. *A priori*, ce taux de réponse peut paraître exceptionnel. Il ne faut cependant pas oublier qu'il s'agit d'une sollicitation ministérielle à laquelle les agents subalternes étaient tenus d'obéir. Les documents manquants ont simplement été perdus ... ou mal rangés. Ce taux de réponses permet au moins d'avoir une couverture assez complète du territoire national ce qui autorise une sorte de photographie de l'état linguistique du pays pour une date donnée. Les réponses sont cependant à analyser avec beaucoup de précautions. À la différence de l'enquête Duruy de 1863, aucun chiffre n'est donné. Il s'agit d'une appréciation subjective d'agents de l'État répondant à un ministre autoritaire dont ils connaissent les attentes et les motivations. De plus, elle s'inscrit dans une politique favorable à la substitution des langues régionales par la langue nationale, comme la circulaire ministérielle le précisait bien.

Nous sommes en présence de 419 rapports rédigés par environ 380 personnes différentes. Même si l'on part de l'hypothèse que nous avons à faire à une population relativement homogène et consciente de ses devoirs, il semble cependant probable que diverses sensibilités s'y soient exprimées. Les rapports sur les idiomes et patois sont de longueur variable : certains se limitent à une phrase « Il n'y a ni

idiome ni patois dans cet arrondissement », d'autres s'étalent sur plusieurs pages en détaillant les divers points demandés par le ministre. La longueur n'a pas de rapport avec la situation linguistique plus ou moins complexe. On trouve ainsi dans un arrondissement aquitain dont la population employait à peine le français la simple phrase : « la langue usuelle des campagnes est l'occitan ». Ailleurs, dans un département considéré comme non-patoisant, l'auteur note toutes les variantes de prononciation du « français corrompu ». Dans certains cas, comme dans le rapport de l'arrondissement d'Albi, l'auteur avoue son incompetence car il n'est pas de la région ou il vient juste d'arriver et il lui faut alors recourir aux notes de son prédécesseur, quand il en existe. Tous les rapports cependant signalent la langue en usage dans les campagnes. Certains illustrent leurs propos par des extraits : ainsi dans l'arrondissement de Louhans, en Saône-et-Loire, pour signaler les nuances entre les quatre patois présents sur le territoire, l'inspecteur primaire juxtapose les versets de la parabole de l'enfant prodigue, parabole aussi reproduite par l'inspecteur d'académie de la Nièvre et qui avait déjà été utilisée dans des enquêtes antérieures et notamment par Coquebert de Montbret et Fortoul lui-même. Le rapport de l'arrondissement d'Avesnes, dans le Nord, traduit des vers de La Fontaine en « patois picard ». Celui de Dunkerque compare des vers du poème germanique Heliand, « écrit en frank ou saxon » avec le flamand local. En tenant compte de toutes les réserves ainsi exprimées, cette enquête de 1856 permet de dresser une cartographie des dialectes et patois en usage, souvent avec un grand sens du détail. J'ai déjà signalé le rapport de Louhans qui répertorie les quatre patois en usage dans son arrondissement ; je pourrai encore évoquer celui de Hazebrouck qui détaille les différences nuances de ce qu'il appelle le « parlé » flamand des diverses communes de l'arrondissement.

Or, un aspect particulièrement intéressant est la désignation de la pratique langagière. Dans ces textes, différents termes sont employés pour définir le parler : langage, langue, idiome, patois et dialecte alors que pour le français, une seule expression est usitée, à savoir langue nationale. C'est dans le *Dictionnaire des synonymes de la langue française* de François Guizot<sup>8</sup>, publié en 1809 et revu en 1822 alors

---

<sup>8</sup>François Guizot, *Nouveau dictionnaire universel des synonymes de la langue française*, Paris, 1822. Article « Langue, Langage, Idiome, Dialecte, Patois, Jargon ». L'ouvrage a été consulté à partir de son édition électronique sur CD-Rom dans : *L'Atelier historique de la langue française. L'histoire des mots du Moyen Âge à nos jours*. Marsanne, 2000.

que le futur ministre de l'Instruction publique n'était encore que professeur de la faculté des lettres de Paris, livre « officiellement recommandé » par le ministère de l'Instruction publique en 1856, qu'on trouve la meilleure explication des divers termes ainsi que leur hiérarchie. En réalité Guizot reprend les définitions de Beauzée dans sa nouvelle édition de l'œuvre de l'abbé Girard, *Synonymes françois et leurs différentes significations* de 1769<sup>9</sup>. La conception linguistique officielle du Second Empire semble ainsi reposer sur les concepts de la fin du XVIII<sup>e</sup> siècle. La hiérarchie de ce dictionnaire est la suivante : langage, langue, idiome, dialecte, patois puis jargon avec les définitions suivantes :

« le mot de langage est le plus général et il ne comprend dans sa signification que l'idée qui lui est commune avec tous les autres, celle de la manière d'exprimer les pensées [...] La langue est la totalité des usages propres d'une nation pour exprimer les pensées par la parole [...] Si, dans le langage oral d'une nation, on ne considère que l'expression des pensées par la parole, d'après les principes généraux et communs à tous les hommes, le nom de langue exprime parfaitement cette idée ; mais si l'on veut encore y ajouter les vues particulières à cette nation, et les tours singuliers qu'elles occasionnent nécessairement dans sa manière de parler, le terme d'idiome est alors celui qui convient le mieux à cette idée moins générale et plus restreinte [...] Si une langue est parlée par une nation composée de plusieurs peuples égaux, et dont les états sont indépendants les uns des autres, tels qu'étaient anciennement les Grecs, et tels que sont aujourd'hui les Italiens et les Allemands, avec l'usage général des mêmes mots et de la même syntaxe, chaque peuple peut avoir des usages propres sur la prononciation, ou sur la déclinaison des mêmes mots : ces usages subalternes, également légitimes, à cause de l'égalité des états où ils sont autorisés, constituent les dialectes de la langue nationale. [...] Tout autre usage qui s'en écarte dans la prononciation, dans les terminaisons, ou de quelque autre façon que ce puisse être, ne fait ni une langue ou un idiome à part, ni un dialecte de la langue nationale : c'est un patois abandonné à la populace des provinces ; et chaque province a le sien. Un jargon est un langage particulier aux gens de certains états vils, comme les gueux ou les filous de toute espèce [...]. Le mot de jargon fait donc toujours naître une idée de mépris, qui ne se trouve point à la suite des termes précédents : et si on l'emploie quelquefois pour désigner quelque langage bien autorisé, c'est alors pour marquer le cas que l'on en fait dans le moment, plutôt que celui qu'il en faut faire dans tous les temps. Le langage se sert de tout pour manifester les pensées. Les langues n'emploient que la parole. Les idiomes se sont approprié

---

<sup>9</sup>Girard, Gabriel. *Synonymes français et leurs différentes significations*. Nlle éd. par M. Beauzée, Paris : Lebreton, 1769. L'ouvrage avait été réédité en 1806. Cet « emprunt » de Guizot lui sera ultérieurement reproché par Lafaye, Pierre-Benjamin. *Dictionnaire des synonymes de la langue française*. 5<sup>e</sup> éd. Paris : Hachette, 1884.

exclusivement certaines façons de parler qui rendent difficiles la traduction des pensées de l'un ou de l'autre. Les dialectes produisent dans la langue nationale des variétés qui nuisent quelquefois à l'intelligence, mais qui sont ordinairement favorables à l'harmonie. Les expressions propres des patois sont des restes de l'ancien langage national, qui, bien examinés, peuvent servir à en retrouver les origines. La question, que j'ai entendu faire si souvent, si le français est une langue ou un jargon, me paraît presque un crime de lèse-majesté nationale ».

Si ce dictionnaire n'est jamais cité dans les rapports, la hiérarchie, les définitions et même les jugements de valeur qui le caractérisent sont employés dans les rapports. Ainsi, le terme de langue apparaît pour désigner le breton, le basque, le catalan, l'occitan et le provençal. Le terme de dialecte désigne le picard ou le normand. Les mots patois ou dialectes sont usités pour désigner les formes dites corrompues comme le gallo ou le corse. Enfin, le dernier terme, jargon, apparaît généralement dans les rapports des départements considérés sans patois (comme l'Oise ou la Nièvre).

Ceci ne signifie pas pour autant que les auteurs des rapports aient toujours une vision très claire des mots qu'ils emploient. Je citerai ainsi un extrait du rapport de l'arrondissement de Digne qui parle du provençal comme d'un dialecte de la langue d'oc puis précise :

« Ceux-là se tromperaient d'ailleurs qui ne regarderaient le provençal que comme un patois barbare, un jargon informe, qu'on peut aisément traduire en français. Ce patois est une langue aussi ancienne que l'espagnol, le portugais, l'italien, etc. [...] Cette langue a fait l'objet de savants travaux qu'il faut consulter pour s'en faire une idée exacte ».

Certains rapporteurs savent d'ailleurs très bien jouer avec les mots et reconnaissent clairement la fonction de l'expression langagière. Ainsi en Bretagne, on constate une différence de traitement entre le breton qui est une langue et le gallo auquel on applique le terme de patois. On reconnaît aussi au breton toutes les qualités d'une langue nationale et notamment sa division en dialectes bien délimités. Mais, la Bretagne appartenant à l'espace politique français, le breton est appelé à disparaître en tant que langue véhiculaire, ce qui n'empêchera pas les érudits de l'étudier, voire de la parler mais il ne saurait être question d'en faire la langue usuelle des départements bretonnants. D'autant plus, qu'au fil du temps, l'ancienne langue celte s'est dégénérée et n'est plus parlée sous une forme corrompue que

dans les campagnes alors que les villes sont sensibles à l'influence du français. De langue, le breton se transforme progressivement en patois dont le maintien n'est pas souhaitable.

On peut s'interroger aussi sur les connaissances linguistiques des auteurs. Il s'agit d'un milieu de fonctionnaires, des inspecteurs d'académie (corps créé sous Napoléon I<sup>er</sup> lors de l'instauration du système de l'Université impériale) et d'inspecteurs de l'enseignement primaire (corps créé en 1834 par Guizot en application de la loi sur l'enseignement primaire). Ce sont donc des administratifs, voire des pédagogues. On peut se demander jusqu'à quel niveau leur culture reflète celle d'une certaine bourgeoisie des notables. Pour le savoir, il faudrait probablement croiser leurs biographies, ce qui est possible mais long, aussi ne l'ai-je pas encore fait pour l'instant. Je me contenterai donc de relever dans leurs rapports les jugements sur l'origine des langues et les annotations relatives aux documents qu'ils utilisent.

La plupart d'entre eux répètent des lieux communs sur l'origine latine du français, sur l'influence celte fort à la mode en cette première moitié du XIX<sup>e</sup> siècle, sur le rôle des invasions germaniques, notamment franque, dans la constitution de la langue française, sur l'influence de la monarchie dans ce processus, etc. Autant d'éléments qui, sous des formes et d'autres, se retrouveront longtemps y compris dans des ouvrages savants comme ceux de Ferdinand Lot. Il y a ceux qui ne se mouillent pas, comme celui de l'arrondissement de Riom, dans le Puy-de-Dôme, qui signale que le patois de son arrondissement est « un mélange de langues et de dialectes divers ». Certains auteurs essayent néanmoins de se documenter. Je suis en train de faire une typologie de ces conceptions. Elle est en cours de constitution et je dois – à ce stade du travail – ne me livrer qu'à un exercice de morceaux choisis révélateurs. Les inspecteurs d'académie – qui doivent rédiger le rapport de synthèse sur le département – s'en acquittent souvent en se documentant sur l'histoire du département ou l'histoire de la langue française. Ainsi celui du Tarn a lu les ouvrages de Compayré sur le département, sur les langues celtiques et romanes et leur influence réciproque. Celui des Hautes-Pyrénées montre un réel intérêt pour la linguistique de son époque ; il a lu les auteurs comme Wilkins, Humbolt, Bopp et Burnouf. Considérant qu'« on ne saurait être parfaitement instruit de l'origine d'une langue si on ne connaît

celle des peuples qui la parlent », il prône un travail sur les sources et l'établissement des documents authentiques. Il se méfie surtout du patriotisme local et des lieux communs : « nous sommes loin, dit-il, de partager la vénération des Méridionaux aquitains pour les poésies par eux si admirées des Goudouli, des Espourrins ». Comme son ministre l'y invite, il se montre prêt « à démolir les vieux débris de cette langue dont la disparition doit laisser place libre à son ingrate fille, jalouse de régner en despote souverain ». Ce dernier avis apparaît de façon récurrente dans divers rapports, sous des formes plus ou moins évasives. Imposer le français, oui, détruire les langues locales – surtout quand elles se sont illustrées dans le passé par une littérature de valeur, non.

Un dernier aspect frappant qui émane de ces rapports est la description du comportement linguistique des différentes couches de la société rurale, comportements qui ne se résument pas au simple problème de la connaissance ou de l'ignorance du français, auquel la vulgate officielle du ministère a longtemps voulu se restreindre. A des degrés variés, une forme de bilinguisme est souvent notée. Connaître le français, le parler et l'écrire, ne signifie pas l'employer en lieu et place du patois. Tout dépend de l'interlocuteur, de son rang social au sein de la communauté rurale. L'inspecteur d'académie du Tarn-et-Garonne l'exprime ainsi :

« La langue française est loin d'être le moyen usuel de communication des pensées : dans les villes, l'ouvrier, plus directement en contact avec l'homme instruit, parle ordinairement français ; dans les chefs-lieux de canton et les centres populeux, l'ouvrier et le paysan l'entendent, mais répondent en patois ; dans beaucoup de campagnes, le paysan ne comprend guère le français. Quelques maires de village m'ont demandé à me parler patois ».

Divers rapports, comme celui de Florac dans la Lozère, reprochent notamment aux notables locaux de n'utiliser systématiquement que le patois quand ils s'adressent à leurs ouvriers, métayers, ou à tous ceux qu'ils considèrent comme des inférieurs. Le français apparaît ainsi comme la langue de la culture et de l'universalité, le patois comme celle du quotidien, du pratique et du concret. L'inspecteur d'académie de l'Aube cantonne le patois à...

« ...l'étroite sphère d'idées dans laquelle se renferme leur intelligence ; ces idées ne s'étendent pas au-delà de leurs besoins et des habitudes de simplicité dont se compose la vie des champs ; leur vocabulaire est donc nécessairement

très circonscrit et ils n'ont de locutions qu'autant qu'il leur en faut pour exprimer leurs sentiments, leurs passions, en un mot, toutes les choses qui leur sont le plus familières eu égard à leur condition : tout ce qui dépasse le degré d'instruction qu'ils ont pu recevoir par les traditions domestiques leur est inconnu, tout ce qui est du domaine de la science ou des intelligences cultivées leur est étranger ».

Il n'y a pas nécessairement de hiérarchie de valeur. Des rapports notent que le français, langue de la pensée, n'arrive pas à traduire des objets ou des faits du quotidien tout comme le patois ne sait pas accéder aux hautes sphères de la pensée. Comme le souligne le rapport de l'arrondissement d'Ussel, le patois est la langue maternelle des ruraux :

« D'où vient, dans les campagnes, que le français n'est encore pour celui qui le possède que la langue des affaires civiles et administratives ? C'est que la puissance de l'idiome maternel est bien grande ; c'est que la langue qui a présidé à la formation de nos premières idées est celle qui amène dans notre bouche avec le plus de force et de spontanéité les mots pour l'expression de nos pensées. C'est encore que l'homme des campagnes qui pourrait parler français a besoin de rester en harmonie avec le milieu qui l'entoure. Lorsque sa femme, sa fille, ses domestiques, ne connaissent d'autre langage que le langage local, il est tout naturellement amené, dans un pareil milieu, à ne répandre les idées que sous la forme de l'idiome connu ».

Son collègue de Dijon ne dit guère autre chose :

« L'éducation de famille, la force de l'habitude et une prédilection bien naturelle qu'ont les habitants de la campagne pour le langage que parlaient leurs pères, s'opposeront encore longtemps, dans ces contrées, à ce que la langue nationale s'y vulgarise. Dans un département du Midi où nous avons exercé nos fonctions, on trouverait inconvenant que les enfants voulussent parler français hors de l'école. Ici, encore on les gourmande parfois, lorsque cela leur arrive : Parle, leur dit-on, comme le pain que tu manges. »

La plupart des rapports, y compris quand ils ne remettent pas en cause l'idéal officiel « une Nation, une langue » et qu'ils acceptent de mener le combat en faveur de l'imposition de la langue nationale, insistent sur la fonction des différents parlers. La question n'est pas une simple question linguistique, c'est avant tout une question sociale. Ce n'est qu'avec l'arrivée du chemin de fer qui facilite les voyages et les contacts avec les autres régions françaises, voire avec la capitale, que la fonction de la parole va se modifier. Tant que l'économie et la

culture locales pourront se contenter de vivre dans leur petit pays, le français n'aura aucune utilité puisque même les édiles locaux emploieront le patois pour s'adresser à leurs « pays », à leurs proches. Mais quand le petit commerçant local, voire le paysan qui va exporter ses biens, ou l'ouvrier qui suit son travail, seront contraints de s'expatrier et d'employer la lingua franca reconnue sur tout le territoire national, ils la préféreront au patois car elle sera devenue plus utile, quitte à conserver la nostalgie de leur langue maternelle et à en regretter la disparition. Par bien des aspects, les problèmes évoqués dans cette enquête de 1856 paraissent bien proches de ceux d'aujourd'hui.



Jean-Paul Chauveau  
ATILF  
CNRS, Université de Nancy  
jean-paul.chauveau@atilf.fr

## Le français d'un rural manceau d'Ancien Régime

*Que parlait-on dans nos campagnes ?*

*La dialectologie française est couramment entendue comme l'étude des dialectes du domaine d'oïl. Ce qui est moins commun, c'est que ces dialectes soient conçus primordialement comme des réalisations du français. Il n'est pas rare, au contraire, que les parlers d'oïl, les parlers dialectaux français, soient envisagés non pas comme des éléments d'une architecture d'ensemble, mais, de manière autonome par rapport à la langue commune, comme des constructions indépendantes dans une zone d'habitat dispersé. Il n'est pas rare non plus, lorsque l'on considère leur position dans l'ensemble des réalisations sociolinguistiquement différenciées, qu'on leur assigne un statut à part, dans une enceinte isolée, d'où on peut les entendre de l'extérieur, mais où ils n'entendent rien de l'extérieur. Jusque dans le cœur même où s'élaborait le français et d'où il se diffusait, le vernaculaire aurait vécu retranché : au 18<sup>e</sup> siècle « dans un Paris où les différentes classes vivaient dans une grande promiscuité et où le vernaculaire était partout, aucun membre du 'beau monde' ne pouvait ignorer la langue du 'peuple'. Il n'en allait évidemment pas de même dans le sens inverse » (Lodge, Anthony, "Les Lettres de Montmartre et l'idéologie normative", in *Revue de Linguistique Romane* tome 59 (1995), p. 444). Ces barrières conceptuelles s'écroulent lorsque l'on peut illustrer l'appropriation dialectale de données du français. Il y a encore moins lieu de les maintenir quand on peut disposer des productions écrites d'un rural du 18<sup>e</sup> siècle, qui n'a pas quitté son village natal et qui n'a pas fréquenté d'école. Ce type d'écrit montre sans conteste que l'oralité dialectale première ne s'efface pas, mais qu'elle compose avec un français banal. Même au fond des villages reculés de l'époque classique, le français était présent et les parlers dialectaux vivaient en symbiose avec lui. Envisager les dialectes français en tant que français populaires provinciaux est une perspective, parmi d'autres, qui mérite d'être examinée.*

Lorsque les enquêtes pour le nouvel atlas linguistique de la France ont commencé, la question de la langue de l'enquête a été assez rarement posée. En tout cas, en domaine d'oïl, il allait de soi que le

français devait être la langue de l'enquête. Personnellement je ne me suis jamais posé la question du choix de cette langue. C'était pour moi une évidence que ce ne pouvait être que le français et non pas le parler dialectal qui m'était familier. Ce que je voudrais envisager ici, c'est le pourquoi de cette évidence, pourquoi la question ne se posait pas. Je donne d'emblée la conclusion à laquelle je suis parvenu : le français s'imposait parce que le français était présent dans l'usage rural en symbiose avec les parlers dialectaux oïliques, à cette époque comme depuis déjà des siècles. C'est l'ancienneté et la quasi permanence de cette situation diglossique qui permet de comprendre l'évidence de l'usage du français, aussi bien pour les enquêteurs que pour les enquêtés. On souligne souvent les difficultés éprouvées par les locuteurs du français à comprendre les dialectophones et à se faire comprendre d'eux, sous l'Ancien Régime. Il n'y a pas lieu de nier ces difficultés, mais il est nécessaire de les relativiser et de ne pas les généraliser à tout l'espace du royaume. L'éclairage que je voudrais donner ne concernera qu'une zone située à la marge occidentale du Bassin parisien, mais ce ne doit pas être une situation exceptionnelle, au contraire<sup>1</sup>.

Je le ferai à partir de l'autobiographie écrite, entre 1809 et 1820, par un rural manceau, né en 1741 et mort en 1820, et qui est restée manuscrite jusqu'à son édition par une historienne, Anne Fillon, qui en a fait l'objet d'une thèse<sup>2</sup>. Son auteur ne destinait probablement pas ce récit de près de cent pages à la publication, mais à l'information de ses descendants. Cela confère à ce document une certaine ingénuité, qui le rend particulièrement intéressant pour mon propos. Mais c'est aussi un écrit tout à fait remarquable, car les autobiographies des ruraux sous l'Ancien Régime ne sont pas légion, surtout de ruraux restés ruraux toute leur vie.

---

<sup>1</sup> Voir Kristol (2009) pour un aperçu des informations sur le plurilinguisme dans les différents domaines galloromans que l'on peut tirer des ouvrages littéraires de la période post-médiévale.

<sup>2</sup> Je cite ce texte d'après Fillon (1984) qui, comme on peut en juger d'après les reproductions de pages du manuscrit, transcrit très exactement le manuscrit, en tête de la thèse sous la pagination : ms 1 à ms 97. J'utilise seulement la pagination entre parenthèses comme référence pour les extraits que j'en cite ou pour les renvois que j'y fais. On trouve le même texte dans Fillon (1996 : 25-122), mais dans une version modernisée à l'usage du public contemporain, donc inutilisable pour notre propos. J'ai déjà évoqué cette autobiographie, sans pouvoir étudier sa langue, dans une communication au *Congrès international de linguistique et de philologie romane* (Innsbruck 2007) sous le titre : "Dialectes et français dans la formation des français expatriés en Amérique".

Son auteur était très certainement une personnalité originale, comme on va le voir, mais sa situation sociale était des plus banales. Il s'agit d'un tisserand manceau qui a passé toute sa vie dans son petit village du Haut-Maine, entre Le Mans et La Flèche, à l'exception d'un tour de France avorté au bout de trois mois (16). Il ne s'est jamais assis sur les bancs d'une école, puisqu'il n'en existait pas dans son village. Voilà une situation typique de la France de l'Ouest, au sud de la ligne Saint-Malo-Genève, où la majorité de la population est dépourvue d'instruction. Mais sa situation ne correspond cependant pas à l'image du rural inculte qu'on a trop communément. Il appartient à une famille de tisserands, où l'on arrondit ses revenus par d'autres voies, notamment par les services ecclésiastiques. Sa famille est liée à l'église paroissiale et à l'abbaye de religieuses, dont l'abbesse est seigneur de la paroisse, par des emplois de sacristain et de chantre, qui se transmettent de père en fils. Ces fonctions exigent bien évidemment de savoir lire, non seulement l'écriture, mais aussi la notation musicale religieuse. Et c'est en famille que ces savoirs s'apprennent :

« quand j'ai eu un peu de Raison mon pere me pris en amitié parce quil aperçu que je comprenais bien se quil me disait et que je lui faisais de jolie Reponces, aussi il ce hâta de me montrer a lire j'aprenais Si fort quil en Etait Etonné, quand je fut plus grand il m'apris a Ecrire, l'arithmetique et le pleinchant » (6).

C'est une tradition familiale qu'il a poursuivie et qu'il engageait ses enfants à prolonger :

« Suivez mon exemple j'ai appris a lire et a Ecrire a mes enfans mon pere m'en avait fait autant faite de même » (61).

Dans la généalogie qu'il établit au début de ses mémoires, il fait remonter cette connaissance de l'écriture dans sa famille jusqu'au 16<sup>e</sup> siècle :

« il paraît que tous ces ancêtres Savoient lire et Ecrire car plusieurs ont signés Sur les Registres mais Louis françois [son père] a eu le plus d'Ecole parce que pierre Son pere Etant sacriste M<sup>r</sup>. le curé frêneau voyant que cet Enfant avait de l'Esprit il lui enseigna tout ce qui pouvait lui Servir, a lire, Ecrire larithmetique et le pleinchant car il chantait au parfait et il apprenait ce quil voulait aussi devint il fameux arithmeticien » (3).

Il cultive par la suite les connaissances que lui a communiquées son père, pendant ses loisirs :

« je passais donc mon temp dans le plaisir de jouer des instruments et à Lire tous les livres que je pouvais me procurer sur tout les histoires Enciennes les guerres, la Geographie, les vies des saints Et prophanes j'aimais aussi beaucoup les chansons et les cantiques » (20).

Bien évidemment c'est par ses relations avec le monde ecclésiastique qu'il se procure les livres. Ces savoirs devaient être exceptionnels dans la population d'un village de quelques centaines d'habitants. Ils lui permirent d'exercer différentes fonctions au cours de sa vie, telles que violonneux, sacristain, agent voyer, percepteur de l'impôt sur les huiles, syndic de la communauté des habitants de sa paroisse, aubergiste (Fillon 1996 : 170-171). La période révolutionnaire, troublée dans cette région par la Chouannerie, ayant fait fuir les élites religieuses et aristocratiques, il se trouva investi de nouveaux rôles. Comme il le dit fièrement,

« On n'avait ici recours qu'a moi puisque j'étais le seul qui savait ecrire et qui entendits un peu les affaires » (58).

Les troubles ne le détournèrent pas d'exercer les fonctions qui lui avaient été confiées et il pouvait signer dans un registre comme « Simon premier maire et dernier syndic royal de La Fontaine » (Fillon 1996 : 224). Sa personnalité originale se révèle en ces circonstances. Très attaché à son église paroissiale et à la religion catholique, il est en même temps adversaire de la Chouannerie et républicain convaincu :

« Jamais le peuple français n'a Eté si heureux que sous le gouvernement Républicain [...] anfin on auret Eté heureux si les prêtres et les Nobles n'avaient pas aportés le trouble par la Contre Revolution » (71).

Ce qu'il faut retenir, c'est que pour lui comme pour ses ancêtres l'acculturation ne se fit pas à l'intérieur d'institutions éducatives, mais par le contact, médiat et immédiat, et les interactions quasi professionnelles avec le monde ecclésiastique. On est typiquement en face d'un autodidacte, qui a tiré parti des maigres moyens dont il pouvait disposer.

Pour que le tableau ne reste pas une simple scène villageoise, fermée sur elle-même, il faut seulement mentionner quelques faits qui montrent que ce modeste rural avait au moins des ouvertures sur le Royaume. Pendant son tour de France rapidement arrêté, il en avait vu la capitale, en passant quelques jours à Paris, et il en avait vu aussi la tête, le roi, puisqu'il avait assisté à Versailles au départ de Louis XV pour la chasse (15). Au fond de son village, l'extension outre-mer de ce royaume ne lui était pas inconnue. Son père, malgré le savoir acquis au contact du curé de leur paroisse, vers 1740 :

« n'eut pas le Bonheur de trouver une place ou faire valoir son talans, cependant un parent voulu l'emmener dans l'Amérique mais il en fut Empêché par un ami et ce fut un bonheur pour lui car la barque fits naufrage dès en partant et tout équipage périt » (3-4),

Ce contretemps ne dissuada pas définitivement son père, ce fut l'amour qui s'en chargea :

« mon pere qui avait toujours le voyage des isles dans la tête n'avait guère eu envie de ce marier Il aurait mieux aimé s'embarquer mais ma mere qui aimait passionnément mon pere fits ce quelle put pour l'arrêter » (4).

Le départ outre-mer était une possibilité sérieusement envisagée par un jeune homme qui ne se voyait pas d'avenir à sa mesure dans le monde clos de ses origines. Un autre contact pouvait aussi ouvrir un autre horizon à notre autodidacte, quoiqu'il n'en parle pas. Le curé du village que, comme sacristain, il a fréquenté pendant toute la première partie de sa vie était un prêtre d'origine bretonne qui avait commencé son ministère au Canada (Fillon 1996 : 159). Ce petit monde rural n'est pas sans ouvertures sur l'extérieur.

Venons-en au texte de ses mémoires. Si l'on fait abstraction de l'absence presque totale de ponctuation et d'une graphie assez bonne mais qui ne vise pas la perfection de l'orthographe académique, on voit qu'on a affaire à un texte qui ne relève pas de la haute littérature, mais dont la langue est un français tout à fait acceptable, seulement marqué de quelques traits qui particularisent la morpho-syntaxe et le lexique.

## Morphosyntaxe

Un bon nombre de ceux-ci relèvent d'un français parlé populaire assez commun, qu'on trouverait tout aussi bien sous la plume de certains de ses contemporains citadins. J'en donnerai comme exemple l'emploi des relatifs *dont*, *duquel*, *que* :

Je pris la route d'Arras dont j'en fut coucher a une lieue (13)

Elle avait toutes les qualités dont ont peut desirer dans une fille (17)

J'ai aussi aimé le tambour j'en ait fait plusieurs dont j'ai vendu (18)

Nannon chapeau etait dans un coridor dont il y avait une porte pour antrer de la dans la chambre et une autre pour antrer dans la chambre du tour (24)

Je voyais des hardes sur la fille qui etoient a moi et dont elle se faisoit honneur de porter pour se parer (35)

Passer sur une muraille auprès duquel jetais (64)

Je les remerciai des qualités quil m'honoraient (68).

On pourrait en dire autant de l'emploi, moins contraint qu'en français standard, du pronom *y* à propos de personnes :

ne pouriez Vous pas m'y faire parler, car je l'aime tant que je ne saurais plus m'en Cacher (21)

Je ne pu jamais y attacher mon amitié pour le mariage quoiquelle fits tout ce qu'une fille peu faire en pareille occasion même jusqu'à sécarter de la bienséance (60)

Si vous voulez avoir un ami ne vous y fiez quaprés l'avoir eprouvé sept fois (62).

L'auxiliaire *avoir* remplace fréquemment l'auxiliaire *être* :

je contai le trajet que j'avais fait depuis que j'avais parti (17)

la Sotise quelle mavait faite le 22 juin a resté sans vangense (53)

les Evénement qui nous ont arivés (56)

si j'avais Resté elle ne serait pas allée a la segrêrie (33)

elle me Repondi quelle avait sorti avec les autres (34)

car il est cause du malheur qui ma arrivé (42)

le malheur qui mavet arivé (42)

si je metais connue je n'aurais pas sorti du couvent (42).

### *Le verbe*

Dans le domaine verbal, ce qui est notable, par rapport au français d'aujourd'hui, c'est son usage courant du passé simple et du subjonctif imparfait. Cela n'a rien d'étonnant pour un homme du 18<sup>e</sup> siècle et pour un homme originaire de l'Ouest où ces deux tiroirs verbaux sont restés usités par endroits jusqu'au 20<sup>e</sup> siècle, quoique ce ne soit plus le cas dans le Haut-Maine où, après le 19<sup>e</sup> siècle, on ne l'a plus rencontré. À quelques exceptions près, que nous allons examiner, l'usage classique des temps du passé est régulièrement observé.

Le plus souvent, mis à part la graphie, les formes sont celles qui sont canoniques.

#### *Passé simple (aux personnes interlocutives) :*

Je fut a la fleche (8), Je trouvai en route un compagnon (8), Je donnez une lettre a un ami (9), Je vit lheure quelle se laisseret tomber par terre (47), J'agetai du bon vin de Clermont (58), Je fits l'antreprise de bâtir une auberge (58), Je vit une bête blanche qui fits une trentaine de pas devant moi puis disparu (64) ;

Nous fûme a la foire (11), Nous fûmes (46), Puis nous joignime son frere (47), Nous fûme sur le soir faire faire nos bans (48), Nous fûmes publié (48), Nous l'achetames (57), nous crûme bien faire (57) ;

Vous eutes (46), Vous mite (46), Vous dites (46), Vous pûte (46), Vous fûte (46), Vous dîte (46).

#### *Subjonctif imparfait :*

Elle me repondit quil n'y avait rien de fait quelle ne le defie bien (23)

Il lui disait quil ne voulait pas quelle en epousa d'autre que lui (27)

elle me dit quil falait que mon pere en parla a son pere (27)

Je dis que je voulais bien que mon pere en parla au sien (27)

Je promis donc que jen parlerais a mon pere afin quil en parla au pere chapeau (27)

Elle lui dis donc...quil fallait quil eu la bonté d’aller lui rendre sa bague (30)

Je ne devais pas trouver mauvais quelle lui demandas son avis (51)

Ce n’est pas quelle aimat le gras mais elle prenait le bouillon par raison (57)

Je demandai a ma femme si elle trouveret bon que nous fissions cette entreprise (59)

L’an 95 St Clement pape ordonna que l’on plaça dans l’Eglise et dans un lieu elevé une chaire pour prêcher (85)

L’an 231 St victor 11<sup>e</sup> pape ordonna que l’on celebra la pâque (86)

L’an 117 St Evariste pape ordonna que les mariages fussent faits dans l’Eglise et en public (86).

*Subjonctif plus-que-parfait :*

Je lui dit si le gas patoy mavait atrapé il m’aurait tué devant vos yeux sans que vous eussies eut pitié de moi (42)

Cependant, par un phénomène d’attraction connu à l’époque classique<sup>3</sup>, le subjonctif imparfait est le plus souvent abandonné dans les contextes où il est gouverné ou simplement sous l’attraction d’un autre conditionnel et, dans ce cas, il est suppléé par le conditionnel :

Ma mere ma dit hier que je me marirais quand je voudrais pourvu que le garçon me conviendrait (24)

je lui demandai si elle avait rendu sa Bague, elle me dit que non et quelle n’en avait pas eut la force attendu que son pere aimet tant patoy quil voudret déjà que nous serions mariés pour venir demeurer avec nous » (26)

---

<sup>3</sup> Cf. « Cette attraction était assez courante dans l’usage classique » (Grevisse / Goose 2008 : 1100) qui cite, entre autres, cet exemple de Vaugelas : « Il sembleroit que cette etymologie seroit bien tirée par les cheveux ».



Je voudrais pour une pinte de mon sang que cela n'auret point arivé (47)

Si vous ne l'aviez point apelé patoy votre gendre au château sénéchal et que vous ne m'auriée point fait reprendre la bague malgré moi nous n'aurions rien a payer (49)

Je croiais que vous me demanderiés en mariage et rien faudret que je me serais offerte et vous savez bien que ce n'est pas aux filles a s'offrir (60).

C'est seulement à la toute fin de son texte, dans un petit mémoire d'histoire ecclésiastique dont il emprunte la substance à des lectures, qu'on voit vaciller l'usage du subjonctif imparfait. A sa place on rencontre un exemple du subjonctif présent, un autre de l'indicatif futur et une dizaine du conditionnel :

Ce fut lui qui institua le jeûne des 4 temps pour supplier le Seigneur qu'il conserve les fruits de la terre (85)

L'an 80 St Anaclet pape ordonna que les prêtres se raseront la barbe (86)

L'an 125 St Sixte pape ordonna que l'on chanterait les Sanctus etc. a la celebration de la messe (86).

Mais on remarque cependant quelques traits particularistes où se manifeste l'affleurement, à la surface du français dont son texte prouve qu'il s'est acquis une bonne maîtrise, de son usage linguistique premier, le parler dialectal de son village.

Dans un seul cas le radical du passé simple est différent de celui qui est canonique en français :

La parole lui manqua quand elle vû quelle ne pouvait plus parler (65).

La présence de la forme française :

Je vit lheure quelle se laisseret tomber par terre (47),

...prouve qu'il s'agit d'un raté de performance. Le maintien d'une telle forme dans un point du Bas-Maine (*je ne le [rvym] pu* « nous ne le revîmes plus » ALF 1154, pt 440) montre qu'on est en face d'une résurgence dialectale.

On a affaire à une autre modalité, plus complexe, du même phénomène avec un certain nombre d'exemples de la première personne du singulier du passé simple :

Je les anvoya (8), Je travailla ( 2×, 9), Je nen trouva point (10), Je ne pensâs aucunement (20), Je lui baisa les joue (23), Je me cacha vite (23), Et moi je la rassura (31), Je lui envoya une lettre (38), Je jeta la vue sur elle du momen quelle la jetait sur moi (39), Je paya la colation (41), Je ne lui en parla pas (56), Je ne manqua pas de la faire bien vite (60), Je ne poussa pas plus loin cet article (60), Il lui fallait un certificat d'honnête homme je le fits et l'anvoya au Mans (61), J'abandonna toute la maison (64).

Ce type de désinence spécifique de la première personne des verbes du premier groupe ne correspond ni à celle du français qui est *-ai*, ni à celle du dialecte qui est *-i*. Elle est idiolectale, mais elle ne peut cependant être négligée, en raison de son importance numérique (une quinzaine d'exemples) et de sa variation graphique, soit *-a*, soit *-âs*. C'est cette variété qui est éclairante, car ces deux graphies sont celles des désinences des 2<sup>e</sup> et 3<sup>e</sup> personnes du singulier. Cela indique que notre auteur ne distingue pas la désinence des trois personnes du singulier du passé simple. C'est justement le cas dans la plupart des parlers dialectaux de l'Ouest, où les personnes, quand elles sont distinguées au passé simple, ne le sont qu'au pluriel, à la seule exception des parlers du Marais vendéen. Dans le Maine, le verbe *envoyer* du premier exemple se conjugue : *j'enveyi, t'enveyi, il enveyi, j'enveyime, vz enveyite, il enveyiran* (Montesson 1859 : 32 ; Dottin 1899 : CXXII). L'indistinction des personnes du singulier dans le français de notre mémorialiste est calquée sur celle du dialecte.

Cela est confirmé par le parallélisme du subjonctif imparfait où la désinence de la première personne des verbes ne correspond, dans un certain nombre d'exemples, ni à celle du français ni à celle du dialecte ancien, mais est identique à celle du passé simple de l'indicatif :

Anfin je ne fut pourtant pas longtems tranquille dans ma vie de garçon quoique je ne pensâs aucunement dans le mariage » (20)

Nannon me priait bien d'aller le jeudi le plus matin que je pourais la trouver dans sa chambre et quelle laisserait les portes ouvertes et que je fut par le cimetièr (40)

Falait donc dire a la st Louis que je vin vous voir le landemain (42)

Turpin vint me chercher pour aller avec eux me disant qu'ils avaient un beau colation qu'il fallait que je fut en prendre ma part (43)

J'étais un fou d'aimer une personne qui avait le cœur si dur de m'avoir vu en danger à cause d'elle et de n'avoir pas voulu me donner du secours quoique je me fut écrié sur elle (44)

Elle craignait toujours que je lui échappâ (51)

Quoique je ne pus m'empêcher de penser dans mon malheur, cependant je pris le parti de ne lui en parler jamais (53)

On n'avait ici recours qu'à moi puisque j'étais le seul qui savait écrire et qui entendait un peu les affaires (58).

Face à des exemples de la forme canonique du français comme celui-ci :

La sottise que j'avais faite est restée sans vengeance quoique je n'eusse jamais approché d'elle que sur le dessein de m[e] venger mieux (53),

...ces attestations convergentes montrent qu'on a affaire à un véritable fait de langue, de nature différente cependant de l'uniformisation des désinences du singulier du passé simple, puisqu'il concerne des verbes de différents groupes. Il faut supposer que ces formes anomales attestent, par calque, la précocité de l'affaiblissement, dans les parlers dialectaux de l'Ouest, de l'emploi du subjonctif imparfait et de sa suppléance par le passé simple, comme on l'a signalé depuis la fin du 19<sup>e</sup> siècle dans les parlers dialectaux du Bas-Maine (Dottin 1899 : CXII), de Haute-Bretagne (Dottin/Langouët 1901 : CXVIII) ou de Basse-Normandie (Lepelley 1974 : 127).

### ***Les pronoms***

La morpho-syntaxe des pronoms présente quelques particularités, dont on n'envisage ici que les plus saillantes.

Dans le dialecte, la distinction de nombre peut être neutralisée pour les pronoms sujets, puisque la forme verbale maintient la distinction de nombre, comme c'est le cas en français pour le relatif. Ainsi, à la différence du français, le dialecte neutralise le nombre pour les pronoms personnels sujets où *je/ nous* est neutralisé au profit de *je* et

*il/ils* au profit de *il*. On constate dans notre texte la même neutralisation entre *quelqu'un* et *quelques-uns* au profit de *quelqu'un* :

Quelqu'uns nous conseillèrent de la nommer Catherine par ce que setait la veille de la fête de la Sainte (56)

Les autres nétaient habillés que de toille barrées noir et blanc et quelqu'un de berluche (66).

Il s'agit d'un archaïsme, puisque les pluriels originels *quelz qu'uns*, *quelcuns* du 16<sup>e</sup> siècle (TLF s. v. *quelqu'un*) semblent avoir perduré en français jusqu'au 17<sup>e</sup> siècle (Brunot 1924 : 700) et dialectalement, au moins dans le Bas-Maine : *kækã* « quelqu'un, quelques-uns » (Dottin 1899 : 286), *køkã* « quelques-uns », *køkãn* « quelques-unes » (Dottin 1899 : CVIII) jusqu'au 20<sup>e</sup> siècle, quoique le FEW (2, 1212b) n'ait pas enregistré ce type de formes.

Par contre, pour les pronoms compléments cette distinction de nombre est favorisée, au point qu'on la voit étendue au pronom réfléchi, devant un infinitif, par une opposition : singulier *se* vs. pluriel *leur* :

il a le moyen de ce monter dans une ferme (30)

[il] lui promettet quil yrait demeurer avec eux sitos quil iroient en ferme mais falait gagner de l'argent avant pour leur monter parce que la fille n'avait Rien et le pere ne pouvait rien donner (26)

il en mourut plusieurs a force de courir pour leurs sauver (72)

puis leurs chefs tranquille et potiron<sup>4</sup> leur dire de leur mettre a genoux et dire chacun un pater et un avé pour le Repos de leurs ames (77).

Dès le plus ancien français est attestée la tendance à remplacer le pronom réfléchi *soi* par des formes régimes du pronom personnel comme *lui*, *li*, *eus*, *eles* (Moignet 1973 : 143), mais c'est seulement au 16<sup>e</sup> siècle qu'apparaît *leur* dans cette fonction. Ce *leur* réfléchi qui a été noté sous la plume de Ronsard (Gougenheim 1974 : 74) est bien attesté en Anjou, du 16<sup>e</sup> au 18<sup>e</sup> siècle, surtout devant l'infinitif et le participe présent (DuPineauR 245). À l'époque contemporaine on a seulement signalé une opposition semblable à Louviers : *y vont leux*

---

<sup>4</sup> Noms de guerre de deux chefs chouans, voir Fillon (1996 : 102).

*batte* « ils vont se battre » (cf. FEW 4, 551b, ILLE). Il est notable que Louis Simon, à la différence de ce dernier cas, n'emploie *leur* en fonction de réfléchi que devant l'infinitif, ce qui est le point de départ de cet emploi. L'archaïsme est patent et il est lié au dialecte manceau où on l'a encore relevé au milieu du 19<sup>e</sup> siècle : « i vont leux assire dans leux chaires » (Montesson 1859 : 301)

### *Les prépositions*

On note une locution prépositive *au droit de* « en face de ; à proximité de » qui n'a été attestée qu'en moyen français et en français pré-classique (DMF s.v. *droit*<sup>2</sup> ; FEW 3, 88a, *directus*), mais qui est bien documentée dialectalement jusqu'à l'époque contemporaine, sous la forme dialectale mancelle [drɛ] du substantif, que notre auteur francise :

Faut que passiez la seine en batteau au droit de l'hotel des invalides pour abreger votre route (10)

Je leur fit la conduite jusqu'au droit de Ligron (41)

Je l'atendi au droit ou j'ai bâtit depuis le soleil [= nom de son auberge] (34).

Une autre locution prépositive est *à l'entour* « autour de » qui n'est pas distincte de la forme de la locution adverbiale, spécificité mancelle (FEW 13/2, 53a, TORNARE) :

Il y avait aussi huit coueurs habillés en blanc veste courte une frange d'argent a l'entour la ceinture 15.

Cette indistinction s'inscrit dans un petit ensemble qui comporte la préposition *proche* « près de » et la locution prépositive *vis à vis* « en face de », qui ont été également connues en français (cf. respectivement FEW 9, 450b, \*PROPEANUS ; 14, 537b, VISUS et note 2) :

Nous demeurions devers cette demoiselle qui etait un peu difficile a contenter (58)

Je fut donc coucher proche angers (9)

C'est une maison qui est vis a vis la cure la rue entre (43)

### ***Les rections***

Si l'on considère les exemples assez nombreux du verbe *attendre* on voit que, mis à part la construction *s'attendre de* + inf. qui est commune jusqu'au 19<sup>e</sup> siècle, les autres régimes du verbe sont anciens : *s'attendre en qn* n'est signalé qu'aux 15<sup>e</sup> et 16<sup>e</sup> siècles (FEW 25, 705a ; DMF), *s'attendre en qch* seulement au 15<sup>e</sup> siècle (FEW 25, 705b ; DMF) et que la construction *s'attendre dans qch* n'a pas été signalée. Voilà une série d'exemples qui montrent l'autonomie de cette langue par rapport au français commun de son temps. Dans la mesure où de telles constructions, notamment *s'attendre dans qch*, sont attestées dialectalement, il n'y a pas à douter que c'est la langue première de l'auteur qui transparaît ici.

*S'attendre en qn* « compter sur qn, se fier en qn ; spéc. compter sur l'amour de qn » :

Je ferai bien mieux que de mattendre en celle la qui n'a rien et qui est inconstante (37)

Ma chere amie eut aussi une peine a essayer raport a une fille du bourg qui s'atendet en moi (50)

Une autre fille qui s'atendait aussi en moi lui dit quelle ferait une folie de m'épouser (50)

*S'attendre dans qch* « compter sur qch, prévoir qch » :

Je devint tout niais ne m'attendans pas dans une telle reponce (25)  
Faut marier ces jeune gens la le lendemain des Rois mais qu'ils ne s'attande pas dans mon héritage, car je mangerai tout avant que de mourir (48)

Je ne matendais pas dans une telle reponce (60)

*S'attendre de* + inf. « se tenir pour assuré de » :

Citos quilz furent arivés au Mans ou ils s'atendaient de passer l'hivert ils furent batus par l'armée de Séphers (73)

*Attendre à* + inf. « laisser passer le temps pour » :

non mon fils faut attendre a voir comment les choses ce tourneront (37).

### *Les adverbes*

La postposition de l’adverbe *assez* « suffisamment » par rapport à l’adjectif qu’il accompagne est un archaïsme qui s’est conservé dialectalement et régionalement (FEW 24, 184a, AD SATIS ; Grevisse/Goose 2008 : 1216) :

Quand je fut fort assés pour apprendre un métier (6)

Et par la suite mon cousin l’abandonna aussi ne la trouvant pas riche assez (19)

mais Bouruet qui voyet nêtre pas *fort* assés pour nous arrêter appella patoy a son aide (35).

### *Lexique*

Comme il fallait s’y attendre, c’est le lexique qui fournit les faits caractéristiques les plus nombreux et les plus précisément localisateurs. Mais ceux-ci ne sont pas des indicateurs absolus du dialecte sous-jacent. Ce dernier ne se révèle parfois que sous un habillage parfaitement français et, assez souvent, on peut hésiter si l’on a affaire à un dialectalisme ou à un régionalisme du français.

Dans quelques cas, le sens du mot n’est pas celui qu’on lui donne généralement en français. De façon claire, il s’agit alors de la francisation formelle d’un mot dialectal.

Le substantif féminin *femelle* se rencontre en français à propos d’êtres humains essentiellement dans des emplois péjoratifs. Lorsque Simon met cet emploi dans la bouche de son amoureuse, il ne lui prête pas une telle intention. En fait il faut comprendre que ce *femelle* est la francisation graphique du *fumelle* n. f. « personne de sexe féminin, femme ou jeune fille » qui était courant et non péjoratif en manceau (FEW 3, 448a, FEMELLA) :

je me croiais dabord Etre assé forte mais comment voulez vous qu’une pauvre fille Resiste a cinq hommes dont mon pere en Etait un, ô que les pauvres femelles sont faible (42).

On rencontre un certain nombre d’emplois de l’interjection *hélas* qui ne surprennent pas comme celui-ci :

Helas que je suis une pauvre fille qui est affligée (40).

Parfois l'emploi semble un peu curieux, comme dans cette réplique de sa future femme. Les deux amoureux sont brouillés, mais notre Louis Simon se rend à son rendez-vous après qu'elle lui a fait savoir qu'elle se mourait d'amour pour lui ; il frappe à sa porte et voici comment elle l'accueille :

Elle dit qui est la je dit c'est moi, hélas mon cher ami vous voilà me dit elle (40).

Le comble est atteint dans les paroles de sa femme lors de leur nuit de noces :

Ma chère amie me dit quand nous fûmes seuls dans notre chambre hélas mon cher ami me voilà donc au comble de mes desirs, voilà donc mes souhaits accomplis hélas que je me trouve heureuse d'être venue à bout de mes intentions car si vous n'aviez pas voulu m'épouser je m'en serais allée loin d'ici car je n'aurais jamais pu résister de vous voir marié à un [sic] autre, puis elle me dit hélas mon cher ami si ce plaisir durait seulement toujours (52).

Evidemment tous ces *hélas* ne sont que la francisation formelle de l'interjection mancelle [øla], qui n'exprime pas seulement la douleur, mais aussi la surprise, l'émerveillement et la joie (cf. FEW 5, 196a, LASSUS) ; à chacun de ces emplois est attachée une intonation particulière que la graphie est bien incapable de reproduire.

Dans le dernier exemple cité se remarque un emploi de *seulement* qui surprend en français :

hélas mon cher ami si ce plaisir durait seulement toujours (52).

C'est la position de cet adverbe qui en fait l'étrangeté, car la phrase française serait : Ah ! si seulement ce plaisir durait toujours ! Mais la position que lui accorde notre Manceau est celle qu'aurait la forme correspondante [smã] dans le parler dialectal [si s pjezi dyre smã tərʒu].



Voici la liste des mots qui, soit par leur forme, soit par leur sens, soit par leur forme et par leur sens, se distinguent du français commun.

*affronter* v. tr. « faire un affront à (qqn), insulter ».

mais comme tout ce que je lui faisais et promettais n'était que par non chalance c'est-à-dire pas de bon Cœur puisque mon dessein Etait toujours de l'afronter c'est-à-dire de l'abandonner je lui dit que cela ne pressait pas et que nous avions toujours le temps (47).

TPC : Ø. Rare au 17<sup>e</sup> siècle en ce sens, le verbe ne s'est maintenu que dialectalement (FEW 3, 820, *frons*).

*ageter* v. tr. « acquérir à prix d'argent, acheter ».

j'agetai du bon vin de Clermont a dix francs la busse (58).

TPC : *ajeter* « acheter ». Variante phonétique, due à l'influence de *jeter*, bien représentée dans les parlers populaires et dialectaux (FEW 24, 66b, \**accaptare*).

*angevine* n. f. « fête de la Nativité de la Vierge Marie, le 8 septembre ».

enfin pour aller au plus court le père fits Reprendre la Bague a sa fille, appelle patoy son gendre puis décidèrent de se marier a l'angevine, patoy dit qu'il allait prendre une ferme pour la toussaints (37).

il Revint Encorre a l'angevine avec ses cousins (41).

Langevine suivante se trouvait cette année le lundi (44).

TPC : *anjuine* « fête de la Nativité de la Vierge, le 8 septembre ». Dénomination ancienne (DuPineauR 48) qui s'est conservée régionalement en Haute-Bretagne, Maine et Anjou (FEW 24, 602b, *Anjou*).

*assemblée* n. f. « fête patronale de village ».

je dinai donc chez lui puis nous fûme a la foire St Ovide s'est une assemblée qui se tiend sur la place Vandôme [à Paris] comme il s'en tiend dans nos campagnes le jour de la fête patronale (11).

le mardi il fut a l'assemblée a st jean (36).

Comme sétait le jour de l'Assemblée Mademoiselle jamin ne manqua pas de venir avec sa mere voit leur cousin turpin il furent boire dans un Cabaret après vêpres (43).

TPC : *assembiée* « assemblée ou fête patronale ». Régionalisme de l'Ouest (FEW 25, 542-3 ; DRF), distinct formellement de l'équivalent dialectal.

*assiéger* (s'-) v. pron. « s'immobiliser ».

Anviron six mois avant que ma sœur mourut ma mere antandi au milieu de la nuit d'un dimanche au Lundi un bruit comme un fons de seille qui serait tombé du plancher sur le pavé et qui tourna un peu de tems avant que de sassièger sur le pavé (63).

TPC : Ø. Emploi et sens originaux qui se rattachent à afr. mfr. *assegier* v. tr. « placer, disposer » comme un certain nombre d'exemples dialectaux parallèles (FEW 11, 411a, *sědĭcare* ; DMF).

*balet* n. m. « porche d'église abrité par un auvent ».

L'an 1819 M. Lelong Curé de cette paroisse a fait ôter l'orloge de l'Eglise et la fait mettre dans le Ballet a ses frays et depend (83).

TPC : *balet* « auvent, portique au dessus d'une boutique, d'un porche d'église ». Type lexical attesté en français du 14<sup>e</sup> au 16<sup>e</sup> siècle (Gdf 1, 564a ; FEW 1, 209a, \**balācon* ; DMF), principalement dans l'Ouest où il s'est maintenu dialectalement jusqu'à l'époque contemporaine.

*basset* n. m. « armoire basse ».

Les deux armoires qui sont dans la Sacristie ont Eté faite en 1755, le Basset a tiroirs a été fait en 1814 (82).

TPC : *basset* « buffet bas, généralement de style Louis XV rustique ». Sens qui n'est attesté que dialectalement dans le Maine et en Anjou (FEW 1, 274b, *bassus*) depuis 1727 (DuPineauR 61).

*berluce* n. f. « droguet de laine sur fil ».

le peuple n'était habillé que de serge sur fil Encore les plus aisés, les autres n'étaient habillés que de toile barrées noir et blanc et quelqu'un de berluce (66).

TPC : Ø. Dénomination d'une production typique de la Normandie et du Maine et dont la forme proprement française est *berluce* (FEW 21, 548b).

*beurrée* n. f. « tartine de pain beurrée ».

mais Citos quelle Etait Couchée elle demandet une Beurée et deux jours après elle se Relevait et soignait son Enfant (57).

TPC : *beurrée* « tartine en général ». Régionalisme de l'Ouest (FEW 1, 664a, *būtyrum*; DRF).

*biguenotte* [bignot] adj. f. « bigote ».

elle n'était pas Bignotte mais elle aimait sa Religion et s'acquittait de ces devoirs avec une grande devotion (55).

TPC : *biguenot*, *-otte* « bigot, dévot outré ». Variante formelle typiquement mancelle, intermédiaire entre le français *bigotte* et Orne, hmanc. *biguennette* (FEW 15/1, 105b, *bî gote*).

*bordage* n. m. « petite exploitation agricole ».

s'était un bordage de cinq journaux de terre, nous y avions deux mere vache nos filles ainées les gardaient ma femme allait a lherbe pour les nourrir dans l'Etable (57).

TPC : *bordage* « très petite métairie, petite ferme en général ». Type lexical qui n'est attesté qu'en Normandie, Maine, Touraine et Vendômois (FEW 15/1, 187b, *bord*). C'est la forme régionale, distincte de la forme dialectale.

*bourg* n. m. « centre d'une commune dans une zone d'habitat dispersé ».

nous demeurions dans la derniere Maison a la droite en descendant le Bourg (21).

elle fut dans le Bourg faire une commission (32).

Le fermier de la segrêrie qui Etait aussi le Cabarettier pris le parti de patoy, Comme Etant de même profession il lui dit il y a une personne dans le Bourg qui est amie de nannon faut la prier d'aller la chercher (33).

Je n'avais que le tems de courir chercher la Sage femme qui Etait dans le bourg souvent en arrivant dans la Maison je trouvais l'Enfant né (57).

(Autres exemples pp. 50, 52 et 67).

*le bas du bourg* « la partie périphérique du centre d'une commune ».

Au bout des six mois mon pere fut demeurer dans la maison ou Etait mon grand pere au devant de la halle a gauche en allant au Couvent et mon grand pere vint avec nous au Bas du Bourg la derniere maison a droite (54).

TPC : Ø. Régionalisme caractéristique de l'ouest de la France (DRF 144-146 ; FEW 15/2, 17a, *\*burg-*).

*busse* n. f. « barrique ».

j'agetai du bon vin de Clermont a dix francs la busse (58).

TPC : *busse* « tonneau de deux cent trente à deux cent quarante litres ; aujourd'hui barrique de deux cent vingt litres ». Dénomination de tonneau attestée depuis le 13<sup>e</sup> siècle, mais régionalisée très tôt ; aux 17<sup>e</sup> et 18<sup>e</sup> siècles elle est signalée comme spécifique de l'Anjou et du Maine (DuPineauR 96, 446 ; SavBr 1741 ; FEW 1, 659b, *\*buttia* ; DMF).

*charte* (*faire la roue de -*) « faire la roue en prenant appui alternativement sur les mains et sur les pieds ».

[Son arrière-grand-père, né en 1669] « il fut aussi tisserand il Etait aussi fameux chantré d'Eglise fameux chasseur et très Leste puisqu'il a une fois fait la Roux de charte depuis la halle jusqu'au Maurier » (2).

TPC : *chârte* « charrette, voiture en général, à âne, à cheval, à bras ». L'auteur connaît bien la forme française *charrette* (67). Mais lorsque le mot entre dans une locution figurée, il l'emploie sous la forme locale, *charte* [ʃart], usitée régionalement et dialectalement (FEW 2, 428b, *carrus* et note 11 ; DuPineauR 106-7 ; DRF 244-245). La locution purement dialectale serait : *faire la reue de charte*.

*closerie* n. f. « petite exploitation agricole ».

Etant survenu une famine arrivée par les gelées de l'hiver de 1709 ils furent obligés d'engager leur Closerie du hallier Celle qui est plus proche de la guilloterie (3).

TPC : *clouserie* « closerie, petite métairie ». Type lexical attesté depuis le 15<sup>e</sup> siècle et caractéristique du quart nord-ouest de la France (FEW 2, 756a, *clausus* ; DuPineauR 116, 447) et encore enregistré par le TLF comme régionalisme de l'Ouest en ce sens. Ce n'est pas la forme spécifiquement dialectale qui est utilisée.

*courée* n. f. « viscères de la poitrine ».

Proche de frénay comme ils allaient pour Egorger un patriote ils le Rencontrèrent dans leur Chemin ils se jetterent sur lui et sans aucune Explications lui ouvrirent le corps lui aracherent la courée puis la porterent chez lui en arrivans ils la jeterent sur la table en disant a sa femme tien voila la courée de ton homme que nous avons seigné la bas fais nous la cuire pour que nous la mangions ensemble la pauvre femme s'Evanouit les domestiques allerent ce cacher les voleurs se chargerent deffets et s'en allerent laissant la courée mais le cœur ni était plus un des celerats lavait mangé (76).

TPC : *courée* « poumons du porc et, par ext., d'autres animaux de boucherie, voire de l'homme ». Signalé comme régionalisme angevin par Ménage (cf. DuPineauR 123), encore enregistré dans les dictionnaires au 18<sup>e</sup> siècle, mais, de Furetière 1704 jusqu'à Trévoux 1771, seulement en référence à Pomey, et encore comme 'vieux' ou 'régional' par la lexicographie française ultérieure jusqu'au TLF. Le mot est resté très vivant seulement dans les parlers dialectaux (FEW 2, 1178b-1179a, *coratum*).

*démarcher* v. intr. « (t. de tissage) ( ? ) »

[A Amiens] j'entrai dans un métier ou on faisait du fort en diable s'était une ouvrage a quatre marches mais on demarchait a toutes les duites des deux pieds ce qui faisait l'ouvrage plate (12).

TPC : *démarcher* « 1. faire ses premiers pas pour un enfant ; 2. marcher en arrière, reculer ». Antonyme de *marcher* v. intr. « enfoncer avec le pied les pédales du métier à tisser » (dp. Enc 1765 ; FEW 16, 528b, \**markôn*) qui n'est pas signalé par ailleurs dans le domaine du tissage, mais qui est peut-être lié au haut-manceau *démarcher* v. intr. « marcher en arrière » (FEW 16, 259a).

*érussée* n. f. « action d'égrener le chanvre ; veillée pendant laquelle on égrene le chanvre et à laquelle on convoque les voisins ».

comme je jouais du violon les filles m'aimoient et m'atiroient dans leur compagnie ce qui faisait que j'allais aux Bouquets aux Erussées et souvent aux veillées du Dimanche jouer du violon pour faire danser la jeunesse (18).

TPC : *érussée* « avanie, mésaventure ». Type lexical qui n'a été repéré dans la lexicographie dialectale, en ce sens, que dans le nord-est de l'Anjou (FEW 10, 595b ; ALBRAM 411\*) où il est connu depuis le milieu du 18<sup>e</sup> siècle (DuPineauR 170).

*fort-en-diable* n. m. « espèce de tissu ».

[A Amiens] j'entrai dans un métier ou on faisait du fort en diable s'était une ouvrage a quatre marches mais on demarchait a toutes les duites des deux pieds ce qui faisait l'ouvrage plate (12).

TPC : Ø. Première attestation d'un type lexical qui n'a été signalé qu'en Haute-Bretagne et en Blaisois (FEW 3, 734a, *fortis*).

*fouace* n. f. « galette des rois ».

elle me fits Entrer dans sa chambre puis elle me dit je vous prie de Recevoir la moitié de ma fouasse que j'ai Reçue ce matin faut observer qu'aux Rois tous les individus du couvent avoient chacun leur fouasse je la Reçu avec plaisir et elle fit en Sorte de me faire Roi (24).

TPC : *fouace* « galette ronde faite de farine de blé et de beurre ; galette poivrée, de pâte à pain, qui se mange chaude au petit déjeuner ». Spécialisation sémantique qui s'accorde avec le sens de « brioche » qu'a le mot dans les parlers manceaux (FEW 3, 647b, *föcācium*).

*grand-doigt* n. m. « majeur ».

il n'avait que le pouce libre, l'index et le grand doigt Se tenoient Ensemble les deux autres doigts Etoient tombés dans la main et s'y tenoient toujours fermés (4).

TPC : Ø. Lexie qui n'a d'équivalents signalés qu'en Lorraine (FEW 3, 77b, *digitus*).

*guérou* n. m. « homme qui se transforme temporairement en loup ».

J'ai bien vu et Entendu autres choses mais Comme il ne faut pas ajouter fois a toutes ces Sortes de Rabatrie je n'en parlerai pas de crainte d'être pris pour un superstitieux comme sont la plupart des gens de Campagne avec leurs Revenand et leurs guérous (65).

TPC : *gairoux* « loup-garou ; sorcier déguisé en loup, sujet d'épouvante autrefois ; revenant, spectre ». Variante phonétique correspondant au fr. (*loup-*)*garou* qui n'a été signalée que dans le Maine et à Alençon (FEW 17, 569b, \**werwolf*).

*huge* n.f. « huche ».

il nous donna un Lit, une armoire, quarante aunes de toile deux coffres une huge mon metier et tout l'Equipage, de la potterie vaisselle les marmite chaudrons et bien autres ustensile de menage (54).

TPC : *huge* « huche ». Variante phonétique de *huche* qui est attestée depuis le 12<sup>e</sup> siècle et qui s'est conservée dans divers parlars dialectaux modernes, notamment, dans l'Ouest, en Bretagne Romane, Maine, Anjou et Vendômois (FEW 4, 519a, HUTICA ; DuPineauR 225).

*importunation* n.f. « action d'importuner ».

Et vous ma cher amie vous ne vouliez pas aller avec eux mais a force dinportunations vous y fûte (46).

TPC : Ø. Dérivé attesté au 16<sup>e</sup> siècle et seulement en Normandie à l'époque contemporaine (FEW 4, 599b, *importunus*).

*joindre* v. tr. « rejoindre, rattraper (qqn) ».

je fut obligé de fuir en courant, il courut après moi mais il vît bien quil ne lui Etait pas possible de me joindre (35).

ils partirent dès le soir après la fille croiant me trouver avec elle mais par bonheur pour moi ils ne purent nous joindre (41).

si je fut après vous ce n'était pas pour faire du mal a personne je ne voulais pas même vous Rattraper comme j'ai dit cidessus je me donnais bien garde de vous joindre (46).

puis nous joignime son frere qui Etait loin devant nous (47).

ils aperçurent la piste des moutons et la suivirent et joignirent le voleur et les moutons vers lantreprise et les ramenerent mais le voleur leur Echapa (65).

TPC : Ø. Sens attesté en français au 17<sup>e</sup> siècle et conservé dialectalement à l'époque contemporaine (FEW 5, 67b, *jüngëre*).

*mal de l'an* loc. nom. m. « convulsions, coliques des petits enfants ».

on me crû mort du mal delan, mais ma mere trempa une Serviette dans de l'eau de la fontaine St martin et m'envelopa avec et jamais je n'en ait eu d'attaque depuis (6)

trois mois après mon frere pierre mourut du mal de l'an âgé de sept ans (64).

TPC : *mal de l'an* « troubles digestifs (diarrhée verte) qui apparaissent chez le nourrisson vers la fin de sa première année ». Dénomination qui n'a été relevée qu'en Normandie, Maine et Anjou (FEW 24, 626b, *annus*).

*place* (*jeter dans la -*) loc. verb. tr. « jeter (qch) sur le sol de la maison ».

aussitos ils ouvrirent armoires, coffres et tiroirs de table et jeterent dans la place sur le pavé tout ce quil trouverent dans les dit meubles (75).

TPC : *piace* « place [sens général] ; sol de la salle commune ». Sens de *place* typique de l'ouest de la France, aussi bien régionalement que dialectalement (FEW 9, 37b-38a, *platea* ; DRF 794-5), qui remonte au plus tard à l'époque classique, du fait de sa présence au Canada (DuPineauR 308). La forme employée n'est pas dialectale.

*rabaterie* n. f. « bavardages stupides ; vétilles ; bagatelles ».

Comme il ne faut pas ajouter fois a toutes ces Sortes de Rabatrie je n'en parlerai pas de crainte d'être pris pour un supertitieux (65).

TPC : *rabâterie* n. f. « rabâchage ». Type lexical qui n'a pas été signalé ailleurs que dans le Maine et en Anjou (FEW 10, 2b, \*RABB-), depuis le milieu du 18<sup>e</sup> siècle (DuPineauR 335).

*racheminer* (*se -*) v. pron. « se remettre en route ».

je vit nannon qui se rachemina avec sa compagnie alors je crus quelle sen allait a la fleche car je netais prevenu de rien (35).

TPC : Ø. Seulement attesté en français du 13<sup>e</sup> au 15<sup>e</sup> siècle (FEW 2, 146a, *cammīnus* ; DMF).

*raisons* f. pl. « paroles, propos, discours ».

il dit que jetais bien traitre quil n'aurait pas crus cela de moi et d'autre Raisons quil ne valaient pas mieux afin de me Rendre Ridicule auprès de la fille (37).

enfin avec bien des Raisons quelle me dit que son père voulait la marier a patoy cela fits qu'en Reprenant la fille en amitié je pris le père en haine (51).

TPC : *raisons* f. pl. « commérages, calomnie ». Maintien d'afr. mfr. *raison* « paroles, propos, discours, récit » qui n'a pas dépassé le 16<sup>e</sup> siècle (FEW 10, 107a, *ratio*), mais qui est poursuivi en haut-manceau contemporain par *raisons* « commérages » (FEW 10, 106a).

*rhabiller* v. intr. « raccommoder les vêtements ».

elle me faisait mes trammes avait soin de son menage Rhabillet et tenet ces Enfants proprement et navions point de servante (57).

TPC : Ø. Spécialisation de sens du fr. *rhabiller* « remettre en état, réparer » attestée en français aux 16<sup>e</sup> et 17<sup>e</sup> siècles et maintenue dialectalement (FEW 1, 367a, \**bilia* et note 25).

*ribergère* n. f. « quatrième dimanche de carême ».

Le jour de la Ribergere ou la mi Carême les garçons du couvent firent un grand diner en morue qu'ils avoient achetée a la St Mathias et je fus invité (25).

TPC : *ribergère* « fête du quatrième dimanche de Carême ; date d'embauche des domestiques [généralement le dimanche et le jeudi de la *Ribergère*] ». Dénomination seulement attestée à l'époque moderne en Haut-Maine, Vendômois, Touraine et Saintonge (FEW 23, 162 ; 14, 335a ; 22/1, 176a). C'est le terme qui vient spontanément à l'esprit de Simon, mais il lui fait aussitôt correspondre la dénomination officielle de cette fête religieuse qu'il connaît bien, en tant qu'ancien sacristain.

*sacriste* n. f. « sacristain d'une église ».

Louis françois a eu le plus d'Ecole parce que pierre Son pere Etant Sacriste M<sup>r</sup>. Le curé frêneau voyant que cet Enfant avait de l'Esprit il lui enseigna tout ce qui pouvait lui Servir (3).

mon pere m'encouraget de chanter me disant quil avait bien chanté la messe tout Seul au pupitre a l'âge de onze ans et comme mon grand pere Etait toujours Sacriste je fut obligé d'aller aussi chanter Etant fort jeune (6).

elle s'en Retourna a anvoise ou elle s'est mariée au Sacriste (19).

TPC : Ø. Archaïsme (FEW 11, 32a, *sacer*) donné comme régionalisme de l'Ouest par le TLF (s. v. *sacristain*).

*sapin* n. m. « variété de pin ».

Les Sapins ne sont ici guère plus anciens que moi quand je suis venu au monde il n'y en avait peutêtre pas un cent dans la fontaine Encorre ce n'était que sur les terres du Maurier il paraît que s'est M. Dorvaulx le grand pere a celui que j'ai vû et qui était de mon âge qui les a fait venir de Bordeaux parce qu'on les



nomme sapins de Bordeaux car il y a de quatre ou cinq espèces de sapin mais le nôtre est le plus profitable parce qu'il croît et végète le mieux, c'est un grand bonheur que le sapin se soit multiplié dans ce pays ici (67).

TPC : *sapin* « toute variété de pin, surtout pin maritime ; tout résineux [par abus] ». Confusion propre au Maine (FEW 11, 214a, \**sappus*).

*truffle* n. f. « pomme de terre ».

J'ai vu le commencement des truffes ou pommes de terre les premières étaient rouges et longues elles étaient âcres puis les truffes blanches ou patates vinrent qui étaient plus douces vers le même temps vint aussi des truffes rondes que l'on appelait truffe de Lion, ensuite est venu la truffe jaune qui sont les meilleures de toutes par leur douceur il y a aussi des truffes noires mais elles rapportent peu je ne dirai rien du bien que nous fait ce fruit tout le monde le connaît dans ce pays ici (67).

TPC : *trufe/truffle* « pomme de terre ». Variante typiquement mancelle du type *truffe*, connu depuis le 17<sup>e</sup> siècle comme dénomination de la pomme de terre et répandu dans les parlers dialectaux (FEW 13/2, 385b, *tüber*).

*turquie* n. m. « maïs ».

Dans le même temps [que la pomme de terre] nous vint aussi le turquis d'abord on ne savait pas le cultiver on en plantait dans les jardins sans rien leur faire puis on essaya de mettre dans les champs et de leur faire du guéret dont on a bien réussi (67).

TPC : *turquie/teurquie* « maïs ; feuillage de maïs ». Dénomination tirée par ellipse de mfr. frm. *blé de Turquie* et répandue dans les parlers dialectaux du sud du domaine d'oïl (FEW 19, 191b, *Türk*).

*viage* (*acheter qch de qn à -*) loc. verb. tr. « acheter qch à qqn en viager, moyennant une rente viagère versée au propriétaire ».

elle demeurait à la Bourne quelle avait héritée en usufruit de M. Letertre prêtre laquelle payait 500 F de rente par M. Fouqueré qui l'avait achetée d'elle à viage (60).

TPC : *viage/viaige* « rente viagère ». Terme régionalisé dès le 18<sup>e</sup> siècle (FEW 14, 541b, *vīta* ; Chauveau 2005 : 40-43).

Ce petit ensemble lexical suffit pour localiser dans l'espace oïlique son auteur comme un manceau. Mais on peut noter que, de cette quarantaine de mots, une dizaine sont des régionalismes du français (*angevine, assemblée, beurrée, bourg, busse, charte, closerie, place, sacriste, viage*) et pouvaient avoir aux yeux de notre rural une légitimité certaine. Une autre dizaine (*affronter, ballet, courée,*

*fouasse, importation, joindre, se racheminer, raisons, rhabiller, truffe*) a été attestée en français plus ou moins longtemps, souvent même jusqu'au 17<sup>e</sup> siècle, de sorte qu'on ne peut exclure que ces mots s'employaient encore dans le français de la province du Maine au temps de Simon. De la vingtaine d'unités qui restent, seule une minorité se distingue formellement de son équivalent français (*ageter, biguenotte, guérou, huge*) ou bien n'a pas d'équivalent formel en français (*berluce, érussée, rabatrie, ribergère*) et était donc susceptible de permettre à ses contemporains de caractériser notre auteur comme un dialectophone s'évertuant à écrire en français.

Cet ensemble de traits morpho-syntaxiques et lexicaux, au fond, n'est pas très fourni numériquement et il n'est pas nécessaire de les identifier tous correctement pour comprendre le texte. Ce sont seulement des indices qui permettent de caractériser la langue première et spontanée de l'auteur, qu'il a apprise à l'intérieur de sa famille, qu'on utilise couramment dans son entourage et qu'il emploie lui-même très probablement avec ses familiers. Ces marques linguistiques affectent la prononciation, la morphosyntaxe et le lexique et permettent de localiser géolinguistiquement la langue de l'auteur comme celle d'un locuteur du parler d'un village manceau du sud de la Sarthe, très probablement le parler de La Fontaine-Saint-Martin où Louis Simon a passé toute sa vie. Mais ces quelques indices sont englobés et inclus dans une langue qui ne se distingue guère du français commun utilisé par les hommes du 18<sup>e</sup> siècle. Car il ne s'agit pas d'une langue mixte, d'un texte qui aurait été pensé en patois et calqué en français, mais bien d'un texte écrit en français par quelqu'un qui avait bien conscience d'écrire, imparfaitement sans doute, la langue normée de son époque. Les indices repérés ne sont que des approximations de ce français qui n'en modifient pas la nature.

La question que pose un tel texte, c'est celle de savoir comment un autodidacte, qui n'a ni fréquenté l'école ni pratiquement quitté son village, peut avoir acquis une telle maîtrise active du français. La personnalité de l'individu, remarquable par le simple fait de vouloir raconter une vie presque tout entière confinée dans l'horizon villageois et dépourvue de grands événements et de perspectives, y a probablement beaucoup de part. Ses dons devaient le faire sortir du commun, mais ils n'étaient pas tels qu'ils l'aient conduit à quitter son milieu natal. Ce qui le distinguait surtout de son entourage, ce devait

être sa compétence passive et active du français écrit, et non sa connaissance du français. Dans la mesure où, ce français, il se l’était acquis par ses contacts réguliers avec le milieu ecclésiastique, cela témoigne de la présence permanente, à l’intérieur d’un tout petit village, d’individus porteurs de la culture savante et porteurs du français qu’ils utilisaient constamment dans leurs échanges verbaux avec les autres villageois. Cela implique que tous les villageois étaient exposés à entendre parler français et à mener des conversations avec des francophones, donc à posséder au minimum la compétence passive du français. On peut valablement supposer que, déjà sous l’Ancien Régime, le français ne fonctionnait pas comme une langue autre pour les ruraux du domaine d’oïl et particulièrement, parmi ceux-ci, pour les ruraux de l’ouest de la France.

De tels contacts, qui ont été certainement effectifs pendant des siècles, ne peuvent pas ne pas avoir eu de répercussions sur les usages linguistiques. Il s’ensuit que, si, pour les besoins de la recherche, nous ne cessons de sérier et de catégoriser les faits comme dialectaux, régionaux, populaires, communs ou standard, etc., il importe d’envisager que ni les locuteurs ni les prises de parole n’étaient enfermés dans ces catégories et que les faits distingués circulaient entre les différents niveaux d’une même langue. Il est également nécessaire d’accorder aux ruraux de l’Ouest français du 18<sup>e</sup> siècle une compétence linguistique polynomique.

## Bibliographie

- Brunot, Ferdinand (1924). *Histoire de la langue française des origines à 1900. Tome 4, La langue classique (1660-1715), Deuxième partie*. Paris : Armand Colin.
- Chauveau, Jean-Paul (2005). "Régionalismes et dialectalismes : quelques exemples manceaux", in Martin-Dietrich Glessgen & André Thibault (éd.), *La lexicographie différentielle du français et le Dictionnaire des régionalismes de France. Actes du colloque en l'honneur de Pierre Rézeau pour son soixante-cinquième anniversaire (Strasbourg, Université Marc Bloch, 20-22 juin 2003)*. Strasbourg : Presses Universitaires de Strasbourg, pp. 31-44.
- DMF = Dictionnaire du Moyen Français. ATILF - Nancy Université & CNRS. Site internet : <http://www.atilf.fr/dmf>.
- DRF = Rézeau, Pierre, éd. (2001). *Dictionnaire des régionalismes de France. Géographie et histoire d'un patrimoine linguistique*. Bruxelles : De Boeck.
- Dottin, Georges (1899). *Glossaire des parlers du Bas-Maine (département de la Mayenne)*. Paris : Welter.
- Dottin, Georges / Langouët, J. (1901). *Glossaire du parler de Pléchâtel (canton de Bain, Ille-et-Vilaine)*. Rennes / Paris : Plihon et Hommay / Welter.
- DuPineauR = Rézeau, Pierre, éd. (1989). *Dictionnaire angevin et françois (1746-1748) de Gabriel-Joseph Du Pineau. Edition critique d'après Paris, Bibl. nat., nouv. acq. fr. 22097*. Paris : Klincksieck.
- FEW = Wartburg, Walther von (1922-2002). *Französisches etymologisches Wörterbuch. Eine darstellung des galloromanischen sprachschatzes*. 25 vol. Leipzig / Bonn / Bâle: Schroeder / Klopp / Teubner / Helbing & Lichtenhahn / Zbinden.
- Fillon, Anne (1984). *Louis Simon, étaminier, 1741-1820, dans son village du Haut-Maine au siècle des Lumières (2 vol.)*. s.l.
- Fillon, Anne (1996). *Louis Simon, villageois de l'ancienne France*. Rennes: Editions Ouest-France.
- Gougenheim, Georges (1974). *Grammaire de la langue française du 16e siècle. Nouvelle édition, entièrement refondue*. Paris: Picard.
- Grevisse, Maurice/ Goose, André (2008). *Le bon usage*. 14<sup>e</sup> éd.; Bruxelles: De Boeck-Duculot.

- Kristol, Andres (2009). *Textes littéraires et sociolinguistique historique: quelques réflexions méthodologiques*. In : Aquino-Weber, Dorothee / Cotelli, Sara / Kristol, Andres (ed.): *Sociolinguistique historique du domaine gallo-roman. Enjeux et méthodologies d'un champ disciplinaire émergent. Actes du colloque de Neuchâtel, 8-9 juin 2007*.
- Lepelley, René (1974). *Le parler normand du Val de saire (Manche). Phonétique, morphologie, syntaxe, vocabulaire de la vie rurale*. Caen: Musée de Normandie.
- Moignet, Gérard (1973). *Grammaire de l'ancien français. Morphologie, Syntaxe*. Paris: Klincksieck.
- Montesson, Charles Raoul de (1859). *Vocabulaire du Haut-Maine. Nouvelle édition augmentée*. Le Mans / Paris: Dehallais, Du Temple et Cie / Dumoulin.
- SavBr 1741 = Savary des Bruslons, Jacques (1741). *Dictionnaire universel de commerce contenant tout ce qui concerne le commerce qui se fait dans les quatre parties du monde...* ; nouvelle éd. ; 3 vol. ; Paris : Veuve Etienne.
- TPC = Bertin, Serge / Beucher, Dominique / Leprince, Jean-Pierre (2004). *Trésor du parler cénoman*. Le Mans : Editions Cénomane.

*Fernand carton*  
*Professeur émérite, Université de Nancy*  
*fernand.carton@free.fr*

## **L'exploitation et la valorisation des données de terrain dans le domaine linguistique picard**

Cette communication est en continuité avec celle que j'ai présentée au Colloque de Lyon 3 en 2006. J'y discutais des limites de l'espace langagier picard, particulièrement vaste et hétérogène. Son *Atlas* peut entrer dans le *Livre des records*. Comme l'écrivait Marie-Rose Simoni en rendant compte du volume I, il a « une longue histoire » : des « nécessités » et des « contingences » ont lourdement pesé sur sa réalisation. Dernier-né de la collection, il fut l'un des premiers à entamer des enquêtes : Robert Lorient (dès 1941), Raymond Dubois (dès 1942) ont répondu à l'appel d'Albert Dauzat. Le *Questionnaire définitif* (1960) permit de lancer sur le terrain une équipe dont j'ai fait partie avec Maurice Lebègue et Claude Deparis. Après le décès de ce dernier et des deux promoteurs, le travail de terrain a été poursuivi avec un questionnaire restreint par Fernand Carton et Maurice Lebègue, désignés en 1982 par le Comité directeur du GRECO 9 (CNRS) avec mission d'achever les enquêtes et de publier l'*ALPic I* (Carton et Lebègue, 1989). Étant donné l'urgence, nous avons dû constituer un *Questionnaire restreint* (réduit au ¼). En 1997 paraissait le volume II, qui portait le total des cartes publiées à 660. La fig. 1 présente la carte du domaine linguistique picard telle qu'on peut en tracer approximativement les limites à partir de l'atlas.

Il me semble utile d'esquisser un bilan des travaux réalisés depuis dix ans. Je me propose de montrer la diversité des types d'exploitation des données de terrain, les divers types de valorisation ; les difficultés rencontrées qui sont de plusieurs ordres et enfin d'esquisser des perspectives d'avenir pour la dialectologie picarde.

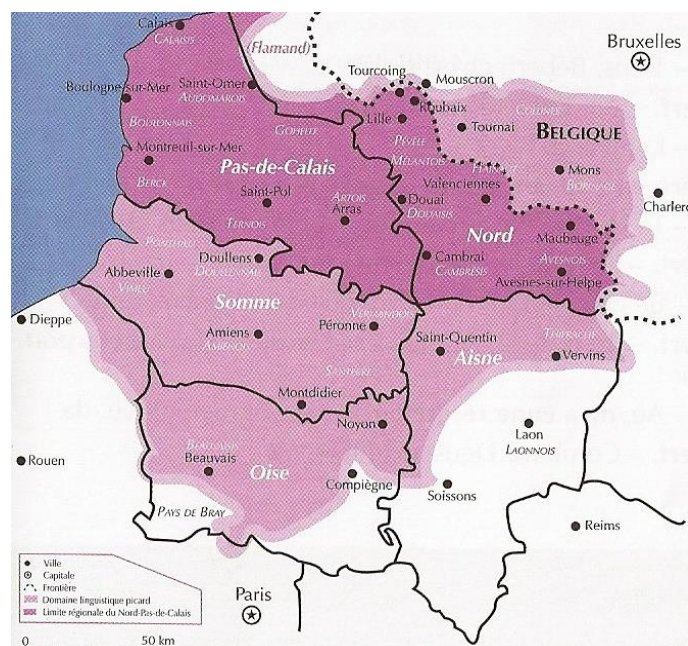


Fig. 1. Le domaine linguistique picard

## 1. Récolte en attente

Le *Questionnaire définitif* (Dubois 1960) comportait 4558 questions dont 1500, soit un peu moins du tiers, sont marquées comme prioritaires. La difficulté qu'il fallait affronter d'abord pour confectionner les cartes était l'hétérogénéité des récoltes car les enquêteurs n'ont pas pu se concerter suffisamment, ce qui posait des problèmes de comparabilité. Certains cahiers présentent des lacunes, surtout pour la partie sud ; nous avons pu travailler avec en moyenne 1100 réponses par point retenu. Nous avons publié 660 cartes, mais il reste environ deux tiers des matériaux recueillis qui n'ont pas encore été publiés. Une grande partie des compléments marginaux est toujours inédite. Il est très souhaitable d'archiver les *données ethnographiques* concernant une vie rurale qui disparaît. Si une opportunité se présentait, une équipe pourrait envisager une publication des données non publiés sous la forme d'un *ALPic III*. Je suggère qu'il soit limité au Nord-Pas-de-Calais, car les cahiers d'enquêtes y sont moins lacunaires. Le Centre d'études picardes (dirigé par J.M. Eloy, Université d'Amiens) dispose des photocopies de la totalité des *cahiers d'enquête*.

## 2. Interprétation de cartes *ALPic* publiées

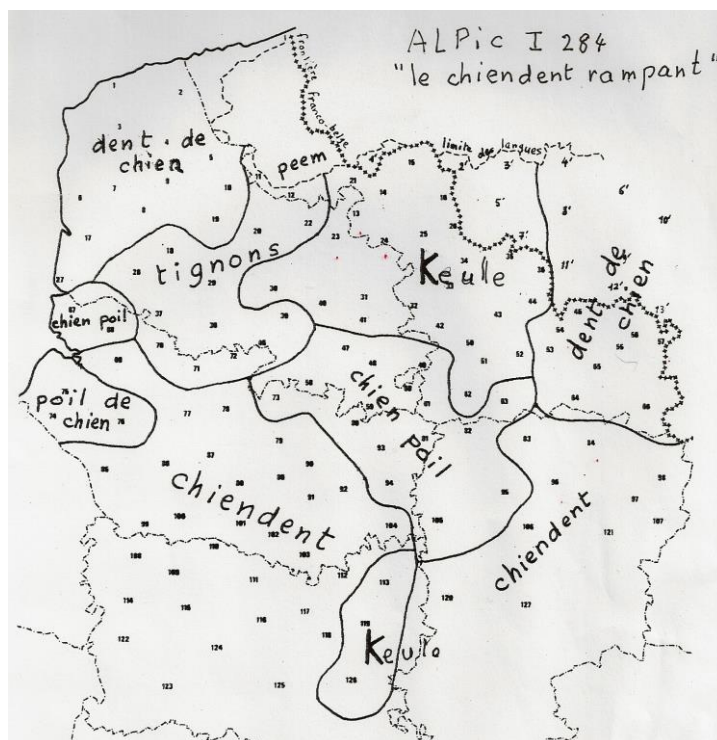


Fig. 2. Interprétation de la carte *ALPic* 284 :  
(Le) *chiendent rampant*

Indépendamment des travaux remarquables menés sur le picard de Belgique, les analyses et interprétations qu'a suscité l'*ALPic* sont de divers types :

- l'interprétation (Carton 1987) de la carte 284 « *chiendent* » (petit et grand) révèle en particulier une butte témoin (fig. 2) qui n'apparaissait pas dans *ALF* 278 : la survivance de l'ancien picard *queule* ; les variantes invalident le classement sous l'étymon CAUDA (FEW 3, 534).
- la carte 81 « *essieu* » confirme l'hypothèse de Fouché (1969 p. 316) : « L'ancien français *aissieu* est une forme dialectale probablement picarde, refaite sur le pluriel *aissieus* » (Carton 1987).



- pour la morphologie, la carte 653 « un peu » montre l'extension, par rapport à la carte correspondante de l'ALF, du type *un molé / un moné* (Carton 2005).
- la comparaison des cartes 188 « chat » et 191 « coq » montre comment se présente des solutions préventives à un conflit homonymique (Carton 1992).
- une approche ethnologique a été menée à propos des noms des vents (Carton 1993).
- l'approche phonologique est représentée par l'étude de 30 cartes où se constatent des palatalisations consonantiques secondaires : c'est une contribution à une théorie des correspondances dialectales (Dawson 2006).

### 3. Indexation de l'ALPic

Notre équipe considère ce travail comme un indispensable « service après-vente ». Le Comité scientifique du Groupe de recherche dit « des atlas » (GDR 9 du CNRS), suivant les recommandations de Pierre Gardette et de Georges Straka, avait souligné la nécessité d'une interprétation linguistique des données fournies par les atlas. Il décida le 4 décembre 1990 de convoquer une réunion de travail afin de fixer les buts et les méthodes d'une indexation. Cette réunion se tint à Nancy le 12 avril 1991. La discussion porta sur le choix entre deux conceptions : 1° index « légers » (minimaux, clés d'atlas) pouvant être réalisés rapidement (comme ceux de Gérard Taverdet) ; 2° banque de données permettant une interprétation scientifique (comme celle que réalise pour l'occitan l'équipe de Jean-Philippe Dalbera). Le principe d'une indexation unique permettant la réalisation des deux types d'index a été posé et j'ai travaillé en ce sens<sup>1</sup>.

André Cherblanc, technicien du GDR 9, avait numérisé en signes phonétiques d'ALPic I (1989) en adaptant le logiciel Access et j'ai procédé à une première lemmatisation : c'est la version zéro, qui

---

<sup>1</sup> Merci à celles et ceux qui nous ont soutenus dans une entreprise longue et parfois ingrate : Guylaine Brun-Trigaud, Jean-Paul Chauveau, Jacques Landrecies, Jacqueline Picoche, Marie-Rose Simoni.

comporte 2354 entrées (2004). *ALPic II*, lui, avait été réalisé à partir de la saisie informatique à laquelle j'avais procédé en 1993-1994. Pour les besoins de sa thèse, Alain Dawson avait récupéré ces données dans un format standard en les transcodant en API ; ce travail lui a permis de réaliser une pré-indexation automatique qui lui a ensuite servi de base de travail pour la lemmatisation du volume II. Après regroupement des formes, il a abouti à 3014 entrées. Puis Alain Dawson a mixé et lissé les entrées des deux volumes. La procédure informatique de saisie n'a pas permis de rentrer les formes notées dans les marges des cartes ; une version ultérieure pourra les intégrer. L'index cumulatif a été publié en 2009 par le Centre d'études picardes (Laboratoire LESCLAP/CNRS, Université Jules Verne, Amiens), dirigé par Jean-Michel Eloy. Chaque entrée est suivie de la référence grammaticale, du titre de la carte, du numéro du volume et de la carte où figure la forme fournie par les témoins, en un ou plusieurs points d'enquête. Suit entre parenthèses la référence au FEW. Exemple : *éclichoire* n.f. « seringue à eau » II 573 (FEW 17, 153a slitan). Var. : *réclichoire*, *églichoire*, *équichoire*, *glichoi* (n.m.).

Les références au FEW<sup>2</sup> ont aidé à la lemmatisation, et peuvent éclairer les homonymies. Il s'agit pour beaucoup de mots, de propositions de classement ; ce travail est à compléter. Par exemple, la série « veillée » en picard représente le latin *sero*.

Pour le mot signifiant « grommeler », nous choisissons l'entrée *berteler* et non *berdeler* car l'étymon est **brittus**. Pour [obéryo] « épervier », que le FEW classe sous le moyen-néerlandais **hobben**, l'entrée est *hobériau*. Les difficultés rencontrées ont été de plusieurs ordres, on peut consulter Carton et Brun-Trigaud (2003). Pseudo-lemmes : la forme en question n'est pas lemmatisée, elle est provisoirement notée en transcription graphico-phonique (ex. *ou* équivaut à [u]).

Afin de faciliter les comparaisons avec le français standard et les autres variétés d'oïl, les deux « lemmatiseurs » ont généralement noté les formes « à la française ». Ils se sont néanmoins inspirés autant que possible des principes du système de notation Feller-Carton, qui permet d'opter pour certaines graphies de l'ancien picard.

---

<sup>2</sup> C'est une équipe composée de Roger Berger, Jacques Chaurand, Denise Poulet et moi-même, qui propose ces références.

Par exemple, l'ancien picard *qu* ou d'autres : *cloque* « cloche », *quên* « chêne », *trau* « trou » mais *tro* « trois » (Dawson 2002).

Conçu au départ comme une contribution à une base de données atlantographiques, il sera mis en ligne, à la disposition de la communauté scientifique internationale. Il comporte 3055 lemmes et pseudo-lemmes sous lesquels sont classées les variantes. Cette version 1.0 en appelle d'autres, plus complètes. Mais si on prend en compte les variantes, cet index rassemble près de 8000 formes, ce qui en fait le plus copieux document lexicographique picard moderne.

#### 4. Archives sonores

Un colloque dont Jacqueline Picoche avait pris l'initiative a révélé l'abondance des enregistrements magnétiques de picard et en même temps leur médiocre signalisation et leur sous-exploitation linguistique et phonétique (Carton 1978). Ce riche patrimoine sonore était conservé dans des cercles restreints ou privés. Les fonds de picard parlé, collectés sans directive ni vue d'ensemble, étaient disparates et dispersés. Nous avons participé à des actions de l'Association française des détenteurs de documents audiovisuels et sonores (AFAS)<sup>3</sup> pour remédier à cet état de choses.

Le travail de repérage est maintenant achevé pour l'essentiel. On dispose d'un répertoire de fonds sonores de divers types : textes littéraires lus par leurs auteurs, récits préparés ou improvisés, enquêtes de terrain, conversations dirigées, chansons ; pièces de théâtre, émissions radiophoniques. Originaux ou copies sont consultables au Centre d'études picardes, à l'Université d'Amiens, là où a travaillé dès les années 70 une équipe animée par des militants comme René Debrie, qui a fait prendre conscience de l'intérêt du patrimoine sonore et de la nécessité d'une sauvegarde aux normes professionnelles.

Un catalogage étendu à l'ensemble du domaine était nécessaire.

L'ensemble des problèmes d'archivage, notamment de leur modernisation, a été discuté à l'Atelier du LESCLAP (Université

---

<sup>3</sup> Constituée en 1979 à l'initiative de responsables de fonds d'archives sonores (notamment Marie-France Callas.

d'Amiens, 23 mai 2008)<sup>4</sup>. Des fiches documentées sont rédigées selon des normes professionnelles par Delphine Vigneux : les métadonnées sont parfois très incomplètes pour des enregistrements anciens et la concertation devrait progresser avec le soutien de la Région Picardie dont les priorités sont dictées par la politique linguistique.

Aucun support actuel ne garantit une conservation pérenne : il faut prévoir à terme le repiquage sur nouveaux supports pour garder une bonne qualité acoustique.

Certains militants régionalistes veulent verser dans le fonds d'archives sonores tout ce qu'on a enregistré. Mais un océan de données mal archivées ou non signalées serait inexploitable. Pour une exploitation et une valorisation scientifique, des choix s'imposent dès maintenant selon des critères de qualité à définir. Il faut clarifier la question de la destination des documents oraux et considérer l'évolution des demandes. Les modalités du questionnement et les exigences techniques diffèrent si l'on enregistre des entretiens épilinguistiques, à visée ethnographique ou linguistique. Un corpus de conversation dirigée doit répondre à des critères différents selon qu'il s'agit d'une étude syntaxique ou d'un ethnotexte.

La qualité technique des prises de son en milieu naturel permet des recherches sur la syntaxe d'échanges verbaux, mais généralement pas une analyse phonétique fine. Une étude spectrographique de timbres vocaliques nécessite un enregistrement plus soigné qu'un tracé prosodique à l'aide d'un logiciel comme Winpitch, qui permet d'analyser le rythme d'un récit quasi spontané à partir de fragments longs. Le « faire savoir » est aussi important que le « savoir-faire ». Nous avons pu commencer la mise en ligne de 36 heures de récits, de dialogues et d'enquêtes de terrain dans le Nord. Cette mise à disposition de la communauté scientifique et d'un large public a été réalisée grâce au Centre de ressources pour la description de l'oral (CRDO/CNRS ; responsable Michel Jacobson). Un script (notation Feller-Carton) et une traduction complète étaient demandés, et n'ont pu être réalisés qu'en partie. On manque de critères internationaux pour ce travail long et ingrat. Des problèmes de copyright se sont posés et ont été résolus par des conseillers du CNRS. Ce fonds est consultable sur :

[http://crdo.risc.fr/archive\\_ouverte/picard](http://crdo.risc.fr/archive_ouverte/picard)

---

4 « Les archives sonores parlées et leur collecte : développement et modernisation des ressources ».

Ils sont également consultables sur le site du Ministère de la Culture (texte, traduction, écoute) :  
 corpusdelaparole.in2p3.fr

Ces archives sonores mériteraient d'être mieux connues. La signalétique devrait être améliorée. Une ressource mal balisée est forcément sous-utilisée.

## 5. Recherches en cours

L'archivage numérique de toutes les monographies concernant le picard moderne a été entrepris par le Centre d'études picardes (Université d'Amiens). On dispose aussi depuis 2008 de la riche Bibliothèque Raymond Dubois, ancien directeur de l'*ALPic* et du *Thesaurus picardicus*. Jean-Michel Eloy a lancé en 2007 le programme de numérisation appelé PICARTEXT. À partir d'une vaste sélection de textes du XVII<sup>ème</sup> siècle à nos jours, une équipe du LESCLAP procède actuellement à l'exploitation informatique de textes picards par un étiquetage qui aura des applications dans divers domaines : morpho-syntaxe, lexicque et phonologie. Voici un exemple d'étiquetage morphosyntaxique codé (XML, figure 3) :

**Exemple d'étiquetage morpho-syntaxique en XML**

**Fin contint d'li, ch'bétail il adjitoait z'érelles pis i r'partoait.**  
(Jean Binand : Contes picards « A ch'tchoin d'nou fu », 2000, 117 p. Editions La vague verte)

Fin	→	<code>&lt;adv&gt;Fin&lt;/adv&gt;</code>
contint	→	<code>&lt;adj&gt;contint&lt;/adj&gt;</code>
d'	→	<code>&lt;prep&gt;d'&lt;/prep&gt;</code>
li,	→	<code>&lt;pron type="pers"&gt;li&lt;/pron&gt;</code>
ch'	→	<code>&lt;det type="art"&gt;ch'&lt;/art&gt;</code>
bétail	→	<code>&lt;nom&gt;bétail&lt;/nom&gt;</code>
il	→	<code>&lt;pron type="pers"&gt;il&lt;/pron&gt;</code>
adjitoait	→	<code>&lt;verbe&gt;adjitoait&lt;/verbe&gt;</code>
z'	→	<code>&lt;det type="art"&gt;z'&lt;/det&gt;</code>
érelles	→	<code>&lt;nom&gt;érelles&lt;/nom&gt;</code>
pis	→	<code>&lt;adv&gt;pis&lt;/adv&gt;</code>
I	→	<code>&lt;pron type="pers"&gt;i&lt;/pron&gt;</code>
r'partoait	→	<code>&lt;verbe&gt;r'partoait&lt;/verbe&gt;</code>

Un autre programme est en cours, qui concerne l'oral. Une enquête approfondie sur la variation phonétique contemporaine est menée dans les régions Nord-Picardie : c'est le projet PhoNoPi mené dans le cadre de la Base de données sur la variation du français oral contemporain dans l'espace francophone (PFC).

## 6. Approche sociolinguistique

Les parlers des « gens du Nord » jouissent depuis quelques années d'un engouement encourageant pour les études picardes. Le substrat dialectal amuse et intéresse le grand public grâce au succès du film de Dany Boon, *Bienvenue chez les Ch'tis* (2008). Cette « chtimania » ravit ceux qui ont à cœur de faire mieux connaître leur identité et d'affirmer leurs valeurs. C'est un « effet collatéral » positif, quel que soit le jugement qu'on peut porter sur ce film.

Mais la promotion du *chti* a commencé depuis plus de vingt ans. Notre étude sur ce désignatif emblématique (Carton 1986) montrait qu'il était lié à un besoin. Né en 1916, dans des régiments reconstitués après la bataille de la Marne, ce sobriquet est formé de trois morphèmes jugés depuis longtemps caractéristiques du picard (*che/le, ti/toi, mi/moi*).

La thèse de Pooley (1995, en anglais), qui utilisait les travaux de William Labov et mes enquêtes dans la région de Lille, avait déjà pour titre *Chtimi : les vernaculaires urbains dans le Nord de la France*. Puis un an avant la sortie du fameux film, il y eut le succès considérable obtenu par *Le chtimi de poche* (Dawson 2007).

Nous observons actuellement, par rapport à la langue régionale, deux attitudes contraires, intéressantes à approfondir pour des sociolinguistes.

D'une part, une tendance à la *dé-standardisation linguistique*. Le « chtimiphone » est a priori rebelle à toute forme de standardisation et de norme graphique. Le *chti* n'est pas un taxon : c'est du picard nordiste francisé qui tient à prendre ses distances avec le picard de la Somme ou de l'Aisne, considéré comme « un autre patois ». Le *chtimiphone* appuie sur son identité propre et refuse d'être identifié avec le *picardophone* du sud. La différence est plus historique et culturelle que linguistique. C'est pourquoi le parler *chti* tend vers une certaine déstandardisation. Alors que le *chti* insiste volontiers sur la

diversité des parlers : « *Chez nous-outes in dit pos comm' cha* ». On pense parfois que les dialectes actuels sont trop francisés pour qu'on puisse en faire une étude linguistique intéressante. Cependant ces parlers composites assument, en plus de leur *fonction identificatrice*, une *fonction ludique* qui n'est pas sans intérêt scientifique. On parle « patois » entre amis ou en famille pour le plaisir de la connivence. Le *chti* sert à se différencier plus qu'à communiquer.

D'autre part, dans le sud du domaine (Somme, Beauvaisis, Vermandois), l'attitude est différente. Amiens et Abbeville sont pour les « Francs Picards » des centres de diffusion dont les animateurs proposent un *standard* écrit et une certaine *normalisation* fondée sur l'usage d'auteurs reconnus. Là, en effet, le grand public emploie le terme de « picard » pour désigner les parlers locaux. Eloy et coll. (2002) ont mené une enquête épilinguistique dans tout le domaine picard, en faisant parler les témoins de leur(s) langue(s). Ils ont fait apparaître le degré d'intégration linguistique de migrants de différentes origines en domaine picard et leurs rapports avec le picard. Kader, un jeune issu de l'immigration, déclarait par exemple : « Moi, en fait, j'ai appris le français avec le picard ». On peut juger ainsi de la prégnance, encore aujourd'hui, de ce substrat dialectal.

La vitalité du *picard* et du *chti-picard* est attesté par le fait que de nombreux militants, chanteurs, animateurs culturels s'expriment dans leur parler sur le site internet *Achteure*. Les Fédérations *Insanne* et *Tertous* regroupent de nombreuses associations qui assurent la promotion de leurs terroirs. Une belle revue trimestrielle est rédigée entièrement en un picard standardisé : *Ch'Lanchron* (Le Pissenlit), et valorise un riche patrimoine. Enfin l'Agence pour le picard travaille à la mise en place d'une formation à destination des bénévoles qui souhaitent participer à la promotion et à la transmission de la langue picarde en milieu scolaire, associatif et social.

## Conclusion

Ce bilan peut paraître motivé par une sorte d'autosatisfaction. Notre souci est surtout de montrer le chemin parcouru par des chercheurs qui ont dû affronter beaucoup de difficultés. En ce qui concerne l'exploitation, il conviendrait d'identifier les divers types de destinataires et d'en adapter les accès : les besoins du public curieux du patrimoine régional ne sont pas forcément ceux des personnes engagées dans des actions de politique linguistique ni ceux des chercheurs. Les programmes de valorisation devraient être évolutifs, mais, dans l'état actuel des choses, il semble nécessaire que soient menées à terme quatre actions fondamentales :

- 1° publication des données d'enquêtes encore inédites ;
- 2° entrée des données picardes dans une banque de données consacrée aux atlas linguistiques de France ;
- 3° diffusion des ressources sonores par internet ;
- 4° balisage de tout ce qui a été engrangé et amélioration de la signalétique afin de favoriser leur utilisation.

« La connaissance est le dernier recours de la nostalgie » (Yves Bonnefoy, *L'improbable*, 1959). Nous espérons avoir montré ici l'ampleur du champ de connaissances qui s'ouvre encore pour les dialectologues.



## Bibliographie

- Carton F. (1978), "Utilisation phonétique d'enregistrements dialectaux : étendue et limites". Actes du colloque *Les archives sonores et la dialectologie*, Amiens, Centre d'études picardes, p.9-24.
- Carton F. (1986), "Aux origines de *chimi*", *Mélanges offerts à Raymond Sindou*, André Chambon éd., vol. 2, p. 108-115.
- Carton F. (1987), "L'atlas linguistique picard. Analyse de la carte *Essieu* », *Linguistique picarde*, fascicule n° 102, p. 1-5.
- Carton F. (1989), "*Chiendent* dans l'Atlas linguistique picard", *Mélanges offerts à G. Tuailon*, (en coll. avec M. Lebègue), Grenoble, Ellug, vol. 2, p.17-28.
- Carton F. et Lebègue M. (1989, 1997), *Atlas linguistique et ethnographique picard*, Paris Editions du CNRS.
- Carton F. (1992), "Les noms du chat et du coq dans l'Atlas linguistique picard. Un conflit homonymique ?", *Source picarde. Hommage à René Debrie*, Amiens, Centre d'études picardes de l'Université de Picardie-Jules Verne p ; 41-47.
- Carton F. (1993), "Le vent d'Écosse en domaine picard", *Lorraine vivante. Hommage à Jean Lanher*, Presses Universitaires de Nancy, p. 301-306.
- Carton F. (1993), "Les dénominations de la crêpe d'après l'Atlas linguistique et ethnographique picard", *Linguistique picarde*, n° 125, p. 2-9.
- Carton F. et Brun-Trigaud G. (2003), "Lemmes, supralemmes... : dilemmes". Problèmes d'indexation de l'Atlas linguistique picard et de l'Atlas linguistique du Centre » in *Mélanges offerts à Xavier Ravier*, Université de Toulouse Le Mirail, Collection Méridiennes, pp. 63-72.
- Carton F. (2005), "*Un peu*. Interprétation de la carte 653 de l'Atlas linguistique picard", *Mélanges offerts au Professeur Lothar Wolf*. Lyon, Centre d'études linguistiques Jacques Goudet, Hors série 2, p. 79-94.
- Carton F. (2006), "La question des limites du picard. Avec considération spéciale des Flandres", *De Franse Nederlanden/Les Pays-Bas Français*, Rekkem (Belgique), 31st Jaarboek, pp. 87-99.
- Carton F. et Dawson A. (2009), *Index cumulé de l'Atlas linguistique et ethnographique picard*, Amiens, Centre d'études picardes.

- Dawson A. (2007), *Le chtimi de poche (Parler du Nord-Pas-de-Calais)*, Paris, Assimil.
- Dawson A. (2002), "Le picard langue polynomique, langue polygraphique", D. Caubet, S. Chaker, J. Sibille (éd.), *Codification des langues de France*, Paris, l'Harmattan, p. 85-97.
- Dawson A. (2006), *Variation phonologique et cohésion dialectale en picard. Vers une théorie des correspondances dialectales*, Thèse, Université de Toulouse II-Le Mirail.
- Dawson A. (2007), *Parlez-moi chti*, Paris, Assimil.
- Dubois R. (1960), *Questionnaire définitif de l'Atlas linguistique picard*, Dijon, Publications de l'Institut de linguistique régionale de l'Université de Dijon.
- Eloy J.-M., Blot D., Carcassone M., Landrecies J. (2002), *Français, picard, immigration. Une enquête épilinguistique*, Amiens, Centre d'études picardes.
- Fouché P. (1969), *Phonétique historique du français*, vol. 2 : Les voyelles, 2ème éd. revue et corrigée, Paris, Klincksieck.
- Pooley T. (1995), *Chtimi. The Urban vernacular of Northern France*, London, Guildhall University.
- Wartburg W. von (sous la direction de), (2003), *Französisches Etymologisches Wörterbuch*. Index du FEW en 2 volumes 1180 + 2370 pp., Paris, Honoré Champion, 2003, sous la direction d'Eva Buchi (ATILF/CNRS Nancy).

*Fabrice Jecic*<sup>1</sup>  
*UMR 8589, CNRS / Paris 1 - Panthéon-Sorbonne*  
*jecic@vjf.cnrs.fr*

**Du texte dialectal aux cartes d'atlas :  
mémoires d'un même terrain beauceron.**  
*D'après l'Heritage perdu d'André Gilbert (1920-2003).*

« Pour pas que les jeunes il l'oubye le patoué, fo pas seul'ment faire de l'«ortograf» un peu mal foutue, y fo mète les anciens mots<sup>2</sup> »... Devoir de mémoire mais aussi exigence d'authenticité, les textes en beauceron d'André Gilbert, ont paru dans *La République du Centre* entre 1980 et 1982. Leur recueil, publié sous le titre *l'Heritage perdu*, a connu un franc succès, sa 3<sup>e</sup> édition est sous presse.

La confrontation du lexique de ces textes avec celui des cartes de *l'Atlas linguistique et ethnographique de l'Île-de-France et de l'Orléanais (ALIFO)* de Marie-Rose Simoni-Aurembou nous fournit des éléments de réflexion quant à une certaine approche des textes dialectaux, non seulement dans leur relation avec le terrain et la conscience linguistique de l'auteur, mais aussi dans leur relation avec les contenus sociolinguistiques des formes graphiques qui les composent.

**Terrain beauceron, terrain dialectal : espace de vie d'André Gilbert**

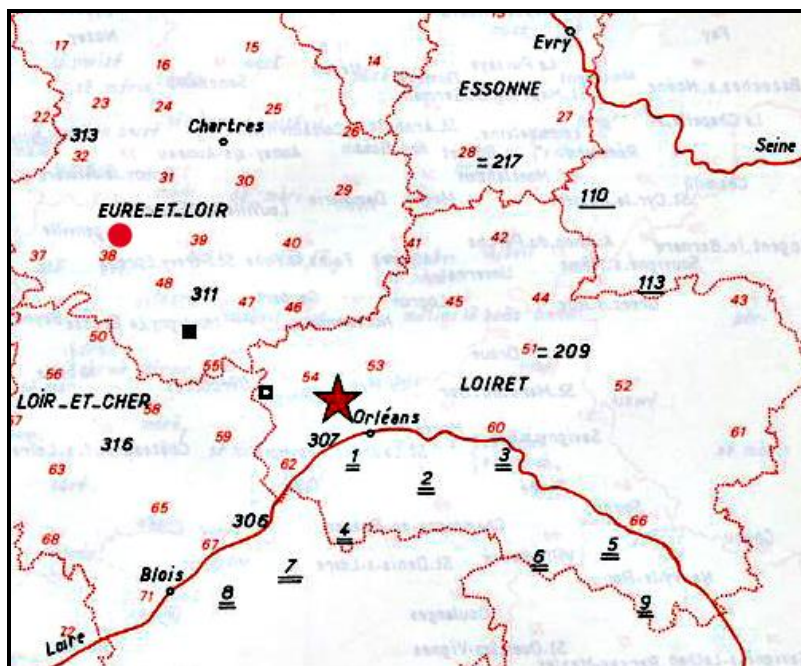
L'auteur a vécu la plus longue partie de sa vie à Bucy-St-Liphard (Loiret) légèrement au sud des points 54 (Gémigny) et 53 (Cercottes), au bord de la Nationale 157, à moins de 10 kms d'Orléans (voir, ci-

---

<sup>1</sup> Histoire des systèmes graphiques du français et de ses variétés. *Laboratoire de Médiévisiologie Occidentale de Paris (LAMOP)*.

<sup>2</sup> Extrait de l'un de ses premiers textes, *Le temps qui s'cayote* 'le temps qui se gâte'.

dessous, la localisation sur la carte, matérialisée par une étoile). Le symbole ■ (entre les points 59 et 54) situe son lieu de naissance : Épièds-en-Beauce. Le symbole ■ (entre les points 55 et 48) indique son lieu de mariage : Lutz-en-Dunois. Le rond rouge ●, près du point 38 (Unverre) localise Frazé (Eure-et-Loir) où il a commencé à exercer son métier de maréchal-ferrant.



*Extrait de la carte du domaine d'enquête de l'ALIFO  
de Marie-Rose Simoni-Aurembou*

- les numéros en rouge sont ceux de l'ALIFO ;
- les numéros en noir sont ceux de l'ALF ;
- les numéros en noir soulignés une fois sont ceux de l'ALCB ;
- les numéros en noir soulignés deux fois sont ceux de l'ALCe ;
- les numéros en noir soulignés d'un trait discontinu sont ceux de l'ALO.

André Gilbert est né le 29 octobre 1920 à Épièds-en-Beauce (Loiret). Il se marie en 1946 à Lutz-en-Dunois (Eure-et-Loir). Père de trois enfants, il a la douleur de perdre prématurément son épouse et revient s'installer à Épièds où il reprend l'atelier familial en 1950. Dans les années 60, la maréchalerie et la réparation du matériel agricole cède le pas à la vente de matériel électrique et de quincaillerie. En 1978, il ferme son commerce à Épièds et travaille encore quatre ans comme coursier dans une entreprise d'électronique

d'Orléans. C'est à partir de ce moment-là qu'il se met à écrire des textes en patois, dont une centaine a été publiée dans *Sept sur Sept*, le supplément hebdomadaire du quotidien régional *La République du Centre*, dont le tirage moyen était de 57.592 exemplaires en 1993. André Gilbert est décédé le 25 mai 2003, dans sa 83<sup>e</sup> année. Ayant fait don de son corps à la médecine, une cérémonie religieuse avait eu lieu à Épièds-en-Beauce, le 28 mai 2003, sans fleurs ni couronnes<sup>3</sup>.

## **Terrain beauceron, terrain des textes : terre d'inspiration**

Au cours d'une discussion (*Entretien*, 1993), je me souviens qu'il m'avait dit apprécier son emploi de coursier : « Tu comprends, Fabrice, c'était un boulot idiot, ça me laissait la tête libre pour écrire et avoir de l'inspiration ». Les prolongements de cet échange m'avaient conforté dans l'hypothèse que je m'étais faite sur sa manière de composer les textes. L'étude comparée des manuscrits, dactylogrammes et articles du journal m'avaient amené à identifier deux types de constructions de ses récits : l'une spontanée, l'autre élaborée.

Les textes spontanés surviennent « quand j'ai pas tous les éléments [les mots dialectaux], que je le fais à la vitesse [...] *Les mouches* [les 'abeilles'], je l'ai fait en un quart d'heure au journal [...] j'avais pas tous les éléments et je l'ai fait quand même [...] ça été fait à la galopée » (*Entretien*, 1993).

Les textes élaborés sont fabriqués à partir de thèmes porteurs et de listes de mots préétablies : « [...] je cherchais un sujet [...] que ce soit un conte, que ce soit une description de vie... des choses de la vie [...]. Je prends les *Poulets des faisances* [poulets donnés au propriétaire en sus du bail], hein, je prenais tous les mots... Je les intégrais dans le texte [...] c'est dans les textes que je veux mettre le plus de mots [...] je me faisais la liste de mots [...] Je me faisais la liste de mots dans la tête » (*Entretien*, 1993).

---

<sup>3</sup> Pour comprendre son personnage on se reportera à *L'haritage perdu* où figurent de nombreux éléments de sa biographie.

18.12.1911

### Annuï on mène au fian (1)

Quanqu'on parle d'écologie, ya 50 ans, on en fésé sans s'en rende compte, on meté pas biocou d'engrais, un peu de gouano, un peu d'cianamide et on m'né à la marne, c'éte du naturel, on avé dé bêtes et pi du fumier, çui d'cochon y l'éte trop fré y aglatisé (2) la terre, çui d'mouton et d'vaches cété l'meilleu, l'ch'vo çaté, chaud on mélangé tout, on curé les barbesio sitou parties au champ au biotemps et pi on m'né apré à la mouesson, cété l'mète charquer qui m'né avec l'atléa a 3 ch'vo, y vidé l'tombereau avec la queroué (3) tu parles si y l'aligné et pi les fésé tous pareils.  
 Ça y v'lé pas qu'on l'traite de pottrassier (4) pour chargé c'éte les

autes charquiés et les briquoyers (5) qui l'fesin ça duré défoué pu d'une semaine pour l'arquarter yavé les briquoyers et des trimards (6) a façon, un coup d'fourche a dréte, un aute à gauche pour arusé l'fête, et y lancin l'fumier tout autour, fallé bin l'aménusé pour pas qu'yé d'couape (7) pour qu'en labourant ça bourre pas dans l'couite ou la rasette.  
 A présent y sont pu généré par ça, la charue a passe bin, pu d'couamelle (8) dans l'champ a fumier d'mouton, pu d'acropi (9) blanchi sous une bousse de vache, mé aussi pu d'bon blé, d'bonne viande, ça y l'save bin, on gagne des sous d'une main et pi on les r'donne d'une aute, engrais de pu en pu, nouvio matériel pour aéré

la terre qui n'ayant pu d'ver de terre et d'venu giate.  
 On fé d'la propagande pour la culture « biologique » comme y dise, cé c'qu'on fésé, mé à présent on va payé pu cher, fère et défère et refère mé en payant cété pas bio hun les mangeux d'sous.  
 A dans huit jours.  
 D.D.

(1) Fian : fumier.  
 (2) Aglatir : rende masse.  
 (3) Queroué : crochet.  
 (4) Pottrassier : mauvais ouvrier.  
 (5) Briquoyers : manoeuvres.  
 (6) Trimards : chemineaux.  
 (7) Couape : masse serrée.  
 (8) Couamelle : coulemelle.  
 (9) Acropi : pissenik.



EN BEAUCE. — Charretier déchargeant le Fumier

Collection ND Phot

On sait que pour mettre le tombereau « à cul », le charretier enlevait la « clé » du tombereau qui retenait celui-ci dans le prolongement des limons, mais sait-on que « perdre la clé du tombereau » signifiait « avoir la colique ? » Expression très pudique, on en conviendra ! Nous devons de la connaître à Muguette Rigaud qui nous a aimablement confié cette carte postale.

Le texte dont il est question ici : *Annuï on mène au fian* ['Aujourd'hui on épand le fumier'] est de ce dernier type, élaboré. Ce que confirme la composition de l'article du journal dont le texte est encore agrémenté par la reproduction d'une carte postale ancienne montrant un charretier déchargeant le fumier. La légende précise l'acte technique avec des expressions locales. Suit le texte agrandi.

### *Annuï on mène au fian*

*Quanqu' on parle d' écologie, ya 50 ans, on en fésé sans s' en rende compte ; on metté pas biocoup d' engrais, un peu de gouano, un peu d' cianamide et on m'né à la marne, cété du naturel, on avé dé bêtes et pi du fumier, çui d' cochon y l' éte trop fré, y aglatisé la terre, çui d' mouton et d' vache cété l' meilleu, l' ch'vo çaté chaud, on mélangé tout, on curé les bargesio sitou les barbies parties au champ au bio temps et pi on m'né apré la mouesson ; cété l' mète*

*charquier qui m'né avec l' atlée à 3 ch'vo, y vidé l' tombero avec l' quéroué, tu parles si y l' aligné et pi les fésé tous pareils.*

*Car, y v'lé pas qu' on l' traite de potrassier ; pour chargé, cété les autes charquiés et les briquoyers qui l' fésint ; ça duré défoué pu d' une semaine, pour l' aquarter, yavé les briquoyers et des trimards à façon, un coup d' fourche à drète, un aute à gauche pour arusé l' fête, et y lancint l' fumier tout autour ; fallé bin l' aménusé pour pas qu' yé d' couape, pour qu' en labourant ça bourre pas dans l' coute ou la rasette.*

*A présent y sont pu géné par ça ; la charrue a passe bin, pu d' couamelle dans le champ à fumier d' mouton, pu d' acropi blanchi sous une bouse de vache, mé aussi pu d' bon blé, d' bonne viande ; ça y l' save bin, on gagne les sous d' une main et pi on les r'donne d' une aute, engrais de pu en pu, nouvio matériel pour aéré la terre qui n' ayant pu d' ver de terre, est d'venue glate.*

*On fé d' la propagande pour la culture "biologique" comme y dise, cé c' qu' on fésé, mé à présent on va payé pu cher, fère et défère et refère mé en payant, cėti pas bio, hun, les mangeux d' sous.*

*A dans huit jours.*

Le corps de l'article, en parler de l'Orléanais, est suivi de notes de renvois glosant des vocables dialectaux, ou ressentis comme tels par André Gilbert. Mais certaines expressions dialectales semblent lui avoir échappé, voir, ci-dessous, le tableau récapitulatif :

<b>Mots retenus et glosés par l'auteur :</b>	<b>Expressions non-commentées :</b>
1 - <i>fian</i> , 'fumier'	a - <i>on mène</i> , 'on transporte, on épan'
2 - <i>aglatir</i> , 'rend(r)e masse'	b - <i>glate</i> , terre ou terrain 'durci, collant'
3 - <i>quéroué</i> , 'crochet' [à fumier]	c - <i>aménusé</i> , prononcé : [amenuse], 'émietter'
4 - <i>potrassier</i> , 'mauvais ouvrier'	
5 - <i>briquoyers</i> , 'manoeuvres'	
6 - <i>trimards</i> , 'chemineaux'	
7 - <i>couape</i> , 'masse serrée'	
8 - <i>couamelle</i> , 'coulemelle'	
9 - <i>acropi</i> , 'pissenlit'	

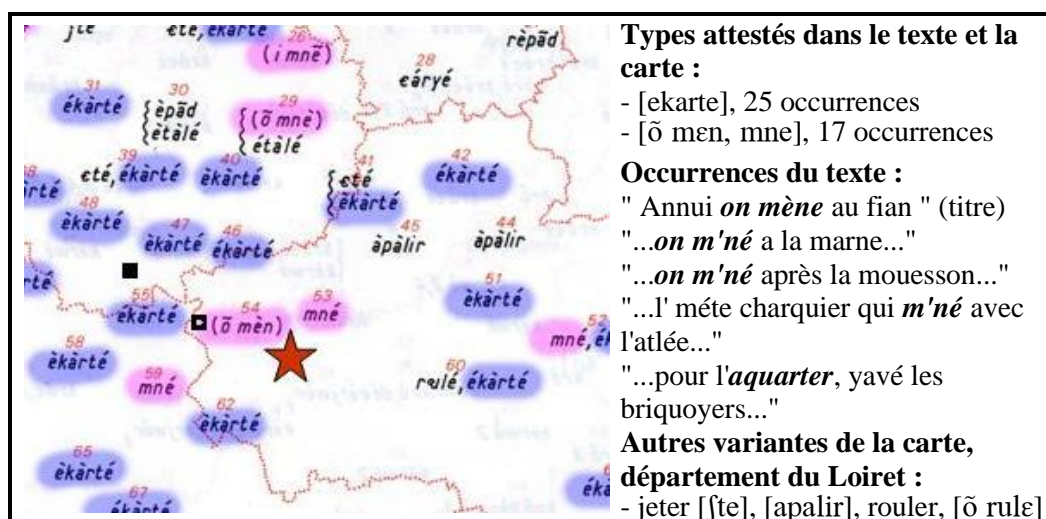
De ce lexique ne sont retenus que les mots et expressions faisant l'objet d'une carte dans l'ALIFO. *Potrassier*, *briquoyers* et *trimards* sont écartés de la présente étude. D'autre part, la question du système

graphique de l'auteur, bien qu'ayant fait l'objet d'une présentation synthétique sommaire lors de la communication orale, n'est pas du tout examinée ici. Cette question est traitée en détail, analyse et théorie, dans l'article "Images de variétés du domaine d'oïl central : dynamiques de représentations graphiques d'auteurs (1911-1997)", Jecic F. (2005), p. 219-249, voir la bibliographie.

## Mémoire du terrain, mémoire du texte

Les extraits de cartes qui suivent mettent en rapport le lexique dialectal du texte avec les attestations de l'*ALIFO*. Les formes des points 54 et 53 sont systématiquement relevées, ainsi que les formes de points proches, si elles correspondent à celles du texte. Dans tous les cas il est précisé le nombre d'occurrences des diverses variantes pour l'ensemble de l'aire de l'*ALIFO*, ce qui donne une certaine idée de l'étendue, plus ou moins importante, d'une expression dialectale donnée<sup>4</sup>.

### On mène, m'né - Carte 25 - épandre (le fumier)

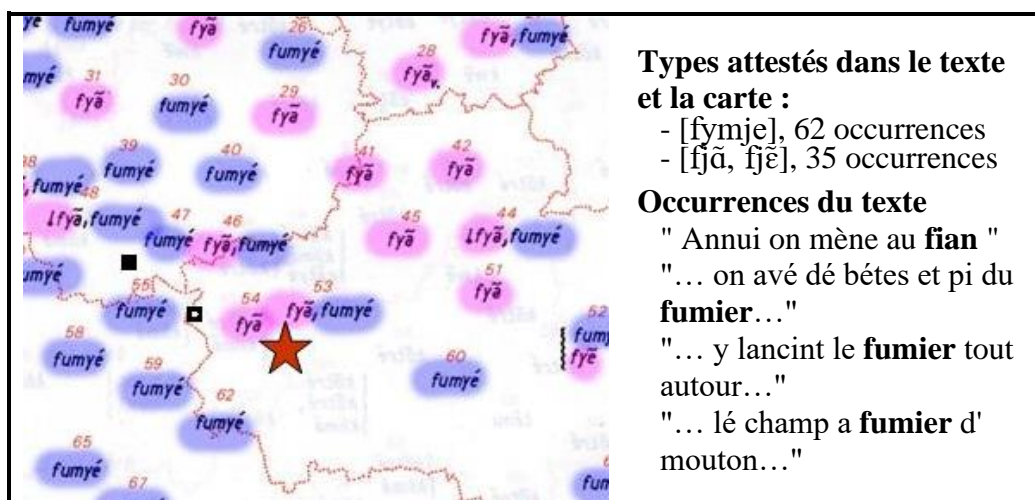


<sup>4</sup> Les transcriptions de l'*ALIFO* utilisent l'alphabet phonétique de l'*ALF* où les valeurs phonétiques correspondent généralement aux valeurs phoniques qu'elles ont dans l'orthographe usuelle (é = e fermé, è = e ouvert, etc.), le tilde ~ indique la nasalisation, le symbole € équivaut au symbole API [ʃ], le symbole °u à la valeur phonique du graphème (ou). Les notations *ALF* apparaissent dans le texte en caractères gras.



Deux expressions concurrentes sont présentes dans son texte et aussi attestées sur la carte parmi beaucoup d'autres expressions différentes<sup>5</sup> : **õ mèn**, **mné**, **mnè** (17 occurrences sur la carte) et **ékàrté**, **èkarté** (25 occurrences sur la carte). La première forme, **õ mèn**, présente dans le titre et à plusieurs endroits du texte sous les formes graphiques *on mène* ou *m'né*, est également attestée aux points 53 et 54 qui correspondent à la localisation du lieu de naissance et du lieu de résidence principale de l'auteur ; les points voisins confirment généralement les variantes **ékarté** ou **èkarté**. L'expression basée sur **mné** (4 occurrences dans le texte) a la faveur de l'auteur, contrairement à l'expression construite sur **ékarté** qui, bien qu'attestée une fois sous la forme dialectale *aquarter*, avec alternance de l'initiale vocalique de *e* en *a*, est ressentie malgré tout comme trop proche de la langue commune<sup>6</sup>. *Rem. ALIFO*, t. 1 : « "transporter" et "épandre le fumier" sont confondus en une seule expression type 'charrier le fumier', on dit 'mener au fumier', 'jeter au fumier' [à 26 points sur 75], ailleurs les verbes sont transitifs ».

### Fian - Carte 23 - fumier



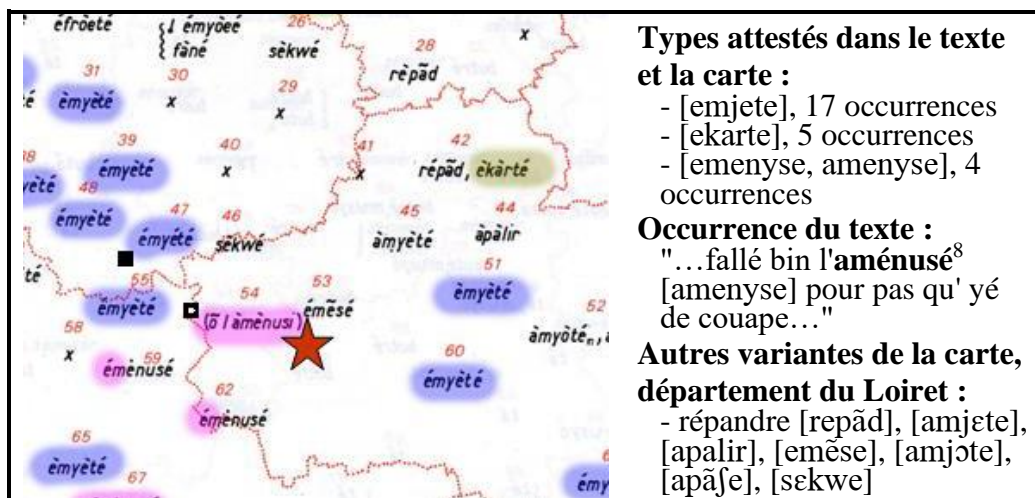
Deux formes concurrentes sont présentes dans le texte, *fian* (1 occurrence) et *fumier* (3 occurrences) ; la première forme, du fait qu'elle apparait dans le titre, profite du pouvoir symbolique que lui confère cette position. Sur la carte, la localisation de l'auteur indique au point 53 les deux formes **fyã** et **fumyé**, et au point 54 la forme **fyã**

<sup>5</sup> Expressions souvent issues de types français : *épand(r)e*, *charrier*, *répand(r)e*, *j(e)ter*, *étaler*, etc.

<sup>6</sup> Jecic F. (2005), *op. cit.*, p. 230-235.

seule. Les variantes **fyã** et **fyẽ** ne présentent que 35 occurrences contre 62 pour la forme commune **fumyé**. Dans *Lectures de l'Atlas linguistique de la France*, le type *fumier* est en expansion évidente alors que le type *fiente* est en recul partout à la fin du XIX<sup>e</sup> siècle<sup>7</sup>. Si les aires de l'Eure-et-Loir et du Loiret confirment la coexistence des deux variantes, en revanche la représentation dominante de la forme de la langue commune est liée à la présence unanime de l'occurrence **fumyé** sur les territoires des départements du Loir-et-Cher et de l'Indre-et-Loire. La *Rem. ALIFO*, t. 1 : « "le **fyẽ** est du fumier de vache en Eure-et-Loir", points 4, 33 et 46 » peut être assimilée à l'expression d'une marque d'usage puisque cette glose provient de locuteurs qui sont situés à l'extérieur du département ou à son extrême limite pour le point 46. Même chose pour la remarque suivante : « "**fyã** est un mot du Perche", point 24 » qui est exprimé par un témoin situé à l'ouest de Chartres sur la route en direction du Perche. Dans ces deux cas c'est une manière de dire que 'là-bas on parle mal'.

### Aménus[s]er - Carte 26 - bien émietter (le fumier)



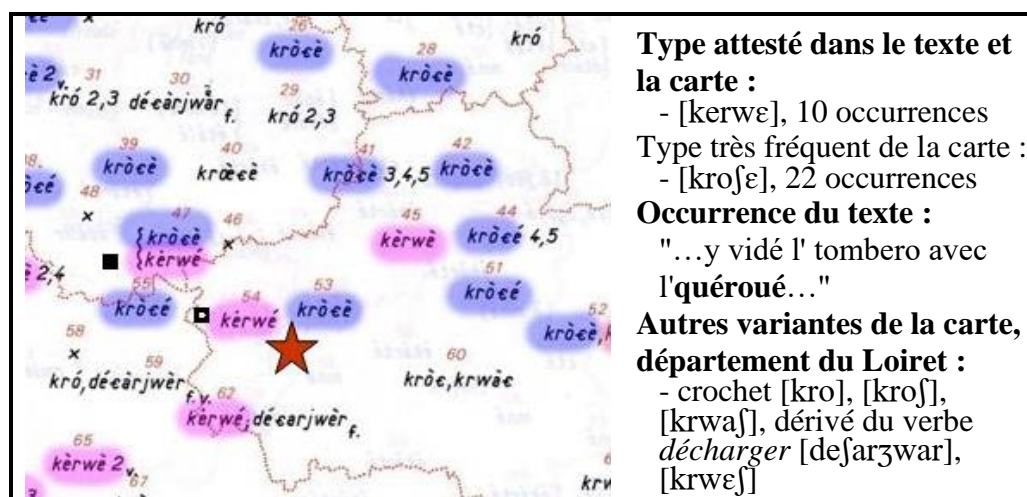
Deux expressions voisines apparaissent dans le texte sous les formes *aménusé* (prononcé avec *s* sourd, 1 occurrence) et *aquarter* (1

<sup>7</sup> Brun-Trigaud G., Le Berre Y. et Le Dû J. (2005), *Lectures de l'Atlas linguistique de la France de Gillieron et Edmont. Du temps dans l'espace*, Carte 509, ALF 618 Fumier, p. 332.

<sup>8</sup> André Gilbert inverse, ou maîtrise mal, la notation du *s* intervocalique. Presque systématiquement il inverse les règles et note -s- simple pour le phonème [s] et -ss- double pour le phonème [z], ainsi il écrit aussi dans le texte *arusé* 'arraché' prononcé [aryse], et bousse 'bousse' prononcé [buz].

occurrence), graphie dialectale du verbe *écarter*<sup>9</sup>. La carte confirme au point 54 une variante de la première forme à travers l'expression **õ l àmènuti** ; des variantes proches, présentant une alternance **a/e** (3 occurrences), sont attestées à des points voisins : **émènuté** (points 59 et 62) et **àmènuté** (un peu plus loin au point 67). *Rem. ALIFO*, t. 1 : « "aménusse le bin" disait la grand-mère de Mme R. qui mettait le fumier dans un grand tablier et l'émiettait à la main ». La carte présente aussi la forme **ékàrté** (5 occurrences), mais loin de la localisation de l'auteur, bien que celui-ci l'emploie par le biais de la forme dialectalisée *aquarter*, présentant une alternance vocalique **e/a**, relativement commune à l'aire dialectale, et déjà mentionnée plus haut. C'est en définitive la forme française **émyété**, variante **èmyété** (17 occurrences), qui est dominante dans la partie ouest et sud de la carte, les autres formes attestées dans l'*ALIFO* étant très dispersées.

### Queroué - Carte 24 - crochet à fumier



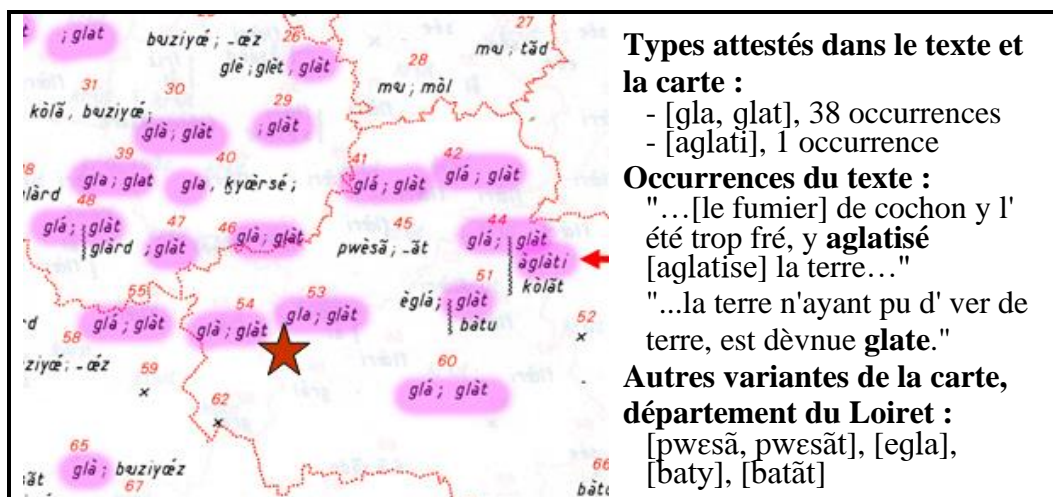
Un seul vocable dans le texte<sup>10</sup> : *queroué*, peut-être à rapprocher de 'croix', par métathèse. Cette forme est attestée sur la carte au point 54 et dans le voisinage proche de la localisation de l'auteur sous les variantes **kèrwé** et **kèrwè** (10 occurrences en tout). La carte atteste généralement des formes du type **krò**, avec chute de *r* final, variante populaire de **kròcè**, plus proches de la langue générale et présentes sur une grande partie de l'aire. Le choix d'André Gilbert indique à la fois

<sup>9</sup> On enregistre ici un sens étendu du verbe pouvant signifier 'épandre' ou 'émietter' (le fumier), cf. *Carte 25*, ci-dessus.

<sup>10</sup> André Gilbert utilise aussi une fois dans le texte le vocable de la langue générale : *fourche*. Ce mot n'a pas été pris en compte dans l'analyse car il n'apparaît qu'une seule fois sur la carte au nord du département des Yvelines, au point 13.

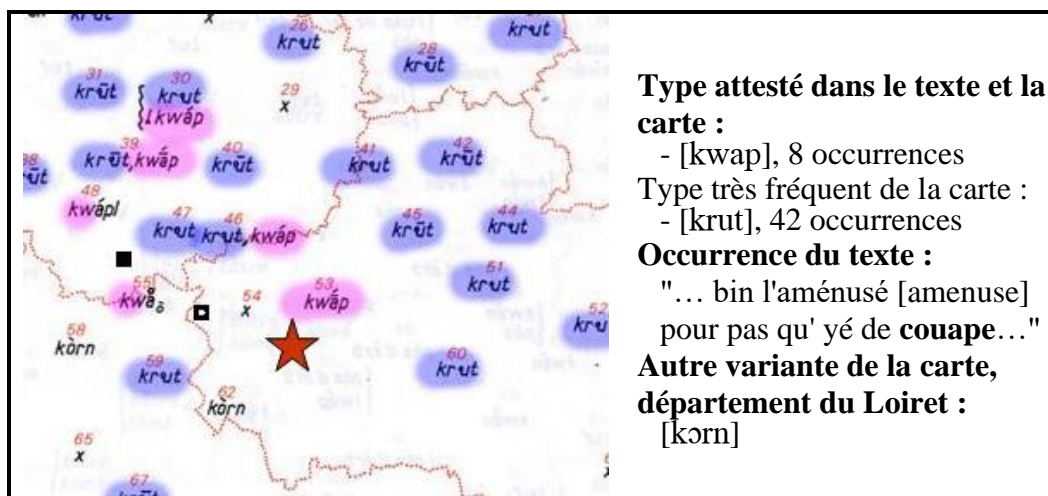
un type très localisé et en même temps le plus éloigné de la langue commune (malgré la présence dans le texte, du mot de la langue générale : *fourche*), ce qui compte pour lui, c'est de sauver le mot de l'oubli.

**Aglatir, glate -Carte 18 - (un terrain) durci, collant ;  
(une terre) durcie**



Deux formes issues d'un même type : *glate* et *aglatisé* (prononcé avec s sourd) sont présentes dans le texte, à raison d'une occurrence de chaque mot. La partie centrale de la carte confirme très largement le type *glate* avec des variantes phonétiques : **gla**, **glat**, **glâte**, **glâte** et même **glàrd** pour désigner une terre ou un terrain de mauvaise qualité. Une seule occurrence du verbe, sous la forme **àglàti** est attestée sur la carte à un endroit relativement proche de la localisation de l'auteur, au point 44. La *Rem. ALIFO*, t. 1, rapporte l'expression d'une marque d'usage : « "on dit **glèt** quand ont cause mal ; on dit **glàt** quand on cause bien", point 26 » à la limite de l'Eure-et-Loir où les deux formes sont attestées. La seconde partie de la remarque concerne l'observation d'une variante morphosyntaxique de la forme verbale : « "on laboure, **sà àglàti** la terre, ça la serre", point 45 ; "si on a été dans les champs par le mou, **sà à glàti**", point 41 » ce qui montre la difficulté à isoler des morphèmes dans un contexte où la segmentation syntaxique n'est pas évidente à réaliser.

**Couape - Carte 20 - croute de terre durcie<sup>11</sup>**

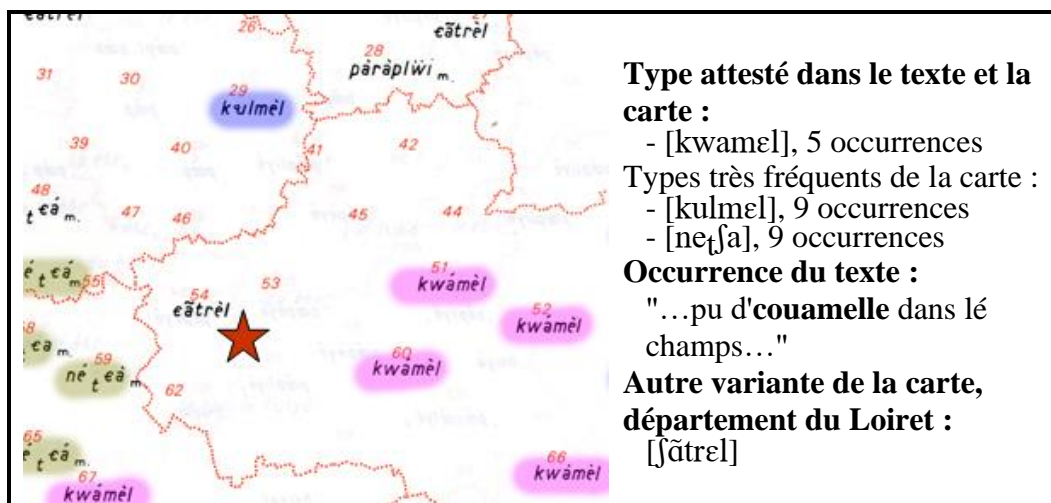


Une seule occurrence dans le texte représentée par le mot *couape*. Sur la carte apparaissent 8 occurrences du même type avec des variantes phonétiques : **kwáp** (avec parfois *a* long), **kwâp** (avec *a* entre [a] et [ə]<sup>12</sup>), **kwápl** (avec consonne finale parasite), **kwâ** (avec amuïssement de la consonne finale) et encore une forme isolée **gwáp** (avec sonorisation de la consonne initiale) au point 69. À l'exception de **gwáp** située au nord du département de l'Indre-et-Loire, à l'extrême sud-ouest de l'*ALIFO*, toutes les autres occurrences sont attestées au centre de la carte sur un axe Orléans-Chartres. Le second type **kr<sup>o</sup>ut** (avec parfois <sup>o</sup>u long), qui correspond à la forme de la langue générale, est nettement dominant sur toute l'aire de l'atlas et compte 42 occurrences. Par l'enregistrement graphique de la forme la moins répandue, André Gilbert préserve le souvenir du mot.

<sup>11</sup> Vous avez probablement été choqué(e) par l'absence d'accent circonflexe sur le mot *croute* dans le titre, il ne s'agit ni d'une coquille, ni d'une erreur, mais d'une nouvelle variante graphique recommandée par l'Académie Française dans la 9<sup>e</sup> édition en cours de son dictionnaire et dont les *Documents Administratifs* du *Journal Officiel* rendent compte dans leur n° 100, du 6 décembre 1990.

<sup>12</sup> La voyelle notée [ɜ] en API est équivalente à la notation ALF /ã/, pour une voyelle située entre [a] et [ə].

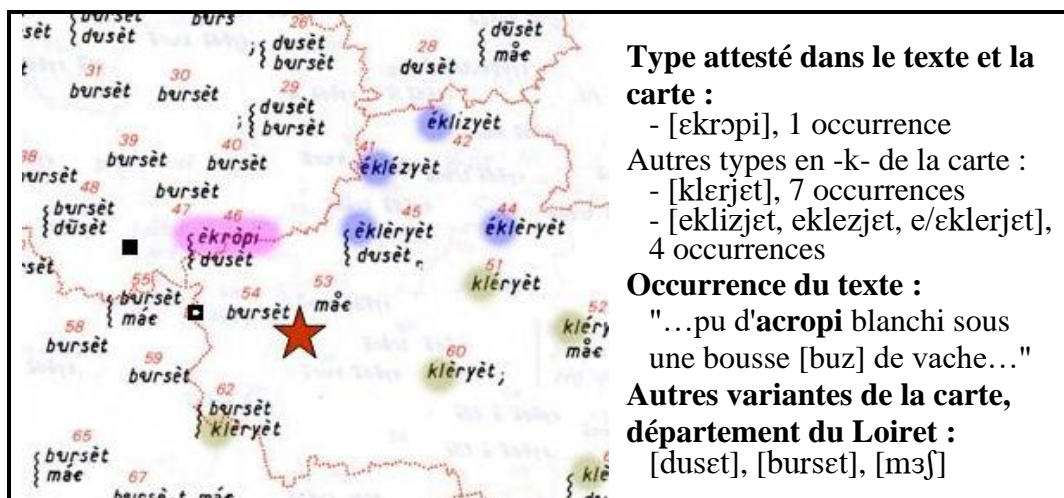
### Couamelle - Carte 311 - coulemelle



Une seule occurrence dans le texte illustrée par le terme *couamelle*. Sur la carte au point 54, lieu de l'auteur, la forme **ĕãtrɛl** est, entre autres, probablement à l'origine du commentaire de Marie-Rose Simoni-Aurembou lorsqu'elle fait observer : « sauf exception, les paysans de notre domaine connaissent mal les champignons, qu'ils n'en mangent pas car ils ont peur de s'empoisonner<sup>13</sup> ». La carte atteste 5 occurrences de **kwámèl** (avec diphtongaison) qui correspondent au mot d'André Gilbert. Cette forme s'avère extrêmement localisée puisque 4 occurrences sont situées dans le département du Loiret, plus précisément dans le Gâtinais (aux points 51, 52, 60 et 66), et que la seule autre occurrence, relativement proche de cette aire se trouve dans le département du Loir-et-Cher (au point 67). Il semble que la cohérence de cet ensemble peut être recherchée dans la continuité qu'assure la Loire entre les points 66, 60 et 67, qui sont situés à proximité de sa rive droite, de part et d'autre d'Orléans. La carte présente aussi 9 occurrences de la forme standard **k<sup>o</sup>ūlmèl** regroupées en deux sous-ensembles distincts, situés au sud-est et au nord-ouest de l'atlas. Enfin, 9 occurrences de la forme **nétɛa**, dénomination analogique dont le petit mamelon brunâtre de la partie centrale du chapeau rappelle celui d'un 'nez de chat'. Cette dernière dénomination est regroupée dans le nord du département du Loir-et-Cher et présente un apax dans le département des Yvelines, au point 8, à l'ouest de Versailles.

<sup>13</sup> En marge de la carte, l'auteure nous fait part aussi de dictons savoureux, tel celui sur la morille, dont elle nous rapporte le propos : « il faut être menteur pour les trouver », témoins des points 9 et 75.

Acropi 'pissenlit' - Carte 278 - mâche sauvage, cultivée



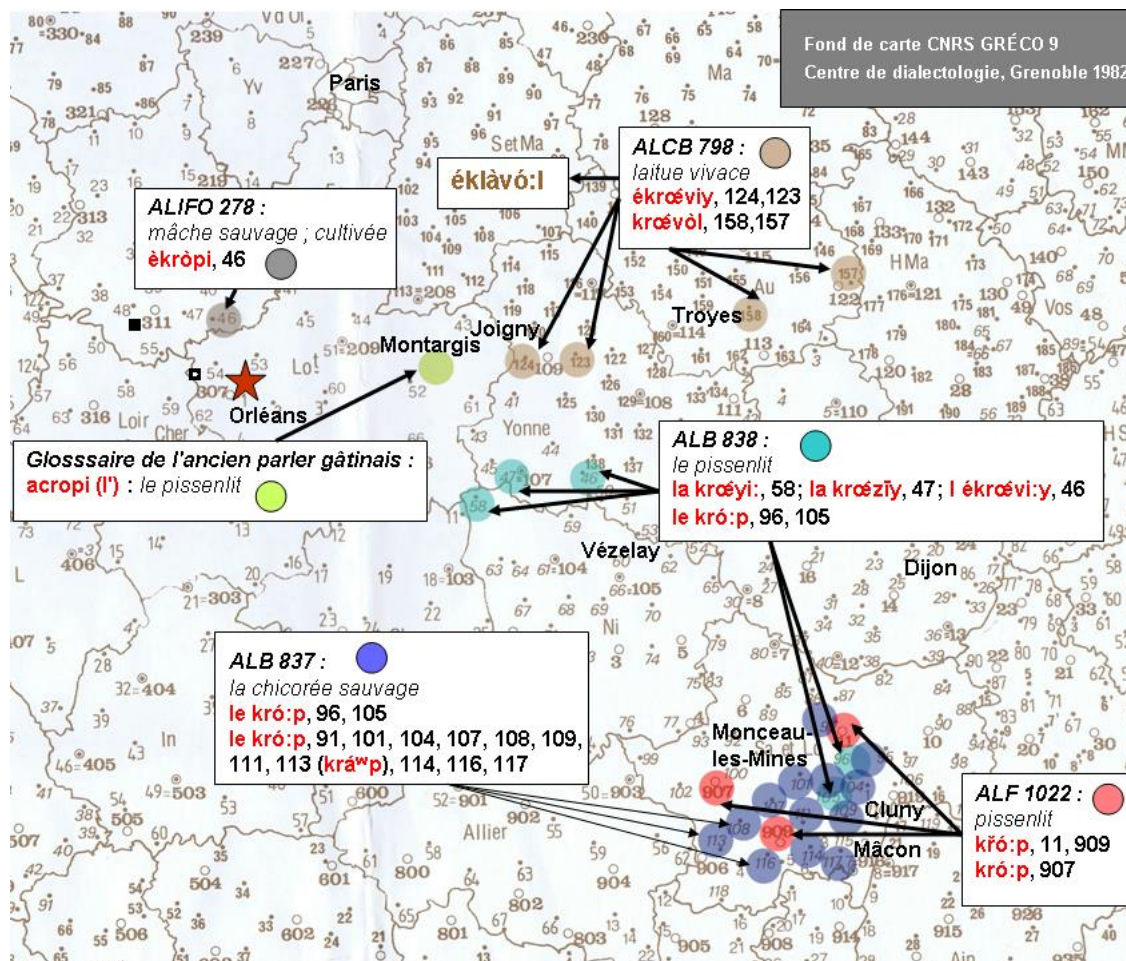
Le vocable *acropi*, 'pissenlit', une occurrence dans le texte, n'est attesté que sous la forme phonétique **èkròpi** (avec alternance a/e de la voyelle initiale) sur la carte *mâche* de l'atlas, au point 46, à quelques encablures du lieu de l'auteur. Ce mot est aussi présent dans le *Glossaire de l'ancien parler gâtinais* de Jean Jourdain et André Bouchier, et dont l'article, accessible par internet, est très sommaire : « **Acropi** (l'), le pissenlit<sup>14</sup> ». En prolongeant les investigations à l'ALF, la carte 1022 *pissenlit*, révèle, en Saône-et-Loire, le type /krop/, avec *o* fermé long, aux points 11, 907 et 909, et avec *r* uvulaire fricatif aux points 11 et 909<sup>15</sup>. Le *r* uvulaire fricatif est produit par les deux témoins les plus âgés, respectivement la 60<sup>aine</sup> et 73 ans, contrairement au jeune témoin, de 35 ans environ, dont le *r* n'est pas fricatif<sup>16</sup>. Soixante-quinze ans plus tard le même type, avec *o* fermé long, est enregistré dans l'ALB, à la fois à la carte 838 *le pissenlit* et à la carte 837 *la chicorée sauvage*. Ces attestations du type /krop/ sur les deux cartes sont probablement à l'origine de la remarque de Gérard Taverdet : « Bien qu'elles soient très connues, on confond assez souvent ces deux plantes (surtout en Saône-et-Loire) ». Si le type /krop/ est très répandu pour désigner le *pissenlit* il l'est beaucoup moins pour dénommer la *chicorée sauvage*. La superposition de tous

<sup>14</sup> Ce *Glossaire* est placé sous la responsabilité de la Société d'émulation de l'Arrondissement de Montargis et de l'Association Epona, site : [www.gatinaisgeneal.org](http://www.gatinaisgeneal.org).

<sup>15</sup> Enquêtes faites en 1898 pour le point 11 et 1901 pour le point 909, cf. Brun-Trigaud G., Le Berre Y. et Le Dû J. (2005), *op. cit.*, p 21. Le *r* uvulaire fricatif est noté *ř* dans l'ALF, *ř* sur la carte de synthèse ; *o* : note le *o* long.

<sup>16</sup> Données provenant de la *Notice de l'Atlas linguistique de la France servant à l'intelligence des cartes* (1902), Paris, Honoré Champion, 58 p.

les atlas régionaux sur un axe Mâcon - Paris s'avère riche en variantes du type /kro/. Voir, ci-dessous, la carte de synthèse.



En effet, tous les enregistrements des atlas ont en commun la séquence phonétique [kro/œ], où la valeur de la voyelle, mi fermée / mi ouverte, varie selon un axe arrière / avant, avec une durée longue pour la voyelle d'arrière dans le domaine bourguignon. Le type /kro:p/ est très stable en Saône-et-Loire et témoigne d'une conservation inchangée entre l'ALF et l'ALB. S'il est possible d'établir un certain nombre de corrélations entre les formes phonétiques, leur richesse et leur diversité montrent que c'est une recherche qui reste à faire (variations vocaliques et consonantiques ; hésitation du genre et variation de la segmentation).



Si l'étude de la variation graphique chez André Gilbert était un prétexte à entrer dans le monde de la dialectologie, l'analyse de *l'Haritage perdu*, du recueil de ses textes, montre le devoir de mémoire que s'était fixé l'auteur. Il avait une conscience aigüe de son parler vernaculaire. La confrontation de son lexique avec les attestations de *l'ALIFO* en témoigne. Tous les mots qu'il a intégrés dans le texte sont des mots de chez lui, ou des mots proches de chez lui. Cette confrontation avec les cartes a aussi montré que lorsqu'il y avait des variantes, il choisissait toujours la forme la plus typique ou la plus rare, la forme la plus éloignée du français. C'était sa mémoire partagée, celle des mots qu'il fallait sauver de l'oubli. Il semble qu'il y soit bien parvenu, même s'il n'avait pas les critères requis de témoin d'atlas.

Si la mémoire du terrain a pu être enregistrée du vivant des derniers locuteurs par les atlas linguistiques, il semble qu'aujourd'hui les écritures dialectales enregistrent à leur façon, et pérennisent, ces parlures en voie d'extinction. En l'absence de témoins vivants, les textes dialectaux constituent certainement une nouvelle ressource à explorer.

Ainsi, à l'instar du travail effectué sur l'oeuvre d'André Gilbert, l'exploitation d'autres textes dialectaux, devrait permettre une certaine reconstruction des parlars disparus.

## Bibliographie

Abréviations des désignations d'atlas avec noms de renvoi à la bibliographie :

ALB, *Atlas linguistique et ethnographique de Bourgogne*, cf. Taverdet G.

ALCB, *Atlas linguistique et ethnographique de la Champagne et de la Brie*, cf. Bourcelot H.

ALCe, *Atlas linguistique et ethnographique du Centre*, cf. Dubuisson P.

ALF, *Atlas linguistique de la France*, cf. Gilliéron J., Edmont E.

ALIFO, *Atlas linguistique et ethnographique de l'Île-de-France et de l'Orléanais, Perche, Touraine*, cf. Simoni-Aurembou M.-R.

ALO, *Atlas linguistique et ethnographique de l'Ouest, Poitou, Aunis, Saintonge, Angoumois*, cf. Massignon G., Horiot B.

Bourcelot H. (1966-1978), *Atlas linguistique et ethnographique de la Champagne et de la Brie*, 3 volumes, Paris, Éditions du CNRS.

Brun-Trigaud G., Le Berre Y. et Le Dû J. (2005), *Lectures de l'Atlas linguistique de la France de Gilliéron et Edmont. Du temps dans l'espace*, Paris, CTHS, Éditions du Comité des travaux historiques et scientifiques, 364 p.

Catach N. (1995), *L'orthographe*, Paris, Presses Universitaires de France, coll. Que sais-je ?, n° 685, 6<sup>ème</sup> éd., 128 p. [1<sup>ère</sup> éd. 1978].

Dubuisson P. (1971-1982), *Atlas linguistique et ethnographique du Centre*, 3 volumes, Paris, Éditions du CNRS.

Gilbert A. (1996), *L'haritage perdu ou l'amour de la terre*, Orléans, Imprimerie Desjardins, 2<sup>ème</sup> éd, 222 p. [1<sup>ère</sup> éd., 1995, Dijon, ABDO], 3<sup>e</sup> éd., revue et augmentée, sous presse.

Gilliéron J., Edmont E. (1902-1912), *Atlas linguistique de la France, avec une Notice et une Table, ainsi qu'un Supplément*, Paris, Honoré Champion.

Jejcic Fabrice, "Images de variétés de français du domaine d'oïl central : dynamiques de représentations graphiques d'auteurs (1911-1997)", pp. 219-256, dans *Images et dynamiques de la langue française. Poitevin-saintongeais, français et autres langues en situation de contact*, Textes présentés et coordonnés par Liliane Jagueneau, Université de Poitiers, Paris, L'Harmattan, coll. *Langue & parole*, 258 p.

- Massignon G., Horiot B. (1971-1983), *Atlas linguistique et ethnographique de l'Ouest, Poitou, Aunis, Saintonge, Angoumois*, 3 volumes, Paris, Éditions du CNRS.
- Pontoire J., Fondet C., éds., (1999), *Glossaire des parlers d'Eure-et-Loir, Beauce et Perche, d'après l'enquête de 1868*, Chartres, Société Archéologique d'Eure-et-Loir (SAEL), 318 p.
- Simoni-Aurembou M.-R. (1973-1978), *Atlas linguistique et ethnographique de l'Île-de-France et de l'Orléanais, Perche, Touraine*, 2 volumes, Paris, Éditions du CNRS.
- Taverdet G. (1975-1980), *Atlas linguistique et ethnographique de Bourgogne*, 3 volumes, Paris, Éditions du CNRS.

*Fernande Krier*  
*Professeur des universités*  
*Université Rennes 2 Haute Bretagne*  
*fkrier741@gmail.com*

## **Les corpus de l'oral spontané<sup>1</sup>**

### **1. Introduction**

La réalité linguistique est imprévisible, variable et continuellement en évolution. Elle génère une linguistique de la « parole », inductive et descriptive, qui manifeste plutôt des tendances et des préférences que des contraintes et des lois, et qui se fonde sur l'évidence empirique et non sur l'introspection et l'intuition du locuteur natif. Les données linguistiques de ce langage réel et authentique, recueillies sur le terrain, sont enregistrées sous forme de corpus, c'est-à-dire d'ensembles d'énoncés, qui représentent un échantillon de la langue ou du dialecte à analyser. Tributaires du contexte situationnel, elles comportent des moyens et formes de l'organisation du discours, qui ne font pas forcément partie des unités et catégories traditionnelles, comme les connecteurs, les particules illocutoires, les exclamations, les interjections et les injures, mais aussi des éléments dialogiques comme le vocatif, l'honorifique et les verbes performatifs, les éléments déictiques (pronoms personnels de la première et deuxième personne, pronoms démonstratifs et possessifs, adverbes de lieu et de temps), sans oublier les signes prosodiques (intonation, rythme, accent, pauses,...). En outre, les corpus de l'oral spontané, c'est-à-dire de l'oralité non planifiée, par leurs séquences discursives, fournissent des signaux d'ouverture, de progression et de clôture du discours (comme *dis donc ! tu vois ! d'accord ?*) dont le linguiste n'aurait pas

---

<sup>1</sup> La présente communication est une version augmentée et enrichie de « Réflexions sur la constitution et l'exploitation de corpus de l'oral spontané », à paraître par ailleurs dans *Actes du 30<sup>e</sup> Colloque international de Linguistique fonctionnelle* (Chypre, octobre 2006).

tenu compte autrement, qu'il interprétera avec les erreurs ou les écarts à la « norme », par exemple des lapsus, hésitations, interruptions et reprises. S'y ajoute que des faits inattendus ou hors norme l'obligent constamment à réviser ses réflexions et le préservent de formuler des généralisations précipitées ou erronées. Ainsi, en partant de la réalité linguistique observée, il parvient à analyser des phénomènes langagiers qui, en général, pris isolément, sont traités succinctement ou tout simplement négligés.

Le recueil des données doit se faire jusqu'à ce qu'elles forment un *échantillon représentatif* du parler, qui est réalisé, lorsque les récurrences classifiées permettent de décrire certaines régularités dans la production d'énoncés, de formes de mots ou de structures phonétiques et d'en déduire des conclusions sous forme de règles.

Les *corpus électroniques* ont l'avantage, qu'une fois réalisés, ils réduisent le temps de recherche et rendent le dépouillement plus facile. Le traitement électronique des données est indiqué pour l'analyse statistique de phonèmes ou de prosodèmes, de lexèmes et de formes grammaticales, pour la saisie lexicographique de toponymes ou d'anthroponymes et pour le recueil de concordances et de collocations. Par conséquent, les versions électroniques sont des compléments utiles pour l'analyse quantitative et qualitative des langues.

Des objections ont été formulées à l'égard de l'approche inductive, à savoir que le matériau enregistré se réduit à une fraction de la langue. Or, il n'a jamais été dit que l'analyse en question était exhaustive (pour cette raison, on parle de corpus « ouverts ») les données restent évidemment incomplètes, ce qui est justifié, car un corpus se crée avec une intention précise et dans un but précis. Les avantages l'emportent de loin : - on n'analyse pas des structures isolées obtenues par introspection, mais des valeurs discursives concrètes obtenues dans un contexte authentique ; - on relève des structures sous-estimées ou non discutées, des phénomènes linguistiques « marginaux » qui peuvent s'avérer intéressants, et qui prouvent que la recherche sur corpus est une lutte permanente pour trancher entre régularités et tendances, voire accidents ; - au lieu d'admettre une norme prescriptive unique, on discernera des normes descriptives, dues aux comportements langagiers réels et concrets des locuteurs.

Cette conception inductive a été fructueuse, entre autres, dans la description de langues et de dialectes menacés ou en voie d'extinction ainsi que dans les recherches sur l'acquisition du langage et la didactique des langues.

## 2. Méthodologie

Pourquoi des descriptions de corpus de l'oral spontané ? Parce qu'elles sont les seules à refléter la réalité du langage, dans le but d'examiner et de corroborer un postulat théorique ; elles offrent inversement aussi un point de départ pertinent pour une ébauche de théorie. Établir un corpus de l'oral spontané exige, en outre, un enregistrement, sans que l'informateur soit renseigné au préalable sur l'objectif de la recherche (= informateur « naïf »), parce qu'autrement, par complaisance, il pourrait produire précisément les éléments linguistiques recherchés ; d'ailleurs, le fait de ne pas être mis au courant de l'objectif est, en principe, indifférent à l'informateur non linguiste. De cette méthode résulte inévitablement un déficit empirique, comme l'absence d'une composante ou d'un de ses traits sémantiques, qui ne font pas partie du répertoire du / des informateur(s) en question. Cette situation de départ défavorable est compensée par le fait qu'aussi bien la spontanéité du corpus que l'authenticité du matériau – par là j'entends le non effacement de faux départs et de reformulations – permettent d'élaborer la fonction réelle et la pertinence communicative spécifique de l'élément étudié.

Pour le recueil de données orales, l'enquêteur est supposé être un excellent *observateur participant* qui, dans une *pré-enquête*, s'est documenté sur le terrain de recherche, a observé la façon de parler des gens, a sélectionné ses informateurs et les a mis à l'aise et en confiance en suggérant pour l'enregistrement des discussions entre personnes qui se connaissent et se trouvent dans un cadre qui leur est familier, bref, il se fait accepter comme membre de la communauté. En ce qui me concerne, j'ai toujours participé comme observatrice à l'entretien sans, pour autant, y intervenir. Il convient de remarquer que les locuteurs, très souvent, après deux ou trois minutes de conversation, ne se rendent plus compte du microphone ; toujours est-il que, pour analyser de l'oral spontané authentique, il est recommandé de sélectionner des extraits du milieu du corpus. Ceci étant, dans de

rare situations, le micro caché pourrait rendre service ; c'est ainsi que, lors d'un de mes enregistrements, un professeur de lycée avait parlé un alémanique « hypercorrect » devant le micro, ce qui ne correspondait pas à son comportement langagier habituel.

L'*enquête* se fera selon l'objet d'étude et les objectifs linguistiques de recherche choisis. – Par *objet d'étude*, j'entends la langue elle-même, qu'elle soit une langue de grande diffusion, une langue minoritaire, un dialecte ou une langue à tradition orale. Cela peut être également une variété langagière marquée par des particularités géolinguistiques : enquête aréale plus ou moins étendue en milieu urbain et / ou rural, comme la zone frontrière du francoprovençal et de l'alémanique dans le Valais, avec treize villages, ou d'un îlot linguistique, comme les deux colonies Walser dans le Val d'Aoste, Gressoney Saint Jean et Gressoney la Trinité.

Dans une situation de bi- ou plurilinguisme impliquant, par exemple, une langue minoritaire, il est indispensable de mener des *enquêtes à grande échelle* à l'aide de questionnaires ciblés selon des critères socioculturels tels que l'âge, le sexe, la parenté, l'école, la profession, etc.

Les *objectifs linguistiques de recherche* de l'enquête peuvent porter sur un point particulier (voir ci-dessous, 3) ou bien avoir des visées plus larges, comme la description phonologique et / ou grammaticale.

Les deux facteurs que je viens de décrire déterminent les autres paramètres, comme le *nombre d'informateurs* : pour analyser un point linguistique donné, on admet en principe que les enregistrements d'une vingtaine de personnes pourraient donner des résultats fiables. Pourtant, des analyses ont été réalisées avec plusieurs centaines d'informateurs. Il n'y a qu'à mentionner l'étude pilote de William Labov (1966), qui, pour son analyse de la vibrante dans les magasins new-yorkais, avait choisi quatre mots avec *r* postvocalique prononcés par 264 personnes en 6,5 heures. Il en est de même d'une analyse de l'allemand parlé par des Juifs allemands émigrés en Israël, qui a été réalisée par la germaniste Anne Betten et ses collaboratrices (1995, 2000) avec 170 informateurs, en vue d'une étude syntaxique. Quant à moi, j'avais étudié, par contre, un idiolecte pour mettre en évidence le parler spécifique d'un locuteur trilingue (Krier 1981-b). Au demeurant, pour enregistrer une conversation, il convient en principe, de se limiter à trois locuteurs au maximum, autrement on risque d'obtenir un charivari vocal. Se pose également la question de

*l'informateur de référence*, à savoir un locuteur qui donne des explications et fait des commentaires sur les données recueillies, tâche qui s'avère très utile, si l'enquêteur ne pratique pas la langue étudiée.

Ceci étant, même un corpus représentatif n'est pas à l'abri de lacunes, et c'est ici que des *questionnaires* adéquats interviendront. Ainsi, pour une enquête phonologique et grammaticale, les enregistrements des données seront complétés par des questionnaires consistant en listes de paires minimales, de mots et de phrases.

La *notation* dépend, elle aussi, de l'objectif de la recherche. Une analyse phonétique et l'étude d'une langue à tradition orale exigent forcément une notation phonétique à l'aide des signes de l'A.P.I. Parfois, on distingue entre notation étroite, phonétique, et large, phonologique. C'est un travail fastidieux, étant donné que cinq minutes d'enregistrement demandent une notation d'au moins une heure et celle-ci doit être affinée par la suite. Aussi réserve-t-on une notation orthographique à l'analyse grammaticale ou pragmatolinguistique de langues et de dialectes plus ou moins connus, agrémentée d'indications du comportement non-verbal, comme RIRE / SOUPIR. Pour une étude de la prosodie s'ajoutent, dans les deux cas de notation, les symboles de l'intonation, des accents, des pauses, etc. Il ne faut pas oublier une traduction mot-à-mot obligatoire du texte transcrit pour un parler peu ou pas connu, suivie au besoin de la traduction intelligible, qui rend le sens du texte.

### 3. Données illustratives

Les exemples suivants sont tirés de quelques-unes de mes analyses antérieures de corpus de l'oralité.

Certaines opérations énonciatives courantes dépendent complètement de l'oral spontané, telles les *formes de débit rapide* (Krier 1981-a), le débit rapide étant caractérisé par des formes plus ou moins réduites, qui sont dues, entre autres, à la chute de phonèmes au niveau du monème : francoprovençal [pa : r] < ['pa : ri] 'père', [dɔn] < [a'dɔn] 'alors', [mɔ̃a] < [mi'ɔ̃a] 'moitié', ou au niveau du syntagme (« formes écrasées ») : [jit] < [jei 'dit] 'j'ai dit', ['neitɪn'du] < ['na jei ɛntɛn'du] 'non, j'ai entendu', avec réduction de syllabes. Ces formes réduites appartiennent au caractère « relâché » de la parole et sont



instructives à plus d'un égard, par exemple sur l'évolution phonétique d'une langue.

L'enregistrement de l'oral spontané est également recommandé pour l'analyse des *rappports entre connecteurs, intonation et pauses* au sein de la chaîne parlée. Le cas classique, souvent décrit, comporte une séquence discursive d'une longueur d'une à quatre phrases, qui commence par un connecteur transphrastique, marquant la progression thématique, et se termine par un contour intonatif descendant et une pause courte ou longue. Cette constellation ne s'est réalisée que pour la moitié des séquences de mon corpus de l'oral spontané (Krier 1995) et, par conséquent, ne reflète guère la réalité. J'ai relevé, entre autres, des séquences sans pause, avec intonation ascendante, le connecteur étant le seul indice démarcatif. Qui plus est, les pauses sont souvent audibles là où l'on ne les attend pas (pauses d'hésitation) et sont absentes là où elles sont attendues dans le continuum de la parole. – Dans une analyse (Krier 1991) et la suivante (Krier 1995), j'ai eu recours aux moyens de la phonétique expérimentale pour vérifier les phénomènes prosodiques. En effet, les *particules illocutoires*, prises isolément, ont un sens lexical assez vague (comme *au juste*, *à vrai dire*). Leur efficacité communicative, qui signale l'attitude émotionnelle du locuteur face à ce qu'il dit, est précisée en concomitance avec l'*intonation* dans le contexte de l'acte de parole (Krier 1991).

La déixis spatiale, complexe à cause de son système à trois dimensions, est indiquée, en principe, par des *adverbes de lieu*, assez nombreux, s'ils se rapportent à la haute montagne. Fidèle à mes habitudes, j'avais enregistré des conversations spontanées, les informateurs n'étant pas informés sur l'objectif de la recherche (Krier 1986). Il en résultait un inévitable déficit empirique, qui devait être comblé par un questionnaire approprié.

Tandis que les exemples précédents se sont référés à des communications privées, non planifiées, l'analyse qui va suivre a trait aux comptes rendus des séances publiques de la Chambre des Députés du Grand-Duché de Luxembourg (Krier 1999). Ils sont rédigés d'après les enregistrements retenus par écrit, après « nettoyage » du texte des phénomènes verbaux erronés, qui sont dus à la phase d'encodage, comme des lapsus et reformulations ; la version corrigée est destinée à la publication. Quant à l'orateur, il parle librement en se servant des notes qu'il a prises lors de la préparation du sujet politique à débattre.

Alors que, précédemment, on avait affaire au langage oral et code phonique, il s'agit ici de l'*oralité planifiée* couplée du code graphique. Quel est l'intérêt du linguiste de recourir à ce genre de corpus ? Il arrive que l'oral spontané se prête mal à un objet de recherche précis, en l'occurrence les *expressions idiomatiques*. Les députés, par contre, aiment aiguïser leurs discours par des formules flexibles, qui jouent un rôle expressif et bien ciblé comme composantes argumentatives et évaluatives du débat. Une telle façon de procéder est moins courante dans le langage quotidien non émotionnel.

Dans ce dernier exemple, le langage oral prime le code graphique, mais l'inverse est aussi possible ! Les corpus de l'oral spontané enregistrés, non marqués par les émotions, comportent rarement des *injures* ou des *interjections*, pourtant supposés caractéristiques du langage oral. Voilà pourquoi le théâtre populaire, qui comporte nombre d'actes de parole expressifs, est une source d'informations efficace pour ce genre de phénomènes (Krier 2006-a, 2006-b). Le style y est en général familier, voire vulgaire ; par conséquent, les injures y sont attestées, tant dans les énoncés performatifs primaires (discours direct au présent de l'indicatif) sous forme de vocatifs, que dans les énoncés constatifs. – Les interjections, quant à elles, sont plutôt expressives (*aïe !*) qu'appellatives (*courage !*), elles apparaissent souvent ensemble avec le vocatif dans des textes qui imitent le langage oral, à savoir les dialogues du théâtre populaire, les contes et les fables. – L'oralité planifiée avec support textuel s'avère donc nécessaire dans la recherche de phénomènes spécifiques.

#### 4. Conclusion

La méthodologie, à laquelle a été consacrée cette communication, se référant à des illustrations de la spontanéité et de l'authenticité de la parole enregistrée et transcrite, peut s'adjoindre, le cas échéant, les moyens de l'oralité planifiée, retenue par le code graphique, mais aussi des questionnaires appropriés et les ressources de la phonétique expérimentale et de l'analyse électronique.

Exceptionnellement, dans l'analyse syntaxique, entre autres, des phrases complexes, caractérisées par un nombre plus fréquent de conjonctions et de prépositions, des corpus de textes écrits seront plus adaptés. Ces derniers permettent, d'autre part, de voir comment, sur le

plan grammatical et lexical, le langage oral se distingue du texte écrit. Il ne faut cependant jamais oublier qu'il y a primauté de l'oral sur l'écrit ; et pourtant, l'analyse de corpus de textes écrits, moins coûteuse en temps et en énergie, est de loin la plus fréquente.

## Bibliographie

- Baker, Paul / Hardie, Andrew / McEnery, Tony (2006), *A Glossary of Corpus Linguistics*, Edinburgh University Press.
- Sprachbewahrung nach der Emigration. Das Deutsch der zwanziger Jahre in Israel*, Ire Partie: *Transkripte und Tondokumente*, Betten, Anne / Grassi, Sigrid (éds.) (1995), IIe Partie: *Analysen und Dokumente*, Betten, Anne / Du-Nour, Miryam / Dannerer, Monika (éds.) (2000), Tubingue, Niemeyer.
- Krier, Fernande (1981-a), "Les formes de débit rapide : essai d'une systématisation", *Actes du 7<sup>e</sup> Colloque international de Linguistique fonctionnelle*, St. Andrews (Ecosse), septembre 1980, 183-188.
- Krier, Fernande (1981-b), "Le français et l'allemand prononcés par un Luxembourgeois", *International Review of Applied Linguistics (IRAL)*, 19, 352-360.
- Krier, Fernande (1986), "Die lokaldeiktischen Ausdrücke im Alemannischen des Wallis", *Zeitschrift für Dialektologie und Linguistik (ZDL)*, 53, 1, 32-44.
- Krier, Fernande (1991), "Modalpartikeln im spontan gesprochenen Walliser Alemannischen", *ZDL*, 58, 2, 147-174.
- Krier, Fernande (1995), "Zur Steuerung von Kommunikationsvorgängen in einem Idiolekt", *Deutsche Sprache*, 23, 3, 269-287.
- Krier, Fernande (1999), "Idiomverwendung in der luxemburgischen Abgeordnetenversammlung", *ZDL*, 66, 3, 280-295.
- Krier, Fernande (2006-a), "Schimpfwörter im luxemburgischen Volkstheater", *Bulletin linguistique et ethnologique de l'Institut Grand-Ducal*, 33, 77-89.
- Krier, Fernande (2006-b), "Interjektionen im Lëtzebuergesch", *Dialectologia et Geolinguistica*, 14, 63-85.
- Labov, William (1966), *The social stratification of English in New York city*, Washington, Center for applied linguistics.
- Scherer, Carmen (2006), *Korpuslinguistik*, Heidelberg, Winter.
- Teubert, Wolfgang / Krishnamurthy, Ramesh (2007), *Corpus Linguistics, Critical Concepts in Linguistics*, volumes I, III et VI, Londres et New York, Routledge.

*Guylaine Brun-Trigaud*  
*Université de Nice, UMR 6039*  
*guylaine.brun-trigaud@unice.fr*

## **Les chaînes de clitiques : traitements cartographiques**

Il ne sera pas question ici de présenter les nombreuses théories sur l'ordre des pronoms clitiques qui divisent le monde des linguistes, le propos sera plutôt d'apporter des faits qui semblent être ignorés par la grande majorité des personnes qui s'intéresse au sujet. La plupart du temps, il est fait référence au français standard, parfois au français parlé ou familier, mais rarement aux dialectes. Pourtant ces derniers pourraient donner un éclairage intéressant sur cette question.

Car, après avoir dépouillé et publié près de 600 cartes de *l'Atlas Linguistique de la France*<sup>1</sup>, la carte 410 "Dis-le-moi" continue de m'interpeller : est-il possible de comprendre et d'expliquer d'une part, une si grande variation dans les données recueillies -pas moins de 270 formes différentes pour les 646 figurant sur la carte- et d'autre part, cette étrange apparence à la fois singulière et pourtant familière.

Les matériaux utilisés seront ceux de l'ALF et des Atlas linguistiques régionaux, en particulier des données inédites issues des carnets d'enquêtes. Dans les exemples qui suivront, la transcription de chacun des auteurs n'a pas été intégralement conservée, pour la clarté des faits, les éléments de la réponse ont été différenciés. Les chiffres qui précèdent sont les numéros de points d'enquête et les informations d'aires dialectales sont purement indicatives.

---

<sup>1</sup> Cf. Brun-Trigaud, *Le Dû, Le Berre* (2005).

## 1. Carte ALF 410 "Dis-le-moi"

Le fait majeur qui caractérise la carte de l'ALF "dis-le-moi" est la répartition inhabituelle des deux zones pour les deux combinaisons possibles des pronoms enclitiques.

Comme en témoigne l'étendue de la zone en vert de la carte 1, c'est la combinaison Verbe à l'impératif + Personne1 au Datif + Personne3 à l'Accusatif, type "dis-moi-le" qui prévaut. Elle occupe massivement l'aire des parlers d'oïl, et laisse une plus grande place à la combinaison Verbe à l'impératif + Personne3 à l'Accusatif + Personne1 au Datif, type "Dis-le-moi", dans l'aire méridionale, en jaune. Ce paysage est assez singulier dans le domaine gallo-roman.

Avant d'aborder le détail, notons quelques cas particuliers :

Tout d'abord, des cas de réduction à "dis-moi" assez sporadiques, symbolisés par les étoiles sur la carte 2...

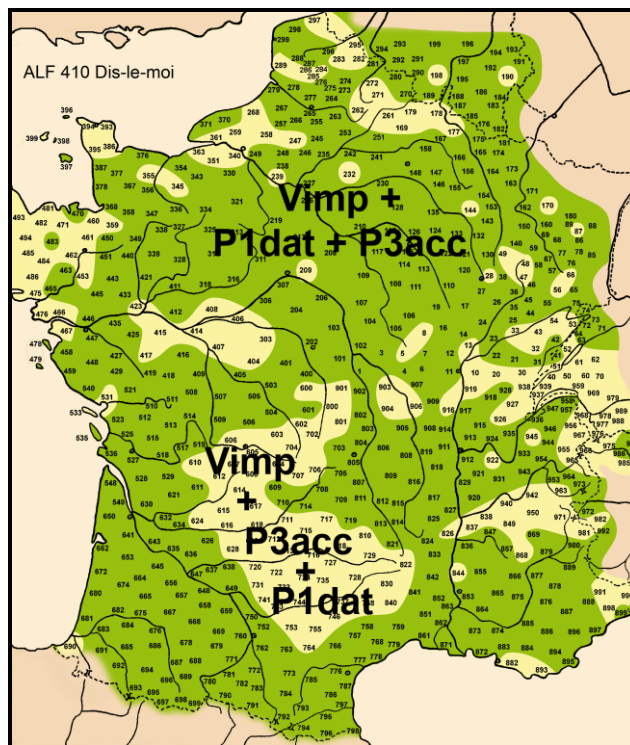
- champ. 114 : di mwa
- lang. 715 : digo me

...et surtout un plus grand nombre de cas de renforcement ou de déconstruction. Il faut entendre par renforcement, des formes du type "dis-moi-le donc", "dis-le-le-moi", etc. et par déconstruction des formes du type "dis-le à moi" dans lesquelles le pronom datif est disjoint ou même qui combine les deux avec "dis-moi-le à moi". Ces formes sont symbolisées par des losanges sur la carte 2.

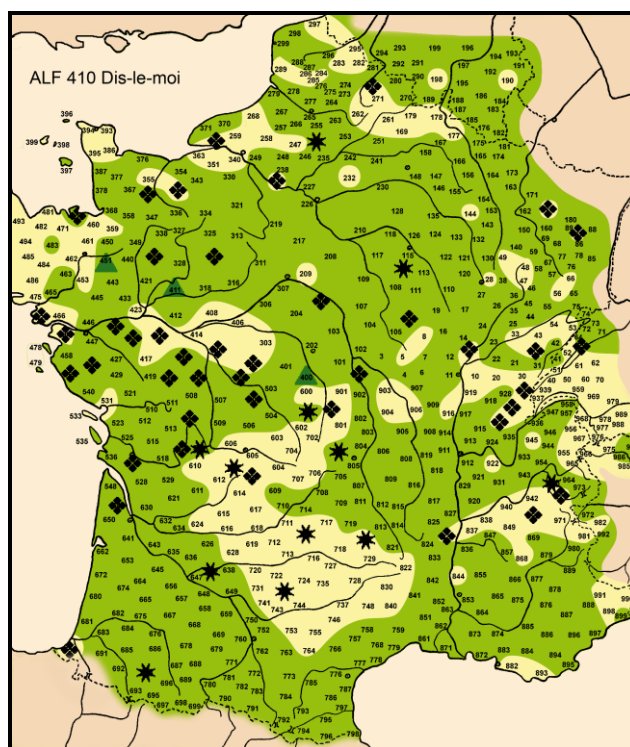
- poit. 435 : di mwɛl dɔ̃ ("dis-moi-le donc")
- frpr. 60 : di lə zu me ("dis-le-le-moi")
- bourg. 13 : di lə a mwa ("dis-le à moi")
- norm. 361 : di m lə a mɛ ("dis-moi-le à moi")

Il faut signaler aussi quelques rares cas de "dis-moi-ça", avec changement de pronom accusatif, symbolisés par les triangles vert foncé de la carte 2.

- tour. 411 di mwa sa ("dis-moi-ça")



Carte 1 : ALF 410 "Dis-le-moi" : répartition des combinaisons de clitiques



Carte 2 : ALF 410 "Dis-le-moi" : autres cas

La majorité de ces formes à exception se trouve dans la partie occidentale du domaine d'oïl ainsi qu'en franco-provençal.

Comme le montre le tableau 1, on constate une étonnante diversité de formes y compris à l'intérieur des mêmes zones dialectales.

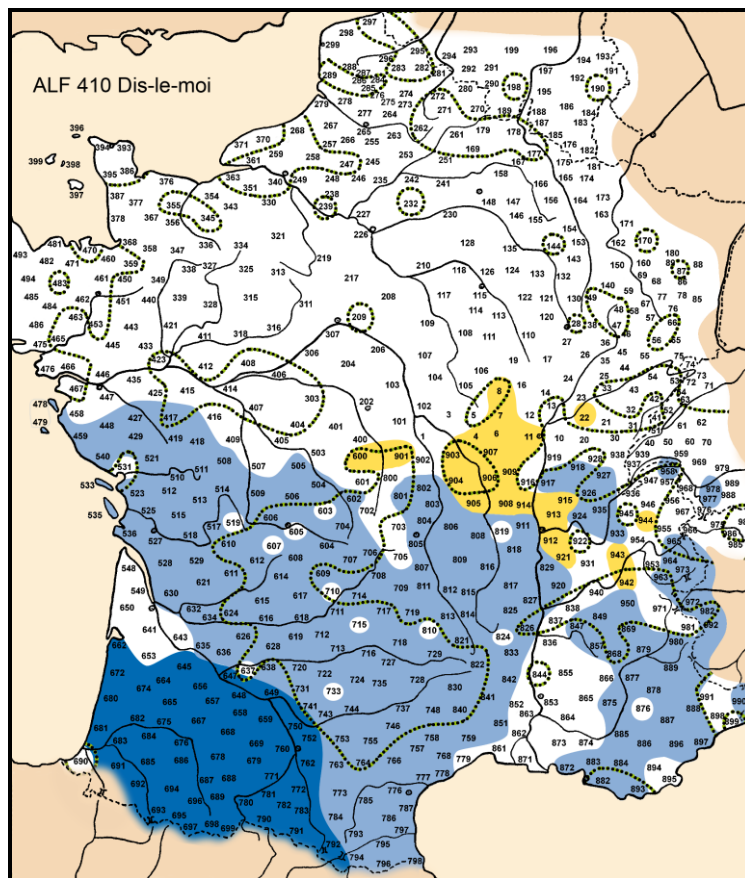
	Vimp + P1Dat + P3Acc	Vimp + P3 dat + P1Dat
pic.	296 di m elle	286 di la mi
norm.	377 di m lɛ	361 di lə mə
orl.	217 di mwə lə	tour. 408 di lə mwa
lorr.	65 di m lo	66 di l me
poit.	525 di m zu	saint. 610 di jɔ mə
frpr.	912 di m zi	942 di zi me
gasc.	750 digo m ɔk	-
auv.	815 djidza m u	801 dzi zo mœ
lang.	776 digo me bo	774 dio bɔ me
prov.	894 dige mi lu	990 di ə me

*Tableau 1*

En effet, aux variations du verbes "dire", dont une partie des formes à l'impératif en occitan est basée sur le subjonctif, et aux variations du pronom datif me, qui en domaine d'oïl s'est confondu avec la forme du pronom tonique, s'ajoute un élément qui introduit encore davantage de perturbations. Il s'agit du pronom neutre à l'accusatif, qui existe sous une forme distincte et variée, dans un large domaine qui déborde des parlers occitans, alors qu'en domaine d'oïl le pronom de la 3e personne et le pronom neutre se sont confondus.

Sous l'influence de ces trois éléments susceptibles de varier, les parlers gallo-romans, se subdivisent à nouveau (cf. la carte 3) :





Carte 3 : répartition des formes du pronom neutre

D'une part, comme le montre l'aire en bleu clair, une grande partie du domaine est occupée par les descendants de HOC sous la forme [ɔ, u].

- auv. 808 di mə ɔ

Aire à laquelle il faut adjoindre la zone gasconne, en bleu foncé, qui présente des formes très variées.

- gasc. 680 : dih m at
- gasc. 667 : digɔ m εk
- gasc. 672 : diz m ik

- gasc. 752 : dijo m ɔk
- gasc. 693 : dizə m mɔt

Cette zone des continuateurs de HOC présente un aspect assez déstructuré, notamment par le fait que dans les interstices, se trouve le pronom [lu] qui dans cette zone, comme ses correspondants d'oïl, est utilisé comme pronom neutre et comme pronom de la 3e Personne.

A propos de la répartition géographique de ce pronom [lu], il semblerait qu'elle se contredit quelque peu : en effet, d'après de nombreuses cartes de l'ALF, une zone rhodanienne comme celle-ci est un signe d'innovation, souvent à l'initiative d'Avignon ou de Marseille, alors qu'une zone médocaine, comme on la voit ici, est plutôt évocatrice d'une zone de conservation, à moins que Bordeaux n'ait joué, pour une fois, ce rôle de centre innovateur.

Enfin, pour être tout à fait complet, il faut dire quelques mots d'une zone de pronom neutre en -i "y", encore très vivant aujourd'hui dans les parlers régionaux où "je le sais" se dit "j'y sais". Ce pronom adverbial "y" est devenu objet direct.

- bourg. 6 di mwɛ zi
- frpr. 942 di zi me

Gérard Taverdet qui l'évoque dans sa thèse<sup>2</sup>, suppose avec Claude Régnier<sup>3</sup> et Gaston Tuillon<sup>4</sup> que cette aire est récente et qu'elle a éliminé les descendants de HOC, ce qui semble avéré par la répartition géographique.

Les trois zones de pronoms sont relativement indifférentes à la répartition des deux aires de combinaison des pronoms enclitiques, si ce n'est la zone gasconne qui ne comporte aucune forme du type "Dis-le-moi". Cependant, il faut constater que cette zone de l'emploi d'un pronom neutre a un air de "déjà vu" avec une limite qui ne franchit pas la Loire.

---

<sup>2</sup> Cf. Taverdet (1980).

<sup>3</sup> Cf. Régnier (1968).

<sup>4</sup> Cf. Tuillon (1969).

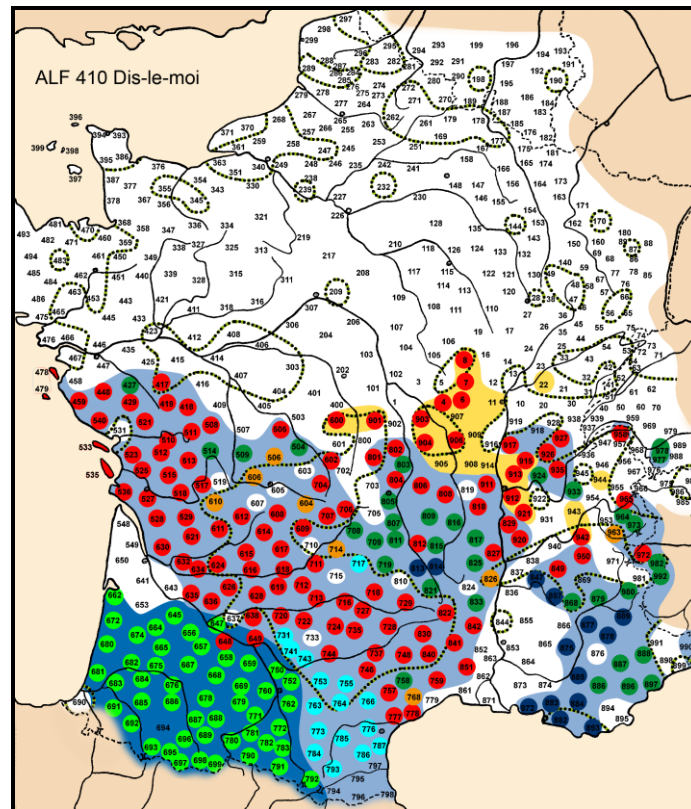
	V + P1Dat + P3Acc "dis-moi-le"		V + P3Acc + P1Dat "dis-le-moi"	
pr. neut.	i	o / u	i	o / u
\\	frpr. 908 di m i	auv. 816 dji m εɔ poit. 427 di m u		prov. 868 di u me
z	frpr. 915 di m z <sub>i</sub>	poit. 515 di m z <sub>u</sub>		
	bourg. 6 di mwε z <sub>i</sub>	auv. 808 di mə z <sub>ɔ</sub> lim. 609 diza me z <sub>u</sub>	bourg. 906 d <sub>i</sub> z <sub>i</sub> me	lim. 602 d <sub>i</sub> z <sub>o</sub> mø lang. 716 di z <sub>u</sub> me
		gasc. 649 digo me z <sub>ɔ</sub> t		
ʒ		périg. 626 di z <sub>ɔ</sub> me z <sub>u</sub>		périg. 617 diza z <sub>u</sub> me
b		lang. 785 dio me b <sub>ɔ</sub> t		lang. 741 digo b <sub>o</sub> me
v		auv. 814 djidza mə v <sub>u</sub>		prov. 893 digo v <sub>o</sub> mi
j		lang. 768 diga mi j <sub>u</sub>		lim. 610 di j <sub>ɔ</sub> mε
autre		prov. 869 di mə m <sub>u</sub>		bourg. 5 di r <sub>o</sub> mwe

Tableau 2

Le tableau 2 donne le résumé des cas possibles pour la zone concernée : un verbe à l'impératif se terminant le plus souvent par une voyelle et un pronom accusatif neutre sous forme vocalique, les deux se combinant avec un pronom datif pouvant ou non subir une élision.

Il faut donc un dernier ingrédient pour faire le "liant" : un élément intervocalique servant à prévenir le hiatus.

Cet élément intervocalique se présente principalement sous trois groupes de phonèmes : [z/ʒ], [b/v] et [j]. D'autres phonèmes, comme [r] ou [m] peuvent servir très marginalement d'élément de liaison.



Carte 4 : répartition des phonèmes intercalaires

La carte 4 montre la répartition géographique de ces phonèmes intervocaliques :

D'une part, les aires respectives de [b] sont figurées par des pastilles en bleu clair dans le Languedoc et pour [v], par des pastilles en bleu foncé, situées en Provence.

L'explication de l'apparition de ces phonèmes réside dans la phonétique syntactique : pour éviter le hiatus, il se développe un élément au contact entre deux voyelles, ainsi [ɔ] > [wɔ] > [vɔ] qui se renforce normalement en [bɔ] dans les parlers languedociens<sup>5</sup>. C'est donc un élément qui se dégage.

Ce type de phénomène s'observe très bien par exemple dans l'aire des continuateurs de HODIE de la carte "Aujourd'hui" (ALF 72), dans laquelle au point 872, on recueille [vwɛj] et au point 787, [bɛj].

<sup>5</sup> Cf. Ronjat (1930-1941) § 508.

- prov. 872 digɔ mi vu ("dis-moi-le")
- lang. 787 diɔ me bɔ ("dis-moi-le")

Mais, il faut très probablement dissocier les attestations auvergnates qui se situent dans une aire où les latérales à l'intervocalique subissent de grandes perturbations. Ainsi dans la carte "Pelle" de l'ALF (992), il a été relevé au point 713 [p'awa], et [p'ava] aux points 813 et 814.

D'autre part, la répartition des formes avec [j] comme élément intervocalique, figurées par des pastilles orange, montre que cette forme est nettement moins répandue et beaucoup plus dispersée.

A contrario de la zone des pastilles rouges qui indique la répartition de l'aire comportant les éléments intervocaliques [z/ʒ], formant une vaste zone en écharpe de l'Atlantique à la Méditerranée, avec une zone secondaire à l'Est.

[z] est l'élément de liaison le plus répandu dans le domaine gallo-roman. Il a fait l'objet de nombreuses discussions chez les linguistes<sup>6</sup>, car, en français standard, on retrouve cet élément en composition dans certaines combinaisons de clitiques, en particulier avec "en" et "y" :

- fr. donnes-en
- fr. % donne-moi-z-en (fr. donne-m'en)<sup>7</sup>
- fr. vas-y
- fr. % mène-moi-z-y (à la gare)

Mais il n'est pas toujours obligatoire notamment dans le cas où "en" et "y" sont suivis d'un infinitif :

- fr. va en chercher un autre
- fr. donne lui en une (donne une pomme à ton frère)
- fr. ose y ajouter du tien

Donc le hiatus est tout à fait possible et admis dans certains cas, comme le montre la répartition des formes sans élément intervocalique, figurées par des pastilles vertes se situant en périphérie de la zone concernée ainsi que massivement dans la zone gasconne.

- prov. 868 di u me

---

<sup>6</sup> Cf. notamment Morin (1979).

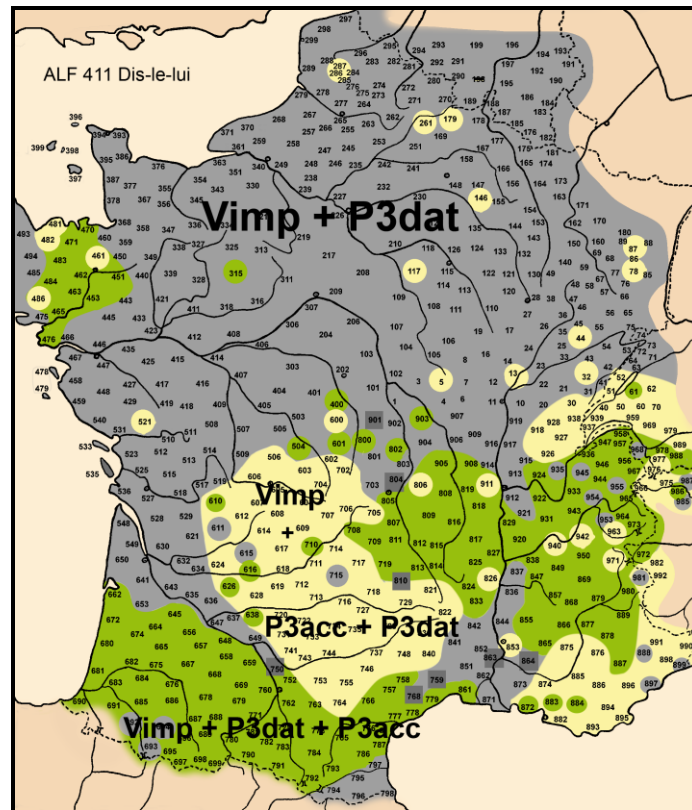
<sup>7</sup> Le symbole % signifie que nous sommes dans du français non standard.

D'où cette interrogation à propos de cette carte, en particulier sur les formes [dimzu, dimzɔ, dimzi] qui ont leurs correspondants [dimu, dimɔ, dimi] localisés à proximité. On peut se demander si dans certains cas, il n'y a pas une sorte de figement de ces formes avec éléments intervocaliques.

Par exemple dans le Bourbonnais, au pt 905, on relève [d̥i m̥ i] pour "dis-le-moi" et [d̥i l̥ zi] pour "dis-le-lui", à côté de [d̥i m̥ i] et de [d̥i l̥ i] au pt 907 voisin.

Un autre exemple en Provence avec [v] : au pt 882, on relève logiquement [digo vo mi] ("dis-le-moi"), [digo vo li] ("dis-le-lui", ALF 411), [aj pa udʒa va li d'ire] ("je n'ai pas osé le lui dire", ALF 1650), alors que la forme [pər va li] ("(j'ai eu du mal) à le leur (faire comprendre)", ALF 761) ne nécessitait pas de phonème intervocalique.

## 2. Carte ALF 411 "Dis-le-lui"



Carte 5 : ALF 411 Dis-le-lui : répartition des combinaisons

Avec la carte ALF 411, "Dis-le-lui", c'est un changement presque total de physionomie, puisqu'une grande partie du domaine est occupée par les formes contractées "Dis-lui" (la zone en gris) ou "Dis-le" (les carrés gris foncé).

- norm. 354 di li ("dis-lui")
- auv. 804 djiza zu ("dis-le")

Cependant on retrouve des similitudes, notamment dans la configuration de l'aire centrale du type Verbe à l'impératif + Personne3 à l'Accusatif + Personne1 ou 3 au Datif (Dis-le-moi / Dis-le-lui) (voir la carte 1).

Autre constat, la séquence non admise en français standard Verbe à l'impératif + Personne3 au Datif + Personne3 neutre à l'Accusatif du type \*"Dis-lui-le" se trouve dans deux grandes aires : l'une en Gascogne, débordant un peu sur le Languedoc et l'autre à cheval entre les Alpes et le Massif Central, dans une sorte d'aire transrhodanienne. Auxquelles, il faut ajouter la petite zone en Haute-Bretagne.

- gasc. 697 diz l ak
- lang. 777 digu zi ju
- prov. 884 digo li vɔ
- gallo 465 di lqi lə ("dis-lui-le")

Enfin une dernière remarque, qui a son importance : dans certains parlars, il n'y a pas forcément de concordance pour la combinaison des pronoms. Par exemple la zone gallo de la carte "Dis-le-moi" est du type Verbe à l'impératif + Personne3 neutre à l'Accusatif + Personne1 au Datif type "dis-le-moi" alors qu'elle est du type Verbe à l'impératif + Personne3 au Datif + Personne3 neutre à l'Accusatif type "dis-lui-le" dans la carte "Dis-le-lui".

- gallo 453 di lə ma ("dis-le-moi") \ di ji lə ("dis-lui-le")

### 3. Les atlas linguistiques régionaux.

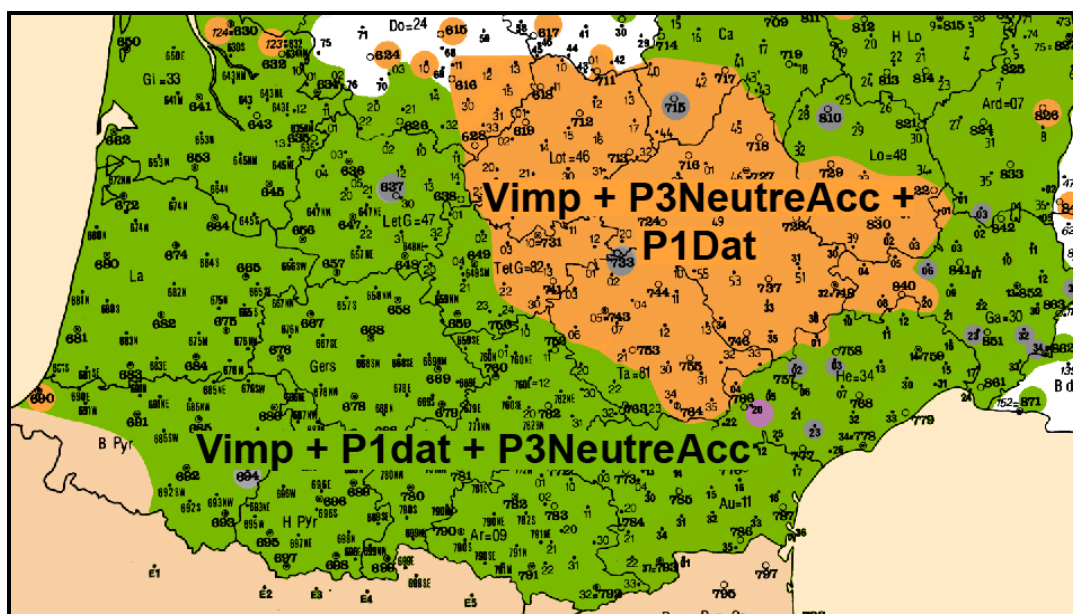
Enfin, qu'en est-il des pronoms clitiques au travers des atlas linguistiques régionaux ?

D'abord, un immense regret : celui de ne pouvoir être renseigné pour certaines régions, comme le Limousin, qui n'avaient pas de questionnaire morphologique et en attendant que la publication des atlas inachevés se poursuive.

Dans l'état actuel des choses, il n'est guère possible de montrer une carte qui nous donnerait une vue d'ensemble, d'une part pour la raison évoquée ci-dessus mais aussi par le choix même des questions, qui ne sont pas toujours exactement comparables d'un atlas à l'autre : "Donne-le-moi, ce couteau" ne peut pas être mis en regard de "dis-le-moi".

Donc nous nous limiterons à faire des zooms sur deux régions distinctes, pour en montrer deux aspects bien différents.

D'une part, dans le Sud-Ouest avec un premier regard sur les données inédites de l'Atlas du Languedoc Occidental et de l'Atlas du Languedoc Oriental, auxquelles ont été adjointes les données publiées de l'Atlas de la Gascogne (c. 2297) et de l'Atlas du Massif Central (c. 1829).



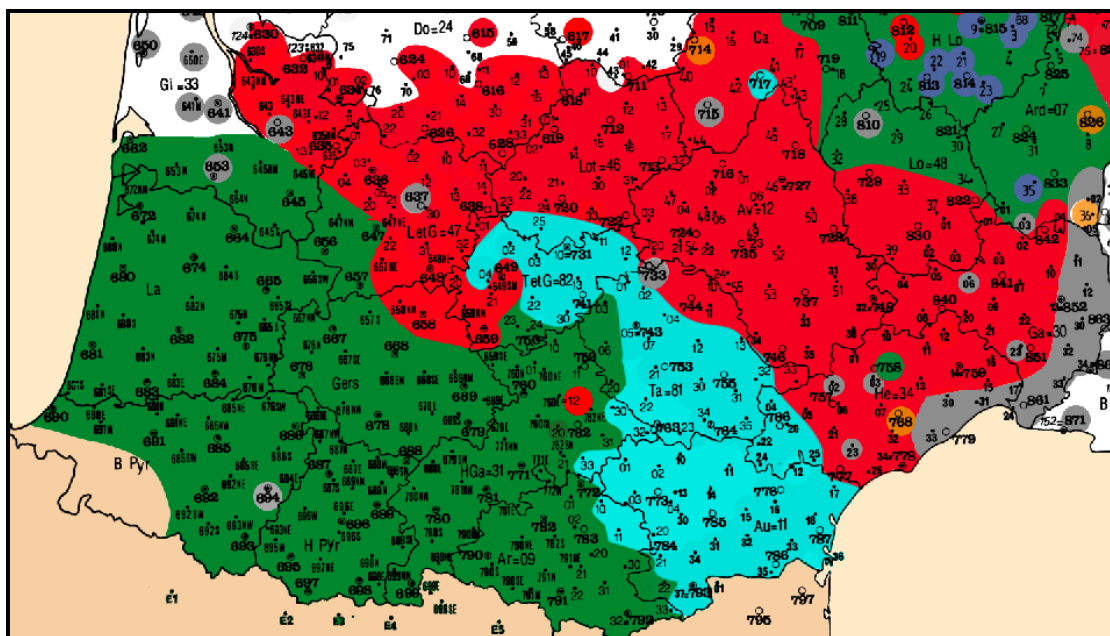
Carte 6 : Sud-Ouest : ordre des pronoms enclitiques



La carte 6 qui récapitule l'ordre des pronoms, témoigne d'une grande stabilité par rapport à l'ALF, avec pourtant en moyenne 70 ans d'écart entre les deux campagnes d'enquêtes.

De même si l'on observe de plus près les différents phonèmes intervocaliques sur la carte 7, les données des atlas régionaux permettent de préciser l'enchevêtrement des faits sans toutefois modifier le grand schéma tracé par l'ALF.

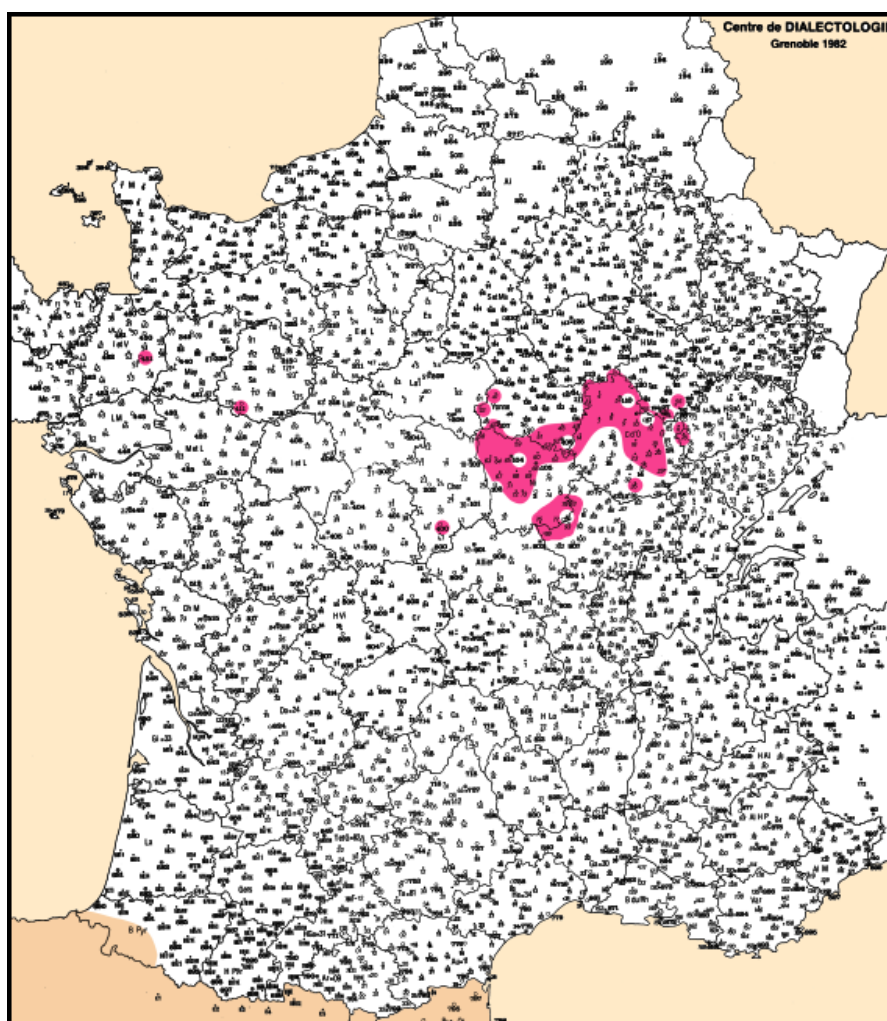
La zone de [b] est toujours bien encadrée par celle de [zʒ] et les zones gasconne et auvergnate qui n'ont pas de phonème intervocalique.



Carte 7 : Sud-Ouest : phonèmes intercalaires

Mais, la perspective est assez différente si l'on porte son regard sur les données du Nord-Est de la France, données disponibles car publiées, mais qui n'ont guère été exploitées jusqu'ici.

D'une part, on remarquera sur la carte 8 une zone relativement vaste d'emploi du pronom démonstratif "ça" à la place du pronom neutre, dans le contexte "Dis-le-moi" relevé dans l'Atlas de la Bourgogne (c. 1731), fait qui n'avait pas été relevé du tout par l'ALF, sauf quelques points épars à l'Ouest.

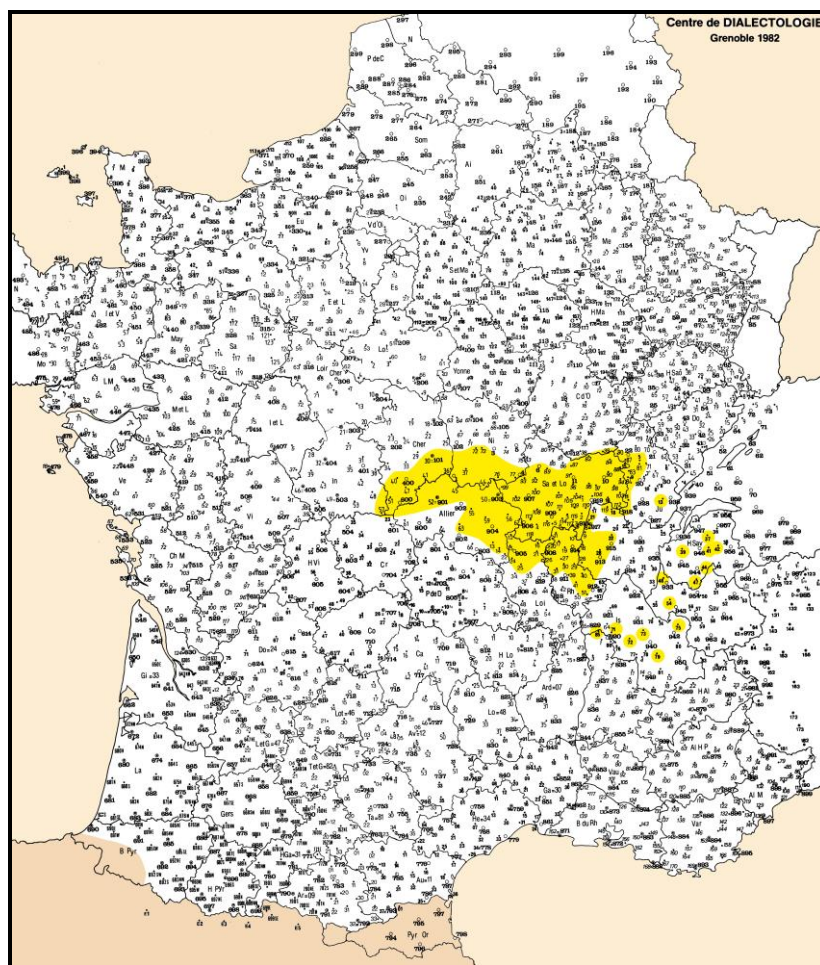


Carte 8 : répartition de ça, dans "dis-le-moi" (ALB et ALF)

Est-ce une nouveauté qui s'est introduite récemment ? Quel rôle ce pronom joue-t-il exactement ? Telles sont les questions que l'on peut se poser.

Car ce pronom "ça" n'apparaît pas dans le contexte "je le sais, on le sait" des Atlas de la Bourgogne (c. 1730), du Centre (c. 1196), du Lyonnais (c. 1238), du Massif Central (c. 1829), de la Franche-Comté (c. 1263) et du Jura et des Alpes (c. 1607).

Par contre, ce même contexte peut nous permettre de visualiser exactement l'aire d'emploi du pronom "y" qui a été évoqué plus haut. On remarquera sur la carte 9 que la zone est très homogène dans sa partie occidentale, tandis que la partie orientale semble plus disloquée. Encore une fois, on peut s'interroger sur le contexte d'enregistrement des données qui a pu jouer un rôle.



*Carte 9 : répartition de "y", dans "je le sais"*

En conclusion, ces quelques cartes démontrent le grand degré de variation dans le comportement des pronoms clitiques et laissent entrevoir un champ d'investigation particulièrement riche. Elles illustrent aussi à quel point l'ALF et les Atlas Linguistiques Régionaux peuvent être une source très féconde et très utile pour qui veut étudier les pronoms clitiques dans l'espace gallo-roman. Sans bouleverser les données fondamentales dont nous disposons pour les langues romanes, ils peuvent apporter leur contribution par l'importante masse de données que les outils informatiques d'aujourd'hui peuvent nous permettre de manipuler aisément.

## Bibliographie

- ALF : Gilliéron, Jules et Edmont, Edmond (1902-1910), *Atlas Linguistique de la France*. Paris, Champion.
- ALAL : Potte, Jean-Claude (1975-1992), *Atlas linguistique et ethnographique de l'Auvergne et du Limousin*. Paris, CNRS (3 vol.).
- ALB : Taverdet, Gérard (1975-1980), *Atlas linguistique et ethnographique de la Bourgogne*. Paris, CNRS (3 vol.).
- ALCe : Dubuisson, Pierrette (1976-1982), *Atlas linguistique et ethnographique du Centre*. Paris, CNRS (3 vol.).
- ALFC : Dondaine, Colette (1972-1991), *Atlas linguistique et ethnographique de la Franche-Comté*. Paris, CNRS (4 vol.).
- ALG : Séguy, Jean (1954-1974), *Atlas linguistique et ethnographique de la Gascogne*. Paris, CNRS (6 vol.).
- ALJA : Martin, Jean-Baptiste et Tuaille, Gaston (1971-1978), *Atlas linguistique et ethnographique du Jura et des Alpes du nord*. Paris, CNRS (3 vol.).
- ALLOc : Ravier, Xavier (1978-1993), *Atlas linguistique et ethnographique du Languedoc occidental*. Paris, CNRS (4 vol.) et carnets d'enquêtes.
- ALLOr : Boisgontier, Jacques (1981-1986), *Atlas linguistique et ethnographique du Languedoc oriental*. Paris, CNRS (3 vol.) et carnets d'enquêtes.
- ALLy : Gardette, Pierre (1967-1976), *Atlas linguistique et ethnographique du Lyonnais*. Paris, CNRS (5 vol.).
- ALMC : Nauton, Pierre (1955-1961), *Atlas linguistique et ethnographique du Massif central*. Paris, CNRS (3 vol.).
- Bonami, Olivier et Boyé, Gilles. "French pronominal clitics and the design of Paradigm. Function Morphology" (<http://www.llf.cnrs.fr/Gens/Bonami/BoBo-MMM05.pdf>)
- Brun-Trigaud, Guylaine, Le Dû, Jean et Le Berre, Yves (2005), *Lectures de l'Atlas Linguistique de la France de Gilliéron et Edmont. Du temps dans l'espace*. Paris, CTHS.
- Dalbera, Jean-Philippe (1986), "Les Pronoms personnels atones dans les parlers des Alpes-Maritimes. Champ et mécanismes de variation". *Actes du XVIII<sup>e</sup> Congrès International de Linguistique et de Philologie Romanes*, T. III, Université de Trèves, Tübingen, 1991, pp. 509-613.

- Escoffier, Simone (1958), *La rencontre de la langue d'Oïl, de la langed'oc, et de francoprovençal entre Loire et Allier*. Lyon, Publications de l'Institut linguistique romane de Lyon.
- Hinzelin, Marc-Olivier, "La position des pronoms clitiques en ancien occitan".  
([http://www.hinzelin.net/resources/Hinzelin\\_AIEO\\_Bordeaux.pdf](http://www.hinzelin.net/resources/Hinzelin_AIEO_Bordeaux.pdf))
- Hirschbühler, Paul et Labelle, Marie (2001), "La position des clitiques par rapport au verbe à l'impératif dans l'évolution du français". *Recherches linguistiques de Vincennes*, 30, pp. 13-38.
- Miller, Philip et Monachesi, Paola (2003), "Les pronoms clitiques dans les langues romanes". In Godard, Danièle. (éd), *Langues Romanes, problèmes de la phrase simple*. Paris, Editions du CNRS, pp. 67-123. ([http://stl.recherche.univ-lille3.fr/sitespersonnels/miller/Miller\\_Monachesi\\_clitiques.pdf](http://stl.recherche.univ-lille3.fr/sitespersonnels/miller/Miller_Monachesi_clitiques.pdf))
- Morin, Yves-Charles (1979), "La morphophonologie des pronoms clitiques en français populaire". *Les Cahiers de Linguistique de l'Université du Québec*, n°9, pp. 1-36.
- Regnier, Claude (1968), "Le pronom personnel régime neutre dans les parlers du Morvan". *Mélanges Gamillscheg*. Munich, pp. 461-476.
- Ronjat, Jules (1930-1941), *Grammaire istorique des parlers provençaux modernes*. Montpellier, Société des Langues Romanes.
- Rooryck, Johan (1992), "Romance enclitic ordering and Universal Grammar". *The Linguistic Review* 9.3, pp. 219-250.
- Taverdet, Gérard (1980), *Les patois de Saône-et-Loire*. Dijon, ABDO.
- Tuailion, Gaston (1969), "Substrat et structure : à propos d'un solécisme du français populaire de Lyon et de sa région", *Travaux de Linguistique et de Littérature romanes*, t.VII,1, pp. 169-176.

*Pierre-Aurélien Georges<sup>1</sup>*  
*Université de Nice, UMR 6039*  
*pierre-aurelien.georges@unice.fr*

## **Les chaînes de clitiques : l'outil informatique au service de l'analyse comparative**

### **1. Le Module MorphoSyntaxique du Thesoc**

#### ***1.1. Présentation***

Le Thesaurus Occitan (ou Thesoc en abrégé) est une base de données dialectales recouvrant tout le domaine occitan. Pour pouvoir figurer dans le Thesoc, les faits linguistiques doivent remplir les deux conditions suivantes :

- les faits doivent être de nature orale et sont saisis dans la base avec leur transcription phonétique API.
- Les faits doivent être précisément localisés, ce qui permet par la suite d'étudier la variation diatopique et de comparer les faits d'une localité à l'autre.

Actuellement, la base contient plus d'un million d'entrées lexicales, correspondant à 831 points d'enquête répartis sur toute l'Occitanie. Une partie de ces entrées lexicales est consultable en ligne sur Internet <sup>2</sup> . Le THESOC contient également des documents multimédia (plus de 1000 enregistrements sonores et plus de 500

---

<sup>1</sup>Ingénieur de Recherche au laboratoire UMR 6039 - Bases, Corpus, Langage, Université de Nice Sophia-Antipolis, CNRS ; MSH de Nice, 98 bd E. Herriot, 06200 Nice.

<sup>2</sup> Cf. Dalbera (1992-).

documents visuels), des données bibliographiques, et des outils d'analyse tels qu'un module cartographique interactif permettant de générer des cartes linguistiques à la volée<sup>3</sup>.

A côté de la base lexicale qui constitue le « cœur historique » du THESOC, signalons également la présence d'un volet de microtoponymie, ainsi que d'un volet plus particulièrement dédié à l'étude de la syntaxe et de la morphosyntaxe, qui a été mis en place depuis quelques années.

Au départ, ce volet était intitulé la « base TEXTES »<sup>4</sup> du THESOC, car il contenait essentiellement des textes et ethnotextes. Puis, au fur et à mesure des évolutions et améliorations apportées, il a été rebaptisé « Module MorphoSyntaxique » afin de lui donner un nom plus évocateur quant à sa finalité d'utilisation et ses possibilités d'exploitation des données, mais aussi pour tenir compte également de la présence de phrases isolées dans la base, qui viennent compléter les textes. Il peut s'agir par exemple de phrases issues de questionnaires d'enquête morphosyntaxique ou de phrases provenant de carnets d'enquête ou de certains atlas linguistiques tel que l'ALMC<sup>5</sup>. Comme pour le reste du THESOC, il est en outre possible d'attacher des documents multimédia à une phrase ou un texte de la base.

## ***1.2. Aperçu des fonctionnalités de la base MMS***

Ce Module MorphoSyntaxique (MMS) est doté d'un certain nombre d'outils linguistiques. Signalons notamment la présence des éléments suivants :

- un transcripteur, qui permet de générer automatiquement une graphie phonologisante à partir de la transcription phonétique. Cette graphie, proche de la graphie mistralienne, permet un premier niveau de « lissage » qui gomme en quelque sorte la variation phonétique, et

---

<sup>3</sup> Pour une présentation générale du THESOC, Cf. Oliviéri (2006).

<sup>4</sup> Cf. Georges (2009).

<sup>5</sup> *Atlas Linguistique du Massif Central*, Nauton (1957-1963)

servira de point de départ à tous les autres traitements linguistiques effectués<sup>6</sup>.

- un lemmatiseur, qui identifie chaque élément lexical d'une phrase ou d'un texte en se basant sur un dictionnaire intégré à la base. Afin de pouvoir gérer convenablement la variation (graphique, dialectale ou flexionnelle), ce dictionnaire est structuré sur deux niveaux hiérarchiques : les variantes sont regroupées sous différents lemmes, ce qui permet tour à tour d'effectuer une recherche ou un traitement, ou bien sur une variante particulière (telle flexion, avec telle graphie, dans tel dialecte), ou bien sur toutes les variantes associées à un lemme donné.

- un analyseur syntaxique, qui se base sur les données issues de la lemmatisation et sur le dictionnaire de la base pour proposer une ou plusieurs structures syntaxiques pour chaque phrase de la base<sup>7</sup>. Les arborescences ainsi générées se situent dans un cadre générativiste, mais les règles syntaxiques sur lesquelles l'analyseur se base pour effectuer son analyse sont modifiables par l'utilisateur.

Ces outils permettent d'automatiser en grande partie le travail d'étiquetage et d'annotation des données de la base. Une fois que les phrases et les textes ont ainsi été annotés, l'on peut utiliser les différentes fonctionnalités de recherches de la base pour sélectionner des données selon divers critères : par lemmes, par variantes, par catégorie syntaxique, par localité, par séquence de catégories syntaxiques ; recherche des phrases contenant tel ou tel fragment de structure syntaxique, etc. On peut ensuite générer et exporter un corpus de travail à partir des résultats de cette recherche. La base dispose d'ailleurs de fonctionnalités d'importation et d'exportation des données aux formats XML et TEI<sup>8</sup>.

---

<sup>6</sup> Rappelons que l'objectif de la base MMS est d'étudier la syntaxe et la morphosyntaxe. C'est pourquoi nous pouvons nous permettre ici de faire abstraction de la variation des réalisations phonétiques pour simplifier la suite des traitements linguistiques opérés. Toutefois, l'information n'est pas perdue puisque, à tout moment, l'utilisateur de la base a toujours accès à la transcription phonétique d'une phrase ou d'un texte, affichée juste à côté de sa graphie.

<sup>7</sup> A partir de maintenant, lorsque nous parlerons de « phrases » au sein de la base MMS, cela désignera toutes les phrases de la base, que ce soit des phrases isolées ou des phrases appartenant à un texte.

<sup>8</sup> Pour une présentation plus détaillée de la base MMS, Cf. Georges (à paraître).



Dernièrement, une nouvelle fonctionnalité a fait son apparition dans MMS : il s'agit de la possibilité d'attribuer à certaines phrases des « étiquettes » définies par l'utilisateur (plus communément appelées « *tags* » en anglais), puis de faire des recherches croisées en fonction de ces étiquettes. C'est cette fonctionnalité que nous allons maintenant mettre en avant dans les exemples suivants pour illustrer en quoi cet outil informatique qu'est le Module MorphoSyntaxique pourrait être d'une grande utilité pour les linguistes syntacticiens ou morphosyntacticiens désirant travailler sur la variation dialectale.

## **2. Traitement des chaînes de clitiques dans MMS**

### ***2.1. Préambule***

Il s'agit ici de montrer quelques possibilités offertes par l'outil informatique à travers un cas concret que nous avons choisi pour l'occasion : les chaînes de clitiques.

Nous ne reviendrons pas ici sur ce qui a déjà été dit à propos des chaînes de clitiques dans le domaine occitan, sujet abordé dans ce même ouvrage par un article de Guylaine Brun-Trigaud (pages 139 à 155) auquel le lecteur pourra se référer pour plus d'informations sur le sujet.

Pour les besoins de la démonstration, nous nous limiterons aux données présentes dans l'ALMC sur les cartes n°1827 à 1838 et plus particulièrement, les réponses aux questions « donne-le-moi », « donne-le-lui », « donne-le-nous », « donne-le-leur », « dis-le-moi », « dis-le-lui », « dis-le-nous », et « dis-le-leur ». Cela nous permet en effet de disposer d'un ensemble homogène puisque toutes ces formes sont des impératifs.

### ***2.2. Démarche suivie***

Tout d'abord, les réponses des 55 points d'enquête de l'ALMC pour ces cartes-là ont été saisies dans la base MMS. Parfois, il y a deux réponses distinctes à une même question dans une même localité. Parfois aussi, il peut arriver qu'il n'y ait pas de réponse à la question demandée dans une certaine localité. Cela ne pose pas de problèmes particuliers pour la saisie dans la base MMS, mais cela

explique pourquoi le nombre de réponses présentes dans la base est différent pour chaque question : il varie entre 38 et 60 réponses par question, avec une moyenne qui se situe malgré tout autour de 55 réponses par question.

Nous avons associé à toutes ces réponses des étiquettes en fonction des critères suivants :

- toutes les phrases dans lesquelles on observe l'ordre « clitique datif ; clitique accusatif » se sont vues attribuer l'étiquette DAT\_ACC
- toutes les phrases dans lesquelles on observe l'ordre inverse, « clitique accusatif ; clitique datif » se sont vues attribuer l'étiquette ACC\_DAT

Notons au passage que la lemmatisation préalable des phrases, bien que facultative, permet ici de gagner un temps précieux puisqu'elle permet alors d'effectuer une recherche de toutes les séquences « pronom clitique datif suivi de pronom clitique accusatif », puis d'affecter d'un seul coup une étiquette donnée à toutes les phrases listées dans les résultats de cette recherche.

Parallèlement à ce premier jeu d'étiquettes DAT\_ACC *versus* ACC\_DAT, nous avons introduit d'autres jeux d'étiquettes :

- toutes les phrases contenant le verbe « donner » se sont vu attribuer l'étiquette « DONNE... »
- toutes les phrases contenant le verbe « dire » se sont vu attribuer l'étiquette « DIS... »

De la même manière, nous avons créé des étiquettes « ...MOI... », « ...LUI... », « ...NOUS... », « ...LEUR... » pour repérer les phrases qui contiennent ces différents pronoms.

À partir de là, il est tout à fait possible de combiner ces jeux d'étiquettes pour former de nouvelles étiquettes si on le souhaite.

Signalons au passage qu'un petit nombre de phrases n'ont pas pu être étiquetées pour diverses raisons : ou bien elles ne correspondaient pas au schéma « Verbe + clitique + clitique » auquel nous avons choisi de nous intéresser, ou bien leur transcription phonétique laissait planer un doute ou une ambiguïté si bien que nous n'avons pas pu analyser leurs constituants avec certitude.

Après avoir étiqueté les différentes phrases, on peut ensuite jongler avec ces jeux d'étiquettes pour effectuer des recherches dans la base ou générer des statistiques. Par exemple, l'utilisateur peut très facilement demander au moteur de recherche de la base la requête

suivante : « afficher la liste des localités dans lesquelles il existe au moins une phrase étiquetée DIS... + ACC\_DAT + ...MOI... et au moins une phrase étiquetée DONNE... + ACC\_DAT + ...MOI... »<sup>9</sup>. Ce que nous avons fait, et la base nous a alors indiqué qu'il y a 18 localités qui correspondent à ces critères de recherche. Il est possible ensuite de consulter la liste des localités en question et d'accéder à toutes les réponses qui ont été récoltées dans ces points d'enquête, notamment pour consulter la transcription phonétique de ces réponses.

Évidemment, si l'on se contente de ne faire qu'une seule requête dans le moteur de recherche, l'investissement, en termes de temps passé, qu'il a fallu pour saisir les réponses dans la base et les étiqueter est supérieur au bénéfice obtenu. En revanche, dès lors que l'on commence à effectuer plusieurs requêtes différentes, que l'on manipule les jeux d'étiquettes pour effectuer différents traitements ou différentes recherches, le bénéfice que l'on en retire devient vite largement supérieur à l'investissement de départ.

---

<sup>9</sup> Pour être précis, il faut également ajouter comme critère : « et ne contenant aucune phrase étiquetée DIS... + DAT\_ACC + ... MOI... ni aucune phrase étiquetée DONNE... + DAT\_ACC + ...MOI... », afin d'éliminer les localités dans lesquelles les deux formes concurrentes (DAT\_ACC et ACC\_DAT) ont été recueillies, et pour lesquelles il existe donc vraisemblablement une variation libre dans l'ordre de ces clitiques.

### 2.3. Résultats obtenus

Ainsi, en procédant de la sorte avec différentes combinaisons d'étiquettes, nous avons pu dresser le tableau suivant en seulement quelques clics de souris :

	<i>“Dis-le-moi”</i>	<i>“Dis-moi-le”</i>
<i>“Donne-le-moi”</i>	18 localités	1 seule localité : <b>MARCHASTEL</b>
<i>“Donne-moi-le”</i>	0	34 localités

	<i>“Dis-le-lui”</i>	<i>“Dis-lui-le”</i>
<i>“Donne-le-lui”</i>	15 localités	1 seule localité : <b>MARCHASTEL</b>
<i>“Donne-lui-le”</i>	1 seule localité : <b>PIERREFORT-PAULHENC</b>	6 localités

	<i>“Dis-le-nous”</i>	<i>“Dis-nous-le”</i>
<i>“Donne-le-nous”</i>	17 localités	2 localités : <b>MARCHASTEL</b> et <b>SAINT-GERMAIN-DU-TEIL</b>
<i>“Donne-nous-le”</i>	0	34 localités

	<i>“Dis-le-leur”</i>	<i>“Dis-leur-le”</i>
<i>“Donne-le-leur”</i>	18 localités	6 localités : <b>MARCHASTEL, MENET,</b> <b>MILLAU, ROCHEPAULE,</b> <b>SAINT-GERMAIN-DU-TEIL,</b> <b>SAINT-BONNET-DE-SALERS</b>
<i>“Donne-leur-le”</i>	1 seule localité : <b>PIERREFORT-PAULHENC</b>	16 localités

Tableau 1. Ordre des clitiques suivant le verbe utilisé

Comme on le voit sur la figure 1, en règle générale, quel que soit le clitique datif utilisé, il semblerait que le choix du verbe « donner » ou « dire » n’influence pas vraiment l’ordre des clitiques datif et accusatif, à quelques exceptions près.

Si l’on regarde de plus près les quelques cas qui semblent faire exception à cette tendance générale, on voit que deux localités reviennent de manière récurrente (en gras sur la figure 1) :

- Marchastel (15), point d’enquête n°32 de l’ALMC
- Pierrefort-Paulhenc, point d’enquête n°43.

Le paradigme de Marchastel (15) constitue ainsi une exception généralisée à cette tendance :

intitulé	réponse (phonétique)
donne-le-moi	b'ajɫɔ lu mi
donne-le-lui	b'ajɫɔ lu i
donne-le-nous	b'ajɫɔ lu nu
donne-le-leur	b'ajɫɔ lu lur
dis-le	diʒɔ zu
dis-le-moi	diʒɔ m u
dis-le-moi	diʒɔ m zu
dis-le-nous	diʒɔ nuz u
dis-le-leur	diʒɔ ju zu
dis-le-lui	diʒɔ ju zu
on ne le sait pas	lɔd u sa pa
on ne peut pas le savoir	lɔ pɔ pa u s'awre
tu veux le savoir ?	bɔ zu s'awre
tu veux le savoir ?	u bɔs s'awre

Figure 2. Séquences de clitiques recueillies à Marchastel

Malheureusement, l’ALMC ne nous fournit ces chaînes de clitiques que pour les deux verbes « donne » et « dis ». Aussi serait-il donc intéressant, à la vue de ces données, de faire une enquête complémentaire sur la commune de Marchastel pour vérifier avec d’autres verbes si l’on constate également cette différence dans l’ordre des clitiques datif et accusatif. Si c’est le cas, cela signifiera alors que le choix du verbe peut influencer l’ordre des clitiques dans ces dialectes-là, et qu’il serait alors bon de pouvoir recueillir aussi de

nouvelles données pour les 54 autres points d'enquête afin de compléter les réponses déjà attestées et de pouvoir vérifier avec d'autres verbes que « donne » ou « dis » si l'ordre des clitiques y reste identique.

En combinant autrement les différents jeux d'étiquettes, nous avons également obtenu les chiffres suivants :

	<i>“Dis-le-moi”</i>	<i>“Dis-moi-le”</i>
<i>“Dis-le-lui”</i>	15 localités	1 seule localité : PIERREFORT-PAULHENC
<i>“Dis-lui-le”</i>	0	7 localités

	<i>“Donne-le-moi”</i>	<i>“Donne-moi-le”</i>
<i>“Donne-le-lui”</i>	19 localités	0
<i>“Donne-lui-le”</i>	0	36 localités

	<i>“Dis-le-leur”</i>	<i>“Dis-leur-le”</i>
<i>“Dis-le-lui”</i>	14 localités	2 localités : MILLAU, SAINT-GERMAIN-DU-TEIL
<i>“Dis-lui-le”</i>	0	5 localités

	<i>“Donne-le-leur”</i>	<i>“Donne-leur-le”</i>
<i>“Donne-le-lui”</i>	19 localités	0
<i>“Donne-lui-le”</i>	9 localités	25 localités

	<i>“Dis-le-nous”</i>	<i>“Dis-nous-le”</i>
<i>“Dis-le-lui”</i>	14 localités	2 localités : MILLAU, SAINT-GERMAIN-DU-TEIL
<i>“Dis-lui-le”</i>	0	5 localités

	<i>“Donne-le-nous”</i>	<i>“Donne-nous-le”</i>
<i>“Donne-le-lui”</i>	19 localités	0
<i>“Donne-lui-le”</i>	0	36 localités

Tableau 3. Ordre des clitiques suivant le paradigme des personnes

On peut y observer qu'en règle générale, le paradigme est uniforme : l'ordre des clitiques datif et accusatif reste le même à toutes les personnes, contrairement à ce qui se passe en français par exemple avec la troisième personne qui se comporte différemment des deux autres personnes. Seules 3 localités sur les 55 contreviennent à cette tendance : on retrouve la localité Pierrefort-Paulhenc (15) qui se distinguait déjà dans la figure 1, ainsi que les localités de Millau (12) et St Germain du Teil (12), respectivement points d'enquête n°51 et 38 de l'ALMC.

On aura remarqué que dans les tableaux de la figure 1 et de la figure 3, le total des 4 cases blanches n'est pas toujours égal au nombre de points d'enquête de l'ALMC (à savoir 55 localités). Cela résulte de trois facteurs :

- Comme indiqué précédemment, certaines phrases n'ont pas pu être étiquetées et n'ont donc pas été prises en compte dans le calcul de ces chiffres.
- Nous avons filtré dans nos critères de recherche les localités pour lesquelles deux formes concurrentes sont attestées, selon les modalités précisées plus haut dans une note de bas de page.
- Lorsque l'on croise deux questions différentes comme c'est le cas ici, les localités qui ne possèdent aucune réponse pour l'une ou l'autre de ces deux questions ne peuvent pas être prises en compte.

Comme on peut le voir à travers ces deux exemples, cette procédure de comparaison assistée par l'outil informatique permet, en quelques clics de souris, non seulement de dégager des tendances sur une aire dialectale et d'évaluer la corrélation ou non entre deux phénomènes linguistiques observés, mais également de faire ressortir les points d'enquête qui se comportent singulièrement et ce de manière récurrente. Cela permet ensuite au chercheur de se focaliser plus précisément sur ces points particulièrement intéressants du point de vue du phénomène linguistique considéré.

On pourra alors étudier en quoi et pourquoi ces points d'enquête se distinguent de la tendance générale : s'agit-il d'une particularité du dialecte local, et si oui, quelle est cette particularité et quel est son fonctionnement ? Ou bien s'agit-il simplement d'un artefact lié aux conditions particulières de l'enquête (utilisation de plusieurs témoins différents sur un même point d'enquête, variation libre qui était passée inaperçue faute d'avoir récolté suffisamment de matériau linguistique,

erreur de transcription phonétique, etc.). Dans ce dernier cas, le « contre-exemple » peut alors être « éliminé » et vient alors renforcer la tendance générale constatée.

Cela est tout particulièrement intéressant pour les syntacticiens qui travaillent dans un cadre théorique générativiste : l'approche « Principes et Paramètres »<sup>10</sup> de la grammaire générative dissocie d'une part des Principes généraux universels communs à toutes les langues, et d'autre part des Paramètres qui sont positionnés différemment d'une langue à l'autre. Cette « conception paramétrique de la variation linguistique »<sup>11</sup> permet de comparer les dialectes et les langues du monde du point de vue de la syntaxe. Dans ce cadre-là, cette procédure de comparaison et de recherche de corrélations éventuelles pourrait aider les chercheurs à dégager de nouveaux principes et à valider ou invalider leurs hypothèses de travail.

#### **2.4. Commentaires critiques**

Il nous faut à présent répondre à un certain nombre de critiques pertinentes qui pourraient être soulevées suite à la démarche que nous venons de présenter.

Tout d'abord, il faut signaler que dans un souci de rigueur scientifique, il faudrait en principe faire la distinction entre pronom neutre et pronom masculin de la troisième personne du singulier, car certains dialectes produisent deux réalisations différentes pour le pronom clitique qui se traduit en français par « le », suivant qu'il s'agisse du pronom neutre, comme dans « je le lui ai dit », ou du pronom masculin de la troisième personne du singulier, comme dans « je le lui ai donné (le livre) »<sup>12</sup>. Il faudrait donc s'interdire de comparer sur un même plan des formes comme « donne-le-lui » et « dis-le-lui », car l'ordre des clitics pourrait peut-être y être influencé par la nature même des clitics (suivant qu'il s'agit d'un pronom neutre ou non). Cependant, il s'agissait ici uniquement de démontrer l'utilité de l'outil informatique et de présenter simplement

---

<sup>10</sup> Cf. Chomsky (1981).

<sup>11</sup> Chomsky (1982 : 349).

<sup>12</sup> Cf. Tuailon (1969). Cette distinction n'est d'ailleurs pas cantonnée à la région lyonnaise, mais se retrouve également dans certains parlers des Alpes Maritimes, comme l'indique Dalbera (1991 : 604), ainsi que dans la région limousine, où j'ai d'ailleurs eu l'occasion de le constater personnellement sur le terrain lors d'une enquête linguistique réalisée à 6 kms au sud de la commune d'Eymoutiers (87), au lieu-dit « Meilhaguet ».



son utilisation et son fonctionnement. Les chercheurs désireux d'utiliser cet outil auront tout le loisir de peaufiner la méthode présentée ici et de prendre en compte ces considérations plus en détail<sup>13</sup>.

En outre, d'aucuns diront que les questionnaires de phrases, technique très souvent utilisée lors des enquêtes linguistiques, entraînent de la part des informateurs des réponses « moins naturelles » qu'un discours libre par exemple, et augmentent le risque de calques sur le français. On peut dès lors se poser la question de la pertinence d'utiliser ce type de méthode d'enquête lorsqu'il s'agit d'étudier la syntaxe ou la morphosyntaxe dialectale. Par ailleurs, l'étude des chaînes de clitiques rencontre un problème supplémentaire à l'utilisation des questionnaires de phrases : d'une part, les chaînes de clitiques s'avèrent finalement assez rares à l'oral, et d'autre part, lorsque l'on demande à un informateur de traduire une phrase contenant une chaîne de clitiques du type « il le lui donne », il n'est pas rare que ce dernier « élimine » l'un des clitiques lors de la traduction, de sorte que l'on perd un élément de la chaîne de clitiques et que l'on se retrouve au final avec quelque chose du type « i b'ajlɔ »<sup>14</sup>.

Ainsi, comme Séguy (1973) l'a fait remarquer :

« la syntaxe proprement dite, en dehors de quelques traits épais, échappe à l'enquête par questionnaire. Si on demande aux informateurs de traduire dans leur dialecte des énoncés en français standard, on n'obtient aucune différence dans les régions centrales, et, dans les régions périphériques, on provoque souvent des calques. Les discours sollicités, comme ceux qui ont servi de tests à X. Ravier, bien qu'ils soient libres, présentent une syntaxe monotone et appauvrie. Le seul moyen est d'écouter du langage spontané, des conversations entre patoisants. »

Pour prendre en compte ces différentes remarques et y apporter une solution, plutôt que de se baser sur des phrases isolées issues de questionnaires de phrases, il est tout à fait possible à la place d'utiliser des ethnotextes, discours libres ou enregistrements de conversations

---

<sup>13</sup> L'ALMC ne nous fournit malheureusement pas suffisamment de données pour pouvoir étudier séparément le pronom neutre et le pronom masculin de la troisième personne du singulier au sein des chaînes de clitiques. Pour bien faire, il faudrait donc envisager d'avoir recours à des données provenant d'autres sources.

<sup>14</sup> Mot à mot : « (il) lui donne »

spontanées, en appliquant exactement la même méthode que celle décrite précédemment : la base MMS contient en effet à la fois des textes et des phrases isolées, et le système d' « étiquettes » qui vient d'être présenté ici s'applique tout aussi bien aux phrases isolées qu'aux phrases contenues dans un texte.

Si l'on s'intéresse plus particulièrement aux chaînes de clitiques, il est vrai que pour arriver à récolter le même nombre d'occurrences de telles chaînes de clitiques, il faudra une quantité de discours libres beaucoup plus conséquente que si l'on utilise des questionnaires de phrases qui orientent la réponse de l'informateur et provoquent l'apparition de chaînes de clitiques avec une plus grande probabilité. On pourrait cependant également envisager une approche mixte, en utilisant à la fois les phrases isolées issues des questionnaires, et les textes issus de discours libres.

D'une manière générale, ces critiques s'appliquent donc surtout à la nature même des données utilisées et à la façon dont elles ont été recueillies en enquête et comment elles sont exploitées, mais ne remettent pas en cause le procédé de traitement de ces données en lui-même, puisque ce même traitement peut être envisagé avec différents types de données.

### **3. Perspectives**

La méthode que nous venons de présenter pourrait être utilisée pour faire des recherches sur différents aspects de la morphosyntaxe des dialectes, quel que soit le cadre théorique utilisé. Un cas a cependant retenu notre attention : l'utilisation de cette méthode dans un cadre générativiste soulève quelques réflexions intéressantes et nécessite de prendre quelques précautions que nous allons maintenant aborder.

#### ***3.1. Utilisation dans un cadre générativiste***

Si l'on se situe dans un des cadres théoriques de la grammaire générative, les fonctionnalités de la base MMS qui ont été présentées ci-dessus pourraient être utilisées autour de l'approche « Principes et Paramètres » de deux manières complémentaires : soit de façon heuristique, pour tenter de rechercher la présence éventuelle de corrélations entre deux paramètres sans avoir *a priori* de position sur

le sujet, soit *a posteriori*, pour confirmer ou infirmer une hypothèse sur la question.

Dans ce dernier cas, je pense notamment à la corrélation établie par Rizzi (1982) entre « sujet nul / *pro drop* » (ou disons plutôt, « pronom sujet non réalisé », de manière plus neutre) et inversion libre du sujet et du verbe<sup>15</sup>.

Cette corrélation est parfois remise en question par certains linguistes, et l'on peut à juste titre se demander si ce principe d'une corrélation entre paramètre du sujet nul et inversion libre du sujet pourrait toujours tenir si l'on étudiait désormais les faits linguistiques à l'échelle de la micro-variation dialectale ?

Oliviéri (2004) a montré que le comportement des sujets post-verbaux de l'occitan n'est pas tout à fait le même qu'en italien. L'ordre des constituants en occitan est certes plus libre qu'en français, mais ne l'est pas autant qu'en italien, malgré une même absence de clitique sujet<sup>16</sup> et une morphologie verbale riche. Cette corrélation potentielle pourrait donc être remise en question s'il se révélait, au cours d'une étude plus approfondie, que l'inversion du sujet en occitan n'est pas aussi libre qu'on pourrait le supposer ; et ce malgré le fait qu'il s'agisse bien, dans sa grande majorité, de dialectes "à sujet nul".

Pour confronter à la réalité des faits dialectaux cette hypothèse d'une corrélation entre paramètre du sujet nul et inversion libre du sujet, l'idée serait donc de pouvoir comparer différentes grammaires pour voir s'il existe des grammaires dans lesquelles on observe l'un de ces deux paramètres sans pour autant observer l'autre. Autrement dit, existe-t-il des grammaires dans lesquelles on observerait par exemple une inversion libre du sujet sans pour autant avoir la présence de sujet nuls, ou bien l'inverse ?

Notez que l'on parle ici de « grammaires » et non pas de « dialectes » : dans une première approximation, on pourrait être tenté de penser que les deux termes sont ici interchangeable, tout du moins dans ce contexte bien précis, et que si l'on compare plusieurs dialectes entre eux, cela revient à comparer plusieurs grammaires entre elles.

---

<sup>15</sup> Pour résumer, en italien on peut dire « *Maria dorme.* » tout aussi bien que « *Dorme.* » tandis qu'en français le sujet doit être obligatoirement réalisé : « *Marie dort.* » mais « *\*dort.* ». Et parallèlement à cela, l'inversion libre du sujet est possible en italien mais pas en français : « *Dorme Maria.* » mais « *\*Dort Marie.* ».

<sup>16</sup> Dans la grande majorité des dialectes occitans, exception faite des dialectes du nord de l'Occitanie, notamment dans certaines zones du Limousin et de l'Auvergne.

Ainsi, la tentation serait grande de vouloir rechercher s'il existe des localités dans lesquelles on observe des phrases avec sujets nuls mais dans lesquelles on n'observe pas de phrases contenant des inversions du sujet.

Mais il faut faire attention : il ne faut pas perdre de vue que les données linguistiques recueillies sur le terrain ne sont que des productions individuelles ; or il se pourrait, d'après ce qui est apparu dans les recherches récentes (Séminaire de recherche, 2008-2009) que chaque locuteur possède en fait sa propre grammaire, son propre idiolecte. En prenant en compte cette hypothèse, si l'on compare des données provenant de plusieurs dialectes sans s'être assuré au préalable qu'il n'y a bien eu dans ces données qu'un et un seul informateur pour chacun de ces dialectes, le risque de mélanger des phrases provenant de différents locuteurs d'un même dialecte, c'est-à-dire de mélanger différents idiolectes, et donc différentes grammaires, nous guette. L'idéal serait-il donc de s'assurer de n'avoir qu'un seul informateur par point d'enquête ? Cela va à l'encontre de l'idée communément admise qu'il est préférable, dans l'absolu, d'avoir plusieurs informateurs par point d'enquête, de manière à lisser les particularités de chaque locuteur et à avoir une meilleure vision du dialecte de cette localité. Par ailleurs, cela compliquerait énormément la tâche, puisque cela signifierait alors qu'on ne pourrait pas compléter les données recueillies auprès d'un premier informateur par des données recueillies auprès d'un second informateur lors d'une enquête ultérieure sur la même localité.

De toute façon, en allant encore plus loin dans la réflexion, même si l'on s'assurait de n'avoir utilisé qu'un seul informateur par point d'enquête, cela ne réglerait pas totalement le problème, car non seulement chaque locuteur possède son propre idiolecte, mais en réalité il possède même plusieurs grammaires : on n'écrit pas de la même façon que l'on parle, et l'on ne parle pas de la même façon avec ses amis ou sa famille que lorsqu'on s'adresse en public à un auditoire. De même que l'on n'écrit pas non plus de la même façon suivant que l'on rédige un article de linguistique ou que l'on envoie une carte postale à sa grand-mère. Dès lors, comment s'assurer que toutes les phrases qui ont été recueillies dans une même localité ont bien été produites non seulement par le même informateur, mais par une seule et même grammaire de cet informateur ?

Il n'est pas impossible de penser, en effet, que certains paramètres comme la présence de sujets nuls ou l'inversion libre du sujet puissent être paramétrés différemment d'une grammaire à l'autre au sein d'un même individu. Prenons un exemple qui se situe dans le même ordre d'idée : à l'oral, dans un registre familier, la majorité des gens diront plus facilement « Marie elle dort » que « Marie dort » et pourront tout aussi bien dire « elle dort Marie » alors que, dans un écrit, avec un registre soutenu, ils se seraient contentés de la forme standard « Marie dort ».

Si l'on utilise des phrases isolées provenant de questionnaires d'enquête, la situation paraît donc insoluble : difficile de s'assurer que toutes les questions ont été posées dans une même situation, un même cadre, un même contexte, avec la même expression, la même intonation, et le même registre de langue. Et il est encore plus difficile de s'assurer que l'informateur y a systématiquement répondu avec le même registre de langue, le même contexte, bref, la même grammaire.

Mais la situation n'est pas désespérée pour autant. En réalité, la solution permettant de s'affranchir de tous ces problèmes est même relativement simple : lors des enquêtes linguistiques de terrain, il faut utiliser une unité de collecte plus grande que la phrase : le « texte », ou plutôt l'ethnotexte. En effet, au sein d'un ethnotexte, on peut s'attendre à une homogénéité beaucoup plus importante, qu'il serait en pratique difficile d'atteindre avec des phrases isolées.

Ainsi, toutes les phrases d'un ethnotexte :

- ont été prononcées à la même époque (unité temporelle)
- dans un même lieu (unité spatiale)
- sont généralement issues du même locuteur, donc du même idiolecte (unité du locuteur), si l'on met de côté les textes qui contiennent des conversations entre plusieurs personnes<sup>17</sup>.
- sont généralement prononcées dans un même contexte, un même genre, voire un même style, un même registre (unité de la grammaire utilisée par le locuteur), sauf lorsque le locuteur change délibérément de registre pour chercher à imiter quelqu'un, à jouer un rôle, ou à donner une tournure spéciale à certaines phrases de

---

<sup>17</sup> Cela étant, on pourrait très bien imaginer de traiter les conversations entre plusieurs personnes de la manière suivante : chaque locuteur se voit attribuer une étiquette X, Y, Z. Chaque phrase du texte est étiquetée en fonction de son locuteur : X, Y, ou Z. On peut ensuite utiliser ce jeu d'étiquettes pour virtuellement faire comme s'il s'agissait de trois textes différents, dans lesquels il n'y a qu'un seul et même locuteur du début jusqu'à la fin du texte. On peut alors traiter ces trois textes virtuels comme les autres textes de la base.

son discours. Dans ce cas, les phrases en question se repèrent en général assez facilement puisque l'intention du locuteur était justement d'introduire une différence qui soit facilement perceptible par son interlocuteur. Ces quelques phrases-là peuvent donc facilement être écartées le cas échéant.

En d'autres termes, pour s'affranchir de tous les problèmes mentionnés ci-dessus, il suffit de revoir quelque peu la méthode utilisée : pour évaluer s'il existe une corrélation entre les paramètres X et Y, au lieu de rechercher les localités dans lesquelles on a recueilli des phrases contenant la caractéristique X et des phrases contenant la caractéristique Y<sup>18</sup>, on préférera plutôt rechercher les textes qui contiennent des phrases contenant la caractéristique X et des phrases contenant la caractéristique Y en leur sein. Ainsi, on peut s'assurer que l'on est bien en train de chercher à déterminer s'il existe une corrélation entre la présence d'un paramètre X et la présence d'un paramètre Y au sein d'une seule et même grammaire.

Pour reprendre notre exemple concret, si l'on veut valider ou invalider l'hypothèse d'une corrélation entre présence de sujets nuls et inversion libre du sujet, il faudra plutôt rechercher si l'on peut trouver des ethnotextes de longueurs assez conséquentes qui contiennent des phrases possédant une inversion du sujet mais qui ne contiennent aucune phrase avec un sujet nul (ou bien l'inverse) au lieu de rechercher s'il existe des localités dans lesquelles on observe des phrases possédant une inversion du sujet mais dans lesquelles on n'observe aucune phrase avec un sujet nul.

Notons qu'en procédant de la sorte, les ethnotextes n'ont alors en théorie pas nécessairement besoin de provenir de localités différentes, puisque l'on compare non plus des dialectes mais des grammaires individuelles. Cependant, on peut s'attendre à ce que la diversité linguistique soit plus importante entre deux locuteurs géographiquement distants qu'entre deux locuteurs du même village. Si les ethnotextes proviennent d'un grand nombre de points d'enquête variés, cela n'en sera donc que plus bénéfique.

---

<sup>18</sup> ou bien : « des phrases contenant la caractéristique X mais aucune phrase contenant la caractéristique Y », selon la nature de ce que l'on cherche à étudier.

### 3.2. *Études diachroniques*

Jusqu'ici nous avons essentiellement traité :

- de la variation diatopique, puisque nous avons comparé des faits linguistiques entre différentes localités.
- de la variation individuelle<sup>19</sup>, puisque nous avons abordé la problématique de l'utilisation de faits linguistiques provenant de différents locuteurs, afin de comparer différentes grammaires individuelles, dans un cadre générativiste.

D'un tout autre point de vue, cette méthode comparative pourrait également être utilisée dans un cadre diachronique. L'hypothèse formulée par Dalbera (1992) suggère que la diversité dialectale (en synchronie) est le reflet de l'évolution diachronique de la langue :

« Mais l'une des directions d'exploitation les plus claires de la description des faits dans l'espace, à notre sens, est le domaine de la reconstruction. L'idée a été maintes fois avancée que, au plan linguistique, l'espace pouvait être exploité comme projection du temps. En d'autres termes, que la variation et la répartition des faits dans l'espace constituaient des indices de première importance pour l'analyse diachronique. A travers la variation, ce sont des stades d'évolution que l'on appréhende ; à travers la répartition, ce sont des indications sur la succession de ces stades que l'on peut se procurer. »<sup>20</sup>

On pourrait alors utiliser cette méthode comme base pour comparer tel ou tel phénomène ou caractéristique linguistique entre les différents parlars de la zone occitane dans le but de rechercher les différents stades d'évolution de la langue.

Par ailleurs, le Module MorphoSyntaxique contient également un champ « date » qui permet de renseigner, pour chaque texte et phrase de la base, à quelle date ils ont été recueillis sur le terrain. En supposant que l'on puisse arriver à avoir suffisamment de données réparties sur une période de temps relativement étendue, on pourrait alors envisager d'utiliser la méthode décrite ici non plus pour comparer des données linguistiques en synchronie entre différentes localités ou entre différents locuteurs, mais cette fois-ci en diachronie.

---

<sup>19</sup> Pourrait-on dire socio-linguistique ?

<sup>20</sup> Dalbera (1992 : 138).

## **Conclusion**

La base MMS dispose d'outils et de fonctionnalités souples, qui peuvent être utilisés de différentes manières et selon différents cadres théoriques, ce qui lui permet d'assurer pleinement son rôle d'outil informatique d'aide à la recherche en syntaxe et morphosyntaxe sur les dialectes occitans.

En particulier, ce système d'« étiquettes » que nous venons de présenter ici ainsi que les différentes fonctionnalités du Module MorphoSyntaxique qui y sont associées permettent d'évaluer rapidement, sur un très grand nombre de données et de localités, la présence ou l'absence de corrélations entre deux (ou plusieurs) paramètres linguistiques qui sont soupçonnés d'entrer en compte dans l'explication de la variation des faits observés.

On dispose donc là d'un outil qui permet non seulement de valider ou d'invalider des hypothèses linguistiques en les confrontant à la réalité des faits linguistiques, mais aussi de faire ressortir de manière heuristique d'éventuelles corrélations qui auraient pu passer inaperçues jusque-là.

Suivant le type des données disponibles et suivant la nature des recherches envisagées, on pourra appliquer la méthode que nous venons de décrire pour comparer des phrases isolées entre différentes localités, des phrases provenant de textes entre différentes localités, des phrases provenant de différents textes que l'on compare entre eux, ou une quelconque combinaison des trois, selon les besoins et les avantages et inconvénients en jeu.

Au-delà des quelques exemples simples et d'une taille raisonnable que nous venons de voir, pour dégager des corrélations d'ordre général plus importantes et plus intéressantes, il faudra bien sûr exploiter un volume plus important de données linguistiques et sur un plus grand nombre de localités. C'est là que nous pourrions alors vraiment mesurer toute la puissance de l'outil informatique, qui permet de réaliser des travaux qui n'auraient tout simplement pas été humainement envisageables auparavant.



## Bibliographie

- Chomsky, Noam, 1981, *Théorie du Gouvernement et du Liage - les Conférences de Pise*, Editions du Seuil, 1991, Paris.
- Chomsky, Noam, 1982, *La Nouvelle Syntaxe*, Editions du Seuil, 1987, Paris.
- Dalbera, Jean-Philippe, 1991, *Les pronoms personnels atones dans les parlers des Alpes-Maritimes. Champ et mécanismes de variation*, Actes du XVIII<sup>e</sup> congrès international de Linguistique et Philologie Romanes, t. III, Université de Trèves (Trier), 1986, Max Niemeyer Verlag, Tübingen, p. 599-613.
- Dalbera, Jean-Philippe, 1992, *Dialectologie et morphologie*, Actes du Congrès International de Dialectologie, Bilbao, IKER-7 p. p. 135-149.
- Dalbera, Jean-Philippe *et al.*, 1992-, *Thesaurus Occitan*, CNRS - UMR 6039 "BCL". Université de Nice - Sophia Antipolis, <http://thesaurus.unice.fr>.
- Georges, Pierre-Aurélien, 2009, "Présentation de la base Textes associée au THESOC", actes du colloque *La dialectologie hier et aujourd'hui* (1906-2006), Lyon, 2006, Brigitte Horiot (Ed.). p. 81-93.
- Georges, Pierre-Aurélien, à paraître, "Le THESOC : bases de données et outils d'analyse consacrés à l'étude des dialectes occitans", actes du colloque *Bases de données, Méthodes, Modèles de description : de nouvelles perspectives pour la recherche sur les langues régionales et minoritaires ?* Tübingen, 2008, Stauffenburg Verlag (DeLingulis).
- Nauton, Pierre, 1957-1963, *Atlas Linguistique et ethnographique du Massif Central*, Editions du CNRS, Paris.
- Oliviéri, Michèle, 2003, "Y-a-t-il des frontières dialectales en syntaxe ?", actes du 128<sup>e</sup> Congrès des sociétés historiques et scientifiques (CTHS) : *Relations, échanges et coopération en Méditerranée*, Bastia, 14-21 avril 2003, à paraître.
- Oliviéri, Michèle, 2004, "Paramètre du sujet nul et inversion du sujet dans les dialectes italiens et occitans", *Cahiers de grammaire n°29, Questions de linguistique et de dialectologie romane*, 2004, p. p. 105-120.

- Oliviéri, Michèle ; Brun-Trigaud, Guylaine, 2006, "Présentation du logiciel Thesaurus Occitan", actes du colloque *La dialectologie hier et aujourd'hui (1906-2006)*, Lyon, à paraître, Brigitte Horiot (éd.).
- Rizzi, Luigi, 1982, *Issues in Italian Syntax*, Dordrecht, Foris.
- Séguy, Jean, 1973, "Les atlas linguistiques de la France par régions", *Langue Française*, Volume 18, Numéro 1, p. 65-90.
- Tuaillon, Gaston, 1969, "Substrat et structure : à propos d'un solécisme du français populaire de Lyon et de sa région", *Travaux de Linguistique et de Littérature Romanes (TraLiLi)*, t VII,1, p. 169-176.
- Séminaire de recherche*. 2008-2009. CNRS - UMR 6039 "BCL". Université de Nice - Sophia Antipolis.

*Jean-Claude Bouvier*  
*Professeur émérite, Université de Provence*  
*bouvier.jc@wanadoo.fr*

## **Le discours des carnets d'enquête des Atlas linguistiques et ethnographiques : l'exemple du maïs et de la truffe dans la Drôme**

Les atlas linguistiques et ethnographiques que nous avons réalisés ou continuons à réaliser sont certainement plus linguistiques qu'ethnographiques. Cela est dû aux compétences professionnelles de leurs auteurs, dialectologues et donc linguistes avant tout, mais aussi à l'organisation de la recherche, conçue sous forme d'enquêtes linguistiques, et au mode de présentation sous forme de cartes des matériaux recueillis. L'ethnographie a une place importante dans la démarche scientifique des atlas linguistiques et plus largement de la géolinguistique, qui s'appuie sur la perception du rapport entre « les mots et les choses » et qui donc est d'une certaine manière au départ des recherches ethnolinguistiques. Mais, dans les atlas linguistiques, l'ethnographie est malgré tout « à la marge », au sens propre de l'expression, puisque les données ethnographiques sont souvent dans la marge des cartes ou dans des compléments d'information situés à la fin des volumes.

Bien sûr, il existe beaucoup de différences entre les atlas de ce point de vue-là et de même à l'intérieur d'un seul atlas entre les thèmes, les concepts abordés. Dans certains atlas, ce sont des transcriptions de véritables textes oraux ethnographiques que l'on trouve, et pas seulement des notations. Selon les thèmes il y aura plus ou moins de photos ou de dessins, selon la curiosité des enquêteurs plus ou moins de détails techniques pour décrire une pratique etc. Mais d'une façon générale, la primauté est tout de même à l'information linguistique.

Le dialectologue n'est pas un ethnographe, encore moins un ethnologue. Et pourtant, étant un homme de terrain, au contact et à l'écoute des réalités culturelles exprimées par le langage, il recueille beaucoup d'informations qui restent souvent enfouies dans les carnets d'enquête ou les enregistrements qu'il a réalisés. J'ai précisément choisi d'explorer mes propres carnets d'enquête remplis pour l'Atlas Linguistique et Ethnographique de Provence (ALP). Ce sont les carnets de mes enquêtes, c'est-à-dire celles du département de la Drôme, que j'ai effectuées pour ma thèse sur les parlers de la Drôme et pour l'Atlas. Je me limiterai à deux thèmes du même chapitre : le maïs et la truffe.

## **1. Le maïs et la truffe dans le carnet d'enquête et dans les cartes de l'ALP**

### *Le maïs*

Le maïs se trouve aux pages 48-49 du carnet d'enquête : il fait l'objet de 13 questions. Deux cartes seulement lui sont consacrées dans l'atlas : « maïs », carte 410, et « épi de maïs », carte 411 (en demi-page). Mais quatre éléments figurent dans la marge de la carte 410 :

- maïs en herbe
- barbe de maïs
- grandes feuilles
- renverser les feuilles.

Un élément seulement dans la marge de la carte 411 :

- fusée sans grains.

### *La truffe*

La truffe apparaît à la page 48 du carnet d'enquête et ne fait l'objet que de cinq questions. Une carte seulement : « truffe », carte 412. Deux notations dans la marge de cette carte :

- arbre truffier : chêne ou autres (noisetier, tilleul...)
- animal truffier : chien ou autre (truie ou mouche jaune).

Ce simple rappel chiffré montre qu'il y a en principe des « réserves » dans les carnets d'enquête. Cela est d'autant plus plausible que certaines des questions qui ont été posées étaient ouvertes et permettaient à l'informateur de développer son témoignage. Mais les indications données ci-dessus montrent bien que ce sont surtout des matériaux lexicaux qui se trouvent dans les cartes et dans leurs marges et que les informations ethnographiques ne sont présentes que sous forme d'esquisses : par exemple les différents types d'animaux renifleurs de truffes.

## 2. Les noms du maïs et de la truffe

Ces noms sont à eux seuls bien intéressants d'un point de vue à la fois linguistique, géolinguistique et ethno-linguistique. Ils mériteraient une étude particulière, mais je me contenterai ici d'un bref résumé.

### *Le maïs*

Deux grands types lexicaux se partagent le domaine de l'ALP :

- le type *blé de Turquie*, qui prend les formes [bla tyr'ke], [bla de tyr], [bla de tyrk] (Alpes- Maritimes), [tyr'ki], [grã tyrk].... ; ou encore [bla de barba'je], [barba'je] (« Barbarie ») ;
- le type *mil*, venant du latin MILIUM et ses dérivés : məj 'nɛgre « mil noir », 'mɛla, 'mɛjo (FEW, VI, 2, 86) ;

Il faut y ajouter, pour être complet, le français *maïs*, qui apparaît ici ou là, mais rarement.

### *La truffe*

On constate pour la truffe une prédominance très nette du type *rabasso* sur l'ensemble du territoire de l'ALP. Il s'agit d'un dérivé du latin RAPA « rave », constitué à l'aide du suffixe augmentatif – ACEU, -ACEA (FEW, X, RAPUM, 68-75). Ainsi la *rabasso* serait-elle à l'origine la « grosse rave » et l'emploi de ce mot pour désigner la truffe, d'une façon exclusive, serait-il un bel exemple d'une taxonomie populaire, distincte de la taxonomie savante, dans laquelle se manifesterait un chevauchement entre espèces voisines, à partir d'un sème commun. Le trait spécifique, le sème, de la « racine comestible » est commun à trois plantes tuberculeuses comestibles, la

rave, la truffe et la pomme de terre et explique donc la désignation *rabasso*.

Mais dans la Drôme, les noms donnés à la truffe expriment surtout la proximité avec la pomme de terre. On observe un parallélisme intéressant dans beaucoup de parlers entre les désignations des deux plantes.

<i>pomme de terre</i>	[ˈtrifo]	[tarˈtifle]
<i>truffe</i>	[ˈtrifo ˈnero]	[tarˈtifle nje]

Dans les deux cas c'est le type lexical venant du latin TUBER « tumeur, tubercule », devenu TUFERA en latin populaire, qui a été utilisé, comme dans beaucoup d'autres régions, *tartifle* s'expliquant évidemment, comme le *tartufo* italien, à partir de TERRAE TUBER (B&W, article TRUFFE). Et, comme on le voit, le sème distinctif pris en compte pour distinguer les lexèmes est celui de la couleur : la truffe est noire ([ˈnero], [ˈnjero] au féminin, [ne], [nje] au masculin).

### 3. Les matériaux présents dans le carnet d'enquête

Ces matériaux se distinguent de ceux qui contiennent les cartes et les marges en ce que, au lieu de fournir des matériaux complémentaires (ce qu'ils font parfois), ils permettent surtout d'approcher et d'approfondir la connaissance de pratiques et d'usages concernant ces deux plantes.

#### *Les truffes*

Ainsi, pour les truffes, avons-nous quelques données lexicales complémentaires très isolées, telles que le nom du « ramasseur de truffes » : le [trifuˈlo], au point 3. Mais ce sont surtout des précisions apportées sur **les techniques de recherche et d'extraction des truffes**. De ce point de vue-là une distinction très intéressante est faite entre le rôle de l'animal et celui de l'homme :

- l'animal, le chien en particulier, exécute trois opérations successives. Il [ˈniflo] « renifle » d'abord, [ˈmarko] « marque » ensuite la truffe, c'est-à-dire indique l'endroit où elle est censée se trouver, et commence à [graˈta] « gratter ».

- l'homme intervient alors avec un instrument, le « petit pic » appelé [pi'kō] ou [pi'tsu], qui lui permet de creuser, [ka'va] ou [tsa'va]. La truffe est alors dégagée avec la main ou arrachée avec le talon (point 7).

On a noté dans le carnet d'enquêtes que le terme de [raba'sja], formé sur le nom de la truffe, la [ra'baso], a la plupart du temps une valeur générique : il signifie « chercher des truffes », mais peut aussi avoir le sens spécifique d'« arracher les truffes ».

En ce qui concerne **la consommation des truffes**, le carnet d'enquêtes nous apporte peu de précisions. Il y a surtout une insistance, très relative, sur le traitement de la truffe une fois qu'elle est arrachée. Elle est [bru'sa per le'va la 'tero] « brossée pour enlever la terre », parfois [la'va a l 'ajgo] « lavée à l'eau », puis vendue sur le marché après toutefois une petite préparation supplémentaire : les fentes éventuelles peuvent être colmatées avec de la terre.

### *Le maïs*

Pour le maïs, les informations du carnet d'enquêtes sont plus riches, parce qu'il s'agit d'une culture qui a été beaucoup plus répandue dans l'espace de l'atlas et qui a tenu une plus grande place dans la vie quotidienne des populations rurales de cette région.

#### *La culture du maïs*

Sur la culture elle-même du maïs le carnet nous apprend peu de choses. Ce sont surtout des détails dispersés sur :

- la taille des maïs, [epō'tsa] « coupés » au dessus du fruit, pour activer la maturité (point 19) ;
- l'entretien de la plante consistant à enlever les petites feuilles qui recouvrent les épis, autrement dit [dɛsfu'la] ;
- le traitement des tiges, quand la cueillette des épis a été faite. Les tiges sont souvent enterrées en labourant et en passant le rouleau, ou parfois coupées, arrachées et brûlées, [ku'pa], [ara'ba] et [bru'la].

#### *Les produits du maïs*

Les informations sont plus nombreuses et précises sur le traitement et l'utilisation à la ferme des produits du maïs. C'est par exemple

l'utilisation des feuilles, des grandes feuilles, celles de dedans qui sont fines. On les suspendait autrefois pour les faire sécher et, quand elles étaient bien sèches, on en remplissait la paillasse du lit, qu'on appelait [kuso de la kwɛjdzɔ] « paillasse du lit » (pts 39, 41), [kuse] (pts 45, 47), [ba'satsa] (pt 67), [pu'sero] (pts 3, 10) ou simplement [pa'laso].

C'est aussi l'opération qui consiste à [egre'na] ou [degre'na], [degru'na], [egra'na], [dezgra'na] « égrener » les épis de maïs. Les techniques traditionnelles qui étaient utilisées étaient de trois sortes, selon le témoignage des locuteurs. On enlevait les grains des épis de maïs :

- avec un couteau (point 7)
- avec une grille, sur laquelle on frottait l'épi pour faire tomber les grains
- avec une bêche, surtout, [lu li'tse pla] « la bêche plate », que l'on plaçait au dessus d'une benne, [yno ba'lasto] ou [kur'njo]. La benne étant posée sur le sol, on disposait une planchette en travers pour faire tenir la bêche et on se tenait debout, à cheval sur le manche.

#### *Consommation du maïs*

Mais les informations les plus intéressantes sans doute sont celles qui concernent la consommation du maïs. Deux spécialités culinaires ont été présentées et décrites par les informateurs : la [pu'lento] et la [be'zãntjo].

#### *La [pu'lento]*

La [pu'lento] est généralement définie comme une « farine de maïs mise en bouillie », une « bouillie de farine de maïs » comme le dit Moutier (Rixte, 2007), ou encore une « soupe faite avec de la farine de maïs » (pt 28). L'aire d'extension de ce mot est plutôt limitée aujourd'hui à la partie médiane de la Drôme, autour de la vallée de la Drôme, mais elle couvre aussi une bonne partie des Hautes-Alpes. Plus au nord, c'est une variante phonétique qui apparaît, mais avec le même contenu sémantique : [pu'lengo] au point 6. Parfois le mot est connu, mais identifié comme appartenant à l'italien : par exemple aux points 19, 67, et on se trouve alors assez proche de l'usage fait en français du mot italien *polenta* et de sa signification usuelle.

D'une façon générale, ce mot de [pu'lento] est senti comme local et il en est de même pour la réalité culinaire qu'il désigne, dans cette



partie médiane de la Drôme. Mais il est certain que, comme l'a dit W. von Wartburg (FEW, IX, 125, POLENTA), les attestations gallo-romanes de *polente* (en langue d'oïl) et *polenta* (en langue d'oc) sont empruntées à l'italien et non héritées du latin, malgré leur ancienneté : *polente* apparaît en français vers 1501, *polenta* en Provence, à Avignon, en 1360, c'est-à-dire en liaison chronologique dans le premier cas avec les guerres d'Italie, et dans le deuxième avec l'arrivée des papes à Avignon.

Du point de vue géolinguistique, ce qui confirme le mieux la provenance italienne de ce type lexical est sa répartition dans l'espace : il est surtout présent dans la partie la plus orientale du gallo-roman, Alpes-Maritimes, Hautes-Alpes, Barcelonnette, dans l'ALP, Savoie, Haute-Savoie et sud-est de l'Isère dans le domaine de l'ALJA (tome II, carte 396).

*La [be'zãntjo]*

La [be'zãntjo] est une autre spécialité culinaire réalisée avec de la farine de maïs. Elle est définie comme une « languette de pâte de maïs » (point 10) ou un « gâteau fait avec de la farine de maïs » (point 43), une « galette de farine de maïs ou de sarrasin » selon Moutier (Rixte, 2007), ou encore une « pâte roulée en forme de boulette, faite de farine de blé noir ou de maïs et cuite à l'eau » (TDF et FEW). Ce mot et cette spécialité sont beaucoup plus étroitement localisés. Il ne semble pas qu'ils soient connus en dehors de la Drôme et même de la moitié nord de la Drôme, le point 47, Châteauneuf du Rhône, étant l'exception qui confirme la règle dans la moitié sud.

Ce mot pose de sérieux problèmes d'étymologie, ce qui explique que le FEW, qui le donne comme dauphinois, d'après le dictionnaire de Moutier, le classe prudemment dans les « matériaux inconnus » (FEW, XXI, 492, Dauph. Besantio). Malgré tout il est tentant de rapprocher ce mot de l'ancien français *besant*, ancien occitan *bezan*, occitan moderne *besant*, signifiant « monnaie d'or » de Bysance et étant également un terme de blason. Ce terme gallo-roman, bien attesté au Moyen Âge, provient évidemment de la forme latine BYZANTIUS « de Bysance » (FEW, I, 669, BYSANTIUS).

Du point de vue phonétique la [be'zãntjo] drômoise viendrait plutôt d'un latin \*BYSANTICA, avec *i* bref, selon une évolution de la séquence *e/i + k + a* bien connue dans cette partie du nord-occitan : URTICA y aboutit à ['urtjo], PERTICA à ['pɛrtjo] ... (Bouvier, 1976,151-153).

L'hypothèse étymologique paraît à première vue bien audacieuse. Car quel rapport, dira-t-on, peut exister entre une galette de maïs et une pièce de monnaie, voire un blason ? Et pourtant... Pour appuyer cette hypothèse, deux rapprochements peuvent être opérés. En termes de blason, le *besant* est, selon Furetière, « une pièce de métal ronde et pleine dont on charge l'écu, qui se distingue du *tourteau* qui est de couleur ». Le *besant-tourteau* est « celui qui est parti, moitié de métal, moitié de couleur ». On remarquera bien sûr que le mot *tourteau*, diminutif de *tourte*, appartient lui aussi au vocabulaire de l'alimentation : il vient de l'occitan *tourteu* défini comme un « petit gâteau de forme circulaire ».

Mais le processus sémantique est du même type pour un autre terme du vocabulaire de la nourriture : le *massepain* en français (d'abord *marcepain*), *massapan* en occitan. Ce mot provient de l'italien *marzapane*, issu de l'arabe *mauṭaban* « roi assis ». C'est à l'origine le nom d'une monnaie qui a été en circulation dans l'Orient du temps des Croisades. Le Christ assis figurait sur cette monnaie. Emprunté par les Vénitiens, sous la forme *matapan*, le mot arabe a désigné une boîte contenant un 1/10<sup>e</sup> de muid, c'est-à-dire une redevance de 10% levée à Chypre. Puis, une fois emprunté par le français et devenu *marcepain*, *massepain*, il a pris le sens plus particulier de « boîte de luxe pour confiserie », puis finalement le contenu de cette boîte, dès 1554 (B&W, article Masepain).

\*

C'est sur une sorte de rêverie aux couleurs orientales que s'achève cette communication. Il reste encore des recherches à effectuer pour faire coïncider cette rêverie sur l'origine de la *besantio* avec la réalité linguistique et historique. Mais la piste est intéressante et elle paraît d'autant plus légitime que l'Orient a toujours été un grand pourvoyeur, réel ou présumé, de plantes et de désignations lexicales de l'Europe occidentale : le maïs, « blé de Turquie », comme l'appellent le plus souvent les Drômois, en est un bel exemple. En tout cas les analyses qui viennent d'être faites et plus largement les données linguistiques et ethno-linguistiques qui ont été présentées montrent qu'il y a beaucoup de trésors à explorer dans les carnets d'enquête des atlas linguistiques.

## Publications citées

- Bloch, Oscar et Wartburg, Walther (von), 1932, *Dictionnaire étymologique de la langue française* (B&W), Paris, Presses Universitaires de France, article massepain
- Bouvier Jean-Claude et Martel Claude, 1979, *Atlas Linguistique et Ethnographique de la Provence* (ALP), Paris, Editions du CNRS, vol. 2, cartes 410-412.
- Bouvier Jean-Claude, 1976, *Les parlers provençaux de la Drôme - Etude de géographie phonétique*, Paris, Klincksieck.
- Furetière, Antoine, 1691, *Dictionnaire universel contenant généralement tous les mots françois tant vieux que modernes, et les termes des Sciences et des Arts*, tome Premier, La Haye et Rotterdam.
- Martin Jean-Baptiste et Tuillon Gaston, 1974, *Atlas Linguistique et Ethnographique du Jura et des Alpes du nord* (ALJA), Paris, Editions du CNRS, vol. 2, cartes 395 « maïs », et 396-400 ; carte 396 « La bouillie ».
- Mistral, Frédéric, 1879-1886, *Lou Tresor dóu Felibrige ou Dictionnaire provençal-français...*(TDF), 2 vol., Vve Remondet-Aubin, Avignon, Roumanille et Paris, H. Champion.
- Rixte, Jean-Claude, éditeur, 2007, *Dictionnaire des dialectes dauphinois anciens et modernes*, par l'abbé Louis Moutier, IEO-Drôme et ELLUG.
- Wartburg, Walther (von), 1922-2002, *Französisches Etymologisches Wörterbuch* (FEW), 26 volumes, Leipzig-Tübingen-Basel-Nancy.

*Jean-Louis Fossat*  
*Professeur émérite, Université Toulouse 2*  
*fossat@univ-tlse2.fr*

## **Des données de terrain à leur traitement**

### *50 ans de recherches*

Un temps de bilan est toujours l'occasion de repenser, à partir d'une archive de plus de 50 années de recueil et d'analyse de données de terrain, les divers itinéraires, les diverses stratégies de recherche des uns et des autres, de dépister les points faibles autant que les points forts de la recherche, dans un mouvement incessant de vérification des hypothèses, de révision des analyses antérieures, de confrontation de points de vue différents, et d'évaluation. Voici donc nos premiers éléments d'un bilan livré à discussion critique. On n'oubliera à aucun moment du recueil et de l'exploitation de données que ce bilan émane d'une personne pour qui la parole en interaction et en énonciation est seule première, et primordiale, et qu'en définitive, tout type de métadonnée qui ne renvoie pas à la parole première contredit ce principe de base.

## **1. Méthodologie de recueil de données de terrain et typologie des sources orales**

### ***1.1. L'instrumentation de départ, en linguistique dialectale opérationnelle et essai de typologie des tâches de production***

On distingue deux formes de recueil, selon le critère de l'instrumentation.

- Avec (A) ou sans (B) enregistrement analogique (prise de son analogique, numérisable).

- Avec (A) ou sans (B) création de ressources langagières numériques (corpus PAROLE) sous licence.
- Avec (A) ou sans (B) prise de vue numérisable synchrone (parole/situation ; archive visuelle et audiovisuelle numérique sous licence).

Dans la première hypothèse (A), une opération de recueil de données de terrain repose entièrement sur la création ou invention d'une archive sonore analogique dont les deux gestes fondamentaux sont le recueil et la transcription indirecte ou différée. Par transcription différée nous entendons transcription contrôlée sur fichier numérique manipulé en station d'écoute avec contrôle pluriel de la transcription in extenso [Fossat-2011 09 Marcel Degeil dit Le Bur ou de Bure 09 Auragnou, à partir d'une source numérisée, et d'un recueil de données par Bruno Besche-Comméngé] ou sélective, selon les objectifs : ainsi, la transcription des données de type atlas, selon les cas, peut être soit sélective (décision lexicale), soit paradigmatique (sur objectif morphologique), soit syntagmatique (sur objectif d'analyse syntaxique et prosodique). C'est notamment le cas lorsque le relevé des données comporte un contexte explicatif définitoire ou des éléments de récit, de manière constante ou épisodique. Dans tous les cas de figure, nous aurons vécu ici, seuls ou en équipes légères, du terrain à l'écoute en station numérique, à des fins d'exploitation de données "hybrides", des heures fabuleuses qui ont laissé des traces indélébiles et dans nos mémoires et dans notre mémoire de laboratoire.

De fait, selon ce critère, les données de terrain d'une opération programmée peuvent se présenter comme des mixtes, combinant, selon les cas, diverses stratégies de recueil de données [Philps-1975 Balaguères, Bethmale-Biros ; Fossat-1970 archive numérique des métiers de l'élevage, de la production et de la commercialisation du bétail ; Besche-Comméngé1985 *Le savoir des bergers de Casabède*, grande première en sémantique dialectale].

On fera remarquer par ailleurs, que dans la perspective dessinée ici, chaque type de données de terrain recueillies enrichit inmanquablement la réflexion du chercheur sur les modes de recueil de ces données, autant que sur les hypothèses et méthodes d'exploitation, notamment des phénomènes structuraux récurrents dans l'usage. Dans ces conditions, toute opération de bilan est

l'occasion de dresser le catalogue chargé des erreurs de production autant que d'exploitation en mode exploratoire. Ainsi, une des erreurs majeures, au terme des opérations d'exploitation des données, serait d'omettre l'identifiant de classe hiérarchique d'une information localisée par rapport à un territoire comme la Gascogne, où la vérification de la carte de gradient de gasconité (Jean Séguy ALG6-1971-1973) a été entièrement exécutée selon des méthodes d'analyse exploratoire des données éprouvées : à toute information localisée, doit être impérativement associée l'identification hiérarchique du groupe territorial déterminé par les applications de type CAHi/j, comme on le précisera plus loin.

### *Évaluation*

L'instrumentation a permis la numérisation des cartes d'origine, reprises en environnement Map-maker (Fossat-Caubet 2012), puis des cartes d'isoglosses établies à partir des tables de critères structuraux d'ALG6, ce qui a permis de constituer des bibliothèques de cartes numérisées, annotées, indexées, accessibles à toutes les catégories d'utilisateurs. D'autre part, elle permet la constitution de bibliothèques sonores, appariées aux collections de cartes muettes, et de revenir aux sources, à partir de ces ressources numériques, à toutes les étapes de l'exploitation des données. Il faut savoir que pour cela, il existe des méthodes, des stratégies et des outillages, pour mieux voir, mieux entendre et écouter, mieux comprendre, bref pour savoir mieux ce qui fait une langue de statut dialectal pour un territoire donné, habité et construit dans l'espace-temps. Loin d'être achevée, l'instrumentation en est encore à ses tout premiers balbutiements : on a fait et on fait encore seulement ce qu'on pouvait et ce que l'on a la capacité de faire, ce qu'on croit savoir faire à présent. On ne le fait pas seuls, et pour soi, mais pour servir encore et toujours, à l'apprentissage et à l'acquisition de connaissances sans cesse approfondies. On verra que les résultats des opérations peuvent varier selon les moyens technologiques utilisés. Quand on n'a pas pu ou su saisir la chance des traitements dans le programme national ADONIS, force est d'opérer avec les moyens du bord, en local, en tenant compte à la fois du degré d'homogénéité des données et d'une typologie d'objectifs : en effet, les problèmes posés par les recueils de paraboles de l'enfant prodige (opérations Bourciez, Sacaze, NALF) sont différents des corpus de parole interactive, des corpus dialogués, des corpus de conversation ordinaire sans intervenant questionneur. Il est par conséquent

indispensable de préciser ce que l'on entend par diversité des objectifs de recueil de données de terrain.

## ***1.2. Typologie des objectifs***

Sans objectif spécifié en phonologie, morphologie et syntaxe, et sans hypothèses de départ sur le mouvement linguistique affectant les structures, sur la variabilité, sur la fluctuation, les données de terrain manipulées risqueraient fort de constituer un magma informe, que des linguistes contemporains confondent parfois avec la notion de chaos par morcellement ; comment entrer dans ces « univers chiffonnés », ou plus exactement selon une expression heureuse « irisés » que sont les univers dialectaux des Alpes/Jura aux Pyrénées et ces « petits mondes » langagiers cantabriques, galiciens, aragonais, ou d'ailleurs ? Dans ces conditions, toute opération de manipulation d'une donnée structurale quelconque localisée, qui ne serait pas rapportée d'entrée aux cartes d'isoglosses réalisées à partir des tables d'ALG6 (Séguy-1971) constituerait au plan scientifique une lacune grave, car ce sont les données de base qui déterminent la nature et les caractéristiques des distributions, ce que des théoriciens sont tentés d'oublier.

### *1.2.1. Recueil de données dans le cadre d'une opération relevant des méthodes de la géographie linguistique.*

Il y a, on le sait, données et données, avec ou sans métadonnées. Selon ce critère, l'enquête comporte (A) ou non (B) des contraintes de type géolinguistique qui définissent la notion de ressources orales géolocalisées, géoréférencées (K. Jaberg, Scheuermeier, G. Rohlfs, Gauchat) : tous nos recueils de données numérisées sont nécessairement de type (A).

- échelle spécifiée par coordonnées GPS (communale, cantonale, infracommunale, intercommunale)
- fichier localité par localité [\*loc] spécifié par n° de nœud de réseau de sociabilité en société rurale ou urbaine, pour une structure donnée (phonologique, morpho-syntaxique)
- fichier informateur spécifié [\*inf] positionné dans un réseau de sociabilité et de communication intra-groupe et non seulement défini par lignage

- fichier requêtes spécifiées [\*q] avec typologie de requêtes en fonction des objectifs (structures et usages).
- fichier des réponses numérisées, positives ou négatives, spontanées ou suscitées dans un plan d'expérience [\*rép mots syntagmes phrases].

W1/n : il ne ferait pas bon ici nier l'existence du mot de parole/langue

B1/n bandes nominales ou verbales avec extraction lexicale

S1/n fichiers de phrases soumises à analyse syntaxique/syntagmatique et prosodique (alpi ; alg sonore).

La graphie populaire de la translittération d'une source orale localisée facilite une transcription phonétique de degré de fiabilité moyen ; nous avons ainsi pu traiter ; en application spécifique, 100 carnets de translittération du célèbre recueil SACAZE pour les cantons de St-Girons, Massat, Castillon ; le corpus des données orales translittérées et transcrites, une fois numérisées, se présentent alors sous forme de fichier texte, comme suit :

\*\*\*t0001 légende de Barbazan \*l Lourdios\_Ichère 64 Pyr. Atlantique groupe\_classification\_xxx\*d 1887 \*réf BM Toulouse Recueil Sacaze ms.1112 JE-2009,167-169 \*mise\_API JLF2011 d'après texte graphique édité par J. Eygun \*ph requêtes sur critères par indexation ; constitution de tables de critères : diphtongue\_xj/xw/hiatus/liaison/aphérèse (J.L. Fossat).

\*01 /ke j\_aβe / awdes\_ 'kɔps / ɛ̃n y βi'ladze deras piren'es y 'umi e ya 'henna ki ɛran hɛra 'prawbez e bj'eɫs /

\*02 /'ets ke nu a'βen ke ya majz'u 'tuta tʃ'ina / y 'ɔrt e ya 'βaka /

\*03 /mes /pɛr ɛsta 'prawbes / 'ets ke sekur'iβan tus'tɛm ak'ets ki ɛran mes 'prawbes 'k ets /

\*04 /y\_'dia / era 'tɛr:a k ɛr aprik'ata dɛ n'ɛw / e ke zel'aβa a\_p'ɛjra\_h'ene /

\*05 / 'tut era 'zen deb bi'ladze ke\_s kaw'haβa ɛ̃m\_minz'an e ɛ̃m\_b'eβen/

\*06 / k ɛran 'r:iʃɛz e i'rus/

07 /a\_r:ɛ̃n'drata\_ dɛra\_'nwet / dyz bja'dzɔs /ki 'biβan de 'lwɛ̃ŋ / ki\_z bulun ar:es'ta ɛ̃n aked' ɛ̃ndr'et / prumu\_k aβen 'r:et e 'hami/ ke



try'kaw ên ya 'pɔrta / aprez ên 'ɲ\_awda / i ên 'd\_awdas ên'gwɛra /  
a'r:ɛs ke nu bulun le'fa\_wz êndra / ni dfa\_jz ar:e ta mǎn'dʒa/

Ces données orales transcrites par nos soins à partir d'une translittération en graphie populaire dite phonologisante, ont pour origine le dossier des enquêtes SACAZE ; mais translittérées ou transcrites en mode API, les données orales ont été rendues muettes, ce qui constitue le paradoxe le plus évident, et la lacune la plus évidente, s'agissant de traitement de données orales. Nos prédécesseurs avaient déjà formé le projet d'une édition phonographique de l'archive orale analogique ; la question qui se pose est donc, après numérisation, l'exploitation optimale et concertée des ressources orales numérisées qui disposent en vue synoptique, de fichiers-textes et de fichiers sonores numérisés.

Comme on peut déjà s'en rendre compte – on y reviendra – un des premiers prétraitements repose sur la segmentation de la totalité en segments identifiés sur critères syntaxiques, prosodiques, morphosyntaxiques, phonologiques, avec annotation systématique par les utilisateurs mêmes et non en haut lieu dans quelque laboratoire que ce soit, ce qui n'est pas sans retombée au plan didactique de la formation à la connaissance ; ou bien on travaille avec des ressources électroniques importantes dans un système international dominant (big data/big system) : c'est la solution de type EUROVIA des opérations ADONIS ; ou bien, par la force des choses organisées, on opère dans de petits ateliers, petites et moyennes entreprises de formation à l'analyse de données, sur des chemins plus discrets, plus modestes, en contexte de pauvreté ; dans un tel contexte, cependant, il est aisé de repérer le graphe temporel de récurrence de tel phénomène structurant (voisement/dévoisement, diphtongue, ..., etc.), comme on peut de toute évidence s'en rendre compte. Donc pas de consensus ex machina.

Dans les trois cas, les données primaires sont constituées par opération d'indexation de mots concrets construits, de syntagmes ou bandes, de phrases ou de groupement de phrases selon des critères explicités par des méthodes d'analyse énonciative ; c'est le cas, notamment, des phraséolexèmes, groupes de mots figés, lexies complexes, expressions idiomatiques, formes étendues, cooccurrences des dictionnaires dialectaux constitués à partir de données de terrain

depuis plus d'un siècle en France par les linguistes romanistes d'Europe.

### BIBLIOGRAPHIE DE RESSOURCES ORALES DISPONIBLES

(données ; classification ; représentation ; interprétation)

Fossat-1996 : inventaire de l'archive numérique du CLid

- 1- bibliothèque des archives sonores 2. Bibliothèque des archives visuelles, sur site [ercvox.free.fr](http://ercvox.free.fr) (enquêtes de terrain post ALG/NALF 1960-1996 occitan, français, langues romanes).
- 2- Fossat-2005 Analyse variationnelle des parlers gascons couseranais et interférentiels (données orales translittérées et transcrites; traitement classificatoire ; représentation cartographique des distributions ; AIEO Reggio Calabria/Messina ; site [umr838 UTM](http://umr838.utm.fr) (MSH Casanova pour la réalisation et la conception des cartes distributionnelles d'ALG6 à partir des tables ; vérification et interprétation : Jean-Louis Fossat ; création des données numériques de départ J.-L. Fossat, sous licence ; droits d'auteur Jean Séguy (tous fichiers non numériques).

#### 1.2.2. Structures et Usage(s)

Selon ce critère, on distingue les données primaires visant à établir des structures (A) et des données primaires qui déterminent tout autant les structures (STR) que les usages (US) en société organisée (B), selon l'opposition en vigueur depuis Hjelmslev (REF).

STRUCTURES (hypothèse A) : OPERATION DE TYPE ATLAS LINGUISTIQUE ET ETHNOGRAPHIQUE DE LA FRANCE PAR RÉGIONS (unités territoriales avant et après 1498)

- Critères phonétiques : table de critères pris en compte (CRIT\_ph/1) ou non (CRIT\_ph/0).
- Critères phonologiques : table de critères phonologiques pris en compte (CRIT\_PHO/1) ou non (CRIT\_PHO/0).
- Critères morphologiques et morphosyntaxiques : table de critères morphologiques pris en compte (CRIT\_m/1) ou non (CRIT\_m/0).
- Critères syntaxiques : tables de critères syntaxiques ou morphosyntaxiques pris en compte (CRIT\_msynt/1) ou non (CRIT\_msynt/0).

-Critères prosodiques : tables de critères prosodiques pris en compte (CRIT\_pros/1) ou non (CRIT\_pros/0).

-Combinaison de catégories de critères (alg6 MS3\_PrCo\_complexe)

-Données de polymorphisme phonétique pour cartographie du polymorphisme phonétique et phonologique (nasalité, couleur vocalique, palatalité, vélarité, densité, compacité) : les ressources numériques créées sous forme de table XL permettent alors la classification des données quantitatives et la fabrication d'une bibliothèque numérique de cartes de distribution quantitative (isoglosses de polymorphisme).

Pour accéder à cet ensemble de données numériques et cartographiques, on s'adressera à [fossat@univ-tlse2.fr](mailto:fossat@univ-tlse2.fr).

### **Exemplification de données de terrain gasconnes d'origines diverses, combinant critères morphologiques, syntaxiques, phonologiques, prosodiques**

-(i) *L'Emile ne s'a pas colhonat ; que s'a balhat còsta coberta*, (40 Montfort-en-Chalosse)

-(ii) *se me lo balha*. (ALG6 MS3PR)

-(iii) *porta-te plan : portatz-vos plan* (40 Montfort, St-Vincent T), (64 Salies-de B.)

-(iv) extraction de dossier données primaires 64 Lescun (Rohlf's, Passet et J. Boisgontier): constat de connexité rythmique et clise asyllabique en gascon occidental.

-*tu, menschido't !* (31 Aspet) *mesfisa't, cara't* [...] (groupe gasc. oriental) ; *cara-t* (groupe gascon aquitain occidental)

-*a força de cercar\_lo* [...] /serk'a\_w] 64 Lescun

-*enta har-los baishar* (las olhas) 64 Lescun

-*enta har-los* [ha\_ws] *pravar* 64 Lescun

-*enta har-los* [ha\_ws] *baishar* (los anhèts) 64 Ossau

-*enta har-las* [h'ar\_los] *baishar* (eras oelhas) ) [eraz gw'ełos]

-*enta virar-lo* [bir'a\_w] « pour LE faire revenir »

**Exemplier languedocien restreint** (Sénégas, Petit et L. Michel 1973), CLASSES SYLLABIQUES, RYTHMIQUES (méthode A avec enregistrement magnétique systématique par J. Boisgontier) avec essai de délimitation d'énoncés de la parole représentatif de structures en usage.

- *digas-i bo, digas-i bo-t*
- *toca-i*
- *agarro-bo't aqui !*
- *va caler i anar*
- *i va caler anar*
- *a calgut le/lo tuar*
- *l'a calgut sannar, lo tesson*
- *l'anèt (angèt véser/veire)*
- *anèt (angèt la véser (veire)*

Comme on peut déjà s'en rendre compte ici, les données recueillies sur le terrain par observation forment des recueils plus ou moins homogènes, sur objectif d'exploitation.

### **Exemplier d'encodage des données de terrain en vue de la classification exploratoire**

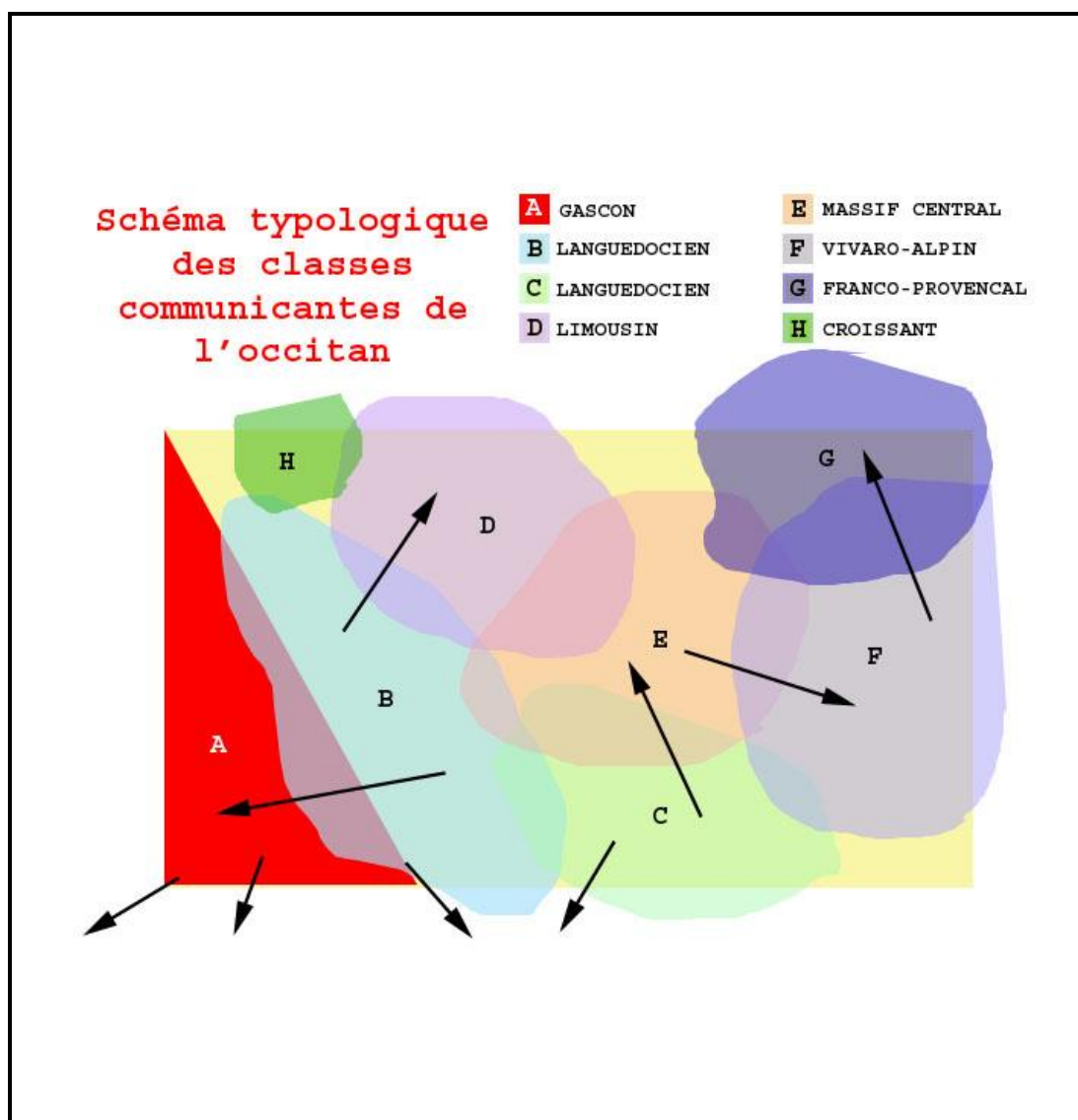
#### (1) Schématisation des observables de terrain

On se réfère ici à une représentation classique de la situation linguistique de la France par régions, qui permet aux utilisateurs de se faire une idée de la fragmentation primaire du domaine linguistique observé de type occitan, articulé au domaine dit franco-provençal alpin et jurassien. Seule l'opération de type NALF permet une classification fine en aires primaires, secondaires et tertiaires de ce domaine. Nos propres données de terrain privilégient toujours la bordure aquitanique de ce vaste territoire, la Gascogne, et l'ancienne Guyenne, avec le Languedoc occidental et méditerranéen, avec une attention spéciale aux vallées pyrénéennes.

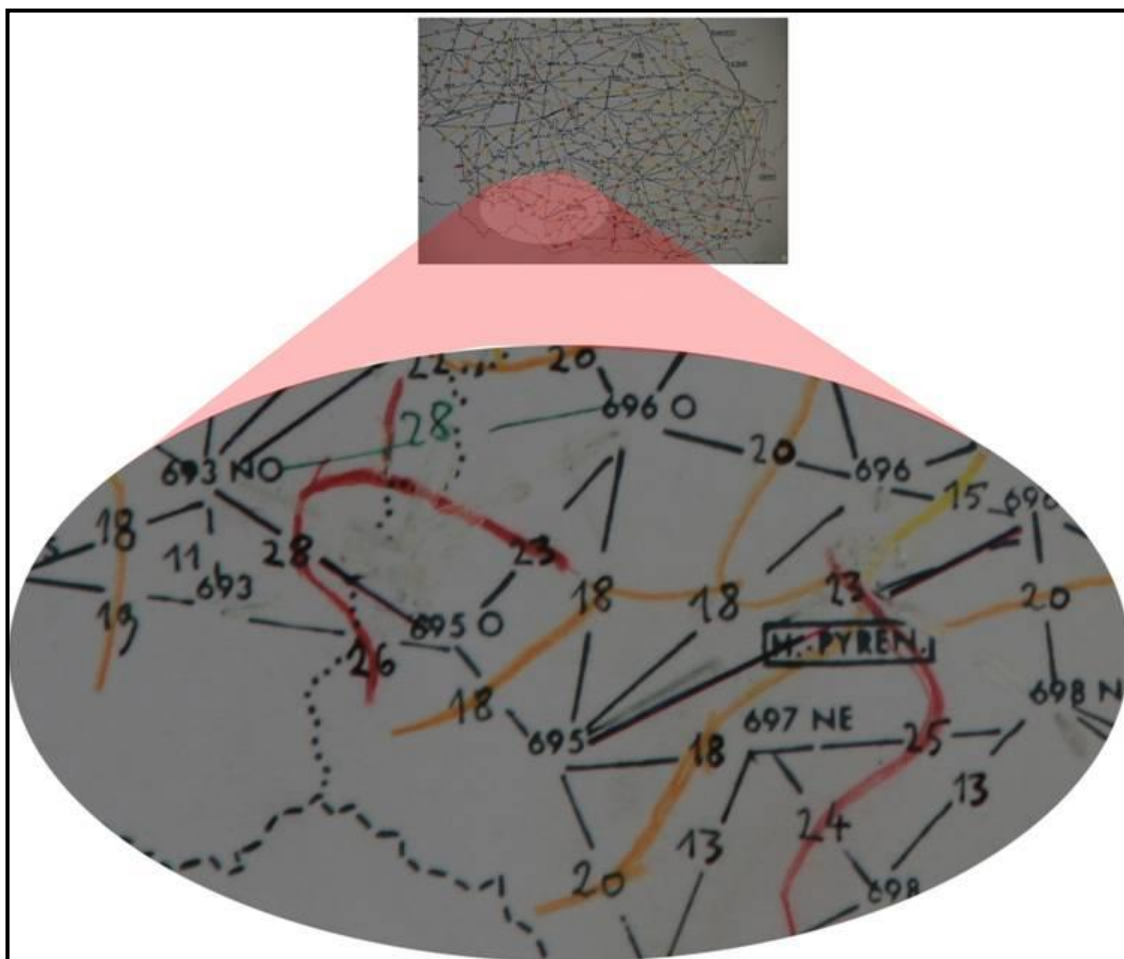
Pour ce domaine, les données de terrain peuvent être de type distinct :

- type géolinguistique NALF.
- opérations de recueil de données lexicales antérieures ou postérieures, parallèles au plan NALF (CORPUS PAROLE ERA352/URL8 INaLF coordination J.L. Fossat).
- monographies confrontées à des dictionnaires régionaux (Dauphiné, Vivarais, Aoste).
- dictionnaires régionaux (Foix, pour les Landes de Gascogne, Arnaudin) établis d'après des données de terrain.
- bases de textes reliées à leur propre dictionnaire (Arnaudin).

S'agissant de déterminer l'extension géographique du terrain des observables, on en donnera ici une représentation globale schématique.



(2) Extraction de variables et cartographie de distances structurales pour tout ou partie de ce domaine d'observable à une échelle donnée, avec attention spéciale à la Gascogne, centrale, occidentale, orientale, et à ses secteurs interférentiels, pyrénéen au Sud-est (Couserans et pays de Foix), et garonnais d'ouest en est. La méthode est rappelée ci-dessous.



*Typologie et gradient (vue partielle-frontières)  
ALG 6*

Au vu de ce schéma, il tombe sous le sens que la méthode mise en œuvre pour le traitement des données de terrain est généralisable à un quelconque élément d'information territoriale géoréférencée ; on évitera donc toute assertion sur l'homogénéité ou non sans avoir achevé l'analyse des données de terrain disponibles, ce qui peut prendre beaucoup de temps et nécessite la conjonction d'énergies ou forces de travail.

(3) Première gasconne de l'encodage des variables structurales (J. Séguy 1973) : vue sur les structures morphosyntaxiques du pronom régime gascon (bloc de 29 variables tables MS3PrCo et vérification du bloc de cartes MS3PrCo d'ALG6).

MS												1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2																				
		1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10														
693		3	1	0	0	2	0	1	1	0	1	2	1	0	2	6	2	0	1	0	1	1	1	1	5	1	1	6	0	1	0	
693	NR	3	0	0	0	3	0	2	1	1	1	2	1	1	0	2	2	1	1	2	1	1	1	0	1	1	1	6	0	1	0	
693	NO	2	1	1	2	1	0	1	1	0	0	2	1	0	2	6	2	0	1	0	1	1	1	5	1	1	6	0	1	0		
694		3	0	0	0	1	0	1	1	0	0	2	1	1	2	4	2	1	1	1	0	1	1	5	1	1	6	0	1	0		
694	Z	3	0	0	0	1	0	1	1	0	0	2	1	1	2	4	2	1	1	1	1	1	1	5	1	1	6	0	1	0		
695		3	0	0	0	3	0	2	1	1	1	2	1	0	2	2	0	1	1	1	1	1	1	0	1	0	1	1	1	0		
695	O	3	1	0	0	3	0	1	1	1	1	2	1	1	0	2	2	0	1	2	0	1	1	1	0	1	0	5	0	1	0	
696		3	0	0	0	1	0	2	2	1	1	2	1	1	0	4	2	0	1	1	1	1	1	1	1	1	1	0	1	1	0	
696	E	3	0	0	0	1	0	2	2	1	2	2	1	0	5	2	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	
696	S	3	0	0	0	1	0	2	2	1	1	2	1	1	0	4	2	1	1	1	1	1	1	1	1	1	1	0	1	1	0	
696	O	3	1	0	0	1	0	1	0	1	1	2	1	1	3	2	0	1	1	1	1	1	1	2	1	1	6	0	1	0		
697		3	0	0	1	1	0	1	1	0	2	1	1	1	0	2	1	1	1	1	1	1	1	1	1	1	0	1	7	1	1	0
697	NR	3	0	0	1	1	0	2	1	1	1	2	1	1	1	2	0	1	1	1	1	1	1	1	1	1	0	7	1	1	0	
698		3	0	0	1	1	0	2	1	1	2	2	1	0	5	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	
698	N	3	0	0	1	1	0	2	1	1	0	2	1	0	5	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	
698	E	3	0	0	1	1	0	2	1	1	1	2	1	0	5	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	

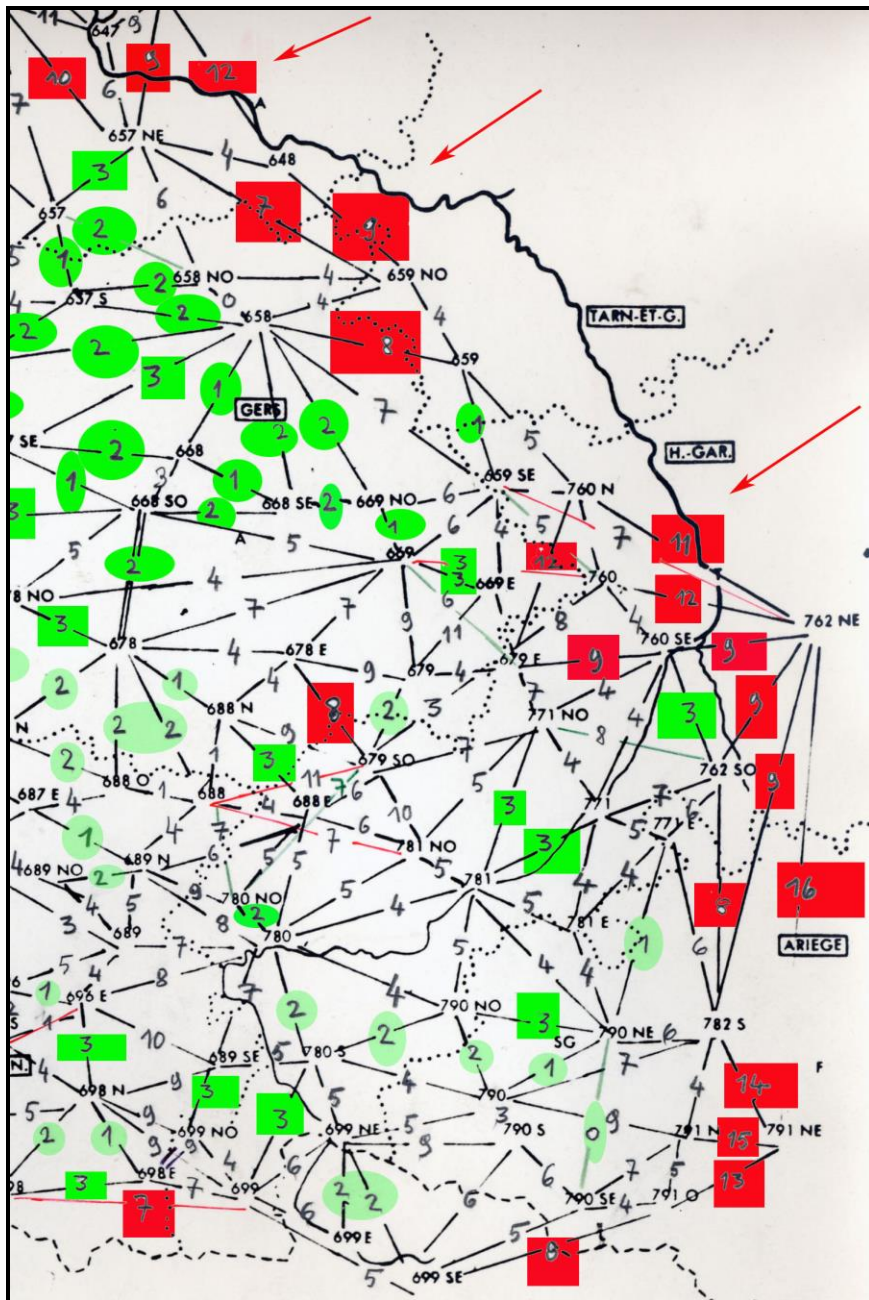
Table 3 : Système morphosyntaxique, Groupes et classes (ALG 6)

Ce tableau de données, de la main de Jean Séguy (fonds numérique Jean Séguy) représente pour l'essentiel le point de départ vérifiable pas à pas de toutes les opérations de manipulation des données de terrain recueillies en hypothèse A (instrumentation avec expérimentation).

De même le tableau de distances interspatiales interponctuelles ci-dessus établi par la main de Jean Séguy 1973 fait apparaître avec netteté l'effet des frontières dialectales en zone interférentielle du gascon oriental et du languedocien occidental.

C'est cette hypothèse (A) qui est à la base de toutes les stratégies d'exploitation de données géolinguistiques dans les opérations de type atlas du NALF et/ou leurs prolongements, avec des degrés et des options divers. Ou toute opération similaire ; le premier objectif à atteindre pour analyser et décrire la variation linguistique passe alors toujours par une étape de tracé de lignes isoglosses simples ou cumulatives. (Séguy, Conférence de Londres 1960 ; Séguy ALG6 tables/cartes isoglosses tracées depuis les tables ALG 6 avec étape de vérification : Fossat-Casanova (2001) isoglosses phonétiques, phonologiques, morphologiques et morphosyntaxiques du gascon ;

Fossat-2001 Classification et isoglosses morphosyntaxiques, Fossat-2008 (Leeds); Fossat-2009 (Lyon), Fossat-2009 (Aachen). Les données primaires intermédiaires mises en exploitation reposent dans tous les cas, qu'il s'agisse d'approche classificatoire dialectométrique ou d'approche cladistique, sur le calcul des lignes isoglosses et des constats faits à cette étape cruciale (Lalanne, Séguy 1960-1973).





### USAGES (Hypothèse A/Hypothèse B)

Les données primaires polymorphes, selon ce critère, présentent un caractère monolingue dialectal (cas 1) ou non (cas 2) ; dans ce dernier cas, les données primaires témoignent d'une réalité multilingue, plus précisément plurilectale, expression consacrée depuis Puech et alii (Lyon) avec ou sans statut polynomique, selon l'expression consacrée depuis Giacomo-Marcellesi et alii.

L'objectif n'est plus alors de recueillir le lexique dialectal seul, isolé en laboratoire par artefact, mais bien d'établir ses relations, simples ou complexes, quand il existe et tant qu'il existe encore dans l'usage social, avec une ou plusieurs langues de force, comme on parlait de maisons de force (Aoste, Aran), et, le plus souvent en situation diglossique, le rapport complexe à une langue dominante, nationale en principe, sur un territoire donné, en France, en Espagne, en Italie, au Portugal ; c'est l'objectif des opérations de type ERA352 (CNRS et UTM2) prolongé à phase d'exploitation par les opérations de l'URL8 de l'Institut National de la Langue Française et des langues de France. Les données témoignent des stratégies des acteurs sociaux, des interactions dans des actes de la production et de la commercialisation de produits consommables avec délimitation thématique aux « petits mondes » de la production animale et du commerce des produits dans les économies locales réticulées. Ces mondes, de manière très prévisible, allaient être frappés brutalement par les forces d'une politique de mondialisation. Les historiens (Max Weber) avaient déjà parlé de cette situation en termes de « fin des paysans », donc de fin des langues historiques dialectales. Les données de terrain leur donnaient à la fois tort et, partiellement, raison. C'était oublier ce que sont les « forces sociales » quand elles sont structurées en réseaux de sociabilité, hier comme aujourd'hui.

Dans cette perspective, les ressources créées à partir des données de terrain sont des unités discursives, d'où sont extraites les unités lexicales catégorisables, comme on le verra plus loin.

Dès la phase de recueil de données, dans cette perspective-là, se pose ici la question cruciale des usages du dialecte dans les centres de commercialisation et à leurs périphéries, en termes de dialectologie des usages sociaux, en milieu rural, urbain et en situation d'interaction pour laquelle a été consacré le terme de situation « ruraine » : ce type d'enquête nous a conduit à vivre quelques expériences fabuleuses dans les sociétés de montagne notamment, du territoire andorran au

territoire cantabrique, dans les mondes de culture pyrénéenne, aquitanique, avec Bruno Besche-Commenge, Claude Costes et François Séguy, voire en domaine basque avec X. Videgain et alii, en domaine cantabrique, avec A. Leal, Claude Costes, François Séguy (Los Montes de la Vega de Pas), et, par la suite, des heures exaltantes, du côté de Benzékri, du côté des dialectométriciens, avec Dennis Philps, Gotzon Aurrekoetchea, Liliane Jagueneau, Marc Peytavi, Jacques Vanpoucke, Choeb Zafar et bien d'autres encore ; se posait alors la question de savoir de quel bord on était : chez les métriciens ou chez Marcel Degeilh et Léonie à Auragnou (09 Biert). La question reste sans réponse, si on la pose en termes brutaux de pensée binariste. Elle trouve sa réponse dans un mouvement de pensée au travail incessant entre données de terrain et méthodes d'exploitation sans cesse en mouvement, avec points de fixité définis en termes de géopolitique linguistique, au sens de Robert Lafont, notamment et entre autres ; nous avons opté pour ce point de vue, ce qui représente un mouvement notable, par rapport aux prétentions de la dialectologie à s'ouvrir aux méthodes de la sociolinguistique rurale ou urbaine (Louis-Jean Calvet, G. Marcellesi).

#### OPÉRATIONS REPOSANT SUR DES STRATÉGIES MIXTES EN HYPOTHÈSE A

De fait, dans ce type d'opération, le recueil de données de terrain - onomastique comprise - met en œuvre des stratégies mixtes.

STRUCTURES DIALECTALES PLURILECTALES mises en évidence (A) : Index D

Exemple retenu : le corpus Sacaze des cantons de St-Girons, Oust, Castillon et Massat (Fossat 2005-2010 Contrat DGLF)

USAGE POLYNOMIQUE LOCALISÉ (A) dans les opérations relatives à l'usage des langues de France.

Exemple de chantier opérationnel : le corpus DRF/DF (opération ERA 352 CNRS/URL8/INaLF)

- index D (tous dialectes de France) avec distribution régionale
- index R (tous usages de France, par régions linguistiques)
- index F en France
- index DF (tous usages de France constituant des mixtes)

-index A (allemand) dans le cas du territoire germanophone en France de l'Est (Alsace, Lorraine, Sarre).

L'erreur majeure des opérations d'indexation ordonnée est de séparer des prétendus niveaux, de fait imbriqués dans des mixtes, ou situations mixtes d'usage, dites de métissage dans des univers caractérisés par leurs stratégies d'interaction.

Ce type d'opération comporte deux types d'enquête (A, B) d'une part ; et d'autre part, pour A comme pour B, une observation à échelle cantonale, voire intracantonale pour les zones A, B, C de mon schéma territorial initial, avec sondages à échelle distendue, pour le nord, le sud et l'est du territoire : c'est le cas, notamment et pour nos données de terrain en zone urbaine (Lyon, Turin, Bruxelles) ou en zone rurale (Isola, Aoste) ; certains secteurs en zone hispanique ont donné lieu à une approche serrée, infracommunale, homologue à notre approche des communautés agro-pastorales pyrénéennes.

J'ai commencé à comprendre ceci à Lyon, aux abattoirs de Lyon-La Mouche, puis à Genève, dans une table ronde de bouchers jurassiens et alpins alors que les industriels de ces mondes me sommaient d'abandonner mes vues folklorisantes, plus encore en vallée d'Aoste, dans une soirée fabuleuse où se mixaient joyeusement les réponses romanes, disons franco-provençales, et tudesques, au sein d'une même cellule familiale ; et tout autant dans une table ronde de bouchers wallons et flamingants entre Namur, Bruxelles et Tourcoing. Dès ces heures-là, j'avais définitivement abandonné l'opposition carrée de Hjelmlev entre structures et usages. Et hier encore, à une journée de rencontre entre paysans éleveurs exploitants et recherche scientifique en Sciences Humaines et sociales, à Ayzac-Ost en Bigorre (Lavedan), J.F. Le Nail, Bruno Besche-Commenge et plus généralement tous ceux qui s'attachent à mieux comprendre comment sont organisées dans le temps-territoire les sociétés de production des mondes ruraux, dans les sociétés de montagne, me confortaient dans ma position ; dès lors avait sonné l'heure du retour aux sources, autrement dit aux données de terrain, à présent toutes numérisées, en ce Noël 2011 : avec un gigantesque merci à A. Meqqori, cet autre infatigable compagnon d'aventure. Alors, avançons d'un pas dans ce mouvement de retour aux gens de terrain qui, précisément parce que gens de terrain, connaissent le terrain comme le fond de leur poche, comme peut s'en rendre compte quiconque a appris à les écouter

systématiquement ; maîtres des données, ils nous laissent le soin d'établir des collections de métadonnées et des notes de bas de page de nos essais successifs d'analyse.

### ***1.3. Typologie des sources orales par rapport aux objectifs et aux axes d'investigation en hypothèse A***

En raison des distinctions fondamentales établies précédemment, nous sommes, après étape de la numérisation des fonds d'archives sonores en présence de deux types de ressources orales, au terme d'une opération minutieuse de numérisation, avec étape d'indexation des ressources numérisées pour être rendues accessibles, - enfin !

#### ***1.3.1. Ressources linguistiques numériques, lexicales, morphologiques, syntaxiques, transcrites (T1) ou non (T0), avec ou sans analyse grammaticale, relevant de protocole de questionnement rigide caractérisé par son inflexibilité relative ou non.***

- Opération NALF (atlas linguistique de la France par questionnement rigide : questionnaire 1-2500 du questionnaire Dauzat, revu et corrigé, augmenté ; le transcripateur est toujours Jean Séguy en station d'écoute des documents analogiques ; le transcripateur est toujours un usager, un utilisateur, dans un plan d'expérience, à partir de données sonores numérisées depuis 2011. Cette opération comporte une annexe de traduction dialectale des Paraboles de l'Enfant Prodigue, entièrement numérisée par les soins de l'équipe LDOR sous la direction de Patrick Sauzet ; son intérêt réside dans l'homogénéité des données sur l'étendue du territoire des Langues de France (recueils Bourciez, Sacaze).

- Opération ERA352 C.N.R.S (Gascogne, Languedoc-Roussillon) par questionnement souple en entretien semi-libre (q 1-300 sociétés agraires, production, commercialisation), en stratégie mixte. Dans ce cas, le transcripateur est toujours un utilisateur des ressources numérisées, dans un dispositif d'exploitation dont le protocole sera exposé plus loin.

Il est exigible de répondre à la double question du volume de données par opération et de la qualité de ces données, par rapport aux objectifs et stratégies de recueil.

1.3.2. Ressources linguistiques discursives transcrites (T1) ou non (T2), et annotées (N1) ou non (N2)

On classera ces ressources numériques indexées selon leur provenance opérationnelle, en fonction des options, stratégies, objectifs et axes de la recherche.

- Opérations NALF : fichiers numériques de la traduction dialectale de la Parabole de l'Enfant Prodigue

- Opérations ERA352/URL8\_INALF/CLID ; archive sonore des ressources discursives, majeures ou non, entièrement numérisée au 31 janvier 2012 ; cette archive sonore numérisée est organisée sur principe de partition régionale (régions en 1498 ; en 1789 ; régions modernes et contemporaines).

***1.4. Les langues en interaction dans les actes d'échange verbal***

1.4.1. Questionnement FRA pour réponses en mode DIAL (avec commentaires partiellement en mode FRA)

1.4.2. Questionnement en mode FRA pour réponses en mode FRA (avec commentaires partiellement en mode DIAL)

1.4.3. Questionnement en mode DIAL pour réponses en mode DIAL.

1.4.5. Question en mode FRA pour réponses polynomiques en L1, L2, L..., Ln)

Exemple pour 1.4.1 : comment prononcez-vous les suites écrites ? (on les montre sur une affiche, une carte, une image : les Tibaous, la Séoune, Péone, Gimonne, Marcaoue, en patois (FRA) – Réponse localisée DIAL. (Fossat-2011 Barcelona ICOS4).

Force est de constater, dès lors, que le mixte NC/NL/NP fait partie intégrante de l'observatoire dialectal, entre structures et usages qui témoignent de l'existence de forces sociales. Mais alors, me dira-t-on, c'est encore une fois de plus la problématique de l'hydre de Lerne qui refait surface, dans une course effrénée aux données de terrain.

C'est bien ce qui a manqué à cette approche depuis 1810, qui nécessite aujourd'hui une révision exhaustive des données d'IGN\_Google\_maps à la lumière des cadastres napoléoniens et de ce qui reste encore ici de structures en usage dialectal, dans la France Rurale et dans les plans de développement urbains (PLU).

### ***1.5. Format des données initiales et évaluation exigible des volumes de données à mettre en exploitation***

#### ***1.5.1. Carnets de terrain constituant PV d'enquêtes de type (B), sans archive sonore***

NALF

Opérations CLID\_ERA 352 URL8

#### ***1.5.2. Bordereaux de transcription sans analyse constituant PV d'enquêtes avec archive sonore analogique cataloguée***

NALF : bordereaux avant et après numérisation

ERA 352 : relevés lexicaux/carnets avant et après numérisation

#### ***1.5.3. Bordereaux de transcription en station d'écoute d'archive sonore numérisée et cataloguée***

NALF : ressources numérisées pour opération TESOC

ERA 352 : ressources numérisées pour opérations CLID/ERCVOX\_2001 sur site ercevox module\_atlas

#### ***1.5.4. Cartes numérisées avec protocole de fouille de données K***

Maquettes de cartes (fonds Jean Séguy)

Brouillons de cartes (fonds J. Séguy NALF ALG 1 2 3 4 5 6)

Collections de cartes numériques par isoglosses : données intermédiaires (Casanova). Traitement informatique MSH Toulouse UMR838 ; Fossat ; tâches de saisie, d'analyse linguistique, d'après la source Jean Séguy ALG6, tables.

Maquettes de cartes (Fonds CLID\_ERA352)

### 1.5.5. Imagerie numérique

NALF : planches (collection J. Allières) associées aux objets de type K (cartes scannées et indexées)

ERA 352 : archive visuelle numérisée (site de l'opération ercvox) : fonds François Séguy, Bruno Besche-Commengé ; J.-L. Fossat ERA 352 et URL 8 mises en exploitation en protocole ERCVOX-2012 (site ercvox remanié en 2012).

### 1.5.6. Archive glossairistique de dictionnaires numérisés (hors ressources orales)

L'exigence est ici de comparer les données lexicographiques géolocalisées par les lexicographes à des données géolocalisées par nos propres opérations de terrain. Depuis [Fossat-1970]

Exemple : NC/NL/NP : *bedat, devesa*

1.5.7. Archive de textes écrits transcrits numérisés (fonds modernes, contemporains à partir d'archive sonore analogique/numérisée), avec confrontation avec les données des fonds historiques de textes juridiques, techniques, dialectaux ou français, en prose, hors ressources orales, mais avec exigence de comparaison à des ressources orales numérisées dans le cadre des opérations programmées antérieures au plan ERSS et CLID erss (1970-1996).

Exemple : documentation BIGERRI (Le Nail ; X. Ravier) confrontée à nos propres ressources orales en Bigorre).

## **2. Méthodologies des divers modes d'exploitation des données de terrain numérisées**

Une fois les données de terrain en nombre (Big Data) classifiées selon des méthodes appropriées, le linguiste descriptiviste a achevé aux trois-quarts sa tâche ; lui reste alors la difficile tâche d'interprétation, risquée hors instrumentation.

## ***2.1. Données numériques de type nalf/alg créées comme ressources réutilisables***

2.1.1. Création d'une base de données cartographiques à partir du fonds de cartes numérisées, en provenance des collections éditées ALG 1 2 3 4 5 6 de Jean Séguy (IEM et CNRS, NALF) : (Fossat-1986)

- volume des données cartographiques numérisées (de format jpeg)
- moteur de recherche création A.Meqqori et J.-L. Fossat
- clés de requêtes sur index de base de données

2.1.2. Opération mémoires de laboratoire géolinguistique NALF/ALG : création de la base de données cartographiques du « fonds Jean Séguy » (Fossat-1996-2011)

- brouillons de cartes pour ALG 1 2 3 4 5 6 : volume de données cartographiques numériques
- maquettes cartographiques élaborées pour ALG123456 : volume de données numériques
- procès verbaux de données primaires cartographiables pour \*loc1, 167 et q. 1, 2500 : volume de données numériques
- cartographie numérisée des distances linguistiques (5 volumes)
- cartographie du polymorphisme phonétique, phonologique, morphosyntaxique en Gascogne (ALG 5, ALG 6)

2.1.3. Création de la base de données numériques secondaires en provenance des tables ou matrices (TM) d'ALG6

*2.1.3.1. Création des data numériques ALG/UTM* (Fossat-Beauchemin-1978)

- Volume 1 TM critères structuraux phonétiques : volume de la base 77 items avec bibliothèque de cartes polymorphes (fichiers XL)
- Volume 2 TM critères structuraux phonologiques : volume de la base, 77 items



- Volume 3 TM critères structuraux morphologiques et morphosyntaxiques : volume de la base, 150 items
- Volume 4 TM critères structuraux de morphologie verbale volume de la base : 48 items
- Volume 5 TM critères lexicaux, volume d'items : 170

Chaque volume de données donne lieu à calcul de distances, et à représentation des isoglosses : la vérification se fait toujours par confrontation aux données éditées d'ALG6 des brouillons de Jean Séguy et des résultats obtenus par méthodes classifiantes d'analyse de données de 1973 à 2001.

Table 6

27b data système morphologique verbe bloc jVARS 548-650  
Groupes et classes (ALG6)

2.1.3.2. *Création d'un index des distances multicritères pour \*loc160* (Fossat-Beauchemin 1989)

Volume des données : 5 volumes de distances – fichier texte/nombres

2.1.3.3. *Report des données nombres sur fonds de cartes ALG* : volume de l'opération (Fossat-1996-2011)

Exemple type de traitement par cartographie quantitative du polymorphisme phonétique attesté par les données de terrain :

traitement de la compacité vocalique en gascon noir, avec ou sans étape classificatoire.

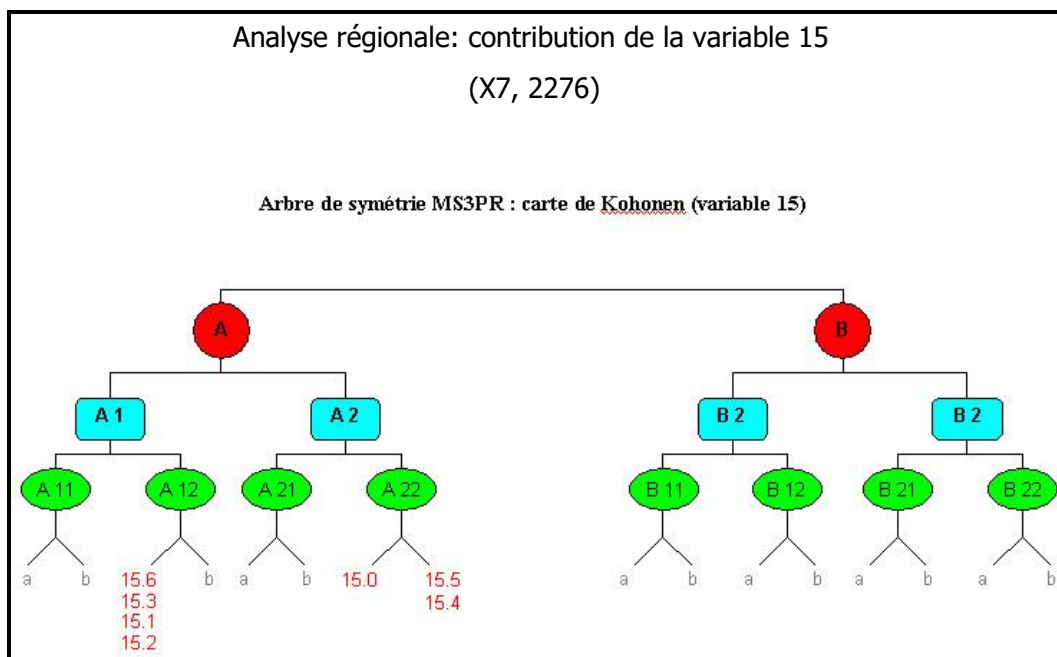
#### 2.1.3.4. Phase classificatoire et cladistique

Exemple retenu : traitement de la morphosyntaxe du pronom régime gascon dans ALG6, à partir des tables, avant et après coopération de classification des J variables sélectionnées.

En colonnes 1, 2, 3, 4, on détermine la hiérarchie des groupes de localités pour la variable prise en compte ou le groupe de variables pris en compte ; en colonnes 5 et 6 on détermine les critères pris en compte par le géolinguiste grammairien, et on renvoie à la carte *n* d'ALG6 (Jean Séguy) ; en colonne 7 on renseigne sur le degré de variation caractéristique de telle variable ; ce qui permet de porter un jugement sur la non homogénéité des variables, sur l'homotypie des résultats ou non.

CLA	N1	N2	N3	JVAR	DOMAINE	K_données	N° mod	OBS.	CODE
A	A1	A11	ab	03_6	PR PERS lé	2253	x9	013	
A	A1	A11	ab	25_2	PR PERS7	2287, 2290	x2		Pb 25_0 1 2
A	A1	A11	ab	03_5	PR PERS lé	2253	x9	012	
A	A1	A11	ab	26_4	PR PERS7	2286, 2288	x8	099	vérif. Code
A	A1	A11	ab	03_4	PR PERS lé	2253	x9	011	
A	A1	A11	ab	24_3	PR PERS	2243	x2	092	
A	A1	A11	ab	18_3	PR PERS	2291			Pb 18_0 1 2
A	A1	A11	ab	57_4	c'est moi Prés.	2503	x6	183	
A	A1	A11	ab	10_3	PR PERS	2233		47	
A	A1	A11	ab	03_7	PR PERS lé	2 253	x9	014	
A	A1	A11	ab	03_8	PR PERS lé	2 253	x9	015	
A	A1	A11	ab	04_7	PR PERS lés	2254	x13	023	
A	A1	A11	ab	04_8	PR PERS lés	2254	x13	024	
A	A1	A11	ab	26_3	PR PERS7	2283, 2286, 2288	x8	098	vérif. code
A	A1	A11	ab	53_5	MS article	2 451		169	
A	A1	A11	ab	05_2	PR PERS prosod	2 239	x4	031	
A	A1	A11	ab	59_1	accord PP	2 500	x4	190	
A	A1	A11	ab	24_2	PR PERS	2 243	x4	0 91	
A	A1	A11	ab	58_2	accord PP	2 498	x4	187	
A	A1	A11	ab	19_2	PR PERS	2310	x4	075	
A	A1	A11	ab	58_3	accord PP	2 498	x4	187	
A	A1	A11	ab	05_3	PR PERS prosod	2239	x4	032	
A	A1	A11	ab	26_1	PR PERS	2283, 2286, 2288	x8	096	vérif code
A	A1	A11	ab	36_1	S COMPAR	2409	x2	127	
A	A1	A11	ab	45_2	PLUR N	2380, 2381	x4	147	
A	A1	A11	ab	04_6	PR PERS lés	2254	x13	022	
A	A1	A11	ab	19_3	PR PERS	2310	x4	076	
A	A1	A11	ab	23_3	PR PERS	2342	x6	085	
A	A1	A11	ab	45_3	PLUR N	2380, 2381	x4	148	
A	A1	A11	ab	04_4	PR PERS lés	2254	x13	020	

À partir de là, on analyse la carte numérique de la partition territoriale réalisée par la ou les variables prises en compte, afin de comparer ce résultat avec la carte éditée initiale (ALG 6 Jean Séguy).



On opère ainsi pas à pas pour les variables des 5 volumes de cartes, avec graphes élaborés en perspective précladistique et préclassificatoire.

On rappelle la très grande simplicité de cette nécessaire vérification des données de terrain par des méthodes douces, mais appropriées à la nature complexe de la variabilité dialectale.

La classification ne fait que valider l'analyse proposée dès 1973 par Jean Séguy, et vérifiée à sa demande par recours aux méthodes appropriées.

Il va sans dire qu'à tout moment, on peut appeler le fichier ou les fichiers sonores numérisés associés à la variable mise en vérification. La vérification s'opère toujours par confrontation des cartes distributionnelles typologiques éditées par Jean Séguy (AKG6) et des résultats issus des opérations d'analyse exploratoire des données (CAH\_i/jVARS).

2.1.4. Opération de base de données relationnelle ALG4 carnets d'enquête numérisés (contribution aux opérations de regroupement dans le cadre du TESOC interuniversitaire (Nice, Toulouse, Montpellier, Aix, Lyon, St-Etienne, Jaca, Barcelona, Torino).

*2.1.4.1. Opération de numérisation des carnets ALG123* (enquêtes de type B sans archive sonore) à charge de CLID\_erss et ercvox, la numérisation des minutes d'enquêtes d'ALG 456 et de l'archive sonore correspondante étant à charge de l'équipe LDOR de l'UTM, sous la direction de Patrick Sauzet.  
Volume de la base la base carnets ALG 123 : 167 carnets de transcription (lacunes ; pertes de fichiers, 2001 : transfert AZF TOULOUSE/sinistres).

*2.1.4.2. Création de base de données relationnelles pour Carnets/Cartes en provenance d'ALG123*, opération coordonnée et exécutée par Jean-Louis Fossat à partir de la collection numérisée des carnets des enquêtes du questionnaire DAUZAT, par ailleurs cartographiées dans ALG123.  
Volume de données \*loc-167 \*q100  
Moteur de recherche base de données relationnelles 4D Société Image/CLID/CLID\_erss  
Transfert de données de la base créée en direction de TESOC (CLID\_erss)

Il s'agit d'un prototype de base de données relationnelles avec incorporation de fichiers sonores et d'iconographie cartographique, avec création de logiciel de représentation cartographique de type CORELDRAW en environnement MAC, avec phase ultérieure de développement en environnement PC.  
La caractéristique essentielle de cette base réside dans la possibilité d'importation des données numériques en direction de logiciels d'analyse exploratoire de données, en perspective qui concilie approche cladistique et approches classificatoires en mode EDA (Eugène Horber 1996, SFC 1996, Toulouse).

*2.1.4.3. Création d'une base de données géolinguistique* cartographie automatisée de classes de variation issues d'une analyse classificatoire : ce point est détaché de cette partie du bilan, car il constitue le point de départ de toutes nos approches classificatoires.

### 2.1.5. Étape spécifique d'analyse exploratoire des données

L'objectif, depuis 1988 (Sherbrooke, Toulouse) est de décrire en analyse exploratoire de données (EDA) l'organisation des structures dialectales grammaticales dans l'espace/temps avec représentation cartographique (K1) ou non (K0).

#### *2.1.5.1. Opération EDA (K0) sur tables UTM/ALG6*

Cette opération repose fondamentalement sur une classification ascendante (CAH) ou descendante (CDH) des données en mode exploratoire, à partir des tables de data (ALG6/UTM), avec utilisation comparative de plusieurs méthodes classificatoires (EDA, ADDAD, ALCESTE, SICLA etc.).

L'avantage des méthodes d'analyse exploratoire des données sur des méthodes statistiques dures, voire brutales, réside dans le fait que ces méthodes laissent en dernier ressort la main à l'expert linguiste, appelé à décider en fin de compte du nombre de classes optimal, après une série d'essais ; elles lui permettent de dépasser la représentation cartographique par isoglosses singulières, au profit de tracés cartographiques cumulatifs de constellations et de structures spatiales.

La CAH/i à partir de la bibliothèque des distances de Hamming ou des distances euclidiennes permet de donner à voir le poids des structures, ou des constellations lexicales dans la partition, dans la réticulation de l'espace dialectal, et ceci dans une stratégie proche du calcul des isoglosses quantitatives et interactives.

La CAH/j sur les distances entre variables permet de valider cette partition (Fossat-208 graphes de CAH/ sur MS3/PR). Elle est une procédure très proche de l'analyse cladistique opérant non sur une liste de lexèmes mais sur un bloc de bandes syntaxiques représentatif des structures syntaxiques et prosodiques en usage ; on s'assurera de cette affirmation en lisant l'analyse syntaxique de Claude Muller, qui règle une fois pour toute la question de la morphosyntaxe prosodique du pronom en gascon.

*Evaluation du volume des résultats*

VOL1 : CHAi/j PHONETIQUE pour 77 objets de structure (première phonétique)

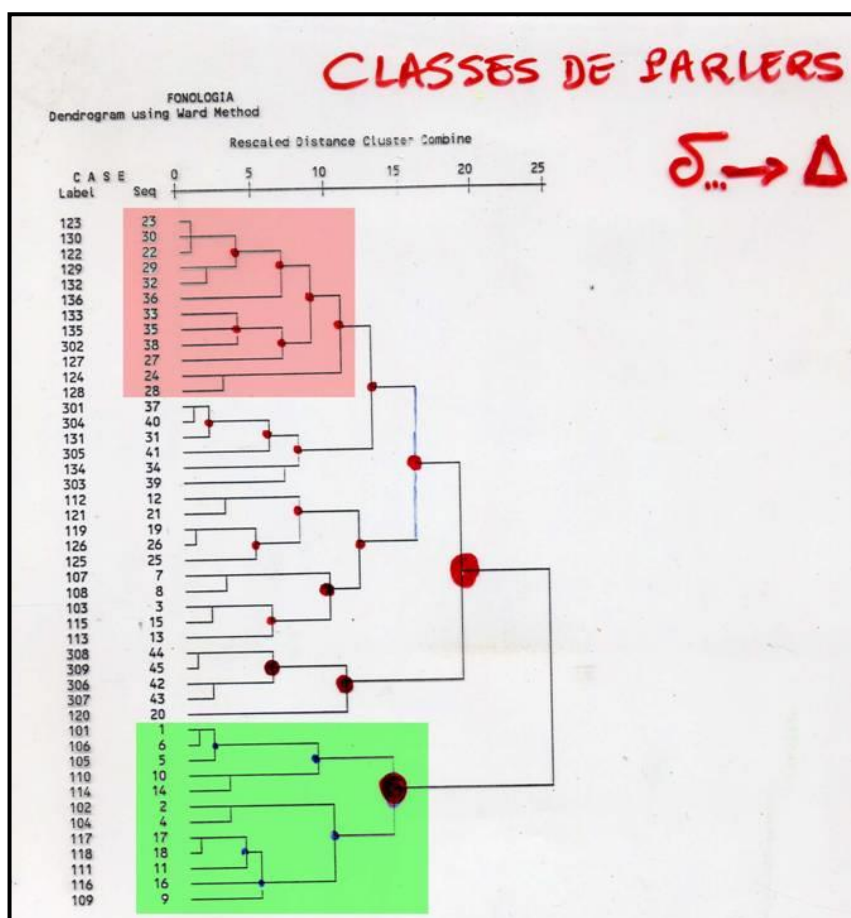
VOL2 : CAHi/j PHONOLOGIE pour 71 objets de structure (première phonologique)

VOL3 : CAHi/j MORPHO-SYNTAXE : pour 150 objets de structure (première morphosyntaxique)

VOL4 : CAHi/j MORPHOLOGIE VERBALE pour 48 objets de structure (première en morphologie verbale de synthèse)

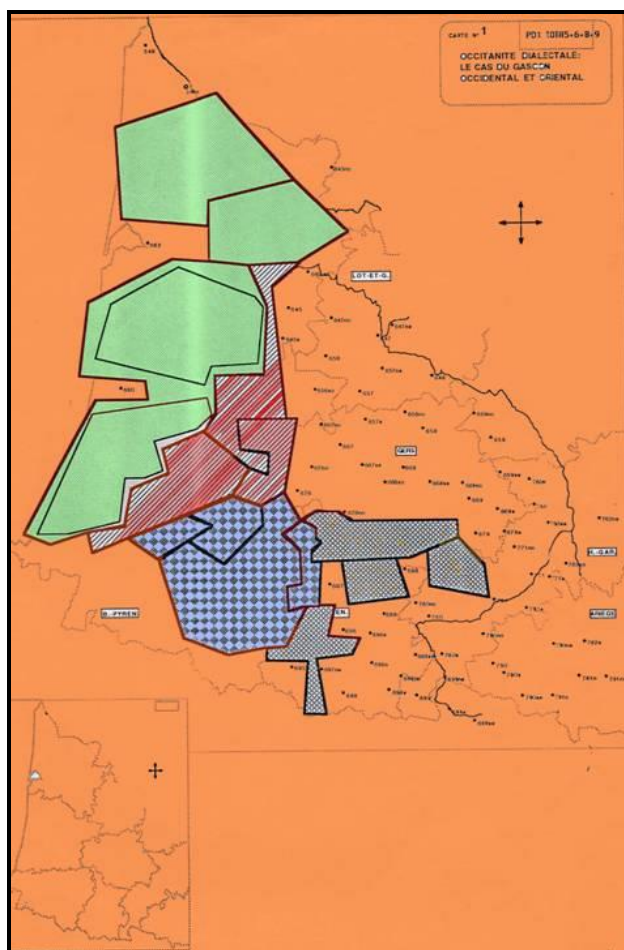
VOL5 : CAHi/j LEXIQUE pour 170 items (première lexicale)

VOL6 : cumul de résultats et analyse en composantes principales  
Exemple retenu : représentation classificatoire des structures phonologiques du gascon.



2.1.5.2. *Création d'une base de données géolinguistiques*  
cartographie automatisée de classes de variation issues d'une analyse classificatoire des groupes et sous-groupes d'une hiérarchie calculée sur une somme de critères : ce point est détaché de cette partie du bilan, car il constitue le point de départ de toutes nos approches classificatoires

Programmation en environnement PC : Zouheir Sandouk  
(Université P. Sabatier, Département d'Informatique, Pr. René Caubet)



**Exemple retenu** : représentation de la partition en classes du gascon pour les critères de phonétique diachronique, à partir d'une classification ascendante hiérarchique. Chaque groupe de l'arbre des distances étant cartographié par ses limites, on obtient l'image de la partition occidentale, comparable à la partition phonologique, morphologique et morphosyntaxique.

L'interprétation de cette carte nécessite le recours au graphe de la partition portant sur les critères comptabilisés pour établir les distances multicritères phonétiques (CAH\_jCRIT\_phonétiques).

### *Evaluation*

La cartographie des résultats de CAH/i, simples ou cumulés, permet de voir les limites territoriales, les forces et volumes des constellations territoriales ; elle permet aussi de voir avec plus de précision la force de réticulation sociale dans l'espace-temps des sociétés rurales dialectalisées.

Seuls les graphes de CHA/J nous permettent de mieux voir par graphes colorés quelles structures dans les usages, donnent à la grammaire de la langue telle qu'elle est parlée dans tel petit monde, par exemple la Charente limousine septentrionale, telle couleur, dans telle constellation territoriale ; on commence à peine à comprendre les relations entre cette couleur et les faits de réticulation des forces sociales ; on sait encore peu de choses sur la façon dont le cerveau « synesthète » des sujets dialectophones traite en mode de perception des produits, cette coloration phonétique, phonologique, prosodique des données ; on sait seulement que dans telle configuration le SP est en /i/ de couleur /i/ ou non ; que telle configuration territoriale comporte des masculins marqués en /i/ ou /u/ avec distribution bien établie.

Seule l'archive numérique empilée qui constitue la mémoire de travail du laboratoire CLId et CLId\_erss dans le prolongement des opérations NALF, ERA352/URL8/INaLF permet de mettre en place une opération de fouille systématique des ressources ou données. Il faut savoir, cependant, que beaucoup de données et résultats ne seront pas accessibles, par exemple s'ils ont été détruits en cours de route, ou ne seront accessibles que dans une configuration instrumentale spécifiée (MAC/PC) ; mais il est aussi utile de savoir que pour lutter contre les forces de destruction, nous nous sommes armés pour sauvegarder ce qui est essentiel à nos yeux. Il faut enfin savoir que bien des outillages utilisés dans les années 1980-2000 sont déjà obsolètes et que les outillages proposés aujourd'hui seront demain obsolètes eux-mêmes.

**Bibliographie** : Philips 1985, Jagueneau 1988, Aurrekoetchea 1991, Fossat 1997 - *Contrat de Région Midi-Pyrénées Atlas sonore numérique et Polarisation territoriale*.



2.1.6. Création d'une disposition d'analyse linguistique des données de terrain (de tout type) en station d'écoute et de transcription de données sonores numérisées : OCCITON2 2011

Tâches linguistiques : Jean-Louis Fossat.

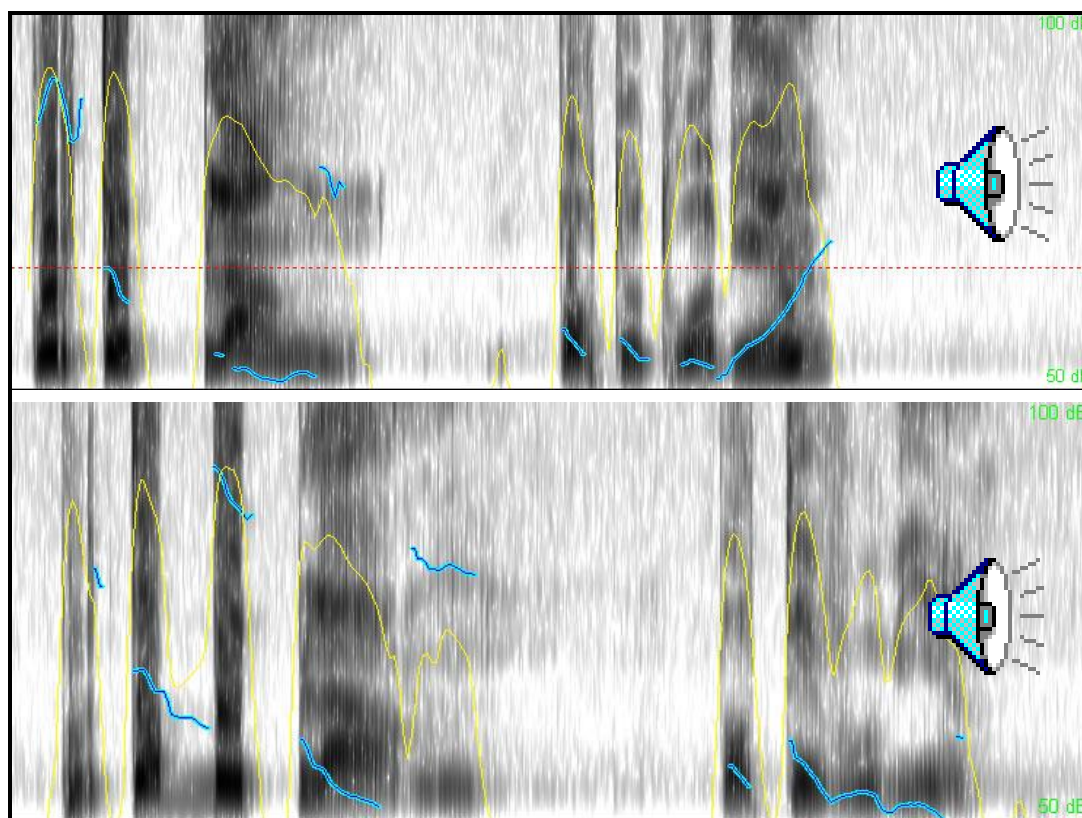
Tâches informatiques : A.Meqqori.

OCCITON2 est une station d'écoute assistée de ressources sonores numérisées. Cet outil constitue l'instrumentation de vérification exigible des données antérieures transcrites, catégorisées, cartographiées, et mises en exploitation.

Il permet en effet, par sélection d'écouter tout ou partie d'une configuration géographique dessinée par CAH\_i ou par CAH/j

Il permet de comparer des bilans de rendement de telle ou telle variable selon le critère typologie des ressources numériques (listes lexicales ou unités discursives).

L'assistant OCCITON2 permet à chaque moment de l'exploitation d'associer à un fichier sonore/wave ou mp3 un fichier texte, un fichier image (carte, table, graphe, arbre, commentaire, annotation catégorielle révisée sur indexation).



Exemple retenu : représentation sonore numérique en environnement PRAAT des bandes sonores *porta-te plan/portatz-ves plan* pour loc 40 MONTFORT en CHALOSSE PROX ALG6 Pomarez.

Il permet à tout utilisateur qualifié professionnellement de corriger telle erreur démontrée, de créer de nouvelles ressources, et de réindexer toute nouvelle base de données géolinguistiques. La multiplication des stations d'écoute constitue un acte pédagogique majeur, sous ces conditions d'appropriation des ressources créées.

Il permet enfin et surtout de confronter à tout moment des opérations d'analyse classificatoire ou cladistique, de retourner constamment aux données terrain telles que recueillies avant d'être mises en exploitation partielle, par extraction et analyse, donc permet de trouver des solutions alternatives en termes de mode d'exploitation des ressources.

Au total, les méthodologies mises en œuvre durant les dix dernières années (2001-2011) permettent de définir les bonnes pratiques et les bons usages de l'instrumentation pour une meilleure appréhension des ressources créées, qu'il s'agisse de :

- (2.1) données de type atlas classées par types d'opérations et de protocoles : structures et usage
- (2.2) données de type discursif et méthodes d'exploitation des sources orales transcrites sous format texte (texte *api* notamment) confronté en mode synoptique à la ressource numérique transcrite et catégorisée en 2.3 par OCCITON2.

### **3. La question lexicale**

#### ***3.1. Principes et méthodes***

##### **LEXIQUE**

L'observable porte sur le nom, la nomination, les relations entre NC/NL et NP dans les structures et les usages.

##### **LEXIQUE-GRAMMAIRE**

L'observable géolocalisé porte alors sur la catégorisation, la transcatégorisation, les phénomènes de délocutivité.

##### **SEMANTIQUE ET PRAGMATIQUE LEXICALE**

L'observable porte alors sur les modes et modalités de désignation.

NOMINATION ET MALNOMINATION DANS LES USAGES :  
*los truca-taulers de Benejac.*

On observe ici la nomination par rapport à l'organisation des sociétés rurales et urbaines de la France, par régions.

STRUCTURES/ ET USAGE

PLURILECTALITÉ ET LEXIQUE DIALECTAL

L'observable repose ici sur la lecture de cartes lexicales, le dessin de configuration selon une méthode généralisée depuis Séguy (ALG4) tant dans les opérations NALF (ALG, ALLOC, ALLOC, ATLAS DU CENTRE etc.) que dans des opérations spécifiques (lexiques agropastoraux, lexique des activités de commercialisation).

POLYNOMIE ET LEXIQUE

Le statut polynomique est déjà observable dans le lexique général dialectal des atlas ; seul le dictionnaire met en évidence de manière systématique le poids de la néologie implantée et/ou adaptée.

LEXIQUE ET ORGANISATION DISCURSIVE

Le lexique est alors extrait des bandes énonciatives, des bandes syntaxiques, prosodiquement caractérisées.

APPROCHE INSTRUMENTALE DU LEXIQUE DE DISCOURS

BM : diaporama (Colloque AIEO Bordeaux 2005) : lexique, prosodie, syntaxe chez Bernard Manciet.

Tout repose ici sur la création d'une base de données prosodiques BM (parler gascon de couleur dite « noire ») en environnement PRAAT de Boersma.

MD Base de données prosodiques de MD Marcel Degeilh dit Le Bur 09 Auragnou (Fossat-2011-2012).

Bé : syntaxe, sémantique et prosodie des petits mots du discours dialectal (Fossat-2010, *Cahiers de Grammaire*).

### ***3.2. Les configurations géolinguistiques lexicales***

*Le lexique dans les opérations géolinguistiques : Structures et usages*

Le lexique dialectal occupe une place centrale dans le recueil de données de type dialectal, des monographies thématiques aux atlas linguistiques, en tant que produit historique, et ne saurait être abstrait de ses conditions de développement, de sa transmission, de ses avatars au cours de son développement et de l'histoire des forces sociales sur le territoire, des origines à nos jours.

Bibliographie :

(1) *OPERATIONS DE TYPE NALF ALG ALLOC ALLOR*

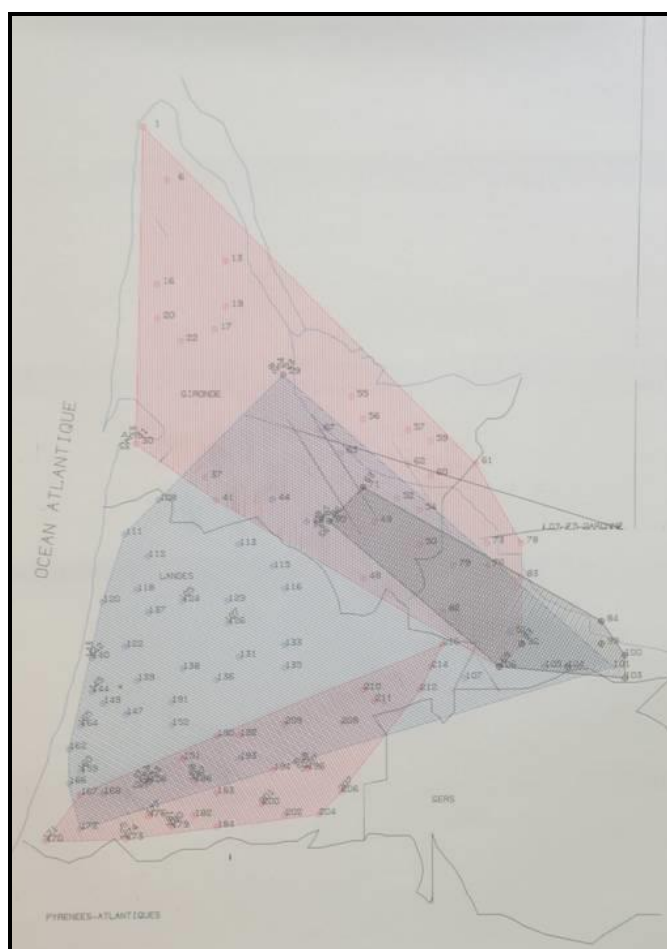
Séguy-1953 atlas pyrénéen des noms dialectaux de plantes : diaporama 2011 (ercvox) : réticulation territoriale (Pyrénées centrales, aragonaises, aranaises, catalanes).

ALG1234 5 6 (J. Séguy) : cartes lexicales du NALF/ALG : index des ressources numériques créées ; lexique dialectal thématique et lexique grammatical catégorisé (opération ercvox).

(2) *OPERATIONS DE TYPE POLYNOMIQUE : COLORATION LEXICALE D, R, F, DF*

(1°) Fossat-1970 : petit atlas lexical DRFDF : diaporama 2012 (opération ercvox) : réticulation territoriale.

(2°) Fossat-1980-86 Lexique et polynomie : création d'un prototype classificatoire CHAUDIEU à partir des ressources lexicales numériques extraites d'une base de données relationnelles.



[Dans ce type d'approche, l'observable cartographié met en évidence, au prix d'une instrumentation appropriée, les usages partagés du lexique dialectal (D), du lexique général français (langue nationale FRA) voire du lexique d'une seconde langue nationale (ALL en Alsace, par exemple) et du lexique innové avec néologie implantée ou adaptée, dans les interactions et les échanges.

Ce prototype a été créé en environnement DBASE3 (Fossat-Ecolia 1996) : il porte sur la coloration des usages lexicaux pour 100 localités du domaine occidental gascon :

- graphes colorés de proximité D et R
  - graphes colorés de proximité R et F
  - graphes colorés de proximité de D et F : DF
  - transfert des données en environnement MAC : à cette étape, en environnement EDA : ou ALCESTE, on dénombre les classes d'usage
  - cartographie des classes d'usage : elle fait apparaître les classes de comportement mixte interprétable en termes d'analyse sociolinguistique.
- La base renseigne sur les usages du lexique dialectal commun en usage technique dans les échanges et interaction.]

(3°) Fossat-1987, création d'une base de données relationnelles généralisée, indépendante des applications dédiées (MAC, PC) :

Conception et analyse ; J. Dendien et J.L. Fossat (INaLF)

Instrumentation : J.L. Hamon (INaLF)

Cette base est dédiée au traitement des archives sonores du fonds CLID entièrement numérisé, dans la section des ressources qui concerne ce mode d'approche DRFDF du statut polynémique des langues de France. Ce mode d'approche fut prolongé par une opération DGLF sur l'implantation terminologique, conduite par Loïk Depecker (DGLF).

Bibliographie : BASE ECOLIA (CHAUDIEU) en environnement DBASE3 et PC (avec problématique de compatibilité (VICTOR/IBM) ; BASE URL8 INF (J. Dendien, J.-L. Hamon, J.L.Fossat INaLF) : base de données relationnelles généralisée DRDF.

### ***3.3. Le lexique général occitan dans le dictionnaire gascon***

#### ***3.3.1. Création des ressources lexicales numériques avec objectif de regroupement***

Palay électronique

D'Estalens électronique

LESPY-RAYMOND

FOIX (dictionnaire gascon des Landes)

LAGARDE

ARNAUDIN (Dictionnaire de la Haute-Lande)

GARY (dictionnaire languedocien)

SAUVAGES

VAYSSIER (dictionnaire languedocien occidental cantalien)

VIGNEAU

PÉRIGOURDIN

GABAL électronique (dictionnaire gévaudanais)

DAUPHINÉ

On dispose parfois de dictionnaires locaux réalisés par des amateurs éclairés voire des professionnels du glossaire ; la plupart exploités par le FEW de Walter Von Wartburg.

### 3.3.2. Création d'un PROTOTYPE DICTIONNAIRE GASCON (bases PALAY/D'Estalenx revue et corrigée, annotée).

Volume de données 68MO (Lettres A, B, C) en environnement 4D (MAC). Description :

- entrée FRA
- entrées de lexique dialectal occitan (lemme et variantes)
- transcription phonétique
- catégorisation
- contextes
- fichiers sonores numériques
- fichiers image numérique : cartes, planches, illustration

Instrumentation : DRAW\_carte\_base option DICTIONNAIRE (Société Image) en environnement base 4D relationnel

Requêtes : exemple -ader et CAT N\_instrumental ; -ader et CA = ADJ/postverbal/passivable

Le prototype de base 4D est disponible pour tout utilisateur qui privilégie un environnement MAC (système 7) et souhaiterait créer sa propre base de données lexicographique, à partir de ressources numériques existantes accessibles.

Quant à moi, j'ai décidé, sous la contrainte, d'abandonner cette base dédiée au profit d'une base revue, corrigée, augmentée, avec objectif de spécification du lexique-grammaire du gascon, en domaine occitan, tâche minutieuse qui m'a occupé tout au long des années 2010-2011,

et m'occupe encore en janvier 2012, avant création d'une base renouvelée.

3.3.3. Prototype de dictionnaire gascon bilingue sur site  
www de CLID\_ERSS repris par site www ercvox

à chaque arbre son fruit	au cerisè, la cerisère (dicton PHRASEOLOGIE PREDIC/ASSERT SYMETRIE_prosodique).
à chaque saint sa chandelle	a cade sén la soue candéle, tous lous sénts que bòlen candéles (dicton PHRASEOLOGIE PREDIC ASSERT SYMETRIE_prosodique).
à demi-perclus	herit (PPadj/statif/herir/aspectuel/herit).
à empattement	a empatamén (NCoP/descriptif/a empattament).
à l'amiable	a l'amiàble (LOC_ADV a l'amiable).
à l'automne, de la montagne; l'été, de la rivière	a l'estiuère / de la ribère / a la castagne / de la mountagne (dicton PHRASEOLOGIE PREDIC/ASSERT SYMETRIE_prosodique) (COMMINGES).
à la dérobee	a la guignaròlo (LOC_ADV/comportement).
à la file	dehù, dehieu (ADV_composé/loc/temps/de hiu).
à le voir si fier, on dirait que c'est un roi qui lui garde les vaches	que sèmbre que lou réy qu'òu goàrde las bàques (PREDIC/ASSERT/ENONCIATIF_que MS3_PrCo PROCLISE).
à part-soi, sous cape	tout tourrén (ADV_composé_manière/tot torrent)(GR_LANDE).
à pieds joints	a pè jùnt (ADV_composé/comportement/manière).
à poignées	a pugnàts (ADV_composé/QUANT_indéf/a punhats).
à Ste-Luce, un saut de puce/ à Noël, un saut de brau	a sènte Luce un saùt de pùce / a Nadàu, un saut de bràu / per sènte Lùcie / un saut de pùcie (dicton PHRASEOLOGIE PREDIC/ASSERT SYMETRIE_prosodique).
à titre d'exemple	en guise d'exèmple (NCoP).
à vieille chatte, peu de badinage	a gâte biélhe / pòc de jouguines (dicton proverbe PREDIC ASSERT ENONCIATIF SYMETRIE_prosod).
à vos souhaits!	dùu (que) b'assiste! (LOC_V PREDIC formule/souhait SP_assistir MS3_PrCo/proclise interjection exclamatif)

Le chemin du traitement des données passe par la classification des données regroupées, ce qui commence à pouvoir être envisageable dès

lors que les ressources numériques lexicographique existent, en dépit de leur non homogénéité, qui constitue un obstacle sérieux.

Chaque ressource lexicographique D1/n est appariée à sa base de textes (T) et à un fichier son, de manière systématique.

Tâches informatiques : A. Meqqori

- moteur de recherches actualisé en janvier 2012
- requêtes sql (2012)
- incorporation de fichiers sonores numérisés
- incorporation d'images numériques (cartes, illustration, planches)
- incorporation des relations NC/NL/NP
- lien avec Occiton2
- lien avec bases de textes dialectaux historiques, modernes, contemporains, anciens.

Exemple ARNAUDIN\_D/ARNAUDIN\_T (contes, proverbes)

#### **4. Base de données textuelles**

Nous avons déjà dit que notre objectif était ici de permettre la comparaison entre ressources orales de type spécifié, et sources écrites numériques accessibles à partir de diverses bases de données existantes.

##### ***4.1. Bases de textes dialectaux juridiques et techniques médiévaux concernant l'encadrement des activités de production et commercialisation des sociétés rurales et des villes.***

CART-ST-GAUDENS  
CART-MONTSAUNES  
CART-GIMONT  
CC-RISCLE  
CC-ALBI  
CART-REQUISTA  
CART-OSSAU

\*t01 CARTULAIRE A OSSAU

\*\*\*t005\_01 Cartulaire A f°59 r° - 60 v° \*d 1219 \*l 40\_Dax \*r  
TCH, 60-63



*So es assaver que a cascune besthie deus soberdiitz boeus, baques,  
de rociis et de geogoes  
que no sera deu rey d'Angleterre soverdiit o de nos Gaston o de  
nostres feuds o darrerfeuds,  
mes sie de homis de autre senhor et teratori que bienque per  
apascentar en las terres soverdiites,  
done et pague au soverdiit rey notre senhor per lo guidoadge tres  
diners morlaas cascun han  
a la mieytat deu Caresme, et nos Gaston soverdit, per lo passadge  
de la notre terre dus dines  
de la medixe monede, ung per anar et aute per lo tornar en lo  
soverdiit termi, exceptatz los frociis  
et las gegoes qui seran deputatz et ordenatz a portar las bitualhes o  
las causes usadisses  
de baques et de cavers, et une baque o un taur deu grey aramat en  
la maneyre qui es estat  
acostumat, lasquoaus besties sien quitis de la pague soverdiite.*

**BASE CART ST-GAUDENS\*r CARTULR SAINT  
GAUDENS 31 Nébouzan \*d 1203 \*l gascon oriental  
commengeois [texte éd. Mondon contrôlé, non remanié, non  
normalisé, non codé]**

**PRE-CODAGE CATEGORIEL : étiquetage et alignement**

01 NOM NP NL NC simple et composé (*sent gaudens, sents  
euangelis, sabuda paraula, Nanfos*)

02 VERBE AUXILIAIRE FORMES PERS FORMES  
FLECHIES IP SP IMPT Fu C PP IF PP G +PT PC V FORT V  
FAIBLE (*sabuda es*)

03 ACTUALISANTS DEFINIS INDEFINIS PARTITIFS  
QUANTITATIFS (*lo comte*) DEMONSTRATIFS

04 ADJECTIF

05 CONJONCTION COORDINATION (*e, ni*)

06 CONJONCTION SUBORDINATION (*que, can*)

07 PREPOSITION SIMPLE OU COMPOSEE (*de sent  
gaudens*)

08 SUBSTITUTS PERSONNELS SIMPLES OU  
COMBINES (PR 1 2 3 4 5 6 7 *en y*)

10 RELATIFS INTERROGATIFS (*qui, lo cal*)

11 POSSESSIFS

12 NEGATION (*no*)

13 ADVERBE (*exament, atal, atals*)

### PRE-CODAGE SYNTAXIQUE DES COMPOSANTS

#### CATEGORISABLES

\*\*\* \*t001A \*r CartulR Saint-Gaudens 31 \*d 1203 \*n faire le  
**graphe de reconnaissance des composants clitiques**

E *can* lo senhor ni son bayle **demanaran** fizansas **diga\_u** lo  
clamant *qui es* e de *que se clama* e ab atant **don lo** fizansas **lo qui se**  
**clama** e **\_l autre** de *cuy se clama don lo* fizansas exament a son  
poder *si auer* no **las pod. valentz** per conoysensa dels prosomes  
**jur** sober sents euangelis *que no pod auer* fizansas per aquet pleyt  
e **\_l** senhor **fassa lo judiar** soler **si mezis**.

\*\*\* \*t001B \*r CartulR Saint-Gaudens 31 \*d 1203

E *can* lo senhor ni son bayle **demanaran** fizansas **diga\_u** lo  
clamant *que es* e de *que se clama* e ab atant **don lo** fizansas lo *qui*  
**se clama** el autre *quy se clama don lo* fizansas exament a son poder  
si **auer** no **las pod. valentz** per conoysensa dels prosomes **jur** sober  
sents euangelis que no **pod auer** fizansas per aquet pleyt e **\_l** senhor  
**fassa lo judiar** sober **si medis**.

La méthodologie d'exploitation des données quantifiées en environnement STABLEX ne diffère en rien de la classification des données dialectales de type géolinguistique des atlas linguistiques, puisque les données de terrain sont nécessairement géoréférencées.

#### ***4.2. Base de textes écrits dialectaux modernes et contemporains***

#### ***4.3. Base de données de textes français (textes scientifiques, XVII-XIX<sup>e</sup> siècle ; archives nationales ; départementales ; archives notariales)***

Lavoisier

Vicq d'Azyr

Manuels de médecine vétérinaire des années 1810-1850 (Maison Alfort ; Lyon ; Toulouse)

## **5. Bases de données prosodiques et stratégies douces d'analyse prosodique**

*5.1. L'homme qui parlait aux mules : grammaire et prosodie du gascon noir (Fossat AIEO Bordeaux 2008 Hommage à Bernard Manciet)*

*5.2. L'homme qui parlait de ses brebis mortes sous la neige de mars (Fossat-2011 Hommage au Bur)*

*5.3. L'homme qui avait vu l'ours (Isidore Farré, Chalabre)*

## **Conclusion**

Ce bilan porte sur près d'un demi-siècle de recueil de données de terrain en vue de leur mise en exploitation accélérée dans la dernière décade, dans une perspective descriptiviste des structures et des usages. On a fait tout le possible, sachant que la tâche à accomplir est gigantesque, pour nos successeurs. Nous aurons connu sur le terrain des heures très précisément fabuleuses, essentielles dans nos histoires de vie ; et sur le métier, au laboratoire, des heures passionnantes, au côté de compagnons sans lesquels rien de cette mémoire de laboratoire n'existerait.

## **Discussion**

On appelle à une discussion serrée de chaque point, pas à pas, de ce premier jet de bilan. Une des questions, à mon sens absolument fondamentale est de savoir quand et comment toutes les données de terrain manipulées pourront être mises à disposition de la communauté scientifique, sans exception aucune, pour être mise en état de fouille dans un environnement approprié ; avec cet objectif constant depuis les années 60 que nous avons agi sur les terrains comme en laboratoire, faisant en sorte que tout type d'utilisateur puisse lui-même vérifier pas à pas les données, les résultats, les procédures opératoires.

## ANNEXE 1

### BILAN 2000-2011

#### 1. Produits définis en méthodes d'analyse de données massives

01	Occiton2	Données massives annotées de type atlas : lexique morphologie verbale syntaxe prosodie	mp3/wav annotation api graphie marquage	moteur sql
02	Occiton2	Données massives type discursif dialectal annoté Lexique morphologie syntaxe prosodie	mp3/wav annotation api graphie7marquage	moteur sql
03	Occiton3	Données massives français parlé annotées	mp3/wav annotation api graphie marquage	oteur sql
04	Gasconsq1	Dictionnaire explicatif et combinatoire annoté	annotation apinew graphie marquage éléments sonores numérisés	moteur sql
05	Occiton4	Corpus de paraboles	mp3/wave annotation apinew graphie marquage	moteur sql
06	Prosodie01	Corpus annotés BM-2005	wav/praat	moteursql
07	Prosodie02	Corpus atlas annotés	wav/praat	moteur sql
08	Prosodie03	Corpus discursifs annotés	wav/praat	moteur sql

#### 2. Production de données massives : archives sonores de la parole (FONDS CLiD 1863\_2001)

Réf. – Préinventaire sur site [ercvox.free.fr](http://ercvox.free.fr)

Corpus dialectaux occitans (type discursif)

Corpus dialectaux français (type discursif)

Corpus dialectaux romans (type discursif : exemple Vega de Pas, Cantabria)

Corpus dialectaux non romans.

Le tout numérisé avec achèvement de l'opération des fonds au 30/01/2012 à l'exclusion de la filmographie AKAI/SONY qui pose encore quelques problèmes d'ordre technique et financier.

01	AS bandes 18 cms GF	wav copie miroir	sauvegarde	exploration
02	AS bandes 15cms MF	wav copie miroir	sauvegarde	exploration
03	AS bandes 12cms PF	wav copie miroir	sauvegarde	exploration
04	AS cassettes	wav copie miroir	sauvegarde	exploration
05	AS/AV	Bandes AKAI ¼ pouce	sauvegarde	exploration
	AS/AV	Bandes sony 1/2pouce	sauvegarde	exploration

### 3. Production de données massives textuelles : textes graphiques dialectaux anciens pour analyse exploratoire de données Réf : corpus numérisés de textes occitans (1996\_1999).

01	Cartulaire GIMONT lat/gasc	texte word	non marqué	NC/NL/NP
02	Cartulaire St_Gaudens gasc	Texte word	non marqué	gascon
03	Corpus ASPE Eygun béarn.	Texte word	non marqué	gascon
04	Corpus Orthez Martinet béarn	Texte word	non marqué	gascon béarnais

### 4. Production de données massives textuelles : textes graphiques annotés en mode apinew/apinewword

01	text Dictionnaire Arnaudin	texte api	non marqué	NC/NL/NP
02	text. Dictionnaire Foix	texte api	non marqué	
03	text. Chavanat 23 Verbe	texte api	marqué	morphologie verbale phonologie
04	texte Chavanat 23	texte api	marqué	morphosyntaxe phonologie
05	Corpus asturien parlé	Texte api	non marqué	morphosyntaxe phonologie

## 5. Production numérique de données visuelles massives mises en relation avec archive sonore numérique

REF. - FONDS FERIA (CLiD)

Réf. - texte de présentation (JLF)

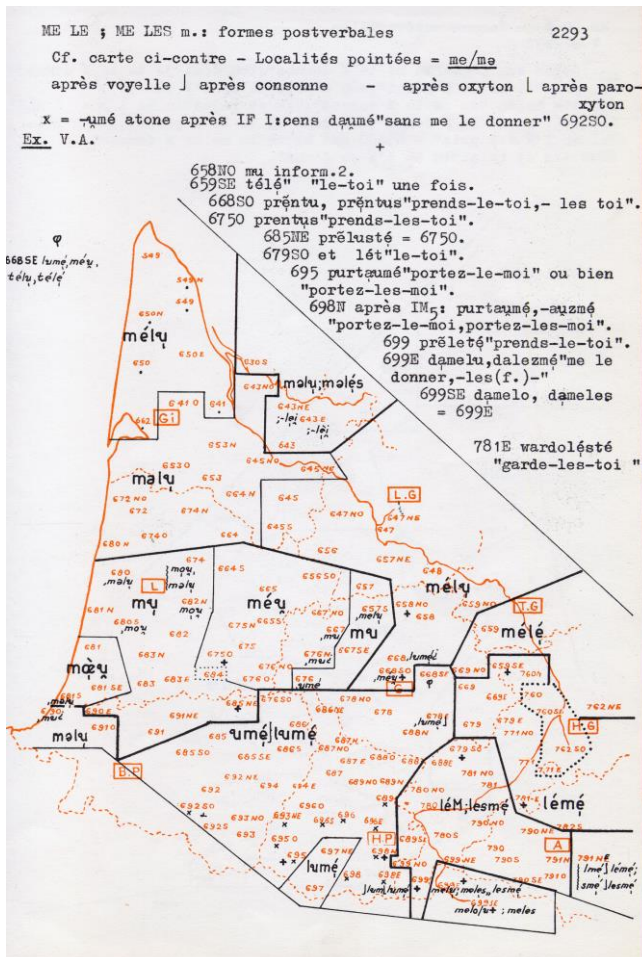
01	65 LUZ	fichiers jpeg	catalogue	moteur sql
02	65 GEDRE	fichier jpeg	catalogue	moteur sql
03	46 FIGEAC	fichiers jpeg	catalogue	moteur sql
04	09 St_Girons	fichiers jpeg	catalogue	moteur sql
05	09 Castillon	fichier jpeg	catalogue	moteur sql
06	09 Seix	fichiers jpeg	catalogue	moteur sql

## 6. Fonds Jean Séguy numérique, données massives indexées

Réf.- Texte de présentation (JLF)

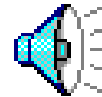
01	Carnets alg123 enq_q_Dauzat	fichiers jpeg	catalogue	moteur sql
02	maquettes cartes	fichier jpeg	catalogue	moteur sql
03	correspondance scientifique	fichiers jpeg	catalogue	moteur sql

## ANNEXE 2 : Archive numérique sonore appariée à représentation cartographique



Carte de données 1

Méthodes d'analyse régionale à partir des fichiers.wav





### ANNEXE 3 : Requêtes aux bases de données textuelles diachroniques

Référence	Loc/ date	Catégorie	ULEX	CONCORD	Objet	PHONOL
JE-2009 9,37,1	AYDIUS 1604	NL	Aydius	au loc d'Aidius	iw	aj'diws
JE-2009 9,37	AYDIUS 1604	NC	aulhes	per goardar las aulhes	aw ɔ u	'awʎəs
JE-2009 9,37,10- 11	AYDIUS 1604	NC	corau	no pot perde lo coelar ny corau ny moleter	aw	ku'raw
JE-2009 9,37	AYDIUS 1604	QUANT	autres	sy d'autres se'n y ban	aw	'awtrəs
JE-2009 9,37		QUANT	auguns	sy auguns pastor deu loc	aw	aw'gyns
JE-2999 9,37,13	AYDIUS 1604	FU3_ poder	poira	e lo loc lor poira prender	uj	puj'ra
JE-2009 7,31,2	AYDIUS 1604	ART_ contract	suus	Suus lous terminis	uʷ/ y	sys
JE-2009 7,31,5	Lées- Athas Osse Arette 1602	NP	Guixarnaud	Guixarnaud de Domecq	aw	gij'ar'nawt
JE-2009 7,31,6	Lées- Athas Osse Arette 1602	NP NC	Lostau	Jaques de Lostau	aw	us'taw
JE-2009 7,31,19	Leées- Athas Osse Arette 1602	NP	Peyroton	Peyroton de Casabonne	ɛj	pɛjru'tũŋ
JE-2009 1,15,3	ACCOUS 1590	NC	dauna	dauna propriary de la maison d'Ussauro	aw	'dawɲə
JE-2009 1,15,14	ACCOUS 1590	NC	eugua	una tersola d'eguas	ɛw	'ɛwga 'jɛwɣwa
JE-2009 1,15,17	ACCOUS 1590	NC	aolhas	una anolha de bacas e detz aolhes balhadas	aw	'awʎəs
JE-2009 2,19,3	Lées Osse 1590	NC	baniu	touta aquera molha, molliar, pachera, baniu	iw	ba'niw
JE-2009 3,23,1	Accous Lhers 1590	NL NP	Eygun	Bernad d'Eigun d'Acos	ɛj	ɛj'gy

JE-2009 3,23,6	Accous Lhers 1590	NL	Larraus	ab terre e bordar de Laraus	aw	la'r:aws
JE-2009 3,23,7	Accous Lhers 1590	NC	ayga	l'ayga aperada Labbadia	aj	'ajga
JE-2009 4,25,6	Lescun 1592	NC NL	peira	peiras molutas	ɛj	'pejras
JE-2009 8,35,3	Osse 1602	ADJ	lauradiis	un tros de terra et camp lauradiis	aw	lawra'dis
JE-2009 2,19,3	Lées Osse 1590	NP	Oussa	Joan pocquet de Lasalla deu locq d'Oussa	ɔw u ɔ	'owsa
JE-2009 11,45,4	Lées Accus, Bedous, Osse, Lescun Anso	NP	Peirolet	Peyrolet de Sancho	ɛj	pejru'let
JE-2009 11,45,4	Lées Accus, Bedous, Osse, Lescun Anso	NP	Audetes	Pedro d'Audetes d'Accus	aw	'pedro d' aw'detas d a'kus
JE-2009 12,49,5	Borce, Etsaut, Urdo Cete 1610	NL	Aubisa	per la baig de Aubisa	aw	aw'biza
JE-2009 12,51,2	Borce, Etsaut, Urdo Cete 1610	NL NCoP	Auloro	per anar ent'Auloro	aw	awlu'ru
JE-2009 12,53,9	Borce, Etsaut, Urdo Cete 1610	NL prépositio nnel NCoP	At Saut	per passar de Borsse at Saut	aw	at 'sawt
JE-2009 12,45,2	Borce, Etsaut, Urdo Cete 1610	NLCoP articulé	Et Saud	Peyrolet de Lambeia det Saud	aw	ɛt'sawt
JE-2009 13,59,3	Aydius 1615	NL NcoP	Petranera	la barera de Patranera	e	'petra 'nera 'p'ejra 'n'era
JE-2009 16,67,4	Aydius 1621	NL NP	Laroi	Joan de Laroi	uj	'jwan de la'r:uj
JE-2009	Aydius 1621	ADJNC NL	vesiau	en un trens de terre vesiau	aw	be'zjaw

16,67,5-7						
JE-2009 16,67,5	Aydius 1621	NC	feus	touta la feus	ɛw	'hɛws
JE-2009 16,67,9	Aydius 1621	ADJNC NL	vesiau	tout so de vesiau	aw	be'zjaw
JE-2009 19,77,16	Lées_Ath as 1604	GN	feugar	ab feugar de Lauttecaze	ɛw	hɛw'ga
JE-2009 19,77,19	Lées_Ath as 1604	GN articulé	deus	a conexence deus juratz	ɛw ɔw u	dɛws
JE-2009 20,83,10	Osse 1664	NP	Larraus	Joan de Larraus	aw	la'r :aws
JE-2009 23,93,9	Accous Joers 1686	GN NC NL NCoP	Ichau	ab terre comune et feugar d'Ichau	ɛw	hɛw'ga d i'ʃaw
JE-2009 23,93,8	Accous Joers 1686	GN NC NL	Haut	ab terre d'Esquerre de Haut et feuga de Laborde	aw	ɛsk'er :ə de 'hawt
JE-2009 23,93,4	Accous Joers 1686	GN NC/ NL/ NCoP	La Cauba	audit locq et parçan aperade a la Caube	aw	a la 'kawbə
JE-2009 23,95,9	Accous Joers 1686	GN NC/ NL	Loustau	l'autre mieytat lodit de Loustau	'je aw	mjej'tat lus'taw
JE-2009 22,10 CC6 AD64	Osse, Lées, Athas 1680	NC/ NL	coila	per lo coila en baig	uj wej	pɛr lu kwej'la kojo'la kajo'lar ɛm_'batʃ
JE-2009 25,107,1 1	Accous Jouers 1688	NP	Jean de Garsaët	Jean de Garsaët	ɛ	'zwan de garsa'et
JE-2009 27,111,1 0	Osse, 1692	NP	Darrerlagleis e	e jou Jean de Darrerlagleise notary	ɛj	e 'zu zwan de dar:e la 'glejzə
JE-2009 27,117,4	Accous, Lhers, Lescun 1705	NP	Pierre de Garsaet	Pierre de Garsaet	e	garsa'et
JE-2009 27,119,1 4	Accous, Lhers, Lescun 1705	NC/ NP	caillau	e a un autre caillau ou pene qui est dou costat de haut ou lou cané deudit de Soubie passe	aw	ka'law
JE-2009 30,129,8	Lées, Osse,	GV	poueira	poueira coustar	uj	puj'a kus't'a

AC64 Osse Registre délib.	Lourdios 1710					
JE-2009 30,131,8 AC64 Osse registre délib.	Lées, Osse, Lourdios, Issaux 1710		corau	la taule de canne de corau a raison de oèit sos bous	aw wej	'taulə ku'raw 'wejt
JE-2009 33,143,1 CC6 AC64 Accous	Accous 1724	NP	Jean d'Apiou	pagarats a Jean d'Apiou deudit locq ... a far taulars ... e fustes au cassiau	iw	ap'iw
JE-2009 36,157,9	Lées, Athas 1738	GN NC/ NP	Peyré	Deu costat deu mouli de Peyré	ej ew	dew kus'tad dew mu'li dè pej're

## Références bibliographiques

- Besche-Commenge, Bruno (1979), "Un carnet de saillie", *Cahiers d'Etudes Romanes*, Université de Toulouse-II Le Mirail, CLiD, Centre de linguistique et de dialectologie 1, 1979 pp. 5-30.
- Besche-Commenge, Bruno (1981), "Le maïs mange, les brebis sont en sève, ou les maîtres de l'herbe. Analyse sémantique du glap et du reish". *Cahiers d'Etudes Romanes* 2-1, Université de Toulouse-II Le Mirail, 1981 pp. 6-25.
- Besche-Commenge, Bruno (1985), *Le savoir des bergers de Casabède*, *Cahiers d'Etudes Romanes* ; N° spécial, [2 volumes]. Université de Toulouse II Le Mirail, Centre de Linguistique et de Dialectologie 1.
- Canobbio, Sabina ; Telmon, Tullio (2008), *Paul Scheuermeier. Il Piemonte dei Contadini (1031-1932* [2 volumes]. Torino, Università degli Studi di Torino, Priuli-Verluccha.
- Fossat, Jean-Louis ; Beauchemin, Normand (1978), ALG6. Matrices dialectométriques et calcul de distances. 1 Phonétique 2 Phonologie 3 Morphosyntaxe 4 Morphologie verbale 5 Lexique. Université de Toulouse-II Le Mirail et Université de Sherbrooke (Québec) (DVD 2012 Fonds d'archives Jean Séguy).
- Fossat, Jean-Louis ; Casanova, Yves (2001), Cartographie distributionnelle d'ALG6 1 Phonétique 2 Phonologie 3 Morphologie verbale 4 morphologie et syntaxe. Université de Toulouse-II Le Mirail 4 volumes ; CD\_ROM, 2001 et DVD 2012 Fonds d'archives Jean Séguy.
- Fossat, Jean-Louis (2008), "Ombres et lumière de la dialectométrie". Colloque international *Methods in Dialectology*, Université de Leeds, 2008 (DVD fonds d'archives J. Séguy, Toulouse Université de Toulouse-II Le Mirail, juin 2012).
- Fossat, Jean-Louis (2010), "Les petits mots énonciatifs gascons : le cas de *bè* énonciatif". *Cahiers de Grammaire*, N° spécial anniversaire. Toulouse, ERSS 30, 2006 pp. 159-174.
- Olive, Francine (1975), *Lexiques techniques agricoles dans le pays de Sault*. Mémoire d'Études Supérieures/ Université de Toulouse-II Le Mirail, 1973 [dactylogramme].
- Philps (Dennis (1985), ADYPIC. *Atlas dialectométrique des Pyrénées Centrales*. 2 Volumes. Thèse de doctorat. Université de Toulouse-II Le Mirail, Centre de Linguistique et de Dialectologie.

- Sénégas, Rose-Marie (1973), *Dialectologie du Minervois*. [2 volumes]. Thèse de doctorat. Université Paul Valéry.
- Sénégas, Yves (1975), *Dialectologie du contact Saint-Ponais-Minervois*. Thèse de doctorat, [2 volumes]. Université Paul Valéry Montpellier.
- Séguy, Jean (1973), *Atlas Linguistique et Ethnographique de la Gascogne*. Volume 6 (ALG6). Cartes, tables, dictionnaire de variables, Paris, Éditions du C.N.R.S.

*Xavier Ravier*  
*Professeur émérite, Université Toulouse 2*  
*ravier.xav@wanadoo.fr*

## **Au sujet de la carte ALG IV, 1115** ***Rouleau de foin***

Le résumé qui vous a été distribué indique dès ses premières lignes que mon propos de ce jour s'inscrit dans la continuité directe de celui qui avait été le mien ici même en 2006<sup>1</sup>. Le commentaire d'une carte linguistique est toujours un exercice productif, à condition bien sûr qu'il ne se limite pas à une énumération des données assortie d'un commentaire restrictif.

La carte 1115 de *l'Atlas linguistique et ethnographique de la Gascogne* est l'homologue, comme je l'avais dit il y a trois ans, de celle que Pierre Gardette avait publiée dans son *Atlas linguistique et ethnographique du Lyonnais* : toutes deux sont intitulées ROULEAU DE FOIN<sup>2</sup>. Il se trouve que la question a été posée lors des deux campagnes au cours desquelles ont été rassemblés les matériaux de l'atlas gascon, mais dans des conditions très différentes. D'abord, elle a été soumise aux informateurs lors des relevés menés avec le questionnaire Dauzat : il se trouve que la seule donnée recherchée à partir de cet instrument était le nom de l'*andain* ; mais les enquêteurs, avec un zèle louable, avaient imaginé de compléter cet *item* en demandant le nom des deux longues rangées de foin sec que l'on fait afin de procéder au chargement, ce qui leur avait permis de recueillir un nombre significatifs de vocables, non seulement des substantifs

---

<sup>1</sup> Lors du colloque de 2006 organisé par Brigitte Horiot sur le thème *La dialectologie hier et aujourd'hui (1906-2006)* : parution en 2009 sous le titre *La dialectologie hier et aujourd'hui (1906-2006). Actes du colloque international tenu à l'Université Lyon III (7, 8 et 9 décembre 2006)*. Ma communication m'avait donné l'occasion d'évoquer le travail de Pierre Gardette, ainsi que l'indique le titre que j'avais choisi : « Pierre Gardette, analyste de ses cartes linguistiques. Réflexions et souvenirs », op. cité, p. 131-143.

<sup>2</sup> La carte de Pierre Gardette porte le n° 33.

mais aussi des verbes désignant l'opération consistant dans le façonnage de ces rangées de foin. En introduisant cet article supplémentaire, les enquêteurs avaient spontanément formulé leur question de manière indirecte, anticipant une pratique qui serait plus tard systématiquement mise en œuvre lors de l'enquête complémentaire de l'ALG. Dans la notice de la carte ANDAIN, Séguy employait lui-même l'expression *question indirecte* au sujet d'*andain* mais on peut penser que ses investigateurs en avaient fait de même pour l'autre objet, le rouleau de foin<sup>3</sup>.

Lors de l'enquête complémentaire, qui a fait l'objet dans les années 1957 et suivantes d'un enregistrement magnétique intégral, on a jugé bon de compléter à propos du rouleau de foin sec les données antérieurement recueillies et publiées en 1956. L'item correspondant fut donc introduit dans le nouveau questionnaire et la seconde récolte confirma ce que la précédente permettait d'espérer : on rassembla donc les données en vue d'une carte spécifique dans toutes les localités du réseau. C'est cette carte qui va être l'objet de mes réflexions : j'ose croire que mon propos sera ainsi en accord avec l'intitulé choisi pour notre rencontre : *Mémoires du terrain*.

La première chose qui frappe dans ce document est la grande diversité des données : même si certains choix lexicaux ont eu la préférence des locuteurs dans certaines zones, par exemple les mots reposant sur la base *and-* au sud-est du Gers, à l'ouest de la Haute-Garonne et à l'est des Hautes-Pyrénées – *anda*, *adanha* -, avec aussi des attestations dans la Gironde, il eût été difficile de dresser dans le cas qui nous occupe une carte avec aires dégagées comme les dialectologues toulousains avaient pris l'habitude de le faire aussi bien pour la Gascogne que pour le Languedoc. Le lecteur pourra se reporter à la liste qui suit le présent exposé, dans laquelle j'ai rassemblé les noms du rouleau de foin tels qu'ils sont mentionnés sur la carte : ces noms sont donnés en orthographe normative occitane, accompagnés de leur transcription phonétique en API.

Profitons au passage de la référence qui vient d'être faite quelques lignes plus haut à la base *and-* pour évoquer, brièvement, le problème

---

<sup>3</sup> La question relative à ce mot consistait en la description que voici de son signifié : « Après avoir formé les andains, on les ramasse en longs tas allongés qu'on charge ensuite sur le char ou la charrette, ou qu'on coupe en petits tas si le foin n'est pas encore sec. Nom de ces longs tas tenant toute la longueur du pré ? ».



qu'elle pose quant à son origine : il est admis, comme vous le savez tous ici, qu'elle représenterait la forme qu'aurait prise le vocable *ambitus* « circuit, pourtour » en latin dit tardif. Je me contente d'une simple question : si *ambitus* est vraiment l'élément originel, ne peut-on pas penser que son évolution vers *and-* s'est faite sous l'influence d'un autre radical de sens voisin mais peut-être absent de la latinité au moins originellement ? L'une des difficultés présentées par la base *and-* est entre autres choses la non-réduction du groupe *nd* à *n*, normalement attendue en domaine gascon : cf. par ex. *brenar* « goûter » < MERENDARE, *brana* « bruyère » < germ. BRAND, etc. ; d'autres problèmes étymologiques relatifs aux vocables de la carte, parmi lesquels : *orde* [urde], *gorruda* [gu`rryde], *estesa* [es`teza] 696 < EXTENSA, etc.

Le dénombrement des formes pleines nous met en présence de 66 occurrences, toutes variantes confondues, pour un réseau de 155 points d'enquête : on est donc en présence d'un taux de variabilité plutôt élevé (pour chaque localité, le seuil à partir duquel ne serait pas entrée dans son lexique l'une des 66 formes relevées est à peine supérieur à 0,4% du matériel rassemblé, ce qui est vraiment très significatif). On voit aussi que du point de vue sémantique, un partage assez net s'est fait entre vocables dont le radical renvoie à une ligne ou un alignement par ex. *anda*, *rua*, *nai*, *còrda* et vocables dont la motivation fondamentale est celle de « rotondité », tels *canon*, *redòla* ; dans cette seconde catégorie entre *artòrcla* < \*RETORCULA : il faut savoir qu'en gascon, ce terme renvoie aussi à l'idée de « tresse ».

D'autres sémantismes, plus ou moins éloignés des deux catégories que je viens d'indiquer, semblent être tout de même entrés en ligne de compte : je pense aux formes *aròla* dont il est à peu près certain qu'elle a un rapport avec ceux des produits de ARARE ; et aussi *taula* « table » (rapprochement avec une planche ensemencée), *coa* « queue », etc.

Permettez-moi ici d'attirer l'attention sur *rudèra* [rry`dèro] : j'avais proposé, lors de notre rencontre de 2006, qu'on le rapprochât du franco-provençal [rwela] (porté sur la carte de Gardette) < ROTELLA « petite roue ». Pour la forme gasconne : le [ll] de la syllabe 3 ne pouvait qu'aboutir à [r], traitement absolument normal en gascon et considéré comme l'une des caractéristiques les plus connues de la diachronie de l'idiome.

La grande variabilité des données que j'ai essayé de mettre en évidence, que doit-on en penser eu égard aux processus par lesquels les mots se sont imposés aux sujets parlants ? Autrement dit, qu'en est-il réellement du langage face aux modes de vie, aux pratiques, aux objets ? La réponse à cette question en apparence banale n'est peut être pas aussi facile à donner qu'il y paraît. En cette affaire, je dirai d'abord que dans l'évocation des manières de travailler, d'user des instruments, rien n'est indifférent, rien n'est un détail : qu'il s'agisse des habitudes prises dans la façon d'amasser une récolte, ou des pièces d'un outil, tous les aspects de l'action ou toutes les composantes de l'engin ont une égale importance. J'en appellerai sur ce point à un exemple connu mais toujours porteur de beaucoup de sens. Depuis le livre, devenu un classique, de André G. Haudricourt et de Mariel Jean-Brunhes Delamarre, *L'Homme et la charrue*, on sait que le passage de l'araire à la charrue a été un processus dont l'histoire est très complexe : ce processus, tout en obéissant à une commune nécessité, a pris des formes diverses selon les parties du monde. Pour y voir clair, ces auteurs ont été, en ce qui concerne la zone de la planète qui est la nôtre, dans l'obligation de se livrer à une relecture minutieuse des sources latines, notamment Pline, *Histoire naturelle*, XVIII, 48, où l'on lit à propos de la pièce appelée en latin *vomer* (cf. gascon *vómet* ['bumet]) les explications que voici :

*Vomerum plura genera : culter vocatur, prædensam, priusquam proscindatur terram secans, futurisque sulcis vestigia præscribens incusuris, quas resupinus in arando mordeat vomer.*

« De plusieurs sortes sont les socs : on appelle coudre (*culter*) [celui] qui entaille la terre très dure avant qu'elle ne soit ouverte et, par ses incisions, impose leur tracé aux futures raies, afin que le soc, incliné et agissant à l'arrière, ouvre celles-ci en labourant. »

De cette explication de Pline, plusieurs auteurs ont conclu que dans le matériel des Romains *culter* « le coudre » et *vomer* « le soc » faisaient partie d'un seul et même instrument. André G. Haudricourt et Jean-Brunhes Delamarre faisaient observer à juste titre que cette manière de voir était en contradiction avec les données iconographiques, celles-ci ayant en commun de représenter les araires démunis de coudre : de telle façon que, même si le début de la description de Pline, *vomerum plura genera*, est susceptible de nous inviter à penser le contraire, il faut admettre que l'auteur parlait en fait

de deux outils distincts, d'une part un instrument consistant pour l'essentiel en un timon avec manche et muni d'une lame verticale agissant comme un couteau, d'autre part un araire avec un soc monté sur un sep. Il se trouve que le premier de ces outils a été utilisé dans les Pyrénées occidentales, où il portait des noms tels que *nabarr* en basque, ou *coudre* en gascon, *cuitre* en aragonais, etc. : on s'en servait avant de recourir à l'araire.

Par conséquent, la terminologie de l'outillage reste un domaine dans lequel la précision quant à la valeur des mots doit toujours mobiliser l'attention de l'observateur. Je viens d'en avoir encore la preuve en travaillant tout récemment et en compagnie d'Éliane Mau à la traduction en français et à la présentation aux lecteurs, à commencer par ceux de la Bigorre, du beau travail de Lotte Paret, *Das ländliche Leben einer Gemeinde der Hautes-Pyrénées dargestellt auf Grund der mundartlichen Terminologie* « La vie rurale d'une commune des Hautes-Pyrénées d'après la terminologie de son dialecte local », l'une des premières thèse dirigées par Gerhard Rohlfs en 1932 à l'Université de Tübingen<sup>4</sup>. Dans la partie de l'ouvrage consacrée aux instruments aratoires, Lotte Paret traite précisément de la *coutría*, prenant soin de bien la distinguer de l'araire. Mais dans son commentaire sur le second de ces outils, elle nomme la pièce appelée *espada* (littéralement « épée »), *Griessäule*, vocable qui peut être rendu par « étançon », ajoutant « die Sech [seç] und Krummholz verbindet » (*relie le coutre et l'age*). En fait, l'excellente et incomparable observatrice qu'était Lotte Paret a été ici la victime d'une petite illusion d'optique : en réalité, l'éstançon ne relie pas le soc et l'age, mais est un régulateur agissant directement sur l'incidence du soc et aboutissant de ce fait au sep de l'araire, si bien que l'ALG le définit comme un « gouvernail de profondeur ». Il n'est cependant pas exclu que l'inadvertance ait joué ici, entraînant la confusion entre deux objets et donc deux termes : Lotte Paret n'aurait-elle pas écrit par erreur *Sech* [seç] « coutre » en lieu et place de *Schar* [ʃɑ:r] « soc » ? Nul n'est à l'abri de cette sorte d'accident. Quoi qu'il en soit, et je ne

---

<sup>4</sup> Lotte Paret, *Das ländliche Leben einer Gemeinde der Hautes-Pyrénées (Arrens), dargestellt auf Grund der mundartlichen Terminologie*, Gedruckt mit Genehmigung der philosophischen Fakultät der Universität Tübingen. Referent : Prof. Dr. G. Rohlfs, 23 juli 1932.

Édition en français : *Arrens 1930. Les mots et les choses*. Lotte Paret, *La vie rurale d'une commune des Hautes-Pyrénées décrite d'après le vocabulaire du dialecte local*, traduit de l'allemand par Éliane Mau et Xavier Ravier, avant-propos et compléments par Xavier Ravier, iconographie par Bernard Ménétrier, Société d'Études des Sept Vallées du Lavedan et Association Guillaume Mauran, Arrens, Argelès-Gazost et Tarbes, 2008.

crains pas de me répéter, rien n'est indifférent dans la description d'un outil : une pièce dormante doit jouir de la même considération qu'une pièce travaillante. Et ce qui est vrai de l'outillage, l'est aussi des pratiques elles-mêmes.

Dans les formes anciennes de l'agriculture, la fenaison constituait une étape d'autant plus importante qu'elle demandait une dépense de temps et de main-d'œuvre comparable à celle qu'exigeraient après elle la moisson et le battage. Par conséquent, la mise en ligne du foin sec était la dernière phase d'un travail continu, astreignant, et elle ne pouvait qu'être en quelque sorte valorisée par ceux qui y prenaient part. Dans ces conditions, il ne faut pas s'étonner que cette opération ait fait l'objet de la grande variété de désignations sur laquelle, précédemment, j'ai attiré l'attention : dans les localités ou les petites zones entretenant des relations entre elles, on a visiblement voulu assortir de dénominations particulières l'opération qui nous occupe. Et à cet égard, il est intéressant de noter que les noms portés par la mise en ligne du foin en vue de son chargement, même s'il est difficile d'en dresser la carte par aires dégagées, s'inscrivent tout de même dans un réseau micro-régional. On voit aussi que le symbolique lui-même a eu sa place en cette affaire : un terme tels que *crapaus* [kra`paws] « crapauds » du point 641 le montre bien ; dans le même ordre d'idées, on assiste à une interpénétration de plusieurs ordres de la réalité : la *canshè(i)ra* de divers points du sud du Gers et de l'est des Hautes-Pyrénées renvoie en fait à la lisière d'un champ ou à une planche de labour, la forme elle-même nous ramenant obligatoirement au radical du lat. CANCELLI « limites », CANCELLARE « délimiter », ce dernier attesté par les *Gromatici veteres* (traité des anciens arpenteurs)<sup>5</sup>.

La variabilité des données, je m'en étais occupé par ailleurs dans des travaux que j'avais publiés en 1983 et 1993 : que l'on me permette d'y revenir. Je l'avais déjà fait dans ma communication d'il y trois ans, présentée ici, à l'occasion du colloque auquel fait tout naturellement suite celui qui nous réunit aujourd'hui. La réflexion que je vais maintenant vous soumettre est évidemment susceptible d'être améliorée, voire mise en cause.

L'idée s'était donc autrefois imposée à moi que dans des investigations comme les nôtres, en ce qu'elles sont étroitement liées au travail de terrain, il faut se résoudre à admettre que les concepts de

---

<sup>5</sup> *Gromatici veteres* (traité des anciens arpenteurs), Lachmann, 1-2, 1848-1852.

*signifiant* et de *signifié* pris dans leur sens le plus saussurien ne suffisent pas à nos besoins ; en conséquence, je suggérais d'ajouter à ce couple célèbre de notions celui de *nommant* et de *nommé* : alors que la dualité *signifiant* / *signifié* possède le statut d'idéalités, dans le sens philosophique du terme, relève en quelque sorte d'un ordre transcendantal, je proposais que l'on usât de *nommé* et de *nommant* pour la relation entre le mot et l'objet qu'il désigne, l'objet pris dans sa particularité. Ainsi, l'étude de la carte des noms de la ruche de mon atlas languedocien occidental m'avait donné l'occasion d'expérimenter ce point de vue et de réfléchir au rapport que, dans une bonne partie du domaine, entretiennent le désignatif générique correspondant au signifié « ruche » et les désignatifs directement inspiré par la forme locale de la ruche, donc autant de *nommants* pour des nommés *spécifiques*, le tout restant d'une parfaite clarté pour les sujets parlants<sup>6</sup>. Autrement dit, le *nommant* et le *nommé* appartiennent à notre vie dans son effectivité, sa quotidienneté, ils se situent donc, du point de vue de la pensée, à l'opposé de l'*eðdoç* platonicien.

Nous venons de voir aujourd'hui que le *nommant*, quand il se rapporte à un outil ou à une pratique, est caractérisé par l'ajustement du lexique à l'objet désigné, y compris dans ses parties : les *nommants* fonctionnent comme en total homomorphisme avec les *nommés*. Une telle situation signifie donc leur inséparabilité. Pour dire les choses autrement, l'essence d'un instrument ou d'une pratique est avant tout la même chose que son existence : cette identité entre essence et existence, s'il s'agit d'un outil, correspond à l'usage qui est fait de lui, renvoie à ce à quoi il sert – et, ce qui revient au même, pour une pratique cette même identité équivaut à une action nécessaire, inévitable, entièrement aboutie. C'est à ce titre qu'un lexique du type de celui dont il vient d'être question doit, selon moi, être évalué à la lumière de la dichotomie *nommé* / *nommant*, sans quoi nous ne nous trouverions que dans le seul monde des choses en soi, peut-être inévitable pour une démarche purement philosophique mais hors de question quand on s'intéresse à la multiplicité des objets ou des façons de faire dont notre espèce a dû se doter. Ce point de vue, vous le savez tous, était à la base de l'une des traditions qui a marqué notre

---

<sup>6</sup> Xavier Ravier, « Atlas linguistiques, ethnographie et ethno-linguistique », *Bollettino dell'Atlante linguistico mediterraneo, Studi in memoria di Carlo Battisti*, Fondazione Giorgio Cini, Pise, 1996, n° 29-35, pp. 17-42. Du même auteur : « Des choses aux mots, des mots au discours : pour une utilisation maximale des données des atlas linguistiques et ethnographiques », *Le Français moderne*, 65<sup>ème</sup> année, n° 1, 1997, pp. 20-34

discipline, celle des *Wörter und Sachen* : si on l'admet, on comprend du même coup que le mouvement scientifique en question ne se réduisait pas à mettre face à face mots et choses, mais cherchait en réalité à discerner par quels chemins, souvent difficiles à déceler, ceux-ci et celles-là s'étaient rencontrés dans l'histoire langagière : Schuchardt a laissé des réflexions capitales à ce sujet. Les outils et ce pour quoi nous les utilisons étaient inscrits dans le destin de l'homme et ont constitué sa première noblesse, même si nos lointains prédécesseurs ont aussi conçu des engins destinés à donner la mort à leurs semblables. De toutes les façons, l'homme ne serait pas l'homme sans l'outil et l'outil ne serait pas sans le langage qui le nomme : pas plus d'ailleurs que ne serait le langage sans l'outil dans sa matérialité.

### **Complément : les données de la carte ALG, IV, 1115 Rouleau de foin.**

Le traitement des données est pour l'essentiel opéré en allant du nord vers le sud et par tranches d'ouest en est (n° des points : ceux de l'ALF constituent la base du réseau et figurent uniquement en chiffres) ; les points qui ont été ajoutés sont situés par rapport à ceux de l'ALF au moyen des lettres N (nord), S (sud), O (ouest), E (est), utilisées seules ou en combinaison. La situation des deux points du Val d'Aran est indiquée par rapport au point 699 (Localité).

Les formes sont données à l'entrée de chaque article en orthographe classique occitane. Le genre est indiqué par m. = masculin, f. = féminin, le nombre, quand cela est nécessaire, par sg. = singulier, pl. = pluriel. On a transcrit en API toutes les formes publiées : ce système ne permettant pas la notation des voyelles d'aperture moyennes en tant que telle, les signes [o], [œ] pour les finales féminines atones doivent être lus « o moyen, œ moyen » : la réalisation nettement ouverte ou fermée des phonèmes occupant la position en question est indiquée par API [ɔ], [ø], la non-utilisation de ces signes correspondant, comme on vient de le voir, à une réalisation moyenne. En revanche, à l'intérieur des formes et dans quelques cas en finale sont mis à contribution [a], [ɑ], [ɔ], [o], [œ], [ø] avec leurs valeurs normales. Les situations particulières font l'objet de remarques.

La question, de type indirect, comportait dans sa formulation l'expression « en longs tas allongés », s'agissant précisément du

conditionnement du foin sec en vue de son chargement : les informateurs ont répondu par le singulier ou le pluriel. Pour des raisons de facilité, les formes obtenues ont toutes été ramenées au singulier dans la liste ci-après, à l'exception de la donnée du point 679 SO : [hur'kaðos], terme dont la valeur singulière est « fourchée » : on a pensé que le recours au pl. était motivé à ce point par le fait que les locuteurs locaux usent du nombre en question en lui donnant une valeur collective (*les fourchées* = la rangée de foin telle qu'elle résulte du maniement de la fourche répété de multiples fois) ; en d'autres termes, nous avons considéré [hur'kaðos] comme un *pluralia tantum*. Il n'est pas impensable que plusieurs des réponses au pl. obtenues d'un nombre non négligeable d'informateurs relèvent aussi de la catégorie des *pluralia tantum* : mais une vérification n'était guère envisageable.

### **Zone a : départements 33, 40, 47, 82**

- ando** f. [ãndu] 548 (la finale f. atone [u] se retrouve en Ariège).  
**aranda** f. [a`rãndo] 549  
**dorga** f. [durgœ] 650, 650 E  
**randa** f. [rrãndœ] 549 N, 650 N [-œs], 645 NO [rãndø], 641 O  
**renda** f. [rrëndø] 634 NO, 643 NE, 643  
**rinda** f. [rĩndø] 630 S  
**crapaus** m. [kra`paws] 641  
**tèira, tièra** f. a) : [tɛjrø] 643 NO, 643 NE ; b) [tjɛro] : 659 NO [tjɛro] (finale entre [o] et [œ] moyens ; dans le début de la forme : hésitation entre [tjɛro] et [ti`jɛroɛ]), 659, 760 N [tjɛros].  
**ranhièira** f. [rə`njɛjrœ, rɑ`njɛjrœ] 653 N  
**rima** f. [rrimø] 681 N, 645 NE, 647, [rrimo] 647 NO  
**endalha** f. [ɛn`daλø, ɛn`dajø] (662)  
**renga** f. [rrɛŋgø] 645 (synonyme de **coa**, recueilli aussi au même point).  
**rua** f. [rryo] 657 NE, 657, 656 [rrywø], 782 S  
**ruia** f. [rryjo] 648

**arrua** f. [a`rrywø] 656 SO, 689 SE [a`rryo] (voyelle finale entre [o] et [œ]), 699 NO [arry`a] (sic pour l'accent tonique ; la forme est féminine), 698 E [-`rryo] et [-`rryɔ], 699.

**nai** : a) Gironde : 645 S, 653, 653 O, 662, 664 N b) Landes : 664, 672, 672 NO, 674, 674 N, 674 O (à ce point a été relevé le verbe **ardolar** [ardu`la] « mettre le foin en rouleaux »). **codrilha** f. [ku`ðriλø] 664 (donné au pl. : [ku`ðriλøs])

**barat** m. [ba`rat] 680, 682 N

**arronça** f. [a`rrũnsø] 681, 681 S, 681 SE, 683, 683 E, 690, 690 E, 692 S

**arrú** [a`rry] m. 665, 665 S, 675 N, 675, 675 O

**canon** m. [ka`nũŋ] 680 S, 682, 683 N.

**arrolha** f. [a`rruλø] 684, 693 [-λo]

**còlina, colina** f. [kø`linø, ku-] 676 O

**cordiòla** f. [kur`djø] 676 O

**meda** f. [`medø] 676 NO, 685 NE

**iròu** m. [i`row] 691 NE ([o] moyen)

### Zone b : départements 64, 65, 32, 31, 09

**redòla** f. [rre`ðo] 657 S, 667

**traça** f. [`trasa] 676 N (donné au pl. : [`trasos])

**còrda** f. [`kørdø] 658, 658 NO, 668, 659 SE

**ruat** m. [rry`at]

**arruga** f. [a`rryγo] 676 (et aussi [a`rryo])

**aró** m [a`ro] m. 676 SO, 686 NE, 687 N [a`rɔ], 694

**aróu** m. [a`row] 686, 686 S, 696 O

**naièra** f. [na`jεø] 685

**naiòla** [na`jø] 685 SE **gabèra** f. [ga`βεro] 685 SO

**nai** m. [naj] 691, 691 O, 692 NE (également relevé *anajar* [ana`žã] « mettre le foin en rouleau »).

**aròla** f. 687 [a`rø] [a`rø] 687 NO, 687 E (et aussi [a`røro] à ce point : le traitement rhotacisant est normal dans la zone à laquelle appartient la localité enquêtée).

**nagía** f. [na`zjo] 693 NO

**lugana** f. [lygana] 693 NE



- artòrcla** f. [ar`tòrkla] 695 O  
**tòrcla** f. [`tòrklo] 698  
**coròla** f. [ku`ròla] 695  
**arredòla** f. [arre`ðolo] 697  
**taulèra** f. [taw`lɛrɔ] 696 S  
**estesa** f. [es`tezα] 696  
**canquèira** f. [kãŋ`kejro] 668 SO (avec palatalisation faible du second [k]).  
**canshèira** f. [kaŋ`ʃejro] ou **cansheira** [kaŋ`ʃejro] : 678<sup>pl</sup>, 678 NO<sup>pl</sup>, 678 E, 688 N<sup>pl</sup> (pluriel aux points dont l'indicatif est suivi de pl. en exposant).  
**canshera** f. [kãŋ`ʃero] 689 N [kãŋ-] ([ŋ] : nasillement) ; 688 O et 688 [kã`ʃero]  
**shera** f. [ʃero] 688 E, 780 NO  
**còrda** f. [`kòrdo] 658 NO, 659 SE (et aussi type **anda**), 760 N, 771 E ; donné au pl. [`kòrðos] à 659 SE (et aussi [ã`ndos], 760 SE, 762 NE ([o] nettement fermé de la syllabe finale  
**cordada** f. [kur`daðo] 679 SO  
**horcadas** f. [hur`kaðos] 679 SO (*pluralia tantum*).  
**ruèla** f. [rɾy`elo] 780 (et aussi **orde**)  
**orde** f. [urde] 780 (et aussi **ruèla**), 780 S [-de]  
**umè** m. [y`mɛ] 699 SE  
**arreg** m. [a`rrɛŋk] 790 NO (forme suggérée par l'enquêteur et répétée sans commentaire par l'informateur).  
**reg** m. [rrɛŋ] 790 NE (le mot désigne aussi l'andain), 787 S [rrɛŋk] (et aussi **rua**).  
**rudèras** f. [rɾy`ðeros] 790, 790 S (termes donnés au pl. par les informateurs de ces deux localités)  
**selhon** m. [se`lu] 791 O  
**airòu** m. [aj`row] 781 NO ([o] moyen)  
**gira** f. [zìros] 781  
**gorrudas** f. [gu`rɾydes] 699 NE.  
**anda** [ã`ndo] f. ([o] final moyen sauf cas particulier) 668 SE, 669, 669 NO, 669 E, 679, 679 E ([o] très fermé ; rouleau de foin et andain ; le mot se trouve aussi sur la carte ALG, II), 344, 760, 771 NO ; à 679 et 771 NO le mot **anda** désigne à la fois le rouleau de foin et l'andain : cela tient au fait qu'ils ne sont pas distingués l'un de l'autre.

**andanha** f. [ãn`dapɔ] 689, 696 E = A [-`dαɾɔ]

## Ordre alphabétique des formes

**airòu** m. [aj`row] 781 NO ([o] moyen)

**anda** [ãndo] f. ([o] final moyen sauf cas particulier) 668 SE, 669, 669 NO, 669 E, 679, 679 E, ([o] très fermé ; rouleau de foin et andain ; le mot se trouve aussi sur la carte ALG, II), 344, 760, 771 NO ; à 679 et 771 NO le mot **anda** désigne à la fois le rouleau de foin et l'andain : cela tient au fait qu'ils ne sont pas distingués l'un de l'autre.

**andanha** f. [ãn`dapɔ] 689, 696 E = A [-`dαɾɔ]

**ando** f. [ãndu] 548 (la finale f. atone [u] se retrouve en Ariège).

**aranda** f. [a`rãndo] 549

**aró** m [a`ro] m. 676 SO, 686 NE, 687 N [a`rɔ], 694

**aròla** f. 687 [a`rɔlœ], [a`rɔlo] 687 NO, 687 E (et aussi [a`rɔro] à ce point : le traitement rhotacisant est normal dans la zone à laquelle appartient la localité enquêtée).

**aróu** m. [a`row] 686, 686 S, 696 O

**arredòla** f. [arre`ðɔlo] 697

**arreng** m. [a`rrēŋk] 790 NO (forme suggérée par l'enquêteur et répétée sans commentaire par l'informateur).

**arrolha** f. [a`rruλø] 684, 693 [-λɔ]

**arronça** f. [a`rrũnsø] 681, 681 S, 681 SE, 683, 683 E, 690, 690 E, 692 S

**arrú** [a`rry] m. 665, 665 S, 675 N, 675, 675 O

**arrua** f. [a`rrywø] 656 SO, 689 SE [a`rryɔ] (voyelle finale entre [o] et [œ]), 699 NO [arry`a] (sic pour l'accent tonique ; la forme est féminine), 698 E [-`rryɔ] et [-`rryɔ], 699.

**arruga** f. [a`rryɔ] 676 (et aussi [a`rryɔ])

**artòrcla** f. [ar`tɔrkɫα] 695 O

**barat** m. [ba`rat] 680, 682 N

**canon** m. [ka`nũŋ] 680 S, 682, 683 N.

**canquèira** f. [kãŋ`kejro] 668 SO (avec palatalisation faible du second [k]).

**canshèira** f. [kaŋ`ʃejro] ou **cansheira** [kaŋ`ʃejro] : 678<sup>pl</sup>, 678 NO<sup>pl</sup>, 678 E, 688 N<sup>pl</sup> (pluriel aux points dont l'indicatif est suivi de pl. en exposant).

**canshera** f. [kãn`ʃero] 689 N [kãŋ-] ([ŋ] : nasillement) ; 688 O et 688 [kã`ʃero]

**còlina, colina** f. [kɔ`linø, ku-] 676 O

**còrda** f. [`kɔrdo] 658 NO, 659 SE (et aussi type **anda**), 760 N, 771 E ; donné au pl. [`kɔrdos] à 659 SE (et aussi [ãndos], 760 SE, 762 NE ([o] nettement fermé de la syllabe finale).

**còrda** f. [`kɔrdo] 658, 658 NO, 668, 659 SE

**cordada** f. [kur`daðo] 679 SO

**cordiòla** f. [kur`djɔlø] 676 O

**coròla** f. [ku`rɔlɔ] 695

**crapaus** m. [kra`paws] 641

**dorga** f. [`durgœ] 650, 650 E

**endalha** f. [ɛn`daλø, ɛn`dajø] (662)

**estesa** f. [es`tezɔ] 696

**gira** f. [`ziros] 781

**gorrudas** f. [gu`rrydes] 699 NE.

**horcadas** f. [hur`kaðos] 679 SO (*pluralia tantum*).

**iròu** m. [i`row] 691 NE ( [o] moyen)

**lugana** f. [lygana] 693 NE

**meda** f. [`medø] 676 NO, 685 NE

**nagía** f. [na`zio] 693 NO

**nai** : a) Gironde : 645 S, 653, 653 O, 662, 664 N b) Landes : 664, 672, 672 NO, 674, 674 N, 674 O (à ce point a été relevé le verbe **ardolar** [ardu`la] « mettre le foin en rouleaux »). **codrilha** f. [ku`ðriλø] 664 (donné au pl. : [ku`ðriλø]).

**nai** m. [naj] 691, 691 O, 692 NE (également relevé *anajar* [ana`ʒa] « mettre le foin en rouleau »).

**naièra** f. [na`jerø] 685

**naiòla** [na`jɔlo] 685 SE **gabèra** f. [ga`βero] 685 SO

**orde** f. [`urde] 780 (et aussi **ruèla**), 780 S [-de]

**randa** f. [`rãndœ] 549 N, 650 N [-œs], 645 NO [`rãndø], 641 O

**ranhièira** f. [rã`njɛjrœ, rã`njɛjrœ] 653 N

**redòla** f. [rre`ðɔlo] 657 S, 667

**renda** f. [`rrëndø] 634 NO, 643 NE, 643

- reng** m. [rrɛ̃n] 790 NE (le mot désigne aussi l'andain), 787 S [rrɛ̃ŋk] (et aussi **rua**).
- renga** f. [ʀrɛ̃ŋgø] 645 (synonyme de **coa**, recueilli aussi au même point).
- rima** f. [ʀrimø] 681 N, 645 NE, 647, [ʀrimo] 647 NO
- rinda** f. [ʀĩndø] 630 S
- rua** f. [ʀryo] 657 NE, 657, 656 [ʀrywø], 782 S
- ruat** m. [ʀry`at]
- rudèras** f. [ʀry`ðeros] 790, 790 S (termes donnés au pl. par les informateurs de ces deux localités).
- ruèla** f. [ʀry`elo] 780 (et aussi **orde**).
- ruia** f. [ʀryjo] 648
- selhon** m. [se`lu] 791 O
- shera** f. [ʃero] 688 E, 780 NO
- taulèra** f. [taw`lɛrɔ] 696 S
- tèira, tièra** f. a) : [tɛjrø] 643 NO, 643 NE ; b) [tjɛrø] : 659 NO [tjɛrø] (finale entre [o] et [œ] moyens ; dans le début de la frme : hésitation entre [tjɛrø] et [ti`jɛrøɛ]), 659, 760 N [tjɛros].
- tòrcla** f. [tòrklo] 698
- traça** f. [trasa] 676 N (donné au pl. : [trasos]).
- umè** m. [y`mɛ] 699 SE

*Jean-Pierre Lai & Michel Contini*  
*Université de Grenoble, GIPSA-lab. UMR 5216*  
*jean-pierre.lai@gipsa-lab.grenoble-inp.fr*  
*michel.contini@gipsa-lab.grenoble-inp.fr*

## **Nouvelles stratégies de recueil de données sur le terrain, exemples à travers le projet AMPER (Atlas Multimédia Prosodique de l'Espace Roman)**

### **Introduction**

Aujourd'hui encore, ce qui demeure une spécificité du dialectologue c'est sa volonté, son besoin et la nécessité qu'il éprouve de travailler sur le terrain. L'Atlas Multimédia Prosodique de l'Espace Roman créé et lancé au tout début des années 2000, en est l'exemple parfait<sup>1</sup>.

De manière conventionnelle, les techniques d'enquêtes sur le terrain consistent le plus souvent à demander à un informateur de répondre à une série de questions posées de manière indirecte. Avec ce nouvel atlas, très particulier, dans la mesure où il ne s'inscrit pas dans la lignée des atlas linguistiques habituels, nous avons dû utiliser un protocole d'enquêtes tout aussi remarquable. C'est aussi la raison pour laquelle, non pas un simple corpus mais plusieurs *corpora* sont nécessaires au recueil de données qui permettront l'analyse intonative d'une variété dialectale romane.

---

<sup>1</sup> L'idée de cet atlas a été proposée dans les années 90, avant de se concrétiser.

## 1. Présentation du projet

La prise de conscience d'une absence d'études comparées d'un fait linguistique majeur comme l'intonation est aussi à l'origine du projet AMPER, qui a vu converger les compétences des dialectologues et des phonéticiens. La présence de ces derniers dans le projet est inévitable car si le paramètre 'prosodie' n'avait pas encore été considéré par les géolinguistes jusqu'à présent, cela était dû au manque de connaissance approfondie de la phonétique acoustique de leur part et à une formation quasi inexistante à l'approche instrumentale, indispensable à l'analyse et à la quantification d'une réalité multiparamétrique.

Aujourd'hui, l'atlas prosodique roman est un projet immense dont l'ampleur ne cesse de croître, d'ailleurs pour avoir d'avantages d'informations sur ce projet, nous vous renvoyons aux articles déjà nombreux<sup>2</sup>. Très brièvement, cet Atlas est une Base de Données<sup>3</sup> qui met à la disposition des membres du projet et aussi du grand public toutes les données recueillies, permettant un accès à une comparaison simple (deux variétés dialectales) ou multiple (jusqu'à quatre dialectes) des intonations des différents parlers romans.

L'intérêt principal réside dans la mise en relief de la variabilité géoprosodique dans l'espace dialectal et dans les variétés régionales des langues nationales. De plus, il est possible de visualiser et d'entendre les schémas mélodiques ainsi que le rythme émergent de toutes ces variétés romanes, avec une étude des segments porteurs des informations prosodiques.

Les paramètres physiques pris en compte pour cette recherche sont la durée vocalique, la fréquence et l'intensité ; il est possible d'établir un classement de ces derniers en fonction de leur degré d'importance dans la phrase.

Aujourd'hui, près de dix ans après les balbutiements du projet, on peut réellement parler d'un véritable réseau de collaboration. En effet, ce ne sont pas moins de 33 universités ou organismes internationaux qui participent à ce projet, dont 14 proviennent d'Amérique du Sud ; le dernier pays entrant dans le projet est le Costa Rica, faisant suite au Brésil, à l'Argentine, au Chili, à Cuba et au Venezuela. Les pays

---

<sup>2</sup> Lai, Romano, Rouillet (1997) ; Contini *et al.* (2002) ; Lai & Romano (2002). Voici également le site présentant le projet : <http://w3.u-grenoble3.fr/dialecto/AMPER/amper.htm>

<sup>3</sup> Rilliard & Lai (2008). Site de la Base de Données AMPER : <http://amper.limsi.fr>

d'Europe membres du projet sont le Portugal, l'Espagne, l'Italie, la Roumanie et la France. Bien entendu, nous parlons plutôt de domaines linguistiques que de pays (ou de langues nationales) à proprement parler : nous avons donc le gallo-roman, le catalan, le galicien, ou encore l'italo-roman pour n'en citer que quelques-uns. Les membres associés au projet, quant à eux, sont déjà au nombre de 115 (professeurs, chercheurs, doctorants).

À ce jour, ce sont plus de 200 articles qui ont été publiés dans des revues internationales ou de nombreux actes de colloques ; pour finir, des réunions annuelles nous permettent d'apprécier l'avancement de nos travaux<sup>4</sup>. En outre, des thèses de doctorat ont été entièrement consacrées au projet AMPER (déjà trois à Grenoble<sup>5</sup> et plusieurs en Europe).

Depuis l'origine du projet nous avons élaboré une méthodologie commune de recueil de données et d'analyse de ces dernières<sup>6</sup>. C'est aussi à ce niveau que résident l'originalité et la force du projet. Une méthodologie rigoureuse est nécessaire si l'on veut comparer par la suite des schémas prosodiques forcément différents mais construits suivant un modèle commun et unique. N'oublions pas que depuis l'origine des études prosodiques (acoustiques), il y a environ un siècle, aucune mesure de comparaison n'est objectivement possible car toutes utilisent une méthodologie différente.

La première condition permettant une confrontation de courbes intonatives réside dans l'utilisation de *corpora* semblables au niveau de la structure phrastique mais aussi dans le maniement de différents types de *corpora*. Si les conditions existent, on fera appel à deux informateurs de chaque sexe par point d'enquête et, dans la mesure du possible, issus du milieu urbain et rural<sup>7</sup>. Nous essayons d'avoir une couverture assez représentative des domaines linguistiques étudiés. Enfin, l'analyse instrumentale se plie à toute une série de programmes écrits<sup>8</sup> à cet effet et spécifiques à notre projet.

---

<sup>4</sup> Plusieurs congrès internationaux ont dédié une journée au projet AMPER : *Congreso de Fonética Experimental* (Santiago de Compostela, 2005 et Granada, 2008), *Congreso Internacional de Americanistas* (Mexico, juillet 2009). Et plusieurs colloques ont déjà été organisés : à Grenoble (2001 et 2004), *Jornadas Científicas* (Santiago, 2003, Ténérife-La Laguna, 2006, Aveiro, 2007).

<sup>5</sup> Romano (1999), Roullet (1999), Lai (2002).

<sup>6</sup> Romano, Lai & Roullet (2005).

<sup>7</sup> Voir Pop & Pop (1959, p. 77) pour les critères de sélection des informateurs, ou encore pour Endo & Bertinetto (1997, p. 30).

<sup>8</sup> Programmes créés par Antonio Romano (co-directeur du projet de l'Université de Turin), ils s'inscrivent dans un environnement Matlab.

## 2. Les *corpora* liés au projet

Les *corpora* incontournables sont au nombre de quatre. Tout d'abord le corpus fixe : il est *ad hoc* et oral car son but est d'obtenir une segmentation plus aisée des voyelles (lieu privilégié de l'information prosodique) mais aussi de contrôler la qualité de production des locuteurs. Il est à la base de toutes nos analyses et obéit à plusieurs contraintes phonétiques, comme l'utilisation d'un maximum d'occlusives sourdes (idéales pour faciliter le découpage vocalique), chaque fois que cela est envisageable, mais le plus souvent évidemment et faute de temps aussi, il faut accepter ce que l'informateur nous donne.

La structure type de la phrase est SVO (Sujet-Verbe-Objet) : bien que présente dans un maximum de parlers romans, elle n'est pas la plus répandue mais elle existe et permet donc toutes les comparaisons possibles. C'est le cas en sarde où la structure syntaxique la plus employée pour l'interrogation est *a lu toca su beccu su cavaddu?*<sup>9</sup> (*le bouc touche-t-il le cheval?*), avec l'élément « a » qui introduit déjà l'interrogation ; cependant, la phrase *su beccu toca su cavaddu?* est aussi parfaitement acceptable.

L'objet de ce choix délibéré réside dans la possibilité de permuter la position syntaxique de S à la place de O et vice-versa, avec, pour chaque élément de la phrase, l'emploi de tous les types accentuels de la variété étudiée (oxyton, paroxyton, proparoxyton), afin de pouvoir identifier le siège de l'énoncé où intervient la discrimination prosodique. Tout l'intérêt sera également d'observer si l'ajout d'un adjectif<sup>10</sup> au S puis à l'O apportera un quelconque changement intonatif (de fréquence, d'intensité ou de durée) à ces éléments. La totalité des combinaisons possibles, compte tenu des exigences, est de 45 phrases, avec les trois types accentuels habituellement observés.

Nous conservons trois bonnes répétitions de chaque phrase dans les modalités interrogatives et déclaratives<sup>11</sup>. Cela veut donc dire que pour présenter une phrase modèle, il faut opérer une moyenne de ces

---

<sup>9</sup> On trouvera aussi les structures syntaxiques suivantes : *su beccu // a lu toca su cavaddu ?* ; *a toca su cavaddu su beccu?* ; *su beccu // su cavaddu a lu tocada?* ; *su cavaddu // su beccu // a lu tocada?* Voir aussi Contini (1976).

<sup>10</sup> Cf. Bruce (1982).

<sup>11</sup> Sous entendu que pour atteindre les trois répétitions jugées excellentes d'un point de vue du respect de l'intonation, il faudra faire répéter plusieurs fois les énoncés, si le besoin s'en fait sentir.



trois répétitions ; ces dernières sont nécessaires à des fins statistiques<sup>12</sup> et de plus l'utilisation d'une seule répétition ne peut représenter la modalité absolue, ce qui entraîne un problème de normalisation (Bellet *et al.*, 1981), et donc la sélection d'une référence. La première solution est phonétique : on parlera de normalisation temporelle de la fréquence individuelle en donnant un contour-type de F0<sup>13</sup>.

Nous envisageons par la suite des tests de validité de la perception. Ils ont essentiellement deux buts : d'abord évaluer la naturalité de la production (contrôle proprioceptif ou exercé par une tierce personne)<sup>14</sup>, ensuite tenter de juger jusqu'à quel point des énoncés désémantisés [c'est-à-dire qu'après l'effacement du contenu sémantique, seules les valeurs de durées, de fréquence et d'énergie persistent] sont reconnus par les locuteurs locaux comme étant leur dialecte, voire celui d'un village voisin. Par la suite, il sera intéressant de constater dans quelle mesure un des paramètres est plus discriminant dans cette reconnaissance dialectale<sup>15</sup>.

Le choix d'un tel corpus fixe s'explique encore par le fait que nous refusons tout corpus de base pris sur le vif, à cause, fondamentalement, de la mauvaise qualité des données et donc du manque de fiabilité. Il répond, d'autre part, aux propriétés souhaitées comme l'acquisition d'un son de qualité, l'absence de bruit de fond ou encore de nuances émotives très marquées.

Évidemment il devient facile de soulever la contradiction engendrée par cette méthode : en effet, comment prétendre obtenir des énoncés naturels, spontanés, alors que nous exigeons des répétitions de phrases préétablies ? Néanmoins, nous revendiquons cette pratique car malheureusement les études prosodiques 'naturelles' ne nous permettent aucune analyse fine<sup>16</sup> et surtout elles échappent à toute méthodologie unitaire qui déboucherait sur une série de faisceaux paramétriques comparables, ce que nous réalisons grâce à ce projet et qui lui donne justement toute sa force. C'est pourquoi un protocole d'enquête s'avère indispensable qui, une fois appliqué par tous les

---

<sup>12</sup> Cf. Boë, Contini & Rakotofiringa (1975).

<sup>13</sup> Cf. Contini & Boë (1973 ; 1975).

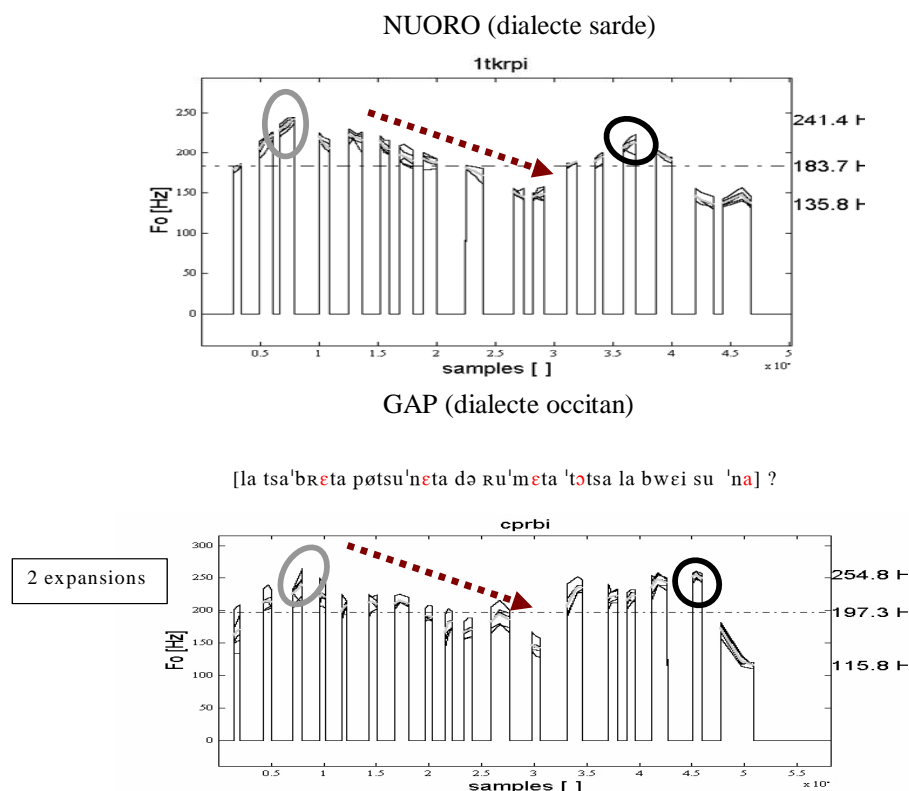
<sup>14</sup> Cf. Rossi *et al.* (1981) ; Park (1989), Lai (2002, pp. 155-161.)

<sup>15</sup> Quelques chercheurs se sont lancés dans cette quête, des manipulations des paramètres (allongement de la durée, augmentation de la fréquence...) permettant justement d'établir le critère prépondérant dans la reconnaissance d'un énoncé.

<sup>16</sup> Sans compter l'absence de structures phrastiques semblables, répétées et donc comparables, nécessaires à notre recherche.

collaborateurs du projet, permettra de pouvoir comparer les données recueillies.

Voici, par exemple, une comparaison entre deux variétés romanes : l'occitan de Gap (Lai, 2005) et le dialecte de Nuoro en Sardaigne (Lai, 2002).

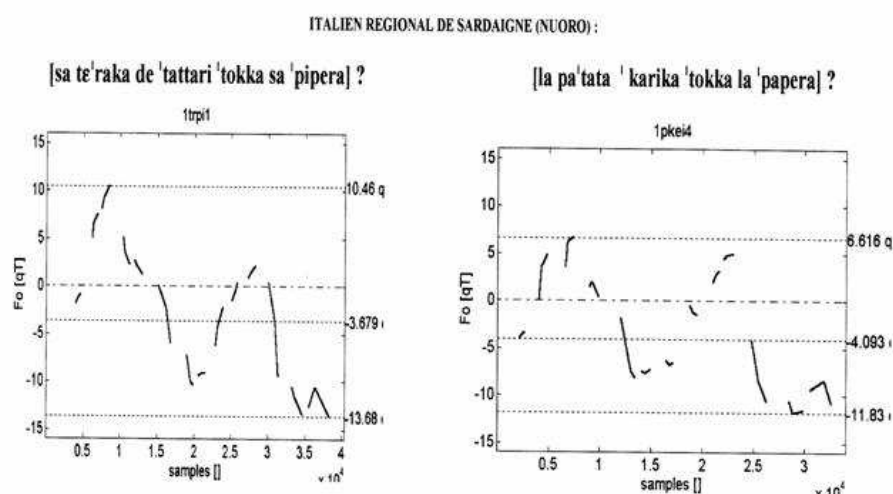


Ces deux dialectes n'ont, à priori, que très peu de rapports et pourtant on note que, dans certains cas, la stratégie intonative mise en œuvre dans les deux parlars est la même, aussi bien pour la partie initiale, lorsque le substantif est un paroxyton (cercle gris sur la tonique), que pour la fin de l'énoncé. Mais seul le pic de F0 dans le syntagme verbal (cercles en noir) a une correspondance exacte, dans la mesure où il tombe toujours sur la prétonique du dernier élément, quel que soit son type accentuel (proparoxyton pour le sarde avec le mot ['pipera] et oxyton pour le gapençais avec [bweisu'na]). Nous avons également pu noter qu'à l'intérieur d'une même aire dialectale, des stratégies intonatives complètement différentes peuvent être utilisées.

Par conséquent il serait très intéressant d'établir assez vite une typologie de ces tactiques prosodiques.

La poursuite de la procédure préconise l'usage de deux autres *corpora* : un corpus libre, comme un récit de vie, dans la variété enquêtée afin de constater si les structures mélodiques du corpus de laboratoire sont bien identiques ou du moins très proches de celles du discours libre<sup>17</sup>. De même un petit corpus fixe dans la variété régionale de la langue nationale nous est très utile (exemple : l'italien régional de Sardaigne) car il permettra ensuite d'évaluer la durabilité éventuelle des faits prosodiques dans ces deux modes expressifs.

Voici un exemple pour le sarde de Nuoro et l'italien régional de la même ville, dans lesquels on retrouve bien des schémas prosodiques identiques<sup>18</sup> : l'énoncé en sarde (graphique de gauche) fait apparaître les pics de F0 sur la tonique du substantif paroxyton et sur la prétonique de l'objet, exactement comme dans la phrase en italien régional, de structure syntaxique identique (à droite).



Pour finir, un dernier type de corpus semi-dirigé préconise le procédé appelé Map Task<sup>19</sup>. Le but est d'inciter l'informateur principal à décrire un parcours sur une carte composé de dessins, ces derniers représentant les mots du corpus fixe. En face de lui, un second informateur vérifiera le tracé du premier sur le sien, ce dernier

<sup>17</sup> En sachant qu'il sera nécessaire de trouver une structure syntaxique identique à celle de notre corpus fixe.

<sup>18</sup> On note toutefois une différence de hauteur.

<sup>19</sup> Cf. Savino & Refice (1997) ; Brown *et al.* (1983).

étant légèrement différent. Ce moyen permet de détourner l'attention des locuteurs qui oublient ainsi le microphone et le stress de l'enquête tout en apportant une meilleure spontanéité. Nous pourrions alors récupérer les données et les comparer avec celles du corpus fixe. Ceci nous amène à la conclusion qu'un système interactif comme le *Map Task* est le résultat, d'une part, de la volonté de recueillir de la parole spontanée (travail du dialectologue et du chercheur de terrain) et d'autre part, de la nécessité d'obtenir de la parole contrôlée incluant une certaine qualité technique, par exemple au niveau des paramètres acoustiques.

### 3. Les techniques d'élicitation des données

Le déroulement de l'enquête doit être soumis à une attention rigoureuse, en veillant également à la naturalité des énoncés prononcés. En outre, il est primordial de rejeter toute lecture continue du corpus car cette méthode, évidemment plus simple, dénature les structures intonatives d'une langue (cf. la longue littérature à ce propos, ce qui n'empêche pas certains auteurs de penser qu'il n'y aucune différence entre la lecture d'un texte et sa prononciation naturelle)<sup>20</sup>. Dans le même esprit, en Italie notamment, de nombreux travaux exploitent des journaux télévisés ou radiophoniques (donc très marqués prosodiquement) en guise de parlé spontané<sup>21</sup>. À ce propos, comme le souligne Guaitella (1992, p. 222), si ces deux formes orales sont gouvernées par des productions différentes, leurs systèmes intonatifs le sont également en conséquence. Elles se différencient aussi de par leurs pauses et leurs vitesses d'élocution<sup>22</sup>.

Traditionnellement, lors des enquêtes dialectales de terrain sur le lexique, il est d'usage de ne pas inciter directement l'informateur à fournir le nom de l'objet du corpus pour éviter une influence trop importante de la langue nationale dans ses réponses. Le risque de produire une forme dialectale altérée est élevé, c'est la raison pour laquelle il faut toujours opérer un questionnement indirect.

---

<sup>20</sup> Cf. Sabio (1996) ; I et J Fonagy (1983 p. 181).

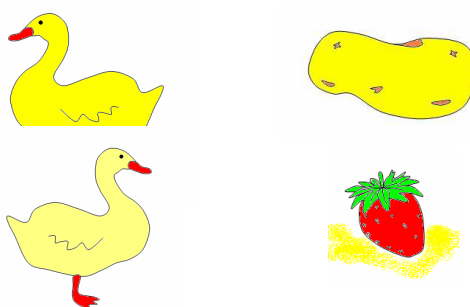
<sup>21</sup> Travaux de l'Université de Naples (Caputo, 1995). On peut voir ici une solution de facilité, il n'y a aucune enquête à mener sur le terrain.

<sup>22</sup> Cf. Pettorino & Giannini (1994).

Or, vu la spécificité de notre corpus, composé de plusieurs phrases d'une structure syntaxique particulière, cela est beaucoup plus compliqué et nous a obligé à opter pour les différentes stratégies suivantes. La première méthode testée, qui a donné ses fruits, consiste à présenter au locuteur des images ou des objets, dans un ordre précis, correspondant aux mots du corpus fixe ; on demande de prononcer les phrases en rapport avec ces illustrations et avec leur mise en scène. Évidemment les premiers essais ne sont pas toujours concluants, mais par la suite l'informateur mémorise les structures des phrases à prononcer.

Cette solution ne s'avère pas toujours efficace, mais elle assure une prononciation qui n'appartient qu'au locuteur et qui n'est pas soumise à une certaine influence de la part de l'enquêteur<sup>23</sup>.

Exemples tirés du questionnaire italien<sup>24</sup> (modalité déclarative et interrogative) :



La pàpera mangia la patata = <i>Le canard mange la patate</i>	La pàpera mangia la patata? = <i>Le canard mange(-t-il) la patate?</i>
La pàpera càtara mangia la fragola = <i>Le canard cathare mange la fraise</i>	La pàpera càtara mangia la fragola? = <i>Le canard cathare mange(-t-il) la fraise?</i>

<sup>23</sup> Évidemment, quelquefois le sens de l'énoncé est très difficile à traduire sous forme de dessins. De plus, une part d'emphase peut aussi apparaître ; par exemple, si on vient de montrer pour la 10<sup>e</sup> fois une patate et un canard qui se touchent, l'informateur, reconnaissant immédiatement la phrase à prononcer, pourra emphatiser sa prononciation et ajouter une émotion qui n'est pas nécessaire : **ha oui ! la patate touche le canard !**

<sup>24</sup> Les images et les énoncés sont issus du corpus d'italien régional du domaine italo-roman expérimenté par A. Romano.

Le contournement de l'élicitation directe - ce qui, dans notre cas particulier, revient à dire la lecture du corpus ou la répétition de notre propre phrase - nous a poussés à explorer de nouvelles techniques d'enquête. Ainsi, il existe une autre possibilité, celle de présenter le corpus sur un ton régulier et monotone afin que la répétition de l'informateur soit totalement dénuée de toute forme d'écho et donc d'imitation (Espuny, 1997).

Une autre pratique, adoptée lors de nos enquêtes en Sardaigne consiste à exiger de la part de l'informateur de reprendre en sarde tous les éléments que nous venons de prononcer en italien avec un ton neutre. Même si cela paraît difficile à admettre, car en principe, un locuteur devrait non seulement parler sa langue mais posséder tout pareillement sa propre intonation, le risque de le voir s'approprier l'intonation voisine - car plus prestigieuse - est bien réel<sup>25</sup>. D'où la nécessité de bien connaître le dialecte étudié ou de toujours avoir un second locuteur dialectophone qui contrôle la justesse des énoncés prononcés.

Il nous est malheureusement arrivé de trouver un informateur qui, au lieu de prononcer naturellement son corpus, tentait de retrouver une intonation authentique à ses yeux : *comme on prononçait autrefois*. C'est la raison pour laquelle il est important de réfléchir conjointement (entre locuteur, chercheur et autres intervenants) afin d'établir un degré satisfaisant de spontanéité. À ce propos, la littérature concernant le rapport entre l'informateur et le chercheur et surtout celui entre la méthode d'enquête et le résultat obtenu, est relativement pauvre<sup>26</sup>.

Une autre stratégie, ne fonctionnant que pour la modalité déclarative, consiste à présenter le corpus dans la modalité interrogative, en demandant au locuteur de reprendre tous les termes de la question dans sa réponse sur un ton déclaratif, obtenant ainsi cette modalité. Un certain contrôle doit s'exercer sur ces résultats car les phrases déclaratives peuvent recouvrir une forme emphatique.

Enfin, en cas extrême, si vraiment le locuteur ne parvient pas à suivre les consignes, il est possible de tenter de lui faire lire la phrase

---

<sup>25</sup> Ceci peut être la conséquence du manque de contextes des phrases proposées.

<sup>26</sup> Nous nous référons ici à l'article de Carpitelli & Iannàccaro (1995) dans lequel les auteurs ont su mettre en évidence, d'un point de vue méthodologique et théorique, les difficultés rencontrées au cours des enquêtes sur le terrain (Cf. Pop (1950) ; cf. aussi Companys (1956) ; Chambers & Trudgill (1987) ; Llisterra Boix (1991) ; Grassi *et al.* (1997)).

à prononcer : cependant, il est préférable d'éliciter uniquement une phrase à la fois plutôt que le corpus entier, comme expliqué précédemment, en imposant la mémorisation et, seulement après un certain laps de temps, enregistrer l'énoncé. Ceci doit rester exceptionnel, et si vraiment le locuteur rencontre trop de difficultés, il est souhaitable d'en questionner un autre.

En réalité le dialectologue est partagé entre sa raison qui est l'étude de la langue naturelle, menée sur le terrain, et les procédés utilisés pour acquérir un corpus de mots ou de phrases exploitables. Ce qui fait dire à Straka (1981, p. 161) que : « un phonéticien peut ne pas être dialectologue, bien que l'observation de faits dialectaux soit d'une extrême utilité pour tout chercheur en phonétique ; en revanche un dialectologue doit nécessairement être phonéticien ».

## **Conclusion**

Les techniques d'enquêtes que nous avons mises en place ne font que démontrer que la quête de tout dialectologue, c'est-à-dire l'obtention naturelle et sans entrave des données de la part des informateurs, demeure un éternel problème. Par conséquent l'objectif recherché à travers ces stratégies d'élicitation est l'obtention d'un contrôle maximal des données, dans le sens d'une vérification et non d'une maîtrise de la parole. Le but final étant pour nous la comparaison des intonations romanes, rendue possible grâce à la mise en place de notre Base de Données, qui nous permettra également des études de calculs des distances prosodiques : « il serait maintenant opportun de pouvoir générer des calculs comparatifs qui mettront en évidence des degrés de distance entre les différents dialectes enquêtés » (Rilliard & Lai, 2008).

## Bibliographie

- Bellet G. ; Contini M. ; Boë L. J. (1981), "Normalisation temporelle et fréquentielle de F0 intra et inter-individuelle", *Bulletin de l'Institut de Phonétique de Grenoble*, 10-11, 265-283.
- Boë L. J. ; Contini M. ; Rakotofiringa H. (1975), "Etude statistique de la fréquence laryngienne. Application à l'analyse et à la synthèse des faits prosodiques du français", *Phonetica*, 32, 1-23.
- Brown G. ; Andersen A. ; Yule G. ; Shillcock R. (1983), *Teaching talk*. Cambridge, U. K., Cambridge University Press.
- Bruce G. (1982), "Developing the Swedish Intonation Model", *Working Papers*, 22, Lund University, 51-116.
- Caputo M. R. (1995), "Gradi accentuali nell'italiano parlato spontaneo", *Studi Italiani di Linguistica Teorica e Applicata*, XXIV, 2, 421-427.
- Chambers J. K. & Trudgill P. (1987), *La dialettologia*. Bologna, Il Mulino.
- Companys M. (1956), "Les nouvelles méthodes d'enquête linguistique", *Via Domitia*, 3, 83-138.
- Contini M. ; Boë L. J. (1973), "Contribution à l'étude quantitative de l'évolution de la fréquence laryngienne dans la phrase énonciative en français", *Bulletin de l'Institut de Phonétique de Grenoble*, II, 77-92.
- Contini M. ; Boë L. J. (1975), "Contribution à l'étude quantitative de l'évolution de la fréquence laryngienne dans la phrase interrogative en français (question totale)". *Bulletin de l'Institut de Phonétique de Grenoble*, IV, 85-102.
- Contini M. (1976), "Contribution à l'étude instrumentale de l'intonation en sarde", *Actes du XIII Congrès de Linguistique et Philologie Romanes* (Québec, 1971), Presses de l'Université de Laval, 229-245.
- Contini M. ; Lai J-P. ; Romano A. ; Roullet S. (2002), "Vers un Atlas prosodique parlant des variétés romanes", in J-C. Bouvier *et al.* (eds.), *Mélanges offerts à X. Ravier*. U. de Toulouse Le Mirail, 73-84.
- Endo R. ; Bertinetto P. M. (1997), "Aspetti dell'intonazione in alcune varietà dell'italiano", in F. Cutugno (org.), *Atti delle 7° Giornate di Studio del Gruppo di Fonetica Sperimentale* (Napoli, 14-15 nov. 1996), Roma, Esagrafica, 27-49.



- Espuny J. (1997), "Aspects prosodiques du discours hétérogène diaphonique". *Estudios de Fonética Experimental*, VIII, 271-295.
- Fonagy I. ; Fonagy J. (1983), "L'intonation et l'organisation du discours", *Bulletin de la Société de Linguistique de Paris*, LXXVIII, 1, 161-209.
- Grassi C., Sobrero A. A. & Telmon T. (1997), *Fondamenti di dialettologia italiana*. Roma, Laterza.
- Guaïtella I. (1992), "Rythme et parole : comparaison critique du rythme de la lecture oralisée et de la parole spontanée", *Travaux de l'Institut de Phonétique d'Aix*, 14, 221-225.
- Lai J-P. ; Romano A. ; Rouillet S. (1997), "L'analisi prosodica di alcune varietà linguistiche d'Italia : proposta per un approccio metodologico unitario", *Bollettino dell'Atlante Linguistico Italiano*, III, 21, 23-70.
- Lai J-P (2002), *L'intonation du parler de Nuoro (Sardaigne)*. Thèse de Doctorat, Université Stendhal-Grenoble3. 2vol.
- Lai J-P. et Romano A. (2002), "Etat d'avancement du projet : Atlas Multimédia Prosodique de l'Espace Roman (AMPER)", in *Bollettino dell'Atlante Linguistico Italiano*, III Serie Dispensa n° 26, 199-203.
- Lai J-P. (2005), "Etude de l'intonation d'un parler occitan : Gap", *Géolinguistique*, hors série n° 3, 65-86.
- Llisterri Boix J. (1991), *Introduccion a la fonetica: el método experimental*. Barcelona, Anthropos.
- Park Y. M. (1989). "Aspects syntaxique et rythmique de l'organisation prosodique des phrases en français : étude acoustique des variables temporelles et mélodiques", *Travaux de l'Institut de Phonétique de Strasbourg*, 21, 1-210.
- Peterson G. E. ; Lehiste I. (1960), "Duration of Syllable Nuclei in English", *Journal of the Acoustical Society of America*, 32, 693-703.
- Pettorini M. & Giannini A. (1994), "Aspetti prosodici del parlato radiofonico", in Salza P.L. (éd.). *Gli aspetti prosodici dell'italiano*. Atti delle 4° Giornate di Studio del Gruppo di Fonetica Sperimentale, (Torino, 11-12 nov. 1993), Roma, Esagrafica, 19-38.
- Pop S. ; Pop D. R. (1959), *Jules Gilliéron, vie, enseignement, élèves, œuvres, souvenirs*. Louvain, Duculot.
- Pop S. (1950), *La Dialectologie : aperçu historique et méthodes d'enquêtes linguistiques*. Université Catholique de Louvain, 2 vol.

- Rilliard A. & Lai J-P. (2008), "La base de données Amper et ses interfaces: structure et formats de données, exemple d'utilisation pour une analyse comparative de la prosodie de différents parlers romans", in L. de Castro Moutinho & R. L. Coimbra (eds), *Actes des I Jornadas Científicas AMPER-POR*, Universidade de Aveiro, 127 à 139.
- Rilliard A. & Lai J-P (2008), "Outils pour le calcul et la comparaison prosodique dans le cadre du projet AMPER-l'exemple des variétés Occitane et Sarde ", in A. Turculet (ed.), *Actes du Symposium International : La variation diatopique de l'intonation dans le domaine roumain et roman*, (Iasi, 20-22 octobre 2008), Iasi, Editura Universitatii « Alexandru Ioan Cuza », 217-229.
- Romano A. (1999), *Analyse des structures prosodiques des dialectes et de l'italien régional parlés dans le Salento (Italie) : approche linguistique et instrumentale*. Thèse de Doctorat, Université Stendhal-Grenoble 3, 2 vol. Publiée à Lille, Presses Univ. du Septentrion (2001).
- Romano A. ; Lai J-P. ; Rouillet S. (2005), "La méthodologie AMPER". *Géolinguistique*, hors série 3, 1-5.
- Rossi M. ; Di Cristo A. ; Hirst D. ; Martin Ph. ; Nishinuma Y. (1981), *L'intonation. De l'acoustique à la sémantique*. Paris, Klincksieck.
- Rouillet S. (1999), *Accent et intonation dans deux parlers francoprovençaux de la Vallée d'Aoste (Sarre et Cogne)*. Thèse de Doctorat, Univ. Stendhal-Grenoble 3, 2vol.
- Sabio F. (1996), "Transcription prosodique et constitution de bases de données suprasegmentales", in Moracchini G. (org.), *Actes du Colloque International de Corte (Corte, 11-14 oct. 1995)*, U. de Corse et de Nice Sophia Antipolis, 1996, 233-250.
- Savino M. ; Refice M. (1997), "L'intonazione dell'italiano di Bari nel parlato letto e in quello spontaneo", in Cutugno F. (org.), *Atti delle 7° Giornate di Studio del Gruppo di Fonetica Sperimentale*, (Napoli, 14-15 nov. 1996), Roma, Esagrafica, 1997, 79-88.
- Straka G. (1981), "Sur la formation de la prononciation française d'aujourd'hui", *Travaux de Linguistique et de Littérature*, XIX, 1, 161-248.

\*

Sites Web: <http://w3.u-grenoble3.fr/dialecto/AMPER/amper.htm>  
<http://amper.limsi.fr>

*Sabina Canobbio*  
*Professeur, Université de Turin*  
*sabina.canobbio@unito.it*

**Enquêtes (ethno-)linguistiques  
et mémoires du terrain : le « Piémont des paysans »  
de Paul Scheuermeier \***

Pour nous, dialectologues italiens – et peut être pour les collègues d'autres pays aussi – Paul Scheuermeier représente un point de repère fondamental, un véritable modèle pour nos activités de chercheurs 'sur le terrain'. On sait qu'il fut le collaborateur principal de Karl Jaberg et Jacob Jud dans la rédaction de l'AIS, l'*Atlas Linguistique Italo-Suisse*. La qualité de son travail est attestée non seulement par les cartes de l'atlas et par cet ouvrage extraordinaire qu'est le *Bauernwerk*, véritable encyclopédie de l'Italie paysanne traditionnelle, mais aussi par ses écrits et par ce que ses maîtres ont écrit de lui. Le témoignage le plus saillant est constitué par le volume d'introduction à l'AIS (Jaberg et Jud 1928), où les auteurs rendent compte des choix méthodologiques constituant la base de la recherche pour l'atlas et de son exécution avec une précision et une rigueur presque inégalées<sup>1</sup>. Dans ce document, comme dans d'autres publications<sup>2</sup>, Jaberg et Jud insistent sur le rôle déterminant joué par Paul Scheuermeier dans le 'chantier' de l'AIS et en tracent un profil professionnel que lui-même

---

\* Je remercie ma collègue Cristina Trincherio qui a eu la patience de m'aider à traduire mon texte en français.

<sup>1</sup> Comme l'annonce déjà le titre - « L'Atlas linguistique comme instrument de recherche. Fondements critiques et Introduction » - ces pages se dessinent comme une réflexion précieuse et exhaustive (à ce moment-là, bien sûr...) sur les questions théoriques de la recherche en géolinguistique en général et de la recherche sur le terrain en particulier. Il s'agit donc d'un texte qui mériterait beaucoup plus d'attention que celle que lui consacrent habituellement les lecteurs des cartes de l'atlas. Ces derniers y trouveraient déjà, également, les traces d'une conscience claire des problématiques de la variation linguistique, pas seulement diatopique, auxquelles la sociolinguistique a ensuite donné forme.

<sup>2</sup> Cf. Jaberg 1936, Jaberg et Jud 1923.

enrichit dans des pages où il raconte sa longue et complexe expérience de recherche<sup>3</sup>, s'exprimant à ce sujet avec une simplicité et une modestie étonnantes. « De ce grand ouvrage, celui qui vous parle ne fut que l'oreille, l'œil, la main et, aujourd'hui, le porte-parole » : ce sont les mots qu'il prononce lors de la communication qu'il fait en 1962, à l'occasion d'un colloque (Scheuermeier 1963) où il présente quelques pistes de lecture possibles du *Bauernwerk*, dont la publication venait d'être achevée. Ces deux volumes extraordinaires, en particulier, ratifient définitivement, par leurs textes et par leurs photographies, ce qu'il affirme de manière très efficace lorsqu'il décrit l'évolution de son parcours de chercheur : « Je partis linguiste et je revins ethnographe » (Scheuermeier 1963), sans toutefois jamais cesser de grandir en tant que linguiste, pourrions-nous ajouter. Certes, cet « attrape-dialectes »<sup>4</sup> était parti récupérer les matériaux nécessaires aux cartes de l' AIS, « Atlas de mots et de choses », en travaillant dans le cadre du courant théorique et méthodologique dessiné par le mouvement *Wörter und Sachen*, que Scheuermeier interpréta, avec Jaberg et Jud, d'une manière jamais mécanique ou simpliste. En effet, il fut bien inspiré par l'intuition très moderne selon laquelle le rapport entre les « mots » et les « choses » ne peut être entendu que comme le rapport entre des « systèmes de mots », des « systèmes de choses » et des « systèmes culturels complexes ». Et c'est autour de ce noyau central que se situe le travail de synthèse réalisé dans le *Bauernwerk*.

En Italie, Scheuermeier entreprit une campagne de recherche longue et presque ininterrompue entre 1919 et 1925, en parcourant la partie la plus dense du réseau des points de l' AIS, qu'il sillonna jusqu'à l'Italie du Centre<sup>5</sup>. Ensuite, entre 1926 et 1935, à une époque où les volumes de l'atlas avaient déjà commencé à paraître<sup>6</sup>, Scheuermeier revint plusieurs fois dans notre pays pour porter à terme des enquêtes, mais surtout pour effectuer, accompagné par le dessinateur Paul Boesch, la série des « approfondissements ethnographiques » qui auraient permis de mettre au point de la meilleure façon, et même d'achever, la description de quelques

---

<sup>3</sup> Scheuermeier 1932a e 1932b, 1936, 1963, 1969.

<sup>4</sup> C'est ainsi que le chercheur s'autodéfinit de façon plaisante dans un de ses écrits (Scheuermeier 1969).

<sup>5</sup> L'Italie du Sud et la Sicile ont fait l'objet des investigations de Gehrard Rohlf's ; la Sardaigne a été explorée par Max Wagner.

<sup>6</sup> Le premier volume date de 1928.

aspects linguistiques et culturels du monde paysan, au sujet duquel un intérêt particulier avait surgi au cours des recherches pour l'atlas.

Évidemment, tout l'immense patrimoine de données linguistiques, de notes, d'observations, de documentation iconographique récolté sur le terrain par Scheuermeier n'a pas pu être élaboré et utilisé pour l'atlas et pour le *Bauernwerk*. Heureusement, tous les matériaux originaux ont été recueillis et jalousement conservés auprès de la Bibliothèque Jaberg de l'Université de Berne et, tout récemment, ils ont été classifiés dans une « Archive AIS », qui est à présent en cours d'informatisation<sup>7</sup>. Il s'agit premièrement des Questionnaires que, dans chaque village visité, il remplit avec les réponses des informateurs, transcrites en alphabet phonétique ; à côté de ces données, il est possible de lire d'autres informations moissonnées au cours des interviews, et de nombreuses observations sténographiées par le chercheur. Dans l'archive bernoise, on trouve aussi les carnets d'enquête, avec des pages et des pages<sup>8</sup> contenant des remarques sur les villages et les communautés, sur les locuteurs interviewés, sur les caractéristiques des parlers locaux et sur les difficultés rencontrées dans leur transcription. Des centaines de pages, dont une petite partie seulement a été ensuite synthétisée dans les brefs « procès-verbaux des enquêtes » insérés dans le volume qui sert d'introduction à l'atlas, déjà cité. L'archive conserve également des centaines de photographies, dont de très nombreuses sont inédites ; chaque photo est accompagnée par un texte contenant un commentaire soigné et une description dont on réduirait l'importance si on la désignait comme une simple « légende » de l'image. Dans ces textes, sténographiés<sup>9</sup>, on rencontre les mots dialectaux transcrits en graphie phonétique, car encore une fois, selon Scheuermeier, seulement *ces mots-là* auraient pu accompagner et signifier les objets et les actions prises par l'objectif.

Parmi les autres inédits de Scheuermeier, l'archive compte aussi un important *Journal* où le chercheur enregistrerait les événements heureux et malheureux de ses journées de travail sur le terrain : ses déplacements d'un village à l'autre, la recherche d'un logement et des

---

<sup>7</sup> Cf. le site WEB [www.italiano.unibe.ch](http://www.italiano.unibe.ch).

<sup>8</sup> Sténographiées elles aussi, parce que, évidemment, le temps était sans doute l'un des problèmes les plus pressants de Scheuermeier...

<sup>9</sup> Et parfois assez longs et articulés ; quoiqu'il s'agisse de simples annotations, seulement une partie a été ensuite employée pour accompagner les photos publiées dans les deux volumes du *Bauernwerk*.

informateurs, les mésaventures, les gênes<sup>10</sup>, les rares pauses de repos et de détente. Enfin, dans l'archive on trouve des centaines de lettres et de cartes postales échangées entre Scheuermeier, Jaberg et Jud pendant leurs longues années de recherches : un épistolaire extraordinaire, témoignant de la mise au point longue et constante de la méthode et peut-être des objectifs mêmes de la recherche pour l'AIS, ajustés progressivement, au fur et à mesure que le travail sur le terrain les suggérait ou les imposait au travail théorique préalablement organisé, l'adaptant aux spécificités de différentes situations locales aussi bien qu'à la maturité du chercheur et au raffinement de sa méthode.

On comprend aisément que l'archive bernoise, avec sa documentation, abritait un véritable trésor pour le savoir dans le domaine de la dialectologie, et pas seulement. Sa redécouverte et sa valorisation ont commencé en 1995, à l'initiative du Musée Ethnographique de San Michele all'Adige, qui a édité la publication de tous les matériaux recueillis au Trentin (*Trentino dei contadini*). Ensuite, on a publié les matériaux concernant la Lombardie (*Lombardia dei contadini*, I, II, III), la Suisse italienne (*Parole in immagine*), le Piémont (*Piemonte dei contadini*, I e II) ; d'autres publications sont en préparation.

Les résultats de ces recherches et la disponibilité de cette quantité imposante de documents longtemps inédits ont une double valeur. Premièrement, grâce à eux, il a été possible de parvenir à une redécouverte de Paul Scheuermeier et à une plus complète définition de son profil de chercheur, qui ressort justement de la lecture progressive de ses propres documents de travail avec une dimension presque « héroïque »<sup>11</sup>.

Deuxièmement, à travers les publications régionales, il est maintenant possible de reconstruire l'immense patrimoine des matériaux recueillis, dont la dimension individuelle a été longtemps sacrifiée afin de les intégrer dans les grands projets de l'AIS et du *Bauerwerk* ; on peut ainsi parcourir à rebours, dirait-on, ces ouvrages de synthèse et redécouvrir les spécificités locales en suivant les pas de Scheuermeier. Des *millions* de pas faits par Scheuermeier... Grâce aux pages de son journal et à ses notes détaillées concernant les

---

<sup>10</sup> La photo n°1, prise à Ostana (*Piemonte dei contadini*, II, p.100), témoigne très efficacement des conditions dans lesquelles le chercheur menait ses enquêtes.

<sup>11</sup> Nous citons la définition donnée par Sanga 1997, qui l'appelle « le héros de la recherche sur le terrain ».

villages et les communautés explorés dans les enquêtes, il nous est maintenant possible de reconstruire les étapes de son tortueux itinéraire, et il nous est finalement possible de comprendre sa fatigue et son engagement dans la longue marche à travers la ‘petite’ et la ‘grande’ histoire, en même temps qu’à travers son histoire personnelle<sup>12</sup>. La dimension quotidienne des communautés et de leurs habitants, qu’il enregistra avec curiosité et participation, se développent dans le cadre d’une Italie qui témoigne – avec des différences d’une région à une autre – les marques de profondes transformations sociales et culturelles. Transformations qui allaient se préciser, aussi bien que les blessures de la Première Guerre mondiale et les conditionnements du régime fasciste, dont Scheuermeier eut l’occasion de connaître les phases et les facettes différentes, plus ou moins inquiétantes et menaçantes, mais au sujet desquelles il laisse un jugement sévère et sans appel. En somme, ces matériaux recueillis à une époque cruciale dans l’histoire de notre pays, et de l’Europe, possèdent une valeur documentaire qui dépasse celle de la linguistique et de l’ethnographie<sup>13</sup>.

À travers les pages du journal, les cartes postales, les photos et leurs didascalies, les notes concernant les villages et les informateurs, publiées récemment dans deux gros volumes<sup>14</sup>, même le « Piémont des paysans » de Paul Scheuermeier nous a révélé sa physionomie, complexe et parfois ambiguë. C’est un autre témoignage du regard aigu du chercheur, de comment il a été à même de saisir et reconstruire la réalité contradictoire d’une région composite fondée au long des siècles sur une histoire tourmentée, impliquant l’union entre des entités profondément différentes. Mais ces deux volumes illustrent aussi à quel point la rencontre entre Scheuermeier et le Piémont n’a pas été simple, à cause de facteurs liés à son milieu physique, humain, socio-culturel et linguistique aussi. Il est souvent difficile pour le chercheur d’interpréter le territoire piémontais et de le comparer aux modèles auxquels il s’était préalablement confronté, et de concilier

---

<sup>12</sup> Pendant sa campagne de recherche au Piémont en 1923 Scheuermeier connut Nellie Nicolet, qui à cette époque-la était encore une des élèves de Jaberg. Elle serait devenue ensuite sa femme. Sur ces aspects de la vie privée de Scheuermeier, voir *Piemonte dei contadini*. I et Canobbio 2008. La photo n° 2 montre le couple s’occupant, le 11 septembre 1923, de l’enquête d’Ornavasso. Avec eux l’informateur local (*Piemonte dei contadini*. II, p. 220).

<sup>13</sup> À ce sujet, voir Mathieu 2008.

<sup>14</sup> *Piemonte dei contadini* I e II, dont on a exclu seulement les lettres qui, par leur importance, feront l’objet d’un travail à part. Nous renvoyons à ce travail pour tout approfondissement sur les recherches de Scheuermeier au Piémont.

ainsi ce qu’il voyait avec l’*idée* du Piémont qu’il s’était fait durant les phases préparatoires de son expédition italienne, sur la base de la littérature disponible et du bagage des renseignements reçus de ses maîtres<sup>15</sup>. Il rencontre surtout les plus grandes difficultés dans la partie occidentale de la région, où il trouve des patois dont le décryptage, la transcription et la segmentation ne sont pas aisés<sup>16</sup>, et trouve aussi des locuteurs utilisant des répertoires complexes, dans lesquels les patois locaux – qui étaient bien vifs à cette époque-là – se confrontent et se superposent continuellement au piémontais, avec un peu d’italien et, dans certaines vallées, avec du français. Parfois c’est justement le contexte géographique et humain qui lui est hostile ; cela arrive surtout dans certaines vallées des Alpes, dans des hameaux misérables et inhospitaliers, habités par des montagnards arriérés et méfiants. En plus, au-delà des difficultés pratiques et linguistiques, le chercheur semble souvent désorienté devant le monde paysan du Piémont, caractérisé en général par une petite, voire très petite propriété et par un travail agricole ne visant qu’à la subsistance, qu’on alternait avec d’autres activités ; et par des paysans qui étaient à l’opposé de l’image du « solide agriculteur bernois » constituant le point de repère naturel d’un Suisse comme Scheuermeier. Un profil dont il a rarement retrouvé les traits chez ses informateurs piémontais.

Scheuermeier explora le Piémont entre 1921 et 1931, quand il y revint pour la dernière fois en vue d’effectuer l’approfondissement ethnographique du village de Montanaro. Pendant cette période, la région étalait un contraste profond entre le retard économique de ses aires ‘faibles’<sup>17</sup> et le développement de ses aires ‘fortes’<sup>18</sup>, entre l’assujettissement à la tradition et les élans vers l’innovation. Un conflit qui était linguistique aussi, et qu’il a souvent remarqué en analysant les rapports entre les jeunes et les vieilles générations, donnant la preuve ainsi d’une attention aux aspects de la variation, vraiment surprenante pour l’époque. En particulier, c’est justement à Montanaro que le chercheur saisit l’occasion d’interviewer à distance un paysan aisé, puis son fils ; ce dernier fut un informateur précieux car, grâce à son attitude positive envers la modernisation, il se posa

---

<sup>15</sup> De Jaberg en particulier, cf. Scheuermeier 1969, p. 331.

<sup>16</sup> À ce sujet, voir Telmon 2008.

<sup>17</sup> Par exemple la montagne.

<sup>18</sup> Où l’industrie s’était déjà affirmée.



justement comme protagoniste et témoin digne de foi de la confrontation dialectique entre l'ancien et le nouveau<sup>19</sup>.

Dans certains cas, le procès de transformation semble déjà très avancé en Piémont et ce facteur limite considérablement la possibilité pour Scheuermeier d'observer des techniques et des instruments traditionnels des travaux des paysans et des artisans, comme l'illustrent certaines photos dont le chercheur même évoque le pénible travail préparatoire. Par exemple, la photo n° 4, prise à Galliate (Novara), dont la didascalie explique que « ce type de baratte n'est plus utilisé et cette femme a dû partir en chercher une [...] ; il s'agit de choses vieilles »<sup>20</sup>.

Naturellement, à la difficulté de repérer les objets et de reconstruire les techniques s'ajoute la difficulté de raviver la mémoire des mots qui originairement les indiquaient. Cet aspect nous rappelle que le procès de désagrégation des compétences linguistiques et des savoirs techniques qui est à présent désormais presque entièrement révolu avait bien commencé déjà dans ces années-là.

Mais la documentation extraordinaire que Scheuermeier recueillit dans les années 20 et 30 offre aussi des éléments qui se révèlent décisifs pour évaluer combien le territoire piémontais paya sa modernisation sur les plans de la culture et du paysage. À ce sujet, je renvoie à la photo n° 5<sup>21</sup>, qui montre la beauté intacte de certains prés autour de Sestrière, dont la destinée aura été, quelques années plus tard, de faire l'objet de dévastations liées à la construction de bâtiments et de structures sportives pour les stations de ski, au nom du développement touristique.

Au-delà des problèmes liés à l'environnement et à la logistique que nous venons d'évoquer, le bilan global de la rencontre entre Scheuermeier et le Piémont paraît en tout cas fort positif et souvent les réalités locales correspondent à la plus rassurante réputation de la région, tout à fait en harmonie avec « le bon idéal piémontais », pour citer ses mots. On compte en fait de nombreuses enquêtes menées dans des villages riants et parmi des communautés laborieuses, où il

---

<sup>19</sup> La scène proposée sur la photo n° 3 est emblématique de cette dialectique. Des balles de foin (déjà pressées à la machine, comme le remarque le chercheur) sont transbordées du chariot traditionnel (mené par des boeufs), sur lequel elles ont été transportées depuis les prés, dans la remorque d'un camion moderne (*Piemonte dei contadini*. I, p.225). L'informateur de l'approfondissement sur Montanaro se distingue aussi pour ses copieuses remarques à propos de la complexité des phénomènes langagiers locaux.

<sup>20</sup> *Piemonte dei contadini*. II, p. 57.

<sup>21</sup> *Piemonte dei contadini*. I, p.101.

fut possible de réaliser la rare alchimie de rencontres humaines et professionnelles heureuses. Les portraits des informateurs qu'on lit dans les notes originales laissent souvent transparaître son admiration pour la ténacité dans le travail, pour le courage montré face aux défis de l'émigration et du changement, pour le désir de s'améliorer par l'étude.

Parmi les personnages les plus frappants, il y a les femmes des villages, qui s'imposent comme le véritable « anneau fort »<sup>22</sup> de ces petites communautés et qui, dans les enquêtes, se révèlent d'excellentes informatrices. Toutefois, au Piémont aussi, se réalise le choix paradoxal de ne prendre que rarement des femmes comme sujets officiels, conformément à l'un des présupposés méthodologiques de Jaberg et Jud, basé du reste sur des considérations que l'on peut en partie comprendre ; par exemple, la difficulté d'approcher des femmes dans des contextes socio-culturels encore fortement conditionnés par des normes sévères réglant le rapport entre les sexes<sup>23</sup>; et, de plus, le fait que les femmes maîtrisaient généralement mal la langue italienne (pourtant nécessaire pour se comprendre pendant l'enquête), sans doute moins que les hommes<sup>24</sup>. En réalité, il suffit de lire les annotations des enquêtes pour comprendre que bien souvent les femmes y ont joué un rôle décisif, même lorsqu'elles restaient dans les coulisses<sup>25</sup>.

En effet, si l'on feuillette les notes piémontaises, on peut reconnaître une grande partie des problématiques classiques et actuelles de la recherche sur le terrain, avec la typologie d'informateurs que le chercheur rencontre, ou qu'il a peur de rencontrer...aujourd'hui comme hier. D'une part, il y a des sujets 'positifs', des locuteurs brillants et disponibles pour une franche collaboration, conscients de leurs compétences personnelles aussi bien que des dynamiques du répertoire linguistique de la communauté. De l'autre, il y a des sujets 'difficiles', timides et hésitants, gênés par le

---

<sup>22</sup> Nous citons le titre d'un livre célèbre de Nuto Revelli (Revelli 1985).

<sup>23</sup> La photo n° 6 est très significative à ce sujet. Prise à Rocca Canavese (*Piemonte dei contadini*. I, p. 116), le chercheur l'accompagne d'une annotation désolée, dans laquelle il remarque que les femmes sont tellement timides qu'elles se sauvent immédiatement lorsque l'on essaie de les photographier avec leurs paniers typiques.

<sup>24</sup> Dans la situation d'italophonie précaire de ces années-là, commune dans l'Italie en général et dans le monde rural en particulier, les femmes avaient généralement moins d'opportunités que les hommes de rencontrer et d'apprendre la langue nationale.

<sup>25</sup> Pour la collaboration des femmes aux enquêtes piémontaises, cf. Canobbio 2008, pp. 308-310.

contexte de l'enquête et prêt à dire n'importe quoi afin de complaire au professeur venu de loin. Ou encore, il y a ceux « qui ont des yeux mais ne voient pas », pour rappeler la définition que Scheuermeier donne de ceux qui ne montrent aucune sensibilité pour les aspects singuliers de leur langue et de leur culture ; enfin, il y a les terribles 'experts' locaux, souvent pédants et puristes, qui affligent le chercheur à coup de précisions continues<sup>26</sup>.

Mais des carnets d'enquêtes, il ressort aussi un autre problème de la recherche sur le terrain, qui n'est pas moins important et sur lequel on ne réfléchit pas souvent : celui de la compréhension ou de l'incompréhension de la part des informateurs de la signification de l'enquête à laquelle ils collaborent. « Il n'est pas facile de faire comprendre à des gens simples que leur méprisé patois peut avoir une signification scientifique », remarque Scheuermeier<sup>27</sup>, et ce sont des mots que nous pourrions écrire nous aussi dans nos carnets de chercheurs ; même si durant les dernières années en Italie on a assisté à des modifications dans l'attitude envers les dialectes.

En somme, la riche documentation publiée nous présente d'une part un Scheuermeier déjà conscient du fait que les données récoltées sur le terrain sont toujours le résultat d'une 'médiation' entre la 'théorie' qui guide le chercheur et celle qui est sous-jacente à la conscience du locuteur, et, de l'autre, que ces données ne peuvent pas être lues sans rapport avec le contexte socio-culturel qui les a produites.

---

<sup>26</sup> À ce sujet, voir Canobbio 2008; pour le rapports entre Scheuermeier et les informateurs locaux, il est intéressant de consulter Sanga 2001.

<sup>27</sup> Dans Jaberg et Jud 1923, p. 492, quand les auteurs lui passent la parole afin qu'il explique son expérience de recherche pour l'AIS.

## Bibliographie

- AIS = K. Jaberg, J. Jud, *Sprach- und Sachatlas Italiens und der Südschweiz*, 8 voll., Ringier u. C., Zofingen, 1928-1940.
- Bauernwerk* = P. Scheuermeier, *Bauernwerk in Italien der italienischen und rätoromanischen Schweiz. Eine sprach- und sachkundliche Darstellung häuslichen Lebens und ländlicher Geräte*, Vol. I, Eugen Rentsch Verlag, Erlenbach - Zürich, 1943 et Vol II, Stämpfli & Cie, Berna 1956 ; édition italienne (à laquelle nous renvoyons) par M. Dean e G. Pedrocco, trad. I. Gaudenzi e K. Dori Egger, *Il lavoro dei contadini. Cultura materiale e artigianato rurale in Italia e nella Svizzera italiana e retoromanza*, Longanesi, 2 vol., Milano 1980.
- Canobbio 2008 = S. Canobbio, "Il Piemonte di Paul Scheuermeier", *Piemonte dei contadini*. II, pp. 301-317.
- Jaberg 1936 = K. Jaberg, *Aspects géographiques du langage (avec 19 cartes). Conférences faites au Collège de France (Décembre 1932)*, Librairie E. Droz, Paris.
- Jaberg e Jud 1923 = K. Jaberg, J. Jud, "Un atlante linguistico-etnografico svizzero italiano", *Le vie d'Italia*, Rivista del Touring Club italiano, XXXI, 1923/5, pp. 483-493.
- Jaberg et Jud 1928 = K. Jaberg, J. Jud, *Der Sprachatlas als Forschungsinstrument. Kritische Grundlegung und Einführung in den Sprach- und Sachatlas Italiens und der Südschweiz*, Max Niemeyer Verlag, Halle (Saale) 1928 ; édition italienne par G. Sanga, trad. S. Baggio, *AIS Atlante linguistico ed etnografico dell'Italia e della Svizzera meridionale*. Vol. I, *L'atlante linguistico come strumento di ricerca. Fondamenti critici e introduzione*, Unicopli, Milano, 1987 [ici on se rapportera toujours à l'édition italienne].
- Lombardia dei contadini*. I = Paul Scheuermeier. *La Lombardia dei contadini 1920-1932*. I. *Lombardia orientale. Le province di Brescia e Bergamo*, par G. Bonfadini, F. Caltagirone, I. Sordi, Grafo, Brescia 2001
- Lombardia dei contadini*. II = Paul Scheuermeier. *La Lombardia dei contadini 1920-1932*. II. *Lombardia orientale. Le province di Cremona e Mantova*, par G. Bonfadini, F. Caltagirone, I. Sordi, Grafo, Brescia, 2002.

- Lombardia dei contadini. III = Paul Scheuermeier: La Lombardia dei contadini 1920-1932. III. Lombardia occidentale*, par F. Caltagirone, G. Sanga, I. Sordi, Grafo, Brescia, 2007.
- Mathieu 2008 = J. Mathieu, "Il significato storico delle testimonianze raccolte da Paul Scheuermeier", dans *Parole in immagine*, pp. 23-28.
- Parole in immagine = M. Frasa, L. Grassi, F. Lurà, Parole in immagine. Le ricerche di Paul Scheuermeier nella Svizzera italiana 1920-1927*, Centro di dialettologia e di etnografia, Bellinzona, 2008.
- Piemonte dei contadini. I = Paul Scheuermeier. Il Piemonte dei contadini 1921-1932. I. La provincia di Torino*, par S. Canobbio e T. Telmon, Priuli & Verlucca, Scarmagno, 2007.
- Piemonte dei contadini. II = Paul Scheuermeier. Il Piemonte dei contadini 1921-1932. II. Le province di Alessandria, Asti, Biella, Cuneo, Novara, Verbano Cusio Ossola, Vercelli*, par S. Canobbio e T. Telmon, Priuli & Verlucca, Scarmagno, 2008.
- Revelli 1985 = N. Revelli, *L'anello forte. La donna: storie di vita contadina*, Einaudi, Torino.
- Sanga 1997 = G. Sanga, "Un eroe della ricerca sul campo", in *Trentino dei contadini*, pp. 31-41.
- Sanga 2001 = G. Sanga, "Scheuermeier in Lombardia. Il ricercatore e i suoi informatori", dans *Lombardia dei contadini. Le province di Brescia e Bergamo*, pp. 7-21.
- Sanga 2007 = G. Sanga, "Gli interlocutori di Scheuermeier", dans *Lombardia dei contadini. Lombardia occidentale*, pp. 27-44.
- Scheuermeier 1932a = P. Scheuermeier, "Observations et expériences personnelles faites au cours de mon enquête pour l'atlas linguistique et ethnographique de l'Italie et de la Suisse méridionale", *Bulletin de la Société de Linguistique*, 33/1, 1932, pp. 93-110.
- Scheuermeier 1932b = P. Scheuermeier, "Il dialettologo in viaggio", dans *Parole in immagine*, pp. 43-45.
- Scheuermeier 1936 = P. Scheuermeier, "Methoden der Sachforschung. Zur sachkundlichen Materialsammlung für den Sprach- und Sachatlas Italiens und der Südschweiz", *Vox Romanica*, I, 1936, pp. 334-369.
- Scheuermeier 1963 = P. Scheuermeier, "Regioni ergologiche della vita agricola italiana", dans *Il mondo agrario tradizionale nella*

*Valle Padana, Convegno di studi sul folklore padano* (Modena 17-19/3/1962), Modena 1963, pp.291-307.

Scheuermeier 1969 = P. Scheuermeier, *Della buona stella sul nostro Atlante*, 1969 [manuscrit original conservé à la bibliothèque Karl Jaberg de l'Université de Berne ; première publication intégrale en traduction italienne par C. Gentili, *Trentino dei contadini*, pp. 329-346].

*Scheuermeier, le Alpi e dintorni* = C. Gentili, G. Kezich e G. Sanga, *Scheuermeier, le Alpi e dintorni. Atti del Seminario permanente di Etnografia Alpina*, (SPEA 4), SM, *Annali di San Michele*, 12/1999.

Telmon 2008 = T. Telmon, « Le osservazioni fonetiche », dans *Piemonte dei contadini. II*, pp.363-381.

*Trentino dei contadini* = Paul Scheuermeier. *Il Trentino dei contadini 1921-1931*, par T. Kezich, C. Gentili, A. Mott, San Michele all'Adige, *Museo degli usi e costumi della gente trentina*, 1997.

## Photographies

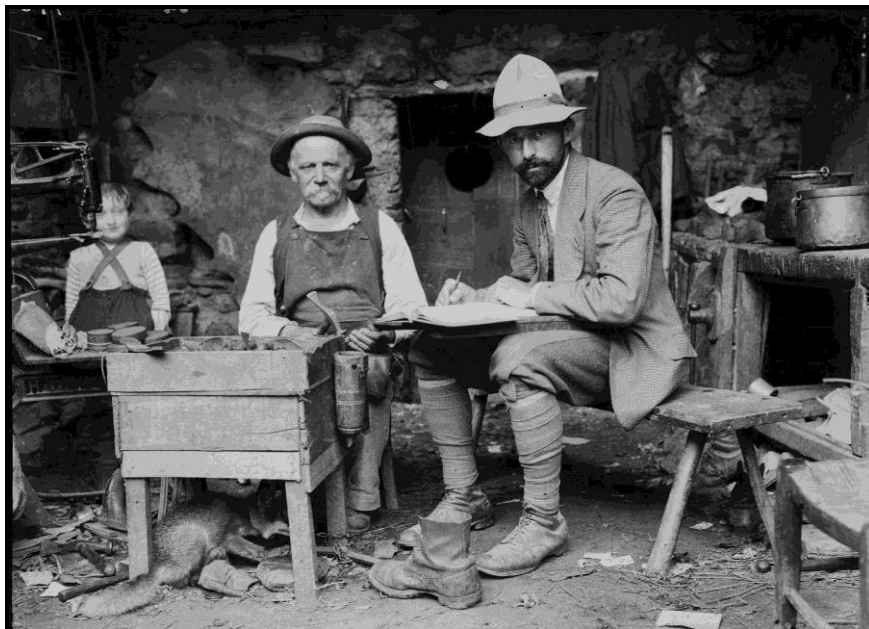


Photo n° 1  
Paul Scheuermeier avec le menuisier d'Ostana  
(n. 798 Archive AIS)



Photo n° 2  
Paul Scheuermeier avec Nellie Nicolet et l'Informateur d'Ornavasso  
(n. 1211 Archive AIS)





Photo n° 3  
Des balles de foin transbordées du chariot au camion sur la place de  
Montanaro en 1932 (n. 6105 Archive AIS)



Photo n° 4  
La vieille baratte de Galliate  
(n. 339 Archive AIS)



Photo n° 5  
Les prés du col de Sestrière photographiés en 1922 par Scheuermeier  
(n. 811 Archive AIS)



Photo n° 6  
La difficulté de photographier à Rocca Canavese, en 1923, les filles  
du village, intimidées par l'étranger  
(n. 1237 Archive AIS)

*Matteo Rivoira*  
*Université de Turin*  
*matteo.rivoira@unito.it*

**L'Atlante Linguistico Italiano (ALI) et l'Atlante  
Toponomastico del Piemonte Montano (ATPM) :  
histoire, méthodes de récolte et de présentation des données,  
et perspectives de développement de deux entreprises  
géolinguistiques turinoises<sup>1</sup>**

**L'Atlante Linguistico Italiano<sup>2</sup>**

L'idée d'un atlas linguistique italien, fait par des Italiens, remonte à 1908. Mais il a fallu plus d'une quinzaine d'années pour que Matteo Giulio Bartoli puisse donner à cette idée une forme plus concrète et présenter au public, en 1924, un programme de travail assez défini. Le travail préparatoire concernant l'élaboration du questionnaire (ALI 1971) et le choix des points d'enquête montre dès le début une démarche théorique différente et une augmentation non négligeable de la quantité des données à récolter, par rapport soit à l'ALF, qui demeure naturellement le modèle, soit à l'AIS qui allait bientôt conclure sa campagne d'enquête. Ce fut Ugo Pellis qui, après avoir contribué de façon déterminante à la révision du questionnaire, commença les relevés en 1925 et continua jusqu'en 1942, quand, à

---

<sup>1</sup> Je remercie pour les précieuses notices concernant l'ALI, le Directeur de l'Institut, Professeur Lorenzo Massobrio, le Co-Directeur, Professeur Giovanni Ronco et le Professeur Temistocle Franceschi, vaillant enquêteur et collaborateur de l'Atlas depuis longtemps, qui a bien voulu corriger certains passages du texte. Je remercie aussi M. le Professeur Christian Abry pour ses conseils précieux et pour avoir bien voulu relire ce texte, ainsi que M.me Christine Tourn qui a eu la patience de revoir le français approximatif de ma contribution. Les erreurs et les imprécisions qui restent sont dues à mon obstination.

<sup>2</sup> Parmi les contributions les plus récentes sur l'histoire et les méthodologies de l'ALI, v. Ronco 2004.

cause de la guerre, il dut abandonner toute activité après avoir effectué seulement une partie d'un plan qui prévoyait d'enquêter aussi dans les territoires italophones en dehors de l'Italie et même dans les colonies italiennes.

Après la guerre, les activités ne purent reprendre qu' en 1952. Bartoli et Pellis étaient entre-temps décédés. Benvenuto Terracini, revenu de son exil en Argentine où il avait dû se réfugier à cause des lois raciales adoptées par le régime fasciste, succéda à Bartoli à la direction de l'ALI. Le réseau des points à enquêter fut réduit en excluant toute enquête en dehors de l'Italie et en diminuant la taille du questionnaire, qui avait déjà fait l'objet d'une révision par Pellis.

Terracini confia la tâche de continuer les enquêtes à Raffaele Giacomelli, secondé par Corrado Grassi, lequel pourrait prendre sa place le cas échéant, ce qui d'ailleurs se produisit à la mort du premier, qui avait déjà enquêté sur 40 points. Suite au renoncement de Grassi après 62 enquêtes, pour accélérer la fin de la campagne d'enquêtes, Terracini se résolut à confier le travail de collecte à plusieurs enquêteurs : Giorgio Piccitto (7 enquêtes) ; Giovanni Tropea (24 enquêtes), Michele Melillo (27 enquêtes) et Temistocle Franceschi (122 enquêtes). Ils terminèrent leurs parcours en 1964. La quantité de matériaux rassemblée au cours des 1009 enquêtes, en interrogeant environ 1700 témoins, est remarquable : plus de 5 000 000 de fiches dialectales, 8 850 photographies, avec de nombreux cahiers contenant les notes des sept enquêteurs. Les premiers essais de rédaction des cartes linguistiques qui furent conduits à partir de ces matériaux (Terracini/Franceschi 1964, Terracini/Franceschi 1965, Terracini 1967) montrèrent tout de suite quelles étaient les difficultés de gestion d'une telle quantité de données récoltées sur une durée aussi longue, par des personnes différentes et suivant (au moins en partie) des démarches différentes. Avant toute chose, il n'existait pas de tableau de conversion contenant les signes phonétiques qui avaient été utilisés, avec une définition univoque de leur valeur. Le classement des fiches nécessitait aussi une révision importante. Et surtout il fallait résoudre les problèmes techniques liés à la réalisation d'une telle œuvre. Un long travail de pré-édition fut entamé sous la direction de Corrado Grassi – qui succédait à Giuliano Bonfante, lequel était devenu directeur au départ à la retraite de Terracini, aux côtés de Giuseppe Vidossi, co-directeur depuis le début avec Bartoli, Bonfante ayant assez vite renoncé à cette charge à cause de divergences avec le

comité scientifique de l'ALI<sup>3</sup>. La responsabilité de ce délicat travail incombait au nouveau comité de rédaction constitué par Arturo Genre, Lorenzo Massobrio, Silvio Campagna e Giovanni Ronco<sup>4</sup>.

Pendant une quinzaine d'années, le comité de rédaction se consacra à la révision des données et à l'élaboration d'une procédure visant à donner une cohérence aux matériaux. Les problèmes techniques furent assez vite résolus lorsque l'*Istituto Poligrafico e Zecca dello Stato (IPZS)* s'intéressa à la publication des cartes, sous la direction d'Arturo Genre. Les techniciens de l'Institut élaborèrent à la fin des années 1980 un logiciel qui permit de transformer les données linguistiques des fiches en données numériques et de publier en 1995, sous la direction de Lorenzo Massobrio, le premier volume de l'ALI (ALI 1995<sup>b</sup>). Depuis lors, six autres volumes sont parus et le huitième va bientôt être imprimé (ALI 1996, 1997, 1999, 2001, 2003 et 2008).

Comme l'a dit Ottavio Lurati (2003) de façon suggestive au début de son discours de présentation du quatrième volume de l'ALI « In un periodo di prepotente prevalere dell'effimero e dell'*istant-book* come è quello in cui viviamo [...] [l'ALI è] un'opera che resta attraverso i decenni ». Malheureusement, son élaboration aussi s'est prolongée durant des décennies. C'est une histoire émaillée de difficultés. Il suffit de dire qu'à un certain moment, une direction peu clairvoyante avait même proposé de renoncer à la publication des volumes, se contentant d'utiliser les archives, ce qui aurait, dans les faits, empêché la communauté scientifique d'accéder aux données ethnolinguistiques. Mais son histoire est aussi une histoire de fidélité et de dévouement de tous les hommes et de toutes les femmes qui y ont travaillé et y travaillent.

Une fidélité aux principes qui ont dirigé la conception de l'œuvre par Bartoli, qui, fort de l'élaboration théorique gilliéronienne lue à travers ses propres intuitions, sut, comme l'a dit Terracini, transformer « alcuni spunti storicistici che indubbiamente affiorano nei primi abbozzi in un criterio di distribuzione metodica dei punti da rilevare tale che tutte le possibilità di accavallamenti e di correnti culturali che può presentare la varietà dei nostri dialetti venisse teoricamente prevista. Un atlante, scrisse Bartoli, deve essere uno strumento di

---

<sup>3</sup> Bonfante voulait inclure dans le réseau des enquêtes la Corse et le Canton du Tessin, ce qui aurait conduit à un prolongement excessif des travaux.

<sup>4</sup> On publia dans ces années-là d'importants instruments de consultation des matériaux des archives de l'ALI (ALI 1971, 1973<sup>a</sup>, 1973<sup>b</sup>, 1980, 1981, 1985, 1995<sup>a</sup>).

lavoro e non soltanto un museo di antichità » (Terracini 1955, 4). Cette vision fut nécessaire pour saisir le caractère dynamique de la langue, surprise dans son mouvement perpétuel (Terracini 1969, 13). Elle s'est réalisée en premier lieu dans le choix de points, qui devaient être de trois types : « massimi », « medi » et « minimi », c'est-à-dire des points où, par leur nature (grandes villes, villes de province et villages ou fractions), on aurait pu collecter des données à même d'illustrer la portée et la direction des courants novateurs. La structure du questionnaire était conçue pour s'adapter à la situation italienne, si complexe et peu homogène du point de vue culturel et linguistique. Le choix des informateurs démontre d'un autre point de vue la démarche du chercheur : Ugo Pellis n'a nullement privilégié les informateurs les plus conservateurs, même si la plupart des témoins se placent dans une classe d'âge assez élevée (entre 55 et 60 ans) (v. fig. 1).

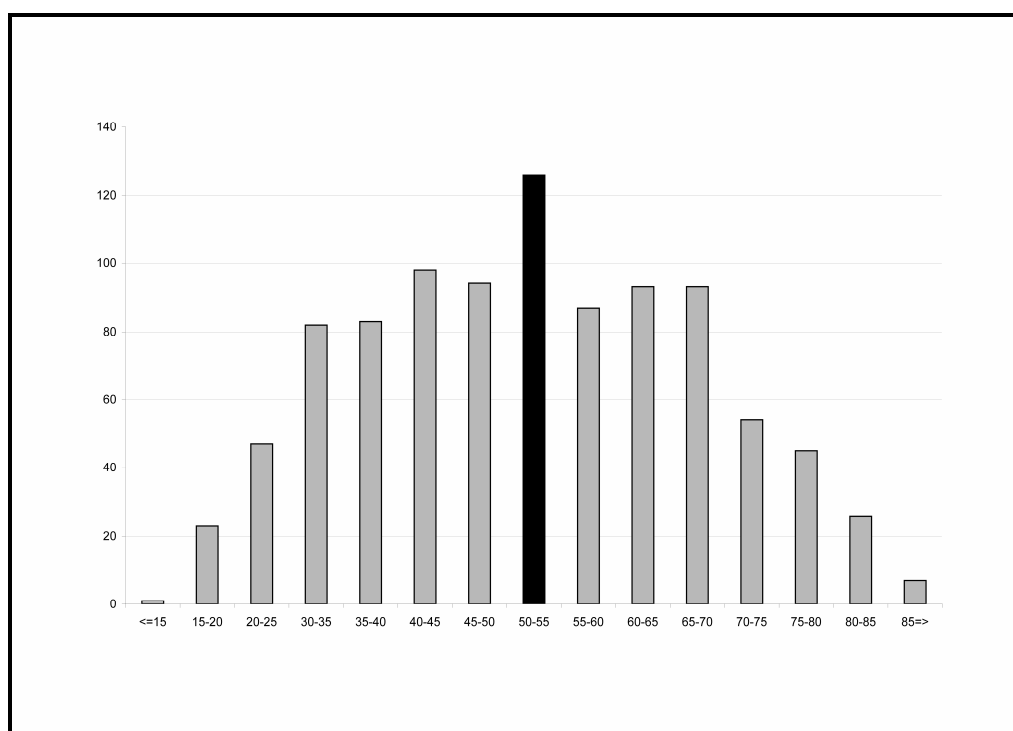
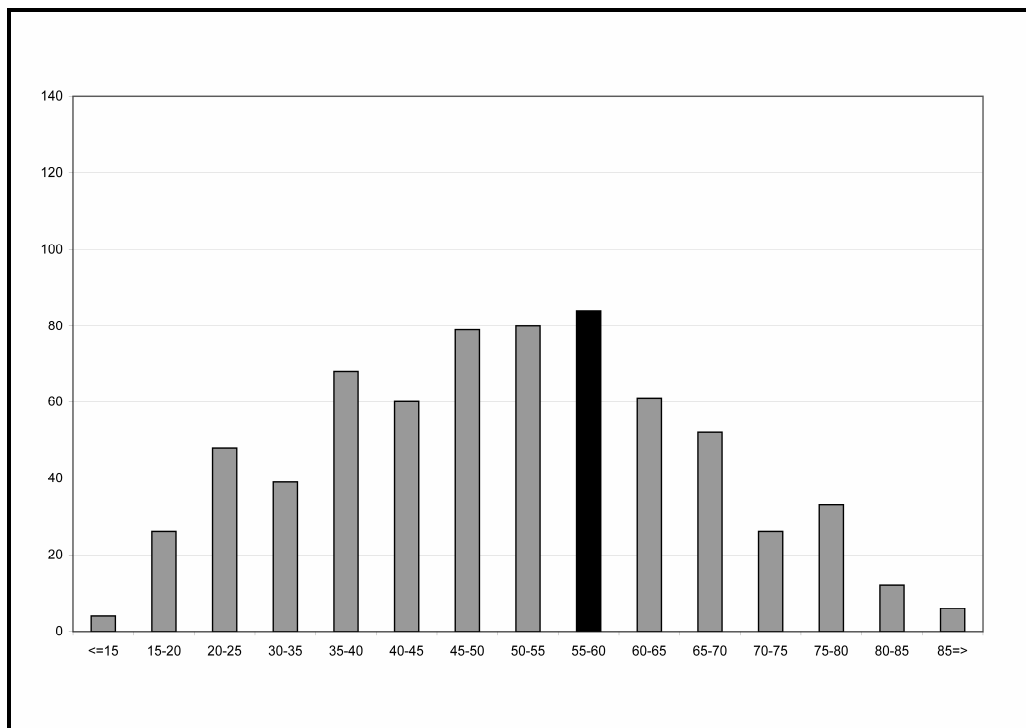


Fig. 1 Classes d'âge des témoins interrogés par Ugo Pellis

Sur ce point, Terracini, dans la seconde campagne d'enquête conduite après-guerre, a innové dans la méthode tout en restant fidèle aux principes. Les enquêteurs, même si le décalage dans le temps pouvait le suggérer, dans une vision de la chose quelque peu naïve,



non seulement n'interrogèrent pas uniquement des personnes âgées, mais on peut voir (fig. 2), qu'ils choisirent des classes plutôt jeunes.



*Fig. 2 Classes d'âge des témoins interrogés par les successeurs de Ugo Pellis*

Mais ce n'est pas là la seule innovation. Terracini introduit de façon systématique une ou plusieurs enquêtes de contrôle visant à vérifier la qualité des matériaux récoltés, favorisant ainsi une vision plus objective du parler enquêté. En augmentant aussi le nombre des informateurs interrogés (Franceschi arrive à interroger jusqu'à dix personnes dans une enquête), et surtout en les choisissant soigneusement par profession, âge, niveau d'instruction, il témoigne d'une attention nouvelle à la dimension sociologique de la variété linguistique. Tout cela dans le cadre de la conception terracinienne de l'unité du point linguistique définie dynamiquement comme résultat de courants – novateurs ou conservateurs – se déployant dans l'espace, dans des directions que la carte linguistique permet de saisir dans leurs complexités.

Du point de vue méthodologique, on remarquera encore l'attention rigoureuse portée à la précision de la notation phonétique

impressionniste – dont la complexité peut paraître parfois déroutante (Canepari 2003, 98), mais dont la pertinence a été à maintes reprises confirmée, y compris par des analyses instrumentales (je pense là aux travaux de Michel Contini sur les parlers sardes [Contini 1987]). Et aussi une attention très aiguë consacrée au rapport entre les mots et les choses, qui se manifeste de façon évidente dans le patrimoine des photos prises au cours des enquêtes et par les illustrations collectées.

Il est évident que la validité de ces principes, nous permet de reconnaître dans les quelques 724 cartes publiées, avec les réponses à plus de 1550 questions, un trésor pour tout spécialiste des langues et dialectes d'Italie. Blasco Ferrer (1999), Lurati (2003), Pfister (2006) et aussi Franceschi (2006), l'ont très bien montré dans leurs études basées sur les matériaux de l'ALI. J'ajoute en plus que, dans le cadre italien, où les Atlas régionaux ayant jusqu'à maintenant publié le résultat de leurs enquêtes sont encore peu nombreux, l'ALI ne se limite pas à être un « musée d'antiquité », mais il demeure cet « outil de travail » préconisé par Bartoli. Cela malgré le retard avec lequel il est en train de paraître et bien que l'on dispose depuis longtemps de l'AIS, par rapport auquel l'ALI présente un cadre plus précis de la situation linguistique italienne de la première moitié du XX<sup>e</sup> siècle, surtout du point de vue de la variation phonétique et lexicale.

Dans un futur proche, le but de l'Institut de l'ALI demeure de continuer la publication, tout en faisant face aux problèmes techniques qui ne cessent de se poser, mais qui, comme ce fut le cas autrefois, peuvent aussi se résoudre par une amélioration de l'organisation du travail. Ainsi, au cours des années 2008-2009, l'adaptation du logiciel élaboré à la fin des années 1980<sup>5</sup> par les techniciens de l'IPZS, a permis de transcrire les réponses et de les placer sur la carte. Pour résoudre le problème posé par les caractères complexes employés dans la notation phonétique, ils ont adopté un codage ASCII en subdivisant chaque graphème dans ses unités minimales, ce qui a permis de réduire le nombre des signes à mémoriser (étant donné qu'il s'agit d'un système de transcription qui a largement recours à des signes diacritiques). Ces unités de signes minimales peuvent être par la suite recomposées sur la page en les plaçant dans un des huit niveaux prévus pour chaque lettre (le standard Unicode n'était évidemment pas encore élaboré, ce qui aurait permis une autre

---

<sup>5</sup> En utilisant le langage *Fortran*, adapté pour le système d'exploitation VMS qui fonctionnait sur des machines VAX.

solution). À ce moment-là, il fallait permettre l'utilisation du système informatique sur une plateforme comme Windows, qui fonctionne sur n'importe quelle machine. Cela a beaucoup ralenti les travaux du septième volume qui, néanmoins, a été imprimé. Désormais, nous pouvons effectuer la numérisation des données directement à Turin, ce qui nous permet ainsi de résoudre à la source de nombreux problèmes d'interprétation des transcriptions. Cette opération, de même, conduit à une meilleure organisation du matériel informatisé, nécessaire pour ses exploitations successives, par exemple pour la création de l'index<sup>6</sup> des formes, au sujet duquel la réflexion théorique conduite au sein de l'ALI n'a pas beaucoup avancé depuis les années 1970 (Genre/Massobrio 1976).

## L'Atlante Toponomastico del Piemonte Montano (ATPM)

L'ATPM est sans doute un des plus importants projets de collecte de données toponymiques et pas seulement au niveau italien, au vu du but qu'il se propose d'atteindre, à savoir la collecte des données toponymiques sur un territoire étendu et complexe du point de vue linguistique, et pour la persévérance avec laquelle ce but est poursuivi.

Le domaine concerné par le projet comprend le territoire des 558 communes piémontaises, regroupées dans les 48 communautés de montagne (*comunità montane*)<sup>7</sup> sur une étendue de 13 965,07 km<sup>2</sup>. Il concerne des zones où sont encore assez vivants les dialectes piémontais, ligure, occitan, franco-provençal, lombard et même germanique dans la partie nord de la région.

L'idée d'entamer un chantier de si grandes dimensions fut conçue par Arturo Genre à Turin, qui la présenta au cours du « Colloque international de Linguistique » qui se tint à Briançon en 1970. Dès lors, le projet mûrit au cours des années 1970 au sein du « Gruppo per l'indagine toponomastica », composé d'un certain nombre de

---

<sup>6</sup> Il ne faut pas oublier que la numérisation des données de l'ALI, comme l'a remarqué Pennisi (1991, 139), est postérieure à l'organisation des archives et elle a comme unique but de publier les cartes linguistiques. Par ailleurs, le souci des techniciens, comme l'a dit Bracaglia (1992, 620) était surtout d'ordre graphique, ce qui fait que ces matériaux peuvent se révéler assez compliqués à gérer.

<sup>7</sup> Une réforme institutionnelle vient de changer légèrement ces chiffres étant donné qu'on a réduit le nombre de communautés de montagne en regroupant différemment les communes, dont certaines n'ont d'ailleurs plus le droit d'être considérées comme « de montagne ».

professeurs et chercheurs de l'Université de Turin et qui trouva sa première formalisation en 1981, lorsqu'il fut accueilli dans le programme régional nommé « Alpi & Cultura », à l'occasion du colloque « Prospettive di ricerca sulle parlate alpine » à Turin.

Sur la base du plan de travail élaboré par Arturo Genre avec Daniele Jalla, du Bureau régional de la Culture fortement impliqué dans cette collaboration, avec la contribution de différentes personnes intéressées, soit au niveau universitaire, soit au niveau local, fut établie une convention officielle entre l'Université de Turin et la Région du Piémont. Au cours des années 1980, commença l'activité des enquêteurs qui reçurent une formation à la méthode de collecte et en 1990, la première monographie toponymique fut publiée. Depuis lors, on a atteint – sous la direction de Lorenzo Massobrio – le nombre de 36 volumes (en 2009) et on peut dire que l'idée conçue par le défunt Arturo Genre vit et même, prospère.

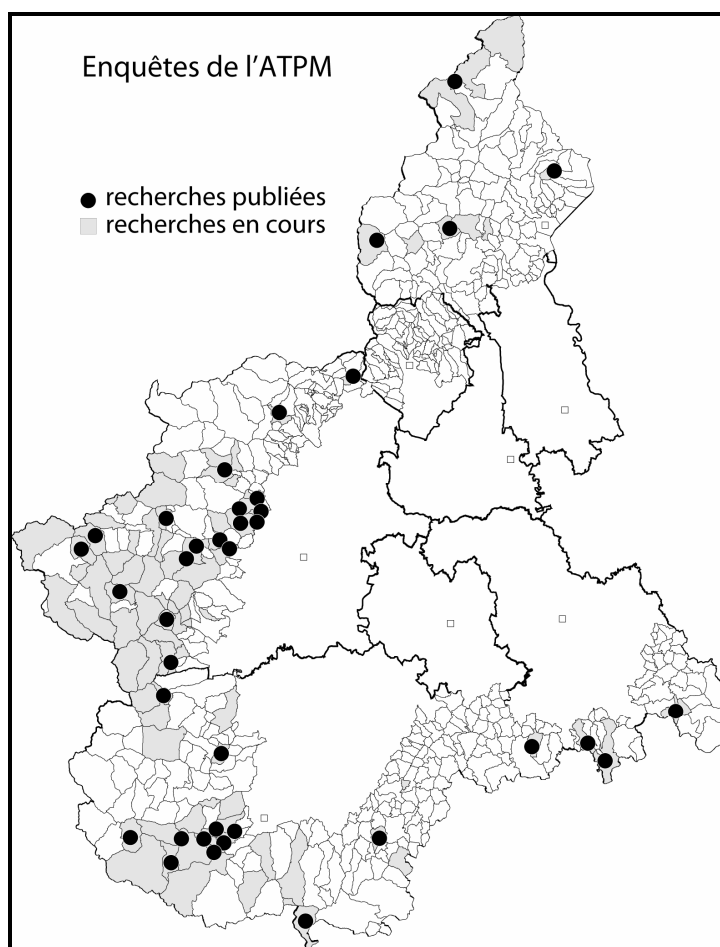


Fig. 3 Domaine d'enquête de l'ATPM (2009)

Mais quels sont les principes inspirateurs du projet de l'ATPM ? Au-delà de la nouveauté représentée par l'idée de mener systématiquement sur un territoire si grand des enquêtes toponymiques orales, ce qui, à mon avis, représente l'innovation la plus intéressante, c'est le choix d'impliquer dans l'action de collecte et de sauvegarde les instituts administratifs locaux (Communautés de Montagne, Communes, Parc Naturels), les associations culturelles ou bien les passionnés intéressés par la sauvegarde du patrimoine culturel. Ce choix se manifeste dans l'utilisation d'enquêteurs non professionnels, dont la caractéristique principale est celle d'être de très bons connaisseurs du milieu dans lequel ils vont enquêter, soit pour ce qui concerne le territoire et la culture rurale, soit pour ce qui concerne la langue. À ce propos on peut lire dans l'introduction des volumes :

“solo dei ricercatori che conoscono a fondo, oltre alla parlata locale, anche il territorio e la gente che vi abita possono raccogliere, accanto ai toponimi principali, nomi relativi a piccole o piccolissime località (...). Solo persone ben inserite e radicate nella realtà locale possono scegliere convenientemente i propri informatori, creando con loro un rapporto non solo di natura scientifica, ma in grado di stimolare una partecipazione attiva agli scopi della ricerca” (Genre/Jalla 1993, 7).

L'élaboration d'instruments assez simples a permis et permet encore à tous ceux qui sont engagés dans la collecte d'être très efficaces, même s'ils ne possèdent pas des notions spécialisées de dialectologie ou de phonétique. Le premier de ces instruments est le système de transcription phonétique précis (Genre, 1995<sup>a</sup>, 1995<sup>b</sup> et 1995<sup>c</sup>), mais suffisamment souple, pour l'enregistrement des données (qui d'ailleurs sont toujours vérifiées par la rédaction de l'ATPM à partir d'enregistrements sur bande magnétique, ou désormais, sur un format numérique). Le second instrument est constitué par les cartes topographiques et les fiches à remplir (elles aussi, désormais numérisées) avec toutes les informations qui doivent accompagner la dénomination du lieu.

Le lien des enquêteurs avec le territoire permet *ipso facto* l'accès à de nombreuses informations qui demeureraient inaccessibles à un enquêteur étranger au réseau local. Les instruments de récolte ainsi conçus présentent en plus un autre avantage qui n'est pas négligeable : ils constituent un filtre qui se place tout naturellement entre l'enquêteur et la culture qui est l'objet de l'enquête. Ils permettent

d'éviter – en partie, du moins – la confusion potentielle entre l'enquêteur et l'objet de l'enquête, ce qui pourrait mener à une interprétation plus ou moins biaisée des informations.

Les archives de l'ATPM comptent aujourd'hui environ 60 000 fiches recouvrant autant de dénominations de lieux. Il ne s'agit pas encore, bien sûr, d'archives homogènes, d'une part, parce que l'uniformisation stricte selon les critères rédactionnels a été faite jusque-là seulement sur les fiches qu'on a publiées dans les volumes<sup>8</sup>, et d'autre part, parce que le territoire, comme on le voit sur la figure 3, n'est couvert jusque-là que d'une façon discontinue et, de plus, la densité des données collectées par rapport à la surface est assez inégale. Il s'agit néanmoins d'archives qui permettent déjà de tirer quelques considérations sur le rôle que peuvent avoir les données toponymiques dans l'étude du processus culturel qui a transformé le territoire en paysage<sup>9</sup>. D'un côté à travers la domestication de l'espace, qui s'explique dans la mise en valeur des sols et des ressources en général, et de l'autre à travers la cristallisation des façons de percevoir ce même espace, qui se manifeste dans l'attribution d'une valeur symbolique aux lieux. L'espace n'est plus neutre étant donné qu'aux lieux se rattache la mémoire de la communauté, ensemble apparemment désordonné de souvenirs singuliers. Opportunément intégrées, ces archives permettent aussi de fournir des clés de lecture de l'espace linguistique très intéressantes – au-delà des sentiers qui sont habituellement battus quand on prend en compte la toponymie – des clés pour mieux préciser la vision de certains phénomènes.

À ce propos, une importante évolution, dans la direction déjà suivie par le *Repertorio Toponomastico Trentino*, et envisagée aussi par d'autres<sup>10</sup> en Italie, est en train de se mettre en place en ce qui concerne la façon d'organiser les archives informatiques (Massobrio/Chiapusso/Monteleone 2007) et l'établissement de la qualité des données fournies. Il s'agit d'élaborer un système SIG (Système d'Information Géographique)<sup>11</sup> à travers un codage

---

<sup>8</sup> Il faut savoir que la structure de l'ATPM est très légère : trois personnes qui travaillent aux archives et sont en charge de toute opération préalable à la publication des enquêtes, de la relation avec les enquêteurs jusqu'à la mise en page du texte et des cartes ; avec une coordinatrice qui s'occupe de tous les rapports administratifs et de la planification de l'activité éditoriale.

<sup>9</sup> Sur ce point, cf. Hunn et al. (2010).

<sup>10</sup> *Repertorio Toponomastico Ticinese* (Vassere, 2006), *l'Atlas Toponomastique de la Vallée d'Aoste* (Favre, 2006), le projet *Flurnamenerhebung in Südtirol* (Kollmann, 2006).

<sup>11</sup> A l'origine GIS (*Geographical Information System*) ou, dans la version italienne SIT (*Sistema*

numérique de toute information territoriale, située dans l'espace à l'aide de coordonnées géographiques. Cela permet le dialogue entre différentes bases de données, par exemple, pour confronter une carte de la diffusion d'un certain type de végétal à une carte toponymique des dénominations phytonymiques locales. Cette démarche permettra, enfin, un réel « regard croisé » (Ranucci/Dalbera 2004), capable de montrer la richesse d'un patrimoine culturel important non seulement pour les linguistes, mais aussi pour les géologues, géographes, biologistes, ethnologues, et plus largement pour toutes les disciplines concernées par les *biens culturels environnementaux* (Genre/Jalla 1993). Malheureusement on est encore assez loin de cela, à cause des difficultés techniques et de l'argent nécessaire pour élaborer un système assez sophistiqué.

En attendant un instrument suffisamment puissant, je me limiterai, à simple titre d'exemple, à proposer quelques notes sur un cas d'étude qui, à mon avis, permettra de saisir la potentialité du patrimoine toponymique collecté jusque-là.

Je commencerai par mettre en relief deux évidences qui sont à la base de mon approche. S'il est vrai que le système toponymique propre à une communauté est constitué par des noms pour lesquelles un processus de "cristallisation" a amené, du point de vue sémiotique, à l'établissement d'une relation directe entre le signifiant et le référent se passant à la limite du signifié, on remarque que ce processus n'a pas complètement effacé ce dernier. En effet, surtout dans le contexte d'une culture traditionnelle s'exprimant essentiellement dans l'oralité, le système toponymique est constitué de noms dont la signification est pour la plupart encore transparente (Caprini 2001; Marrapodi 2006, Bessat/Germi 1993)<sup>12</sup>.

La seconde considération concerne ce même rapport en évolution, considéré à la lumière de la vision du monde des usagers de ces signes, à savoir la façon dont ils classent l'univers qui les entoure, ou – mieux – aux façons de le nommer qui se sont surajoutées les unes

---

*Informativo Territoriale*).

<sup>12</sup> Caprini (2007: 230) va plus loin en affirmant que c'est seulement « l'obbligo di *trascrivere* toponimi e antroponimi che porta alla loro cristallizzazione in forme che con il passare del tempo si rivelano opache ». L'affirmation, même si on peut la partager sur un niveau général, doit être néanmoins atténuée, à la fois parce qu'on a pu remarquer qu'il y a des toponymes qui, tout en n'ayant jamais été écrits, ont subi un processus de "cristallisation" qui les a rendus obscurs et, par contre, il y en a d'autres, que l'on peut trouver écrits depuis des siècles, et qui pourtant conservent leur signifié originel. Comme l'écrit d'ailleurs Ranucci (2004, 205), « si le toponyme relève d'abord, nous semble-t-il, de l'oralité, il est également et depuis longtemps un fait de l'écrit ».

aux autres. L'enquête toponymique orale, conduite à travers une conversation qui a pour objet le territoire – c'est-à-dire la façon de le percevoir<sup>13</sup> – permet en effet de saisir avec précision les catégories taxonomiques activées dans cette perception. Elle permet aussi d'essayer de le traduire dans la langue nationale, surtout si la tâche de l'enquêteur, qui est aussi traducteur, est accomplie par quelqu'un qui dispose d'un accès privilégié à cet univers culturel.

Le cas sur lequel je vais focaliser mon attention est celui des désignations des thalwegs et des cours d'eau permanents ou intermittents qui les empruntent, et des canaux artificiels. Le choix n'a pas été fait au hasard. Il s'agit en effet d'un cas qui offre deux importants atouts de mon point de vue. Il peut en premier lieu être considéré comme paradigmatique par rapport aux considérations que je viens de faire et, de plus, il a déjà été étudié par Jean-Claude Bouvier dans deux contributions, l'une consacrée au domaine provençal français, à partir des données de l'ALF et de l'ALP (Bouvier 1974), et l'autre spécifiquement aux vallées piémontaises décrites dans l'*Atlante Linguistico ed Etnografico del Piemonte Occidentale*, de même que par l'ATPM (Bouvier 1985). Cela permet en même temps de tester la validité heuristique de mon approche et de montrer que la ressource toponymique peut ajouter quelque chose à la lecture des cartes des atlas régionaux dans un domaine très spécifique comme celui du lexique géographique.

Bouvier (1974) a abordé l'analyse de la microstructure lexicale des noms relatifs aux cours d'eau en considérant en même temps trois types de réalités: les cours d'eau naturels, les canaux artificiels (pour l'irrigation ou pour l'alimentation des moulins à eau, usine de forge, etc.) et les vallons. Une fois qu'a été construit un « axe sémiolexical » qui regroupe les trois concepts, il a été possible de comprendre le fonctionnement des systèmes actuels et d'en dégager les possibles motivations sur la base des variations qu'on reconnaît dans l'espace. Par la lecture des cartes linguistiques dans le domaine français, puis dans le domaine piémontais, Bouvier a pu montrer comment à un microsystème original dont on peut reconnaître les traces dans des régions marginales du domaine considéré se substituent d'autres formes de relations entre signe, concept et réalité géographique. Dans ce système l'axe sémantique est ainsi découpé : le type /'riju/, /'rjeu/ (<

---

<sup>13</sup> Sur ce point, cf. Lurati (2004, 10).



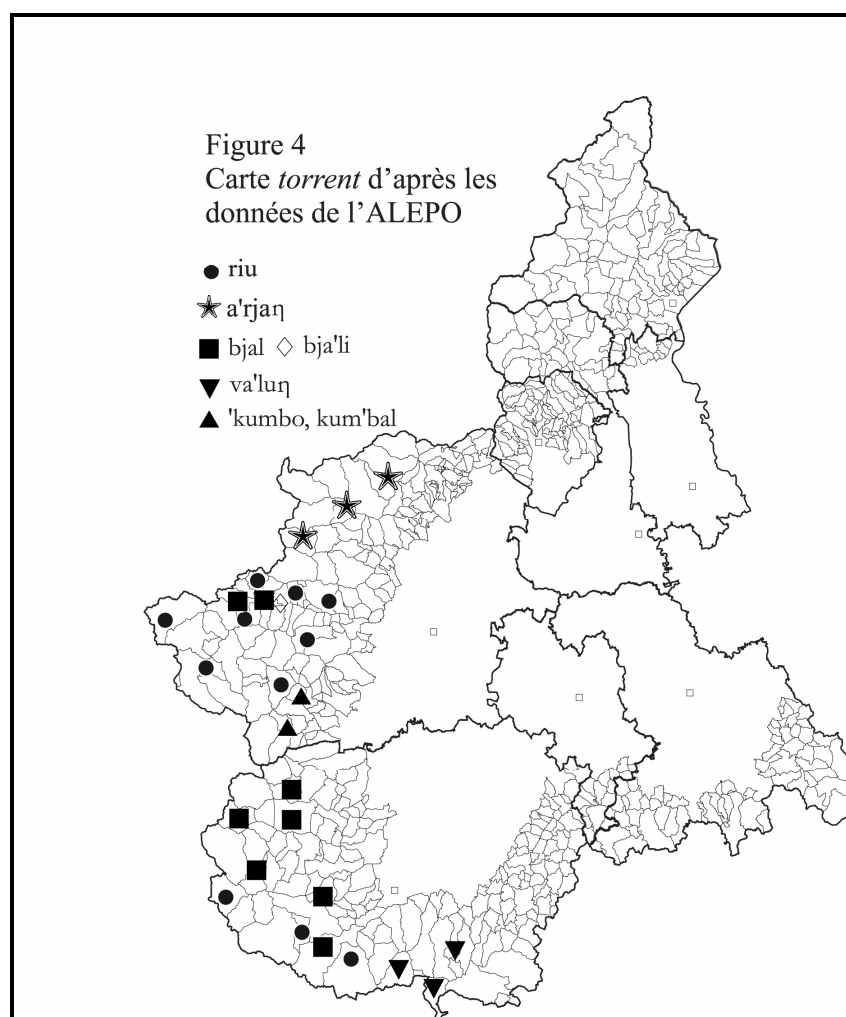
lat. RIVUM) indique le ‘ruisseau’, le type /bjal/, /be'al/, /be'zal/ (< gaul. BEDO) tout canal d’arrosage ou d’alimentation d’origine artificielle et les types /va'la/, /va'lun/ (< lat. VALLIS ou VALLUM) d’un côté et /'kumbo/, /'kumba/ (< gaul. KUMB-) de l’autre indiquent le vallon ou le ravin<sup>14</sup>. A partir de ce schéma, nous avons, en résumant, un recul du type /'riɥ/ et une avancée du type /bjal/. Dans la partie septentrionale du domaine provençal, c’est ce dernier qui va prendre la place de /'riɥ/ pour indiquer le ‘ruisseau’ — grâce à une superposition partielle des deux concepts ‘petit ruisseau’ et ‘canal d’arrosage’, étant donné l’étroite relation entre les deux —, tandis que pour indiquer le ‘canal’, on a recours à un autre matériel lexical. Dans la partie méridionale, par contre, les deux concepts qui vont se superposer sont celui de ravin et celui de ruisseau : ici /va'la/ vient indiquer le ruisseau, tandis que /bjal/ indique toujours le canal artificiel. Dans ce cas, la superposition est favorisée par le fait naturel — bien noté par l’école de géographie humaine de Grenoble — que dans une Provence à fort *ravinement* « en définitive les notions de ‘ravin’ et de ‘ruisseau’ se laissent assez mal dissocier l’une de l’autre » (Bouvier 1974, 69).

Une situation tout à fait analogue se retrouve sur le versant italien des Alpes dites sèches (une présentation circonstanciée se trouve dans Cabouret, 1999, 25) (fig. 4 et 5). Ici aussi, le domaine de /'riɥ/ est limité aux vallées occitanes (Vallées de la Germanasque, du Cluson et de Suse) et à la haute Vallée de la Stura de Demonte, ce qui n’apparaissait pas dans Bouvier (1985), à cause des données partielles utilisées (les matériaux n’étant pas encore entièrement disponibles). Dans ces vallées, le système est clairement dans son état de départ : /'riɥ/ ‘ruisseau’, /bjal/ ‘canal’ et /'kumbo/ ‘vallon’. Dans les vallées comprises entre ces deux domaines on a, en gros, les deux évolutions possibles. Bouvier montrait comment dans la Vallée du Pélias (Villar Pellice) on avait le même type de rencontre « sémiolinguistique » qu’en Provence, aboutissant à une neutralisation entre le concept de ‘cours d’eau intermittent’ et ‘cours d’eau permanent’ : ce qui entraîne la

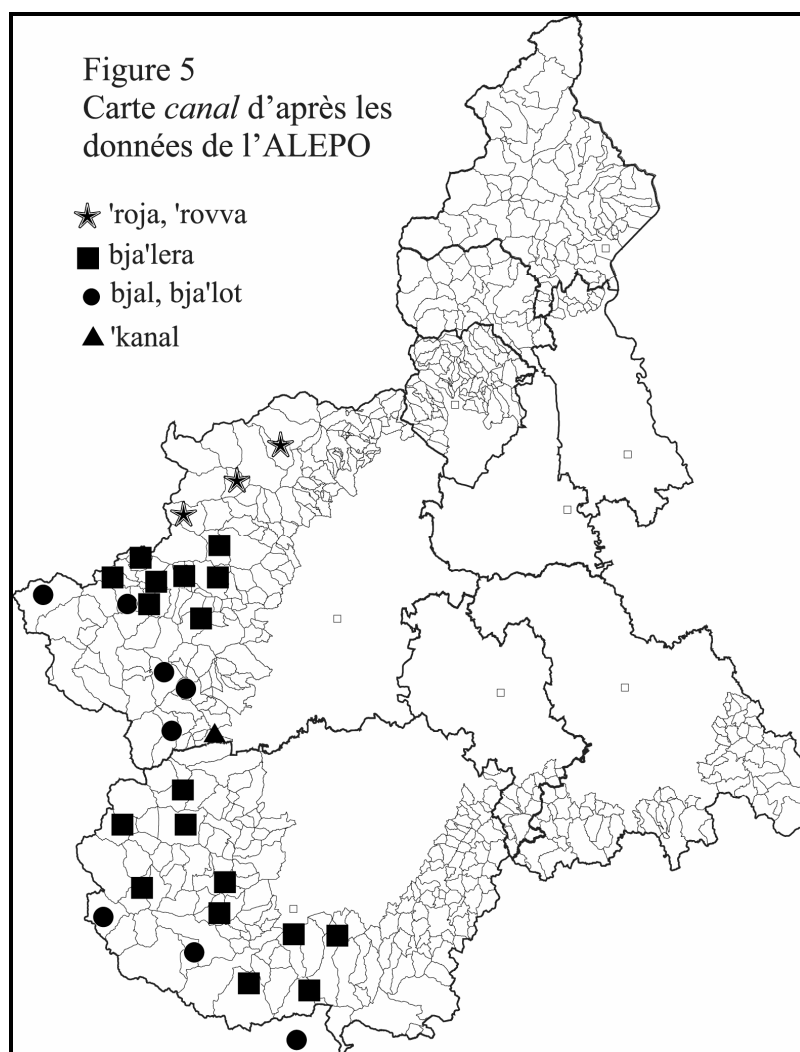
---

<sup>14</sup> Ce schéma simplifié ne tient pas compte d’autres termes, appartenant à des couches linguistiques plus anciennes ou de diffusion plus limitée, comme *nant* ‘torrent’, mot savoyard avec une diffusion au niveau toponymique plus étendue, ou *doron* ‘torrent’, avec les variantes *doire*, *dore* etc., les deux étudiés par Bessat et Germi (Bessat/Germi 2001 : 137 et ss.).

spécialisation des successeurs de CUMBA, à l'origine 'vallon', 'ravin', dans le sens de 'ruisseau'. Sur la base des données disponibles il s'agissait apparemment d'une exception (même si on avait /'kumbo/ 'torrent' pas très loin, à Pramollo). Ainsi tout au sud on retrouvait la même situation qu'en Provence, où /va'la/, /va'luŋ/ indique le 'ruisseau'. Dans les autres vallées centrales, par contre, c'est /bjal/ dans le sens de 'ruisseau' qui l'emportait, et par conséquent, on introduisait /bja'ljero/, /bja'lera/ etc. au sens de 'canal'<sup>15</sup>. Tout cela n'est bien sûr que la version très simplifiée du raisonnement, qui faisait face, comme tout géolinguiste peut facilement l'imaginer, à une plus grande complexité de la réalité langagière.



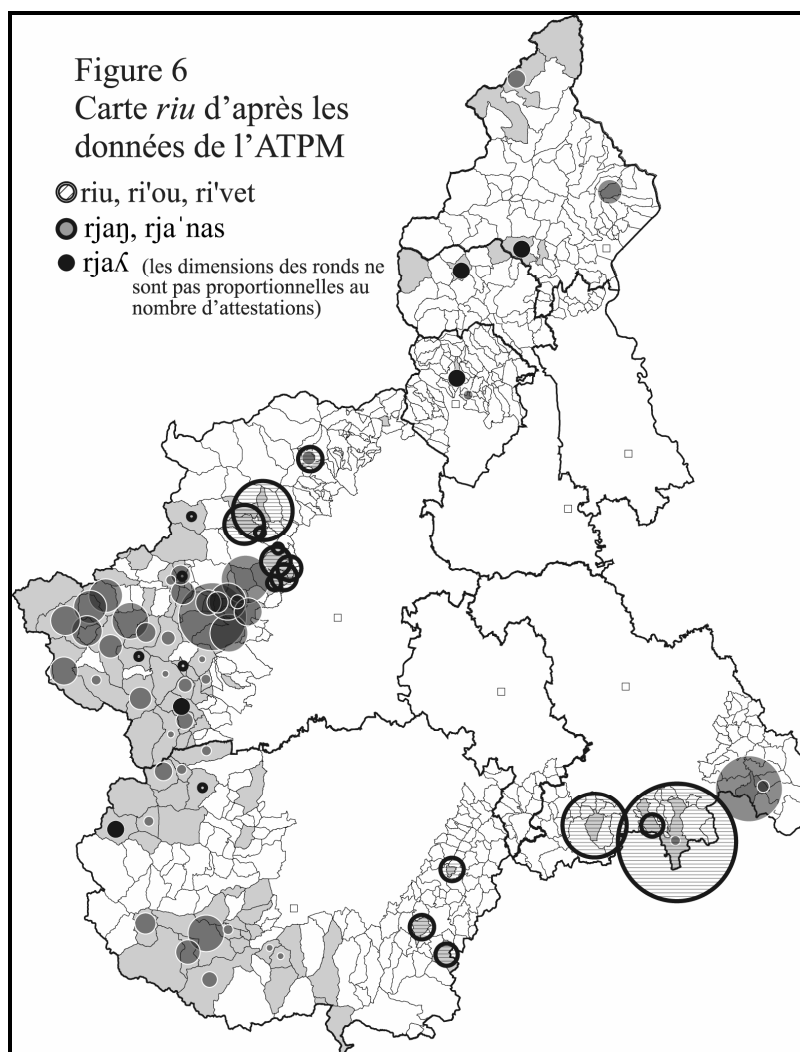
<sup>15</sup> Sur l'origine et la diffusion piémontaise des continuateurs de BEDO, cf. Massobrio 1982.



Dans ce cadre, l'apport du matériel collecté par les enquêteurs de l'ATPM contribue, à mon avis, à confirmer les hypothèses élaborées par Bouvier et, en même temps, à être plus précis dans plusieurs directions. D'abord, comme on s'y attend en toponymie, on trouve dans les noms de lieux des attestations qui témoignent d'une extension d'une ancienne couche lexicale plus étendue qu'aujourd'hui. C'est le cas de /riʊ/ qui, bien que présent seulement dans certaines vallées, est tout de même attesté un peu partout, signe que jadis sa vitalité devait être plus grande<sup>16</sup>. Une situation semblable se retrouve pour le type /va'lun/, dont la diffusion, loin de montrer sa provenance 'italienne' (Calleri 1985), semble, elle aussi, appartenir à une couche lexicale

<sup>16</sup> La présence de termes à la forme diminutive comme *rivet* nous permet d'exclure dans la plupart des cas qu'il s'agisse d'un calque du latin des documents, où le ruisseau est un *rivus*.

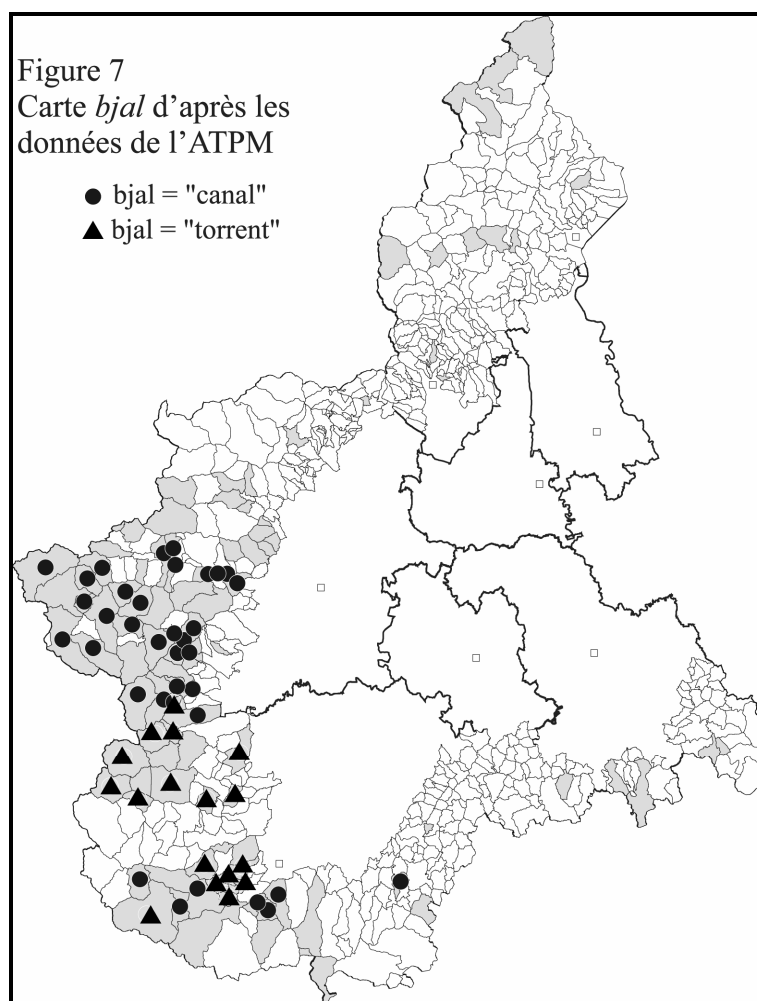
plus ancienne, remplacée par /'kumbo/. On notera, en passant, que les attestations de /'riu/ sont d'habitude encore transparentes pour les patoisants.



Mais le réseau plus serré des relevés de l'ATPM (auxquels s'ajoutent les attestations toponymiques fournies par un certain nombre de publications à diffusion locale<sup>17</sup>) nous permet aussi d'insérer le cas isolé représenté par le /kum'bal/ 'ruisseau' de Villar Pellice dans un cadre beaucoup plus étendu. Il ne s'agit pas d'une

<sup>17</sup> Artusio *et Alii* (2005), Bernard (2004), Bruna Rosso (1980), Di Maio (1988), (2000), (2001), Garcin/Souberan/Di Maio (2003), Garibaldo (2002), Guiguet/Gallizio/Di Maio (2003), Hirsch (1955), (1958), (1963), (1980), (1983), Ottonelli (1990), Paseri (2008), Sibille (2004), Sibille/François (2006).

« exception qui confirme la règle », mais du représentant d'une aire qui s'étend de la basse Vallée du Cluson, jusqu'à la haute Vallée du Pô et à la basse Vallée Varaita (où, comme dans une partie de la Vallée du Pélias, l'on a soit /kum'bal/ 'petit ruisseau encaissé', soit /bjal/ 'torrent, ruisseau').



Dans les trois cartes proposées par Bouvier (1985), si celles qui sont relatives au 'torrent' et au 'canal' montrent d'une façon assez claire ce qu'on vient de résumer et se laissent compléter aisément par les données toponymiques, celle qui est relative au 'ruisseau' pose quelques problèmes. On voit ici qu'en de nombreux points le terme est le même que celui de 'torrent' (c'est le cas d'une partie de la Vallée de Susse et des Vallées vaudoises) ou le même que celui qu'on utilise pour 'canal', avec plusieurs cas où la suffixation à laquelle on recourt pour maintenir une distinction signale, à mon avis, une difficulté de

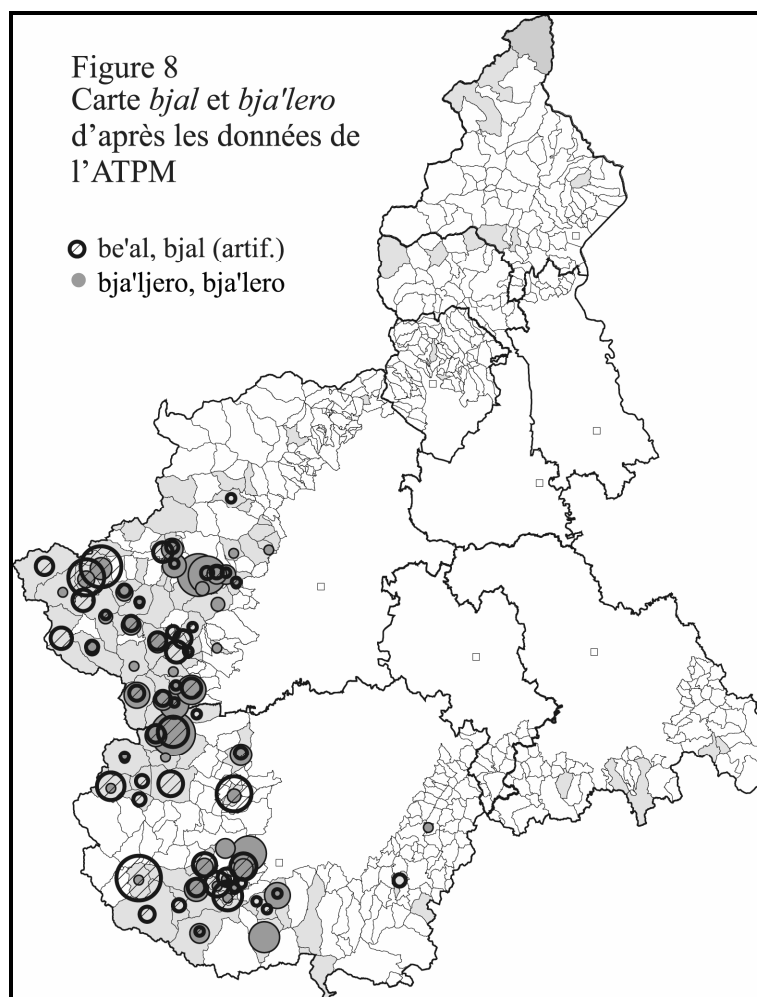
l'informateur à saisir correctement la question. Apparemment, le questionnaire utilisé par l'ALEPO est très précis : on demande comment l'on appelle le 'fleuve', le 'torrent', le 'ruisseau (dans les prés)' et le 'ruisseau creusé par l'homme' ou, plus loin, le 'canal d'arrosage'. Mais, si pour répondre à la question 'fleuve' ('fiume')<sup>18</sup> on utilise, comme on pouvait s'y attendre, soit un calque de l'italien, soit le toponyme du cours d'eau le plus important (/ˈstyro/, /vaˈratʃo/, /po/, /ˈpeli/, etc.), lorsqu'il s'agit de distinguer entre un 'torrent' ('torrente') et un 'ruisseau (dans les prés)' ('ruscello nei prati') on a du mal à distinguer les deux réalités<sup>19</sup>. Je doute que les catégories taxonomiques suggérées soient toujours pertinentes et il me semble que parfois l'enquêteur n'a pas été capable de saisir le décalage entre l'hypothèse de classement suggérée par le questionnaire et la réalité. La distinction entre 'torrent' et 'ruisseau' proposée par le questionnaire repose apparemment sur l'idée qu'est pertinente la distinction entre les traits +/- grand, +/- impétueux. Cette hypothèse est d'ailleurs vérifiée dans certains cas, ceux où l'on a distingué (ou pas) les réalités sur ce plan (exemple : là où l'on oppose /bjal/ à /bjaˈlot/, ou /kumˈbal/ à /kumbaˈlot/, ou même là où l'on ne distingue pas et on n'a que /riu/). Mais en réalité la spécification 'dans les prés', au lieu de servir à éclaircir le concept, pour certains, semble l'avoir rapproché de celui de 'ruisseau creusé par l'homme', pour le simple fait que sur les pentes les plus abruptes de nos vallées, seuls les canaux entretenus peuvent demeurer au milieu des prés sans les ravager, alors que tout ruisseau naturel a, dans le temps, creusé sa combe. Là où on a répondu /bjaˈljero/, comme le suggèrent certaines notes, on a indiqué les canaux artificiels et le fait que les quelques 200 attestations de ce type dans les archives de l'ATPM sont toutes traduites par canal d'arrosage<sup>20</sup> (sauf dans le cas de Boves, point

<sup>18</sup> L'italien 'fiume' indique un cours d'eau d'une certaine importance indépendamment du fait qu'il se jette dans la mer (fleuve), ou qu'il ne s'y jette pas (rivière).

<sup>19</sup> On le voit d'ailleurs dans les commentaires des informateurs : à Novalèse à la question 'torrent', il termine par : « per torrente non c'è, consideriamo torrenti la Cenischia, abbiamo solo *gurdzè* che è il ruscello » ; à Lemie : on signale qu'un ruisseau dans les prés est le même qu'un ruisseau creusé par l'homme ; à Entraque à la question 'ruisseau (dans les prés)' : « *na bjalère faite da l omə* » ; à Sampeyre, à la même question on précise : « *na bjero (fatto dall'uomo)/ən bjalot (naturale)* ».

<sup>20</sup> Dans ce cas le suffixe -ARIA a une valeur d'agent, donc *bjal* + -ARIA « ce qui sert à arroser ». C'est pourquoi on ne peut pas le considérer avec les autres, à valeur diminutive ou augmentative, ou bien péjorative, très utilisés pour préciser les dimensions ou la mauvaise position ; ainsi -ALIS

d'enquête de l'ALEPO) semble le prouver (v. fig. 8, où les attestations du type *bjaljero* sont présentées en concomitance avec celle de /bjal/ dans le sens de canal).



Le décalage entre les catégories taxonomiques de la culture de l'écrit, bien que reformulées à la lumière des connaissances ethnolinguistiques qui font désormais partie du bagage des faiseurs d'atlas, n'a en tout cas pas été éliminé par la démarche adoptée par l'enquête toponymique. On le voit assez bien au moment où les chercheurs, avec leurs informateurs, essayent de rendre en italien et à l'écrit l'exact signifié du mot /kum<sup>1</sup>bal/ dans une des vallées où ce terme désigne le thalweg, le ravin et, en même temps, l'eau qui coule dans le fond (mais qui d'ailleurs pourrait aussi ne pas y être). Et je dis

---

modifie CUMBA en le changeant de série.

bien en italien, c'est-à-dire dans une autre langue, avec d'autres termes, et à l'écrit. Car c'est parfois plutôt le passage de l'oral à l'écrit qui pose des difficultés, étant donné que le propre de l'écrit est une classification plus rigide, alors que, dans une situation orale, on peut utiliser de catégories plus souples — ce qui ne signifie pas qu'elles soient incohérentes ou qu'elles créent de l'ambiguïté dans le contexte communicatif. Cela est d'autant mieux vérifié dans les termes relatifs aux accidents géographiques. Dans ce cas en effet, la tentative de traduction d'un terme donné va souvent se superposer à la description analytique d'une des réalités auxquelles le signe linguistique peut se référer. Autrement dit, dans un milieu où tout ruisseau, torrent, vallon ou rigole a, ou a eu, son nom, le rapport entre signe, signifié et référent subira – en partie, du moins – ce processus de recodage dont on a parlé plus haut qui entraîne les appellatifs géographiques une fois qu'ils sont liés à un endroit précis.

Pour conclure, il resterait – entre autres – à considérer l'histoire des courants culturels qui ont participé à la définition actuelle. Mais pour aller plus loin avec une assise suffisante, il faudrait pouvoir disposer de données toponymiques plus complètes, non seulement du côté italien des Alpes, mais aussi dans la plaine piémontaise et, encore plus, dans les Alpes françaises<sup>21</sup> comme le soutenait déjà en 1993 Arturo Genre (Genre 1993, 24), suivi en cela, mais dans une perspective encore plus englobante, par Jean-Claude Bouvier (Bouvier 2003 [1998], 89), qui préconisait une toponymie romane comparée.

---

<sup>21</sup> Pour ce domaine, on dispose des deux volumes de l'*Atlas Toponymique Savoie, Vallée d'Aoste, Dauphiné, Provence*, intitulés *Noms du paysage alpin* et *Noms du patrimoine alpin* de Hubert Bessat et Claudette Germe (Bessat/Germe 2001 et 2004). Avec Bessat/Germe (1993), ces ouvrages offrent un cadre d'ensemble très bien tracé, qui laisse ouvertes de nouvelles possibilités d'investigation.



## Références bibliographiques

- AIS = Jaberg, K./Jud, J. (1928-1940), *Sprach- und Sachatlas Italiens und der Südschweiz*, 8 volumes, Zofingen.
- ALF = Gilliéron, J./Edmont, E. (1902-1910), *Atlas Linguistique de la France*, 18 volumes, Paris.
- ALEPO = Telmon, T./Canobbio, S., *Atlante Linguistico ed Etnografico del Piemonte Occidentale*, en cours de réalisation (3 voll. publiés), Università di Torino.
- ALI 1971 = Bartoli, M. *et al.* (1971), *Questionario dell'Atlante Linguistico Italiano. I,a – Testo*, Genre, A. *et al.* (a c. di), Torino, Istituto dell'Atlante Linguistico Italiano.
- ALI 1973<sup>a</sup> = Bartoli, M. *et al.* (1973), *Questionario dell'Atlante Linguistico Italiano. II – Indice*, Genre, A. *et al.* (a c. di), Torino, Istituto dell'Atlante Linguistico Italiano.
- ALI 1973<sup>b</sup> = Bartoli, M. *et al.* (1973), *Indice delle inchieste dell'Atlante Linguistico Italiano*, Genre, A. *et al.* (a c. di.), Torino, Istituto dell'Atlante Linguistico Italiano.
- ALI 1980 = Bartoli, M. *et al.* (1980), *Questionario dell'Atlante Linguistico Italiano. I,b – Illustrazioni (Parte Generale)*, Genre, A. *et al.* (a c. di), Torino, Istituto dell'Atlante Linguistico Italiano.
- ALI 1981 = Bartoli, M. *et al.* (1981), *Questionario dell'Atlante Linguistico Italiano. I,b – Illustrazioni (Parti Speciali)*, Genre, A. *et al.* (a c. di), Torino, Istituto dell'Atlante Linguistico Italiano.
- ALI 1985 = Bartoli, M. *et al.* (1985), *Questionario dell'Atlante Linguistico Italiano. III – Indice inverso*, Massobrio, L., Nosengo, M.C. (a c. di), Torino, Istituto dell'Atlante Linguistico Italiano.
- ALI 1995<sup>a</sup> = Bartoli, M. *et al.* (1995), *Atlante Linguistico Italiano. Verballi delle inchieste*, Massobrio, L. *et al.* (a c. di), Roma, Istituto Poligrafico e Zecca dello Stato, 2 volumi.
- ALI 1995<sup>b</sup> = Bartoli, M. *et al.* (1995), *Atlante Linguistico Italiano*, vol. I: « Il corpo umano I », Massobrio, L. *et al.* (a c. di), Roma, Istituto Poligrafico e Zecca dello Stato.
- ALI 1996 = Bartoli, M. *et al.* (1996), *Atlante Linguistico Italiano*, vol. II: « Il corpo umano II », Massobrio, L. *et al.* (a c. di), Roma, Istituto Poligrafico e Zecca dello Stato.
- ALI 1997 = Bartoli, M. *et al.* (1997), *Atlante Linguistico Italiano*, vol. III: « Indumenti e abbigliamento », Massobrio, L. *et al.* (a c. di), Roma, Istituto Poligrafico e Zecca dello Stato.

- ALI 1999 = Bartoli, M. *et al.* (1999), *Atlante Linguistico Italiano*, vol. IV: « La casa e l'arredamento », Massobrio, L. *et al.* (a c. di), Roma, Istituto Poligrafico e Zecca dello Stato.
- ALI 2001 = Bartoli, M. *et al.* (2001), *Atlante Linguistico Italiano*, vol. V « La casa e l'arredamento », Massobrio, L. *et al.* (a c. di), Roma, Istituto Poligrafico e Zecca dello Stato.
- ALI 2003 = Bartoli, M. *et al.* (2003), *Atlante Linguistico Italiano*, vol. VI « L'alimentazione », Massobrio, L. *et al.* (a c. di), Roma, Istituto Poligrafico e Zecca dello Stato.
- ALI (2008) = Bartoli, M. *et al.* (2008), *Atlante Linguistico Italiano*, vol. VII « La vita in famiglia e le età dell'uomo », Massobrio, L. *et al.* (a c. di), Roma, Istituto Poligrafico e Zecca dello Stato.
- Artusio, L. et Alii (2005), *Disiounari ousitan. Roubilant – Roucavioun*, Roccabruna/La rocha.
- ATPM = *Atlante Toponomastico del Piemonte Montano*: [1] *Gaiola*, Alessandria, 1999<sup>2</sup>; [2] *Aisone*, Torino, 1993, [3] *Mombasiglio*, Torino, 1993; [4] *Quassolo*, Torino, 1993, [5] *Chianocco*, Alessandria, 1995; [6] *Roccasparvera*, Alessandria 1995; [7] *Givoletto*, Alessandria 1997; [8] *La Cassa*, Alessandria, 1997; [9] *Val della Torre*, Alessandria, 1997; [10] *Vallo*, Alessandria, 1997; [11] *Varisella*, Alessandria, 1997; [12] *Demonte*, Alessandria, 1997; [13] *Ostana*, Alessandria, 1998; [14] *Pont Canavese*, Alessandria, 1998; [15] *Parco naturale Alpe Veglia e Alpe Dévero*, Alessandria, 2000; [16] *Rittana*, Torino, 2000; [17] *Avigliana*, Torino, 2001; [18] *Sant'Antonino di Susa*, Torino, 2001; [19] *Valloriate*, Torino, 2001; [20] *Salbertrand*, Torino, 2002; [21] *Coazze*, Torino, 2002; [22] *Mezzenile*, Torino, 2003; [23] *Rorà*, Torino, 2003; [24] *Chiusa San Michele*, Torino, 2004; [25] *Pramollo*, Torino, 2005; [26] *Moiola*, Torino, 2005; [27] *Roccaforte Ligure*, Torino, 2005; [28] *Briga Alta*, Torino, 2006; [29] *Sambuco*, Torino, 2006; [30] *Exilles*, Torino, 2006; [31] *Rimella*, Torino, 2007; [32] *Alagna*, Torino, 2007; [33] *Falmenta*, Torino, 2007; [34] *Morbello*, Torino, 2008; [35] *Capanne di Marcarolo*, Torino, 2008; [36] *Venasca*, Torino, 2008; [37] *Tagliolo*, Torino, 2009; [38] *Massello*, Torino, 2009.
- Bernard, G. (2004), "Bellino: passeggiata tra i nomi di luogo", dans *Valados Usitanos*, n. 77, p. 45-70.
- Bessat, H./Germi, C. (1993), *Noms en mémoire de l'Alpe*, Grenoble, Ellug.

- Bessat, H./Germi, C. (2001), *Les noms du paysage alpin*, Grenoble, Ellug.
- Bessat, H./Germi, C. (2004), *Les noms du patrimoine alpin*, Grenoble, Ellug.
- Blasco Ferrer, E. (1999), "L'ALI e la linguistica italo romanza", dans *Romance Philology*, n. 52, p. 57-87.
- Bouvier, J.-C. (1974), "Les dénominations du « ruisseau » dans les parlers provençaux", dans *Revue de Linguistique Romane*, n. 38, p. 59-74.
- Bouvier, J.-C. (1985), "Le ruisseau alpin – essai d'analyse sémiologique des dénominations du ruisseau dans les parlers de *L'atlante linguistico e (sic) etnografico del Piemonte occidentale*", dans Canobbio/Telmon (éd. par), p. 53-70.
- Bouvier, J.-C. (2003 [1998]), "Pour une toponymie romane comparée", dans, *Colligere atque tradere. Études d'ethnographie alpine et de dialectologie francoprovençale. Melanges offerts à Alexis Bétemps*, BREL.
- Bracaglia, G. (1992), "La cartografazione computerizzata delle carte dell'Atlante Linguistico Italiano", dans Ruffino (éd. par), p. 299-301.
- Bruna Rosso, A. (1980), *Dizionario del dialetto occitano di Elva*, Gaiola.
- Cabouret (1999), *L'irrigation des prés de fauche en Europe occidentale, centrale et septentrionale : essai de géographie historique*, Paris, Karthala Editions.
- Calleri, D. (1985), "Il rendimento di tre termini geografici nell'Alepo: vallone, indritto e inverso" dans Canobbio/Telmon (éd. par), p. 43-52.
- Canepari, L. (2003), *Manuale di fonetica*, München, Lincom.
- Canobbio, S./Telmon, T. (1985) (éd. par), *Atlante Linguistico ed etnografico del Piemonte occidentale (ALEPO), Materiali e saggi 1984*, Torino.
- Caprini, R. (2001), *Nomi propri*, Alessandria, Edizioni dell'Orso.
- Caprini, R. (2007), [Compte rendu de] Marrapodi, G. (2006), *Teoria e prassi dei sistemi onimici popolari: la comunità orbasca (Appennino Ligure centrale) e i suoi nomi propri*, QuRION, Roma, in *Rivista Italiana di Onomastica*, n. 13/1 (2007), p. 228-232.
- Contini, M. (1987), *Étude de géographie phonétique et phonétique instrumentale du Sarde*, Alessandria, Edizioni dell'Orso.

- Di Maio, M. (1988), *Vaïi, gias e vastére. Toponomastica nel Massiccio del Marguareis-Mongioie*, Torino, Valados Usitanos.
- Di Maio, M. (2000), *Guida dei toponimi di Bardonecchia e frazioni. Parte 1a: Bardonecchia*, Pinerolo, Alzani.
- Di Maio, M. (2001), *Guida dei toponimi di Melezet, Les Arnauds, Valle Stretta*, Pinerolo, Alzani.
- Favre, S. (2006), "Obiettivi raggiunti e problemi aperti della ricerca toponomastica valdostana", dans Flöss (éd. par), p. 37-46.
- Flöss, L. (2006), (éd. par), *Genius Loci. I nomi di luogo dalle fonti antiche alle banche-dati attraverso la tradizione popolare*, Atti del Convegno di Trento 2004, Trento.
- Franceschi, T. (2006), "Sèdano – prezzémolo (Ombrellifere); origano – persa – basilico (Labiata)", dans *Bollettino dell'Atlante Linguistico Italiano*, III Serie, n. 30, p. 45-56 [maintenant aussi dans *Temistocle Franceschi, Pagine sparse*, (éd. par A.M. Mancini et M.V. Miniati), Alessandria, Edizioni dell'Orso, 2009, avec le titre "Una fratellanza occulta: sèdano e prezzemolo" et quelques corrections].
- Garcin, A./Subeiran, L./Di Maio, M. (2003), *Guida dei toponimi di Rochemolles*, Pinerolo, Alzani.
- Garibaldo, D. (2002), *Guida dei toponimi di Millaures*, Pinerolo, Alzani.
- Genre, A. (1993), "Les recherches en microtoponymie au Piémont: l'Atlante Toponomastico del Piemonte Montano", in *Nouvelles du Centre d'Études Francoprovençales "René Willien"* n. 27 [La recherche toponymique dans les Alpes Occidentales. Actes de la conférence annuelle du Centre d'Études Francoprovençales "René Willien" de Saint-Nicolas (19-20 décembre 1992)], p. 21-32.
- Genre, A. (1995<sup>a</sup>), "ATPM: Norme per la trascrizione: occitano e francoprovenzale", (dattiloscritto inedito).
- Genre, A. (1995<sup>b</sup>), "ATPM: Norme per la trascrizione: piemontese", (dattiloscritto inedito).
- Genre, A. (1995<sup>c</sup>), "ATPM: Norme per la trascrizione: walser", (dattiloscritto inedito).
- Genre, A./Jalla, D. (1990), "L'Atlante Toponomastico del Piemonte Montano", dans *Nouvelle Revue d'Onomastique*, n. 15-16, p. 169-179.
- Genre, A./Jalla, D. (1993), "L'Atlante Toponomastico del Piemonte Montano", introduction à ATPM [2] et reprise dans les volumes

successifs.

- Genre, A./Massobrio, A., (1976), "Progetto di pubblicazione dei materiali dell'Atlante Linguistico Italiano con particolare riguardo all'Indice", dans *Bollettino dell'Atlante Linguistico Italiano*, III serie, n. 1, p. 78-92.
- Guiguet, D./Gallizio, S./Di Maio, M. (2003), *Guida dei toponimi di Savoulx e Costans*, Pinerolo, Alzani.
- Hirsch, E. (1955), "Die Flurnamen von Fenestrelle nach dem *Registro delle mutazioni catastali 1769-1822*", in *Beiträge zur Namenforschung*, n. 6, Heft 1, Heidelberg, p. 52-64.
- Hirsch, E. (1958), "Die Flurnamen von Pinasca und Villar Perosa nach Gemeindebüchern des 18. bzw. 19. Jahrhunderts", in *Beiträge zur Namenforschung*, n. 9, Heft 2, Heidelberg, p. 154-160.
- Hirsch, E. (1963), "Die Flurnamen von Sauze di Cesana", in *Beiträge zur Namenforschung*, n. 14, Heft 3, Heidelberg, p. 140-148.
- Hirsch, E. (1980), "Die Flurnamen von Oulx", in *Beiträge zur Namenforschung. Neue Folge*, n. 15, Heft 3, Heidelberg, p. 333-341.
- Hirsch, E. (1983), "Die Flurnamen von Bardonecchia", in *Beiträge zur Namenforschung. Neue Folge*, n. 18, Heft 3, Heidelberg, p. 284-300.
- Hunn, E.S. & Mai Johnson, L. (éd. par.) (2010), *Landscape ethnoecology. Concepts of Biotic and Physical Space*, Berghahn Books.
- Kollman, C. (2006), "Südtirols geografische Namen: von der Erhebung zur Aufbereitung für die Kartographie", in *Flöss* (éd. par.), p. 47-59.
- Lurati, O. (2003), "Intervento alla Cerimonia di presentazione del quarto volume dell'Atlante Linguistico Italiano (Torino, 14 ottobre 2003)", dans *Bollettino dell'Atlante Linguistico Italiano*, III Serie, n. 27, p. 336-341.
- Lurati, O. (2004), *In Lombardia e in Ticino. Storia dei nomi di luogo*, Firenze, Franco Casati Editore.
- Marrapodi, G. (2006), *Teoria e prassi dei sistemi onimici popolari : la comunità orbasca (Appennino Ligure centrale) e i suoi nomi propri*, Quaderni della Rivista Italiana di Onomastica, Roma, Società editrice romana.
- Massobrio, L. (1982), "Bjò 'bedale', ligòs 'acquidoccio' e ruza 'roggia'", dans *Bollettino dell'ALI*, III serie, Dispense 5-6, p. 5-10.

- Massobrio, L./Chiapusso, G./Monteleone, P. (2007), "L'Atlante Toponomastico del Piemonte Montano", conference présentée au colloque *Verso una multimedialità sostenibile. Ripensare le tecnologie per le scienze umane*, Università di Torino (<http://www.cisi.unito.it/convegno20anni/interventi/mas-sobrio.pdf> [25/02/2009]).
- Ottonelli, S. (1990), "Una ricerca toponomastica a Chianale, in alta Val Varaita [(1)-2-3]", dans *Valados Usitanos*, n. 35, p. 47-57; 36, p. 48-54; 37, p. 57-63.
- Paseri, G. (2008), *I nosti post. Toponimi dei luoghi di Melle*, Valados Usitanos.
- Pennisi, A. (1991), "L'informatica per la dialettologia", dans RID: *Rivista Italiana di Dialettologia*, n. 15, p. 137-164.
- Perulli, S. (2008) éd., *Ugo Pellis. Un fotografo in movimento*, Udine, Società Filologica Friulana.
- Pfister, M. (2006), "Presentazione dell'Atlante Linguistico Italiano (ALI), volume 6", dans *L'Italia dialettale*, n. 66-67 (Serie terza, II-III), p. 123-135.
- Ranucci, J.-C. (2004), "Micro-toponymie des Alpes-Maritimes: strates motivationelles", dans Ranucci/Dalbera (éd. par), p. 203-224.
- Ranucci, J.C./Dalbera, J.-Ph. (2004) (éd. par), *Toponymie de l'espace alpin : regards croisés*, Corpus, Les Cahiers n. 2.
- Ronco, G. (2004), "Au delà des dictionnaires : les atlas linguistiques", dans *International Journal of Lexicography*, vol. 17, n. 4, p. 441-455 et id. « Erratum », dans *International Journal of Lexicography*, vol. 18, n. 1 (2005), p. 151-152.
- Ruffino, G. (1992) (éd. par), *Atlanti linguistici italiani e romanzi. Esperienze a confronto* (Atti del Congresso Internazionale. Palermo, 3-7 ottobre 1990), Palermo, Centro Studi Filologici e Linguistici Siciliani.
- Sibille, R. (2004), *Guida ai toponimi e alla storia di San Marco di Oulx*, Pinerolo, Alzani.
- Sibille, R./François, O. (2006), *L'Adreyt di Oulx. Il territorio e la storia delle comunità di Amazas, Soubras e Vazon*, Pinerolo, Alzani.
- Terracini, B. (1955), "L'atlante linguistico italiano nell'opera di Matteo Bartoli e Ugo Pellis", dans *Bollettino dell'Atlante Linguistico Italiano*, Nuova Serie, n. 1, p. 3-9.
- Terracini, B. (1967), "Per l'edizione dell'Atlante Linguistico Italiano", dans *Bollettino dell'Atlante Linguistico Italiano*, Nuova Serie, n.

15-16, p. 61-65.

Terracini, B. (1969), "Discorso introduttivo", dans *Gli atlanti linguistici. Problemi e risultati*, Atti del convegno internazionale, Roma 20-24 ottobre 1967, Roma, Accademia Nazionale dei Lincei.

Terracini, B./Franceschi, F. (1964), *Saggio di un Atlante Linguistico della Sardegna. I. Carte; II. Testo*, Torino, Stamperia Editoriale Rattero.

Terracini, B./Franceschi, T. (1965), "Per l'edizione dell'Atlante Linguistico Italiano", dans *Bollettino dell'Atlante Linguistico Italiano*, Nuova Serie, n. 11-12, p. 43-53.

Vassere, S. (2006), "I programmi di informatizzazione e cartografia informatizzata nel Repertorio toponomastico ticinese", dans Flöss (éd. par), p. 103-109.

*Melissa Barkat-Defradas  
Praxiling, UMR 5267 CNRS  
(Université Montpellier 3)  
melissa.barkat@univ-montp3.fr*

*Élisabeth Motte-Florac  
Praxiling, UMR 5267 CNRS (Université de Montpellier 3)  
& Université Montpellier 1  
elisabeth.florac@wanadoo.fr*

## **Dialectologie arabe et atlas linguistique. Les apports du multimédia.**

### **1. Un atlas linguistique du monde arabe**

#### *1.1. L'arabe*

Aux environs du IV<sup>e</sup> siècle, l'arabe n'est parlé que par les tribus nomades de la péninsule arabique. C'est à partir du VII<sup>e</sup> siècle qu'il connaît une extension géographique très importante, l'expansion de l'Islam ayant comme effet le déploiement sur un vaste territoire de cette langue dans laquelle s'est faite la révélation coranique. Même dans les régions où il ne peut s'instaurer comme langue de communication (comme au Pakistan, en Turquie ou en Indonésie), l'arabe s'impose peu à peu comme langue de culte. Actuellement, la langue arabe, dans sa très grande diversité dialectale, concerne quelque 250 millions de locuteurs et couvre un immense territoire, tant en Asie (dans certaines républiques du Sud de la Russie, de l'Iran et de la Turquie) qu'en Afrique où le monde arabe parvient à ses limites naturelles, les déserts. Sont donc arabophones, en totalité ou en grande partie : l'Iraq, la Syrie, le Liban, la Jordanie, la Palestine, certains territoires israéliens, le Koweït, Bahreïn, le Qatar, les Émirats



Arabes Unis, l'Arabie saoudite, le Yémen, Oman, l'Égypte, la Libye, une large partie du Soudan, l'Érythrée, la Somalie, la zone située à l'est du lac Tchad, la partie Ouest du Mali, la Mauritanie, le Maroc, le Sahara occidental, l'Algérie, la Tunisie. Le domaine s'étend jusqu'en Europe (Malte et Chypre). Il convient également de noter qu'en France, l'arabe est aujourd'hui la deuxième langue la plus parlée par les populations immigrées.

Si tous les parlers arabes présentent des traits qui les rapprochent, surtout en ce qui concerne la phonologie et la morphologie, le domaine arabophone connaît une variabilité dialectale très importante qui est due à la vastitude de son étendue, à la coexistence possible de plusieurs langues sur une même aire dialectale (comme au Maroc, le berbère, l'arabe, l'espagnol et le français), à la nature des substrats anciens sur le terrain (berbère, copte, araméen) et/ou à l'influence des diverses langues (sémitiques ou non) en contact avec l'arabe. C'est pourquoi l'intercompréhension concerne surtout les parlers appartenant à une même zone géographique. Ainsi, les parlers de la partie orientale du monde arabe se différencient assez nettement des parlers de type maghrébin, surtout si l'on considère la phonétique et le lexique (Barkat 2001, Barkat-Defradas 2007, Embarki 2008). Toutefois, aucune entité linguistique appelée « arabe maghrébin » ou « arabe oriental » n'a été reconnue à ce jour. Nous avons ainsi montré à propos de certaines plantes aromatiques, que le pourcentage de lexique commun aux différents parlers arabes varie de façon significative en fonction des paires de parlers considérées, et ce, indépendamment de la nature du vocabulaire (Barkat-Defradas & Motte-Florac 2007a).

Traditionnellement, les distinctions permettant des regroupements géographiques et sociologiques sont établies en fonction du traitement de critères linguistiques tels que :

- au niveau phonétique : (i) le maintien ou la disparition des trois fricatives interdentes ث /θ/, ذ /ð/, ض /ðˤ/ ; (ii) la réalisation sourde ou sonore de l'occlusive uvulaire ق /q/ ; (iii) l'organisation des systèmes vocaliques : nombre de segments vocaliques ; présence ou absence des diphtongues /aj/ et /aw/ ; fusion ou non de certains timbres ; rôle phonologique de la voyelle centrale schwa /ə/ ;
- au niveau morphologique : (I) la construction du morphème de la

négation /ma-iʃ/ vs /miʃ/ ; (II) la forme des paradigmes verbaux de l'inaccompli n-/j- vs b-.

Bien que l'évolution récente des parlers ne permette plus de considérer ces macro-discriminants comme véritablement pertinents dans le cadre de la classification dialectale (Barkat 2001, Barkat-Defradas, 2007, Embarki 2008), il reste néanmoins possible de distinguer, à l'instar de K. Versteegh (1997), cinq groupes de parlers principaux : (1) les parlers de type arabe, (2) les parlers de type mésopotamien, (3) les parlers de type levantin, (4) les parlers de type égyptien et (5) les parlers de type maghrébin. Les quatre premiers groupes de parlers sont parfois rassemblés dans un ensemble oriental et opposé à un ensemble occidental constitué par le groupe des parlers de type maghrébin et les parlers de quelques pays ou régions spécifiques de la Méditerranée occidentale (Andalousie, Sicile, Malte...). Chacun des cinq groupes de parlers définis par K. Versteegh (1997) a été subdivisé en différents sous-groupes par ce même auteur.

## *1.2. Les atlas linguistiques*

Selon G. Bonnerot & F. Joly (2012), « un atlas est un recueil ordonné de cartes, conçu pour représenter un espace donné et exposer à son sujet un ou plusieurs thèmes » ; deux catégories majeures peuvent être distinguées : les « atlas de référence » qui « sont des répertoires de lieux et de toponymes » et qui, « composés de cartes et d'index associés », incluent peu de commentaires ; et les « atlas thématiques » qui sont « descriptifs, qualitatifs et explicatifs du monde, ou d'un État, ou d'une région ». La constitution de tels atlas impose que les données soient cohérentes et ordonnées autour d'un objectif. Par ailleurs, les atlas thématiques présentent la particularité de comporter pratiquement toujours des notices explicatives fournissant des indications indispensables sur la construction de chaque carte et des commentaires sur l'analyse spatiale proposée.

Dans la tradition de la géographie linguistique, l'objectif que se fixent les linguistes est de dégager les variations linguistiques régionales, surtout en matière lexicale. C'est pourquoi les atlas linguistiques sont constitués d'un ensemble de cartes qui renseignent sur les particularités morpho-lexicales et/ou les variations de

prononciation dans un parler, ou présentent les isoglosses qui ont pu être déterminés dans un continuum linguistique. Généralement imprimés sur papier, ils rendent compte, dans un réseau plus ou moins représentatif de localités, des résultats lexicaux ou phonétiques obtenus pour un certain nombre d'items. Tous les atlas linguistiques partagent donc le même rapport entre trois types d'éléments : des termes d'enquête, des points géographiques et des formes dialectales. En conséquence, la manière de collecter les données géolinguistiques consiste à se rendre sur les lieux, y interroger un certain nombre de locuteurs natifs tout en contrôlant certaines variables dont on sait qu'elles peuvent être à l'origine de variantes linguistiques dites « conditionnées », et fixer les réponses ainsi obtenues à l'aide d'une transcription phonétique appropriée.

Malgré leur qualité et le très grand intérêt des atlas linguistiques réalisés ou en cours de réalisation (*e.g.* Nouvel atlas linguistique de la France, Atlas linguistique roman, Atlas linguistique de la Péninsule ibérique), de nombreuses critiques ont été émises à leur encontre, en particulier : le manque de précision des isoglosses ; l'évacuation des aires de transition ; la difficulté de superposer les informations recueillies à différents niveaux d'analyse.

### ***1.3. Les atlas linguistiques des parlers arabes***

À ce jour, aucun ouvrage synthétique n'a été publié pour la description des différents systèmes phonologique, morphologique et syntaxique. De même, font défaut les dictionnaires présentant l'essentiel de la description du lexique dans les différents parlers arabes. Ceci ne signifie nullement que les travaux sur l'arabe dialectal sont inexistantes, bien au contraire, mais ils demeurent dispersés et sont souvent limités à certains aspects du système et/ou à des zones dialectales précises (*e.g.* Égypte (Behnstedt & Woidich 1985), Syrie et Palestine (Bergstraber 1915), Tunisie (Mejri 2005), Algérie (Bouhadiba 1992)). Cette situation tient certainement à l'idée partagée par les autorités académiques – et largement répandue – selon laquelle les études sur les parlers arabes représenteraient une remise en question de l'arabe standard en tant que facteur d'identité nationale. Du point de vue de sa structure interne, l'aire linguistique arabe est

structurée par une situation diglossique<sup>1</sup> : l'arabe littéral (*i.e.* arabe classique et arabe standard moderne) – écrit et parlé par ceux à qui il a été enseigné – côtoie un grand nombre de parlers. Si l'utilisation de l'arabe littéral est limitée au domaine intellectuel (*i.e.* liturgie, presse, radio, domaine juridique, enseignement...), les parlers arabes trouvent leur usage dans les échanges de la vie courante. Ils fournissent une image fidèle, exhaustive et vivante de la culture des différentes populations arabophones.

Cette absence d'outils permettant une évaluation de la distance linguistique existant entre les différents parlers arabes a motivé notre projet d'un atlas linguistique qui fournirait les informations pertinentes pour analyser les différents parlers arabes. Tout en s'inscrivant dans la tradition des atlas linguistiques habituels, cet atlas s'en démarquerait en fournissant d'autres composantes du système linguistique, qui sont importantes pour l'analyse des parlers, comme le phonétisme.

## 2. Les avantages d'un atlas linguistique multimédia

Les atlas linguistiques traditionnels, dans la mesure où ils sont imprimés sur papier, ne peuvent livrer qu'une partie descriptive de l'information ; et pour aussi exhaustive qu'elle puisse être, elle ne saurait fournir aux dialectologues les informations suffisantes pour analyser les parlers arabes. Le désir de combler ce manque est à l'origine du projet de base de données multimédia ARABER présenté en 1998 par M. Barkat. Pour engager ce projet ambitieux et en faire un atlas linguistique accessible en ligne et donc ouvert à la très vaste communauté de chercheurs concernés par la langue arabe (mais qui pourrait aussi s'envisager comme un atlas participatif plus large), il était nécessaire de commencer par une première étape dont le but serait de repérer les difficultés propres d'une part à cet exercice, d'autre part aux parlers arabes. C'est pourquoi un atlas linguistique

---

<sup>1</sup> Le mot « diglossie » a d'abord été synonyme de « bilinguisme » avant d'être utilisé par le linguiste W. Marçais en 1930 dans « La diglossie arabe » pour caractériser la situation linguistique du monde arabe. Aujourd'hui, ce terme marque davantage l'état dans lequel se trouvent deux systèmes linguistiques coexistant sur un territoire donné, dont l'un occupe – le plus souvent pour des raisons historiques – un statut sociopolitique supérieur. C'est pourquoi la situation diglossique est généralement une situation conflictuelle.

multimédia a été envisagé sur un thème assez réduit, celui des plantes aromatiques. La limitation du thème devait également permettre une publication numérique des données sur disque optique numérique. En optant pour ce format, il devenait possible de stocker un très grand nombre de données, de les visualiser sur un écran d'ordinateur et donc de s'affranchir de certaines des contraintes liées aux cartes et aux types de données et de requêtes.

### ***2.1. Les cartes***

Les Systèmes d'Informations Géographiques (S.I.G.) qui se sont beaucoup développés depuis les années 1980, permettent comme le soulignent G. Bonnerot & F. Joly (2012), « de rassembler, d'organiser, de localiser, d'analyser et de gérer un lot sans cesse croissant et mis à jour de données cartographiables. » Parmi les nombreux problèmes posés par les atlas linguistiques sur support papier, un des plus importants est celui de l'échelle ; celle-ci doit être correctement déterminée pour pouvoir intégrer sur une même carte l'ensemble des données. Or dans le cas du monde arabophone, l'étendue du territoire et la diversité géolinguistique rendent ce choix particulièrement difficile. L'un des avantages de l'atlas multimédia est sa grande flexibilité puisque le numérique permet de passer instantanément d'une échelle à l'autre (aire dialectale, pays, ville) et de construire des cartes adaptées à chacune des requêtes.

### ***2.2. Les données***

Dans un atlas multimédia, la quantité d'informations stockées peut être considérable, contrairement à ce qu'il est possible d'attendre d'un atlas linguistique ordinaire. Or, le champ de la dialectologie arabe traditionnelle dispose, depuis les premières monographies rédigées au milieu du XIXe siècle, d'une importante masse de données. Certes, la quantité ne saurait être considérée, en tant que telle, comme une amélioration. Mais il est fondamentale que les linguistes puissent disposer, pour leur analyse, des renseignements nécessaires pour comprendre les différences lexicales apparaissant sur les cartes, en particulier données linguistiques (lexicales ; phonétiques – sons, sonagrammes – ; narratives – proverbes, récits, etc. – enregistrements audio), de données biologiques pour des objets naturalistes

(photographies), et de données ethnographiques et ethnoscientifiques sur les savoirs et savoir-faire traditionnels (fichiers audio, vidéo). De fait, sont indispensables : de nombreuses données pour l'analyse des phytonymes et de leur motivation (Motte-Florac 2002) ainsi que de leur évolution (Motte-Florac *et al.* 2012), des enregistrements audio pour toute analyse phonétique, des photographies pour une détermination correcte des taxa, et des vidéos pour une description fidèle des savoir-faire. Ces données ne pouvant être disponibles que sur des fichiers de nature différente, seule la numérisation les rend accessibles. Notons par ailleurs, que cet accès est immédiat, aisé, et qu'il autorise aussi l'affichage de plusieurs informations en parallèle.

### **2.3. Les requêtes**

L'atlas multimédia permet de faire des recherches indexées en faisant varier non seulement les territoires (par région, pays ou ville) mais aussi les entrées, puisqu'elles peuvent être engagées à partir du signifiant (*i.e.* le nom vernaculaire d'une espèce dans le cas des savoirs naturalistes locaux) ou à partir du signifié (*i.e.* le nom scientifique de la (ou des) espèce(s) scientifique(s) correspondant à l'espèce vernaculaire, ou encore les noms donnés à cette espèce dans d'autres langues).

Alors que le support papier rend extrêmement difficile (voire impossible) le croisement de diverses données, ces requêtes donnent accès à toutes les informations qui sont rassemblées sur un item déterminé (données linguistiques, biologiques et ethnoscientifiques, comme nous l'avons vu précédemment).

### **2.4. Autres avantages**

Riche d'informations sur une langue (ici, les parlers arabes), un atlas multimédia constitue un outil de sauvegarde linguistique, d'autant plus indispensable que le contexte actuel est celui d'un avenir incertain, voire critique, pour les langues vernaculaires<sup>2</sup>. Facile d'accès, ces données peuvent aussi circuler *via* Internet ; la mondialisation, dans le même temps qu'elle repose largement sur la généralisation d'un modèle culturel dominant et sur l'usage de

---

<sup>2</sup> Claude Hagège, *Halte à la mort des langues*, 2000.

quelques langues de grande diffusion, favorise les communications à l'échelle internationale et, en conséquence, contribue à ouvrir à une audience nouvelle des langues encore peu décrites, parfois même méconnues.

Au-delà de la seule linguistique, un atlas linguistique multimédia constitue aussi une remarquable source d'informations, qui peut être mise à profit pour des objectifs et dans des registres extrêmement divers. Ainsi, dans le domaine éducatif, il présente un intérêt pédagogique par sa possible utilisation dans des contextes divers (en particulier pour l'Enseignement des Langues et des Cultures d'Origine). Dans le domaine des ethnosciences, il contribue non seulement à la sauvegarde des TEK (Traditional Ecological Knowledge)<sup>3</sup>, mais aussi à un renforcement de leur exploration (les langues vernaculaires comme les parlers arabes, représentent une voie d'accès incontournable vers les connaissances naturalistes, *lato sensu*, autrefois transmises uniquement à travers la tradition orale) et un perfectionnement des données (les erreurs dans les dénominations vernaculaires étant malheureusement assez fréquentes<sup>4</sup>). Notons que ces données sont particulièrement intéressantes dans la mesure où elles ouvrent sur des perspectives très concrètes dans le domaine de la gestion durable de l'environnement des populations autochtones ; les savoirs anciens sur l'alimentation, la thérapeutique, l'hygiène..., constituent d'importantes données pour un éventuel développement économique. C'est le cas, par exemple, des savoirs traditionnels qui ont permis, au Maroc, de développer l'économie locale autour de l'huile d'argan, en adaptant les utilisations traditionnelles aux exigences des consommateurs des pays du Nord. Notons également qu'un atlas multimédia contribue de façon non négligeable à la sauvegarde du patrimoine culturel immatériel, notion apparue dans les années 1990 lorsque l'attention a été attirée sur la disparition des cultures traditionnelles (savoirs, savoir-faire, littérature orale, musique

---

<sup>3</sup> Formulation académique anglaise – sans équivalent français – pour désigner toutes les connaissances traditionnelles (savoirs, savoir-faire, comportements) concernant les ressources de l'environnement naturel d'une société humaine.

<sup>4</sup> Les fréquents manques de précision ou inexactitudes dans la transcription des dénominations vernaculaires par les scientifiques sont, comme les erreurs de détermination scientifique par nombre de linguistes, extrêmement préjudiciables lorsqu'il s'agit de l'appréciation de la valeur des savoirs traditionnels. C'est d'autant plus important que cette appréciation est déterminante pour les possibilités de développement économique des populations locales.

et danse traditionnelles, etc.)<sup>5</sup>. Par ailleurs, au niveau social, un tel atlas témoigne du rôle essentiel tenu par les femmes dans la transmission d'un grand nombre de connaissances et, par là même, contribue à revaloriser leur image.

Enfin, au niveau économique, la numérisation évite les dépenses occasionnées par une impression sur papier, puisque le travail peut être réalisé virtuellement sur un écran d'ordinateur. Ceci n'empêche nullement la possibilité d'avoir accès, si nécessaire, à une version imprimée, grâce aux fichiers PDF (disponibles, par exemple, dans l'atlas multimédia sur les plantes aromatiques pour la plupart des informations).

### **3. Un exemple d'atlas linguistique multimédia : les plantes aromatique dans l'environnement arabophone**

L'atlas linguistique multimédia *Les plantes aromatiques dans l'environnement arabophone* fait partie d'un ouvrage en deux volumes. Le premier présente les caractéristiques phonétiques de ces parlers<sup>6</sup>. Le second, qui présente l'analyse d'un exemple particulier de lexique, celui des plantes aromatiques, contient l'atlas linguistique multimédia sur DVD<sup>7</sup>.

La réalisation de ce dernier a imposé non seulement de repenser la construction même de cet outil d'analyse linguistique, mais également de reconsidérer la méthodologie traditionnellement utilisée lors des enquêtes de terrain.

#### ***3.1. Objectifs et limites du travail***

##### **→ Thème : les plantes aromatiques**

Un travail ayant déjà été réalisé en dialectologie arabe sur la comparaison de mots de la vie courante (*e.g.* Synyukov 2004, Ibrahim 2009), nous avons souhaité constituer un corpus à partir de mots

---

<sup>5</sup> En 2003, la *Convention pour la sauvegarde du patrimoine culturel immatériel* a été adoptée par l'UNESCO.

<sup>6</sup> Barkat-Defradas M., *Les parlers arabes modernes*. Vol. 1. *Aspects phonétiques* (à paraître).

<sup>7</sup> Motte-Florac É. & Barkat-Defradas M., *Les parlers arabes modernes*. Volume 2. *Aspects lexicaux : l'exemple des plantes aromatiques* + Atlas multimédia : *Les plantes aromatiques dans l'environnement arabophone* (à paraître).



spécifiques. Ces mots devant dans le même temps intervenir de façon importante dans le quotidien de l'ensemble des populations arabophones, notre choix s'est naturellement porté sur des noms de plantes. Beaucoup d'entre elles interviennent dans la vie de tous les jours. Cependant leur nombre étant trop élevé, nous avons restreint notre choix à quelques espèces (les données devaient être limitées pour pouvoir être stockées sur un unique disque optique numérique). Les plantes aromatiques présentent l'avantage d'être parfaitement connues dans l'ensemble du monde arabe où elles sont utilisées dans tous les registres de la vie quotidienne (alimentation, hygiène, thérapeutique...). Si ces nombreuses utilisations ne sont guère étonnantes pour des plantes qui sont spécifiques à l'environnement naturel des pays qui bordent la Méditerranée (Motte-Florac 2011), il est important de noter que leur importance tient aussi au fait que les senteurs constituent un élément culturel fondamental dans le monde arabe. J. Cherif Abderrahman (2001:113) explique cette importance des odeurs<sup>8</sup> en rappelant que « le parfum faisait partie des éléments les plus appréciés par le prophète Mahomet, car il était censé exercer un effet bénéfique sur la préservation de la santé du corps. »

Outre ces considérations environnementale et culturelle, le choix a été arrêté sur les plantes aromatiques en raison de l'intérêt de leurs phytonymes dans une étude menée sur le choix des dénominations, leur évolution et les paramètres qui les déterminent. En effet, parmi les plantes aromatiques de l'univers méditerranéen, un certain nombre font partie de l'environnement naturel, mais beaucoup d'autres ont voyagé, occasionnant des emprunts multiples (parfois sources d'erreurs) et de nombreuses variations.

Pour constituer le corpus et pour que le travail puisse être réalisé dans un temps raisonnable compte tenu de l'importance des données bibliographiques et de l'étendue du territoire à couvrir, seules vingt-cinq plantes ont été retenues. Plusieurs critères sont intervenus dans leur choix. Les espèces devaient être connues et utilisées dans l'ensemble du monde arabe et par toutes les personnes, quelle que soit leur classe sociale, leur classe économique, leur classe d'âge. La moitié d'entre elles environ devaient être des herbes aromatiques de l'univers méditerranéen et l'autre moitié des épices, importées et utilisées dans ces contrées depuis plusieurs millénaires, les arabes

---

<sup>8</sup> Importance qui se retrouve également dans le lexique (David *et al.* 2006).

ayant détenu le monopole de ces produits pendant de nombreux siècles. Les espèces suivantes ont été sélectionnées : ail, anis, basilic, cannelle, cardamome, carvi, coriandre, cumin, curcuma, fenouil, gingembre, girofle, laurier, menthe, muscade, myrte, nigelle, oignon, persil, piment, poivre, romarin, safran, sauge, thym.

➔ **Les sources : recherches bibliographiques et enquêtes**

Les données présentées proviennent de recherches bibliographiques (dictionnaires, travaux ethnographiques, livres de recettes, ouvrages sur les plantes médicinales, etc.) et de résultats d'enquêtes. Celles-ci ont été réalisées de deux façons différentes. Certaines ont été menées dans plusieurs pays arabophones (Égypte, Tunisie, Maroc, Algérie...). Dans ce cas, les informations ont été recueillies auprès de personnes compétentes (sur les marchés, dans les magasins spécialisés comme les marchands d'épices) comme auprès de personnes les plus diverses. De nombreux échantillons ont été achetés pour pouvoir vérifier la détermination botanique des espèces. D'autres enquêtes ont été menées en France auprès de locuteurs de passage ou auprès d'immigrés de première génération. Les données ont été recueillies pour un ensemble de parlers représentatifs de chaque groupe dialectal. Des photographies ont été présentées pour les plantes les plus communes (*e.g.* oignon, ail) et des produits frais (pour les herbes aromatiques) et secs (pour les épices) pour toutes les plantes susceptibles d'être peu ou mal connues.

### 3.2. *Présentation de l'atlas*

Une page d'accueil permet d'accéder à une « Introduction » (présentation générale de l'atlas multimédia, de ses objectifs et de la démarche méthodologique) et à la « Bibliographie » (détail de toutes les références bibliographiques citées et des enquêtes réalisées) (fig. 1). À partir des autres boutons, il est possible de rentrer dans la base de données par une recherche à partir du signifiant, nom de plante en arabe standard ou dans l'un des parlers, écrit en arabe (*e.g.* بَسْبَاس ) ou en translittération (*e.g.* besbās). Il est également possible d'y rentrer à partir du signifié ; dans notre cas, il peut s'agir soit d'un nom vulgaire ou vernaculaire en français (*e.g.* fenouil) ou en anglais (*e.g.* fennel), soit d'un nom scientifique (*e.g.* *Foeniculum vulgare*).

The screenshot shows a web interface for a multimedia linguistic atlas. The main title is "Les plantes aromatiques dans l'environnement arabe" with a sub-title "Fenouil". A navigation bar includes "Généralités", "Botanique", "Usages Alimentaires", "Usages Thérapeutiques", "Usages Cosmétiques", and "Divers". A sidebar on the left lists menu items: "Accueil", "Introduction", "Index des noms français de plantes", "Index des noms anglais de plantes", "Index des noms arabes de plantes", "Index des noms scientifiques de plantes", "Tables des illustrations", and "Bibliographie". The main content area is titled "Raqqa" and features a table with two columns: "Syrien" and "(Raqqa)". The "Syrien" column contains a small map of Syria. The "(Raqqa)" column contains a large spectrogram of the word "jemrah" and a phonetic transcription "j e m r a h" with vertical lines separating the syllables. Below the transcription is a play button and the text "ECOUTER".

Figure 1. Données linguistiques de l'atlas linguistique multimédia :  
*Les plantes aromatiques dans l'environnement arabophone*  
(Motte-Florac & Barkat-Defradas, à paraître)

Les plantes aromatiques dans l'environnement arabophone quitter X

## Fenouil

Généralités Botanique Usages Alimentaires Usages Thérapeutiques Usages Cosmétiques Divers Littérature Orale Bibliographie

Accueil  
Introduction  
Index des noms français de plantes  
Index des noms anglais de plantes  
Index des noms arabes de plantes  
Index des noms scientifiques de plantes  
Tables des illustrations  
Bibliographie

**Classification**  
Apiales  
Apiaceae  
*Foeniculum vulgare* Miller  
= *Anethum foeniculum* L.  
= *Anethum piperitum* Ucria  
= *Foeniculum capillaceum* Gilib.  
= *Foeniculum dulce* Miller  
= *Foeniculum officinale* All.  
= *Foeniculum piperitum* (Ucria) C. Presl

**Écologie et culture**  
Le fenouil sauvage supporte bien la sécheresse. Il pousse spontanément en climat méditerranéen, en plein soleil, dans les friches, les coteaux secs et pierreux, au bord des chemins (Frachon, 1990 : M).  
Il se récolte d'août à novembre.  
Le fenouil bulbeux se cultive dans les sols sableux et bien drainés.

**Description botanique**  
**Fenouil sauvage**  
Plante herbacée très parfumée, bisannuelle ou vivace, de 80-250 cm de hauteur, glabre, à port léger.  
Racine pivotante, charnue, fusiforme, presque toujours bifide, blanche.  
Tiges droites, cannelées, grêles, anguleuses, brillantes, couvertes d'une pruine blanchâtre.  
Feuilles alternes, de couleur vert glauque, très découpées, bipennatiséquées, à lanières nombreuse, allongées et filiformes, à gaine charnue et sucrée ; feuilles supérieures à gaine plus longue que le



Image 4 of 10  
Elisabeth Motte Florac

**Origine géographique et aire de répartition**  
Plante méditerranéenne (Haudricourt & Hédin,

Figure 2. Données botaniques de l'atlas linguistique multimédia :  
*Les plantes aromatiques dans l'environnement arabophone*  
(Motte-Florac & Barkat-Defradas, à paraître)

Notons que pour tous ces noms, deux listes sont disponibles : une « liste réduite » n’affiche pour chaque espèce que le nom le plus communément admis et le plus général<sup>9</sup> ; la seconde est une « liste complète » présentant tous les noms mentionnés dans le DVD pour la langue concernée. Une requête sur l’une de ces deux listes permet d’accéder à la page de « Généralités » de l’espèce (ou des espèces) correspondante(s).

Pour chaque espèce scientifique sont développés, en plus des noms scientifiques<sup>10</sup>, tous les noms vulgaires et vernaculaires en français, anglais et arabe. Notons que sur cette même page, s’ouvre un diaporama de photos de l’espèce (individu entier et/ou différentes parties selon les cas) qui permet de faire une rapide vérification morphologique (fig. 2).

Pour les dénominations en arabe, il est possible d’accéder aux cartes de répartition géographique (pour chaque dénomination ou pour chaque espèces scientifique) et un curseur permet de changer l’échelle pour accéder au niveau souhaité. Chaque nom géolocalisé est une porte d’entrée vers un grand tableau dans lequel sont recensés, pour une espèce donnée, tous les noms en arabe standard (en écriture arabe ou avec la translittération utilisée par l’auteur) et tous les noms dans les différents parlers. Ces derniers sont classés en fonction de leur localisation géographique : groupe de parlers, puis pays, puis ville. Pour les noms enregistrés lors des enquêtes, trois types de données sont disponibles (fichier audio, transcription phonétique, sonagramme ; fig. 1). Des informations complémentaires sont également fournies sur l’éventuelle motivation des noms ou leur étymologie, ainsi que les références bibliographiques ou celles du corpus d’enquêtes.

À partir des boutons correspondants (fig. 1), il est possible d’accéder aux autres informations :

- « Botanique ». Pour éviter les erreurs (nombreuses) dues à une expertise trop simpliste des non-initiés (qui peuvent considérer certaines espèces banales et évidentes alors même qu’elles sont

---

<sup>9</sup> Outre le fait que plusieurs noms vernaculaires (*e.g.* fenouil, cannelle, menthe) se déclinent en plusieurs possibilités qui peuvent n’être connues que de certaines personnes (cuisinières, marchands, thérapeutes...), les noms diffèrent souvent selon les auteurs et les lieux.

<sup>10</sup> Pour les noms scientifiques, deux listes sont également proposées. La « liste complète » comprend tous les synonymes apparus au cours des révisions successives de l’espèce, ce qui permet d’éviter d’éventuelles erreurs.

fort complexes pour les botanistes ; *e.g.* *Foeniculum vulgare*), des informations détaillées sont fournies pour chaque plante : classification détaillée, synonymies scientifiques, origine géographique, données sur l'écologie et la culture, description botanique. Ces informations permettent également de mieux appréhender créations, emprunts et changements.

- « Usages alimentaires », « Usages thérapeutiques », « Usages cosmétiques », « Divers » sont des boutons (fig. 1) qui donnent accès aux informations ethnoscientifiques (sur la production des espèces, leurs utilisations, leur commerce, etc.). Celles-ci offrent des données qui sont indispensables pour comprendre les choix de dénominations et l'évolution des phytonymes dans l'espace et dans le temps.
- « Littérature orale » permet d'accéder à la vision symbolique que les sociétés peuvent avoir des espèces. Ici, elle nous sera transmise essentiellement à travers les proverbes.

#### 4. Interprétations et analyses

L'interprétation des cartes et les analyses réalisées à partir des nombreuses données rassemblées dans le DVD constituent la partie imprimée du tome 2 (*Les parlers arabes modernes. Volume 2. Aspects lexicaux : l'exemple des plantes aromatiques*<sup>7)</sup>). Les commentaires qui y sont rassemblés concernent la mosaïque des zones dialectales avec ses similitudes et ses différences, la motivation des phytonymes, leur évolution et l'érosion des parlers et/ou des substrats linguistiques sous l'effet de la globalisation linguistique, les problèmes de traduction des noms vernaculaires et les erreurs de détermination, les homonymies générales ou locales concernant des espèces différentes (*e.g.* basilic et myrte, fenouil et aneth) et les raisons de ces homonymies, la variabilité phonétique (qui se manifeste au niveau d'un groupe dialectal, d'un pays, d'une région, d'une localité, d'une population), etc. (*cf.* Barkat-Defradas & Motte-Florac, 2007b).

## Références bibliographiques

- Barkat M., 1998, "Araber: a Database of Areal Characteristics in Arabic", in *4ème rencontre de l'Association Internationale de Dialectologie Arabe*, Malte, 1998.
- Barkat M., 2001, *Détermination des parlers arabes*, Lille, Éditions du Septentrion.
- Barkat-Defradas M., 2007, "Some reliable acoustic cues for the discrimination and the classification of modern Arabic vernaculars", Conférence Internationale *Typologie des parlers arabes modernes, traits, méthodes et modèles de classification*, Montpellier 14-15 mai 2007.
- Barkat-Defradas M. & Motte-Florac É., 2007a, "Évaluation de la proximité dialectale arabe en fonction de la nature des vocabulaires", Communication affichée à la *Conférence Internationale « Typologie des Parlers Arabes Modernes : traits, modèles & méthodes de classification »*, Montpellier, 14-15 mai 2007.
- Barkat-Defradas M. & Motte-Florac É., 2007b, "Atlas linguistique multimédia *Les plantes aromatiques dans l'environnement arabophone* et analyse des données", communication orale, *132° congrès National des Sociétés Historiques et Scientifiques : Image e imagerie*, Arles, 1—21 avril 2007.
- Behnstedt P. & Woidich M., 1985, *Die ägyptisch-arabischen Dialekte* (vol. 1 & 2). Wiesbaden, Reichert.
- Bergstraber G., 1915, "Sprachatlas von Syrien und Palastina", *ZDPV* 38:169-222.
- Blalock M., 1960, *Social statistics*. New York, McGraw-Hill.
- Bonnerot Guy & Joly F., 2012, Cartographie, *Encyclopaedia Universalis* (consultable sur Internet, <http://www.universalis.fr/encyclopedie/cartographie/>)
- Bouhadiba F., 1992, "On phonemic isoglosses in Western Algeria", *Cahiers de dialectologie et de linguistique comparative* 3(1-2):1-32.
- Cherif Abderrahman Jah, 2001, *Los aromas de Al-Andalus: la cultura andalusí a través de los perfumes, especias y plantas aromáticas*. Madrid, Alianza Editorial.
- David S., Barkat-Defradas M., Mouby C., 2006, "Odeurs et accès au lexique : lexicographie comparée en français et en arabe".

- Communication orale, Colloque *Les mots du goût*, Lyon, 1-2 décembre 2006.
- Embarki M., 2008, "Les dialectes arabes modernes : état et nouvelles perspectives pour la classification géo-sociologique", *Arabica*, 55(5/6):583-604.
- Hagège C., 2000, *Halte à la mort des langues*. Paris, Odile Jacob.
- Ibrahim Zeinab, 2009, *Beyond lexical variation in Modern Standard Arabic: Egypt, Lebanon, Morocco*. Cambridge Scholar Publishing.
- Marçais W., 1930, "La diglossie arabe", *L'Enseignement Public*, 97:401-409.
- Mejri, S., 2005, *Atlas Linguistique de Tunisie: les questionnaires*, Paris, Maisonneuve et Larose, coll. Lettres du Sud.
- Motte-Florac É., 2002, "Motivation et permanence des phytonymes ; l'exemple des *Plantago* poussant en France", in Colombel V. de & Tersis N., *Lexique et motivation ; perspectives ethnolinguistiques*. Leuven, Paris, Dudley (MA), Peeters, pp. 223-256.
- Motte-Florac É., 2011, *Épices et herbes aromatiques en Méditerranée. Un art de vivre*. Montpellier, Poisson soluble, cédérom.
- Motte-Florac É. & Barkat-Defradas M., 2007, "Flux de plantes et de savoirs en Andalousie, plaque tournante entre le monde arabe, l'Europe et au-delà. La circulation des phytonymes", *Bulletin de Liaison de l'Association des Amis du Musée de la Pharmacie*, 32 : 15-30.
- Motte-Florac É., Barkat-Defradas M. & El Ghannami R., 2012, "Transfert de plantes. De la Méditerranée arabophone aux Amériques, les herbes aromatiques et leurs dénominations" in Simoni-Aurembou M.-R. & Dalbera-Stefanaggi M.-J., *La langue française : vecteur d'échanges culturels*. Paris, Éditions du CTHS, p. 73-85.
- Synyukov Leonid, 2004, *Traitement cognitif de la distance linguistique par approche pluridisciplinaire*. Université Lyon 2, Mémoire de Maîtrise de Sciences Cognitives.
- UNESCO 2003, *Convention for the Safeguarding of Intangible Cultural Heritage*.  
<http://unesdoc.unesco.org/images/0013/001325/132540e.pdf>.
- Versteegh K., 1997, *The Arabic Language*. Amsterdam, J. Benjamins.



*Gérard Taverdet*  
*Professeur émérite, Université de Bourgogne*  
*taverdet.gerard@neuf.fr*

## **La fin de la dialectologie ?**

C'est évidemment avec un parfum de nostalgie qu'on parle de la dialectologie traditionnelle et, au cours de ces deux journées, nous l'avons tous plus ou moins fait. Colette disait qu'elle appartenait à un pays qu'elle avait quitté ; mais nous, nous appartenons à un pays qui nous a quittés. Il est sûr que les nouvelles générations n'auront jamais les difficultés (et surtout les joies) du dialectologue traditionnel qui avait la mission de découvrir ce qui pouvait rester dans nos campagnes profondes de l'ancienne langue, issue des parlers médiévaux et peut-être même des premières vagues de la romanisation. Souhaitons d'autres plaisirs à nos successeurs. Aujourd'hui, nos craintes ne sont plus celles de voir disparaître les parlers « invariables » (ou historiques) de la France, puisqu'il y a déjà longtemps que nous sommes assurés de l'inéluctable, du moins dans la France d'oïl<sup>1</sup>, même si, parfois, quelques agréables surprises viennent nous prouver le contraire. Le danger immédiat est de voir les résultats de la dialectologie sombrer dans les oubliettes des bibliothèques, ce qui priverait du même coup la recherche historique d'une part essentielle de sa documentation. Nous avons laissé à notre titre un point d'interrogation que l'on jugera peut-être téméraire, voire optimiste. Il convient aussi de réfléchir aux différents sens du mot *fin* ; *fin* c'est bien sûr le terme ultime ; c'est peut-être aussi le but, mais il n'appartient à personne de fixer un but précis à la recherche ; la recherche existe, elle se contente d'être et c'est à chacun de nous de

---

<sup>1</sup> Mais on peut avoir d'heureuses surprises ; pour notre traduction des *Bijoux de la Castafiore*, nous avons choisi le parler le mieux attesté, mais aussi le moins vivant (ce choix a été fait avant tout parce que c'était le parler de notre famille) ; la réponse des lecteurs a été sans appel pour les pessimistes.

lui fixer une fin. N'oublions pas aussi qu'il y a sur nos cadastres de nombreux lieux-dits nommés *La Fin*, qui sont des territoires destinés aux assolements ; c'est peut-être le meilleur sens de *fin* pour la dialectologie actuelle, une montagne de résultats actuellement laissés en jachère<sup>2</sup>, surtout dans les universités, dans l'attente de circonstances plus favorables.

Même s'il ne nous appartenait pas de fixer un but à notre recherche, outre le but immédiat qui voulait noter les anciennes langues, nous espérons tous que nos enquêtes et nos atlas étaient destinés à une meilleure connaissance de la langue française, voire de l'histoire générale de notre pays. Ce travail d'exploitation n'est pas encore fait, soit parce que nous n'avons pas eu le temps de l'accomplir nous-mêmes, soit parce que nous n'avons pas trouvé ou su trouver les nouvelles générations capables de le réaliser, non par manque d'intérêt, mais peut-être aussi parce que ces nouvelles générations n'ont pu rencontrer les conditions matérielles dignes de les attirer vers ce genre de travaux.

### **La carte « poupée » de l'ALB**

Pour montrer, plutôt pour essayer de montrer, toute l'importance de la documentation atlantographique, nous partirons de la carte "poupée" de l'ALB (III, 1656). Cette carte présente un certain nombre d'avantages ; elle nous permettra entre autres de revoir l'étymologie de quelques mots, même si nous devons mettre en doute ce que nous avons pu lire dans de nombreux ouvrages. Peut-être pas de tout bouleverser, mais de nous poser d'autres questions. Elle nous permettra surtout de parler de Guignol, car il eût été regrettable que ce colloque se terminât sans que fût évoqué le nom de la célèbre marionnette lyonnaise.

Comme la plupart des cartes des atlas linguistiques, la carte « poupée » est partagée entre plusieurs types (dont les aires peuvent parfois se chevaucher, d'autant plus facilement qu'il s'agit d'une notion marquée par l'affectivité ; on peut même parfois noter plusieurs

---

<sup>2</sup> Voir notre *Microtoponymie de la Bourgogne*, tome IV, p. 635 (Dijon 1990).

réponses sur le même point) ; mais il s'agit là d'un fait connu de tous ceux qui ont pratiqué les enquêtes. Bref, cette carte n'est marquée par aucune ligne directrice, contrairement à ce qui se passe pour d'autres notions où il est facile d'opposer un type méridional à un type septentrional (qu'on pourra bien sûr qualifier de façon plus précise, selon les cas)<sup>3</sup>.

Dans un premier temps, nous verrons rapidement les différents types qui se partagent la région, malgré la francisation, puisque de nombreux points ne semblent pas posséder de types particuliers. Nous avons quelque mal à nous exprimer autrement qu'au présent ; mais nos auditeurs pourront corriger d'eux-mêmes et utiliser l'imparfait et tout autre temps du passé.

a) un type *p-p*, proche du français *poupée* : *paponne*, surtout dans le nord et l'est de la Côte-d'Or ; on note d'autres dérivations, comme *popette* (surtout en Charolais) ou *popine* (une fois en Bresse ; en dehors de l'ALB, type signalé également à Ménetreuil<sup>4</sup>).

b) le type *fille* est présent dans l'est du département de l'Yonne (deux points). Probablement un usage venu de la langue des petites filles qui jouent à la maman et qui considèrent leur poupée comme leur fille. Cette formation ne semble pas poser de graves problèmes étymologiques. Mais elle peut éventuellement nous aider à comprendre l'évolution sémantique des autres types.

c) le type *catin* (de *Catherine*) est très bien représenté dans la Nièvre et l'Yonne ; il a parfois une connotation péjorative (mais minoritaire). Une forme voisine *catiche* dans les environs de Dijon (points 15 et 16). De même [kátó:] au point 15. À Tannay<sup>5</sup> : *catin* « poupée avec laquelle jouent les petites filles ». Voir aussi *ALCe* II, 986 ; voir *FEW* 2-54a.

d) un type *boille* [bòy] a été noté deux fois dans les environs de Chalon-sur-Saône (points 39 et 95). Nous n'avons pas recherché les origines de ce type.

---

<sup>3</sup> C'est le cas des cartes "aulne" où s'opposent nettement deux types (voir *ALB* I, 555).

<sup>4</sup> Maryse Giovacchini, *Le Patois de Ménetreuil*, Dijon 1985.

<sup>5</sup> Dominique Perreau, *Le Patois de Tannay*, Dijon 1997.

Après ces formations dont la plupart sont bien connues, nous trouvons trois autres types que, pour l'instant, nous sommes obligés de séparer, tout simplement parce que les dictionnaires étymologiques les ont séparés.

e) le type *gouine* est bien représenté en Côte-d'Or, avec quelques légères variantes phonétiques ([gwèñ]) ; sa valeur péjorative n'est pas systématique ; par exemple, en 10, 21, 23, 97, aucune notation péjorative ; en 97, la *gouine* est une poupée de chiffons. À Ménetreuil, [na gwin] « une vieille poupée », comme à Saussey<sup>6</sup> ; [gönüc], entendu au point 96, a une valeur péjorative, ce qui n'est pas le cas en 109 ; ces deux mots pourraient être considérés comme des dérivés de *gouine*.

f) en 103, la [gönõ] est une vieille poupée. Cette réponse est unique, mais on sait par ailleurs qu'elle se prolonge largement sur l'aire francoprovençale. Il ne s'agit donc pas d'une réponse accidentelle, liée à une erreur du témoin.

g) enfin, au point 21 (il s'agit de Saffres, point très conservateur, du moins à l'époque de l'enquête), nous avons obtenu deux réponses (de la part du même témoin) ; la première est [gwèñ] qui est évidemment une variante phonétique de *gouine*, type que nous avons signalé plus haut ; la seconde est [gwèñó:l], nom féminin et qui désigne plus particulièrement une vieille poupée.

Nous serions tentés dans un premier mouvement de classer dans la même étymologie ces trois derniers types, mais notre démarche nous est interdite par la totalité des dictionnaires étymologiques ; nous suivrons ici le *Robert historique* et le *Trésor de la langue française*, puisque ces trois mots ont eu la chance d'être conservés, à des titres divers, par le français commun.

---

<sup>6</sup> Robert Féral, *Le Patois de Saussey*, Dijon 1986.

## Autres formes en *gon-*

Aux formes présentées par l'*ALB*, nous avons ajouté la liste des formes que nous avons relevées dans les glossaires publiés par l'Association ABDO, principalement entre 1980 et 2000. Nous avons indiqué également les formes en *gon-* (et variantes légères) présentes sur les cartes de l'*ALB* (hors de la carte "poupée"). Nous n'avons procédé ici à aucun classement étymologique.

**Gonaiche**, à Suin<sup>7</sup>, « jeune fille peu sérieuse ».

**Gonelle** : René Lapierre<sup>8</sup> : « jeune femme coureuse ».

**Se gôner**, **A** - verbe : s'habiller, *ALB*, III, 1526 ; « s'habiller mal », *ALB*, III, S. 23 ; Boussand : « habiller grossièrement » ; Mireille Fougeu<sup>9</sup> : [gó:né] « mal habillé » ; à Mercurey<sup>10</sup>, *gôner* « mal habillé » ; « gôniâ » « mal vêtu » ; Farruggia<sup>11</sup> : **mal gôné** « mal attifé ».

**Gôner**, **B** : Mongenet<sup>12</sup> : **se gôner**, « se salir copieusement » ; Alex : [gó:ñá], « salir, mettre dans un état dégoûtant » ; « saccager » (en parlant d'un champ) ; *ALCe*, **goné** « borbier » (I, 45) ;

**Gonelle** « averse » (*ALCe*, I, 11).

**Goniots**, nom masculin pluriel, « gros nuages », I, 12 ; ce terme est également signalé par R. Rouffiange<sup>13</sup> [gó:ñyò] « gros nuage de pluie ou d'orage » ; étymologie inconnue selon le *FEW* (XXI, 8a) qui cite à Louhans *gouniau* avec le même sens ; à Sagy, [na gu:na] « bande de nuages » ; Boussand<sup>14</sup> : **goniau** : 1) nuage d'orage ; 2) personnage grossier ; *goniot* désigne également les personnages à grosse tête qui animent le Carnaval de Chalon-sur-Saône (très vivant en français

<sup>7</sup> Gérard Taverdet, *Glossaire de Suin*, Dijon 1990.

<sup>8</sup> René Lapierre, *Glossaire de Rully*, Dijon 1988.

<sup>9</sup> Mireille Fougeu-Fontaine, *Le Patois de Chaumont-le-Bois*, Dijon 1986.

<sup>10</sup> Joseph Ancelin, *Glossaire de Mercurey*, Dijon 1986.

<sup>11</sup> Farruggia, *Le Parler nivernais*, Dijon 1999.

<sup>12</sup> Georges-Henri Mongenet, *Le Patois de Poiseul-la-Ville* (Côte-d'Or), inédit.

<sup>13</sup> Robert Rouffiange, *Le Patois et le français rural de Magny-lès-Aubigny* (Côte-d'Or), Dijon 1983.

<sup>14</sup> Alain Boussand, *Saint-Julien-de-Jonzy, patois et vie rurale*, Dijon 1983.

régional) ; **goniauder** « avoir une allure de *goniau*, au sens 2) ; **gouniô**, René Lapierre : « casanova de village ».

**Gonu**, nom masculin, « rebouteux », *ALB*, III, 1679.

**Gougne**, nom féminin, « truie », *ALB*, II, 1128.

**Gouja**, nom masculin, « mortier », *ALB*, III, 1382.

**Goujard**, nom masculin, « serpe » *ALB*, I, 574 ; « croissant pour tailler les haies », *ALB*, I, 575<sup>15</sup>.

**Gouine**, nom féminin, « jeune fille », *ALB*, III, 1614 ; « vieille poupée », *ALB*, III, carte 1656. A. Demongeot signale *gwainy* (lire [gwèñ]), avec le sens de gouine, prostituée. S. Petit signale également à Serley *la gouine*<sup>16</sup>.

**Gouné**, Alex<sup>17</sup> « jupon »<sup>18</sup>.

**Gounalle** : René Lapierre : 1) haricot nain ; 2) lesbienne ; à Mercurey, « poupée mal habillée ».

**Guenette**, nom féminin, « brebis », *ALB*, II, 1111 ; « vieille brebis », *ALB*, II, 1112. Demongeot indique une forme *gairleut* [gèrlœt], avec le sens de mauvaise brebis (maigre ou vieille) ; Mongenet : *guenette* « mauvaise brebis ».

**Guenâ** : Mancey<sup>19</sup> : « niais, naïf, qui a l'air stupide, lent, traînard ; Etym : bressan : *guenard*. À Sagy : *guenard*, sobriquet collectif des habitants du village<sup>20</sup>.

---

<sup>15</sup> Ces formes ne sont là que pour mémoire ; elles sont évidemment issues de GULBIA.

<sup>16</sup> Samuel Petit, *Le Patois de Serley*, Dijon 1999. Ce même auteur indique dans le même village [là bøy] ; ces deux mots n'ont aucune valeur péjorative.

<sup>17</sup> Paul Alex, *Le Patois de Naisey* (Doubs), Paris 1965.

<sup>18</sup> Ce terme (\**gonnel*) devrait être classé sous *gunna* ; à remarquer la voyelle brève, alors que le verbe *gôner* (quel que soit son sens) a toujours une initiale longue. Cette initiale longue est constante, aussi bien à Dijon qu'à Lyon (comme nous l'a rappelé R. Chaudenson).

<sup>19</sup> Charles Millot, *Dictionnaire de patois de Mancey*, Tournus 1998.

<sup>20</sup> Voir P.-M. Alix et F. Boivin, *Notice sur la commune de Sagy*, Louhans 1901.

**Guenuche**, nom féminin, « poupée », *ALB*, III, 1656 ; cette forme est également signalée par A. Demongeot, avec le sens de « femme de mauvaise vie, dévergondée »<sup>21</sup>. Mongenet : *guenuche*, « femme dévoyée » ; à Belleruche<sup>22</sup>, *guenitse*, n.f., « jeune fille qui se laisse courtiser et espère le mariage ».

**Guignol**, nom masculin, « épouvantail », *ALB*, II, 735 (points 79 et 81).

### **Le Trésor de la langue française**

Nous suivrons principalement ce dictionnaire (*TLF*), ne serait-ce que parce qu'il est un des plus récents et qu'il a eu le loisir de prendre en considération les solutions proposées par les ouvrages antérieurs, principalement le *FEW*.

Pour **gouine**, le *TLF* signale principalement le sens du français moderne (« femme homosexuelle », sens apparu seulement à la fin du XIX<sup>e</sup> siècle) ; il rappelle le sens ancien « femme de mauvaise vie » ; le sens de poupée que nous avons relevé sur le terrain est totalement absent (il l'est également des autres dictionnaires). Pour l'étymologie, ce terme serait issu de l'hébreu *goy*, « le non juif » (voir *FEW* 4, 189b)<sup>23</sup> ; cette interprétation suppose qu'une (faible) minorité numérique puisse imposer à la majorité des termes péjoratifs, surtout des termes qui désignent cette majorité elle-même. D'autre part, que faire du *n*, qui apparaît dans toutes les formes. Il reste enfin le sens péjoratif, certes omniprésent dans les emplois modernes, mais pas nécessaire dans les emplois anciens. On pourra comparer avec l'évolution sémantique d'un des concurrents de *gouine* sur le terrain : *catin* est issu de CATHARINA, « la pure » (avec un croisement, il est vrai, avec *Hécate*, la déesse de la lune, ce qui explique les formes des pays slaves comme *Ecatherina*) ; la *catin* de nos petites filles est devenue elle aussi un des noms qui ont désigné la prostituée, alors que

---

<sup>21</sup> André Demongeot, *Il y a cent ans, Rochetaillée-sur-Aujon (Haute-Marne), sa vie et son parler*, Dijon, 1986. Dans le même ouvrage, on note [góy] ou [góyō] avec le sens de femme malpropre, souillon.

<sup>22</sup> Auguste Comby, *Dictionnaire du patois de Belleruche (Loire)*, Dijon-Lyon 1994.

<sup>23</sup> Le *Robert historique* se montre sur ce point beaucoup plus réservé ; il ne propose pas d'autre solution, mais l'article utilise beaucoup le conditionnel et les *peut-être*.

l'étymologie ne la prédestinait nullement à cette fâcheuse évolution sémantique. Le sens général du français (bien que légèrement vieilli) semble d'ailleurs assez récent ; *catin* est issu de *Catherine*, tout simplement parce que ce prénom eut autrefois une très forte fréquence<sup>24</sup>.

Pour **guenon**, le *TLF* fait un rapprochement avec *guenipe* ; et ce serait une racine d'origine gauloise, \*WADANA (voir *FEW* 14, 111) ; les dictionnaires étymologiques proposent ensuite des explications difficiles qui disent que la *guenon* serait ainsi nommée à cause de sa queue. Nous n'avons jamais vu de poupées à queue même pas dans la région francoprovençale où ce terme de *guenon* est assez répandu. D'autre part, la datation pourrait être revue ; le sens de *guenon* « sorte de singe, puis singe femelle » est généralement attesté à partir du début du XVI<sup>e</sup> siècle. C'est la date de la découverte des Amériques et on décrit des animaux jusque là inconnus ; les découvreurs ont donc le choix entre les noms déjà employés par les indigènes et leur propre vocabulaire ; or, nous avons trouvé un terme bien antérieur, dans le *Siège de Barbastre* ; dans cette chanson de geste du XIII<sup>e</sup> siècle, éditée par Bernard Guidot, nous trouvons au vers 5400 le terme *gagnon* ou *gaignon*<sup>25</sup>.

Pour **guignôle**, le troisième type de notre carte, il est difficile de ne pas faire un rapprochement avec le Guignol lyonnais ; il était en effet difficile d'éviter, ici à Lyon, que l'un d'entre nous ne s'aventurât point sur le chemin difficile de l'étymologie de *Guignol*. Il faut dire que l'occasion était trop belle. Au cours d'une réunion où l'on devait parler de la dialectologie, il était tentant de voir ce que les résultats de nos enquêtes pouvaient encore apporter à la connaissance de l'histoire de la langue. Il est vrai aussi que la gloire de Guignol dépasse largement le cercle étroit des linguistes. On peut, dans un premier

---

<sup>24</sup> Un jeune poète local, Paul Boccard publie en 1890 (Boccard est mort l'année suivante) un poème bucolique en patois intitulé *Tu m'boudes, Catiche*. Le prénom n'a manifestement aucun sens péjoratif ; on le retrouve d'ailleurs dans nos noms de la poupée (voir Paul Boccard, *Dreuleries...*, rééd. ABDO Dijon 1995).

<sup>25</sup> Il s'agit d'un passage où l'on voit les héros s'en prendre à une statue de Mahomet ; on la fait tourner comme un *gagnon* ; et une traduction par *poupée* ou *pantin* serait tout à fait possible. En fait, B. Guidot a retenu une autre leçon, mais ces deux manuscrits éliminés sont de la fin du XIII<sup>e</sup> siècle ou du début du XIV<sup>e</sup> siècle, dans tous les cas très antérieurs à la découverte des petits singes à queue des Amériques. Ce terme de *gagnon* apparaît également chez Chrétien de Troyes (*Yvain* 646) sous la forme *gaignon* ; on lui donne dans les traductions le sens de gros chien.



temps, dresser un rapide bilan étymologique de *Guignol*. Il faut reconnaître que ces mots, avant Wartburg, n'ont pas immédiatement interrogé les spécialistes de l'histoire de la langue : ainsi, Littré possède bien une entrée *Guignol* ; il en donne une définition exacte, mais il ne signale même pas l'origine lyonnaise de la marionnette, ce qui aurait pu être considéré comme une amorce d'explication étymologique. Le *Larousse du XX<sup>e</sup> siècle* apporte une importante information encyclopédique, mais rien sur la question linguistique ; on apprend cependant qu'il aurait existé un adjectif lyonnais *guignolant*, dont aurait été tiré le nom de la marionnette (mais le dictionnaire réfute cette explication peu plausible) ; il nous apprend également que le *Guignol* lyonnais a été inspiré par les pupazzi italiens, ce qui peut déjà nous mettre sur la voie d'une explication. Plus récemment, les étymologistes (*FEW* 17, 591a et 594a) nous ont proposé une véritable étymologie : le *guignol* est celui qui cligne de l'œil, du verbe *guigner* (du germanique \*WINGJAN), verbe assez répandu dans la région francoprovençale avec le sens de « remuer »<sup>26</sup> : C'est le point de vue qui sera repris par le *Trésor de la langue française* : *Guignol* est appelé ainsi, à cause des regards furtifs qu'il lance de tous côtés.

Avant d'aller plus loin, il conviendrait d'examiner la forme notée pour l'*ALB* ; nous l'avons notée au seul point 21 (Saffres) qui est un point particulièrement conservateur, du moins à l'époque de l'enquête, comme nous l'avons dit plus haut ; le terme est féminin et désigne plus particulièrement une vieille poupée. La coexistence avec [gwèñ] montre nettement que les termes sont apparentés (ou, du moins, sentis comme tels par les témoins) ; on pourrait certes penser à un emprunt très récent au français d'origine lyonnaise *guignol* ; mais la forme locale, phonétiquement évoluée, et le genre féminin, font penser surtout à une évolution purement locale<sup>27</sup> (dont l'expansion géographique a pu être autrefois plus importante). Nous penserons donc à un dérivé de *gouine*.

---

<sup>26</sup> On pourra se référer à la carte 907 de l'*ALB* « la bergeronnette » ; dans les points francoprovençaux (Bresse), on trouve le type « guigne-queue », mot à mot « qui remue la queue » ; cette explication avait été avancée auparavant par Gilbert Salmon au cours d'un colloque à Montpellier (mai 1983) ; mais nous n'avons pu retrouver les actes publiés par Paul Fabre.

<sup>27</sup> Il est évident que certains noms locaux de l'épouvantail, notés sous la forme française *guignol*, viennent directement du nom lyonnais. Voir *ALB* II, 735.

Les **gones** : nous venons donc de trouver sur une même carte trois termes sémantiquement très voisins ; et, toujours parce que nous sommes à Lyon, nous ajouterons le terme de *gone*, non traditionnel de l'enfant dans cette ville (le terme figure depuis quelque temps dans les dictionnaires courants, comme le *Petit Larousse*, mais on signale toujours qu'il s'agit d'un terme lyonnais). On a fait un rapprochement avec le cas latin GUNNA (*FEW* 4, 326)<sup>28</sup> ; le gone serait donc ainsi nommé à cause de ses vêtements sales et en désordre<sup>29</sup>.

## Pour une autre solution

Nous nous trouvons donc devant une série de mots paradoxaux ; ces mots, à l'exception de *gone* que nous avons ajouté, sont réunis par nos cartes linguistiques ; mais ils sont partagés par les étymologistes entre plusieurs solutions ; sémantiquement, les résultats de l'*ALB* sont assez simples : ces mots désignent la poupée ; certes le sens est souvent péjoratif ; mais c'est souvent le sort des mots patois que de devenir péjoratifs ; pour rester dans le vocabulaire des jouets, on pourrait citer la *fiarde*, la toupie ; qui oserait entrer dans une boutique de jouets pour acheter une *fiarde*, alors qu'on utilise volontiers l'expression *ronfler comme une fiarde* ? En revanche, les solutions des étymologistes tournent volontiers autour de l'idée de saleté, idée très secondaire dans les formes dialectales ; le *goy*<sup>30</sup> n'est pas forcément sale, mais les termes issus (ou prétendus issus du mot) le sont, souvent à la suite d'évolutions récentes. À ces termes qui tournent autour de l'idée de saleté, on pourrait même ajouter la racine GON « porc »<sup>31</sup> à laquelle le

---

<sup>28</sup> Il nous semble évident que tous les termes liés à la notion de vêtement sont liés à cette racine, comme le bourguignon *gôné* « mal habillé » ; *vos êtes jâr dreûlemant gôné po in homme de vote âge*, dit la Castafiore au Capitaine Haddock (*Lés ancarpions de lai Castafiore*, p. 22, bande 3, image 1). Voir plus haut, note 18.

<sup>29</sup> Pour une étude plus complète de *gone*, on peut se reporter au *Dictionnaire des régionalismes de France*, de Pierre Rézeau (2001) ; l'explication historique est surtout consacrée aux attestations du mot dans les parlers actuels ; on le rattache cependant à GUNNA, à la suite du *FEW* (4, 325) qui considère que *gone* est refait sur le verbe *goner* « mal vêtir » ; mais le lien sémantique avec *gunna* « vêtement » est loin d'être évident.

<sup>30</sup> On peut s'étonner que les dictionnaires étymologiques n'établissent jamais de lien entre les *goïm* et les païens ; et pourtant, PAGANUS, de même que GENTILIS et ETHNICUS qui ont eu moins de succès, sont bien des calques de *goïm* ; en cette année consacrée à saint Paul, on pourrait peut-être y réfléchir, bien que ce ne soit pas une question de dialectologie.

<sup>31</sup> Nous ne reviendrons pas sur ce point ; nous avons dit par ailleurs que le rebouteux (le *regueugnou*) n'avait rien à voir avec le cochon ; cette série de formes devrait être classée sous

*FEW* a rattaché un des noms bourguignons du rebouteux, comme si cet homme faisait un travail de cochon et qu'il n'était pas estimé de plupart de ceux qui faisaient appel à ses services. Dans tous les cas, une meilleure analyse sémantique serait nécessaire.

On pourrait penser alors à une racine \*GON- qui a désigné soit la marionnette, soit de façon plus générale, l'enfant. Nous ne voudrions pas ici réécrire complètement les travaux antérieurs ; mais on pourrait peut-être penser à une racine apparentée à des racines bien attestées en gaulois « fils »<sup>32</sup>.

## Pour conclure

Nous ne voudrions pas imposer ici, dans ce qui doit rester de simples réflexions, de nouvelles étymologies ; si le lien paronymique entre *gouine* et *guignol* nous semble évident pour les patoisants de Saffres (du moins pour ce qu'il en reste, puisque notre principal témoin, Marcel Boucherot, aurait aujourd'hui largement plus de cent ans), le lien historique reste encore à établir ; certes on pourra dire que les atlas l'ont déjà fait en partie et que la géographie a permis de mettre en relief des liens qui avaient échappé à des chercheurs longtemps habitués à travailler sur des listes et non sur des cartes (ce qui n'est plus vrai des rédacteurs actuels du *FEW* qui ont lu les cartes, du moins pour les derniers fascicules de la lettre A, ce qui leur a permis de remettre parfois en cause les solutions de la première mouture). Le mot a une autre dimension que sa forme et son sens ; il a aussi son espace et cette notion d'espace doit entrer dans les études qui lui sont et, surtout, qui lui seront consacrées.

De façon plus générale, on voit l'immense intérêt que peut représenter pour l'histoire de la langue française la prise en considération des résultats de la dialectologie, peut-être en dépassant le jeu de la simple hypothèse étymologique qui, souvent, consiste à réunir ce que nos prédécesseurs ont séparé ou, dans d'autres cas de

---

CUNEUS. On pourra se reporter à notre étude "Les noms du rebouteux en Bourgogne", in *Mélanges Lepelley – Annales de Normandie*, Caen 1995, p.187-192.

<sup>32</sup> Peut-être apparenté à des mots comme *genos* ou *gnatos* ; voir X. Delamarre, *Dictionnaire de la langue gauloise*, Paris 2001.

figure, séparer ce qu'ils ont réuni<sup>33</sup>. Nous ne sommes certainement pas à la fin de la dialectologie ; nous sommes certainement dans une période de jachère, comme nous y autorisent les emplois de *fin* encore présents en toponymie ; *les Fins* étaient autrefois des terrains généralement éloignés des villages ; aujourd'hui, par suite du développement urbain et industriel, ces *Fins* retrouvent une nouvelle vie. Souhaitons que la dialectologie soit comme ces *Fins* de nos cadastres. Peut-être convient-il de lui fixer de nouveaux buts, de nouvelles *fins* ?

---

<sup>33</sup> C'est précisément ce que nous avons fait dans une étude sur les *doux* (sources), à paraître dans un prochain numéro des *Cahiers de la Société française d'Onomastique*.

*Francis Manzano*  
*Université de Lyon (Lyon 3, Jean Moulin)*  
*Centre d'Études Linguistiques (EA 1663)*  
*francis.manzano@neuf.fr*

**Du magasin dialectal de la France : une approche  
géodiachronique et sociolinguistique des terrains**  
*Du français classique à la dialectologie*

Dans le même ouvrage, Jean Le Dû, Robert Chaudenson, Jean-Philippe Dalbera, rappellent des explications utiles si l'on veut comprendre pourquoi et comment se sont développées la dialectologie romane en général, la dialectologie française en particulier.

J'aimerais à mon tour revenir sur la spécificité française, en soulignant de mon côté le lien profond qui existe entre dialectologie et histoire ancienne du paysage sociolinguistique de la France. Si la dialectologie « française » (plus justement gallo-romane, voire occitano-romane, celtique etc. pour nombre d'entre nous<sup>1</sup>) est devenue un modèle depuis le XIX<sup>e</sup> siècle, c'est bien parce qu'elle s'inscrivait au sein d'un paysage exemplaire, concret et symbolique en même temps, articulé autour d'une normativisation très haute de la langue française, elle-même sortie du terrain dialectal du très Ancien Régime.

\*

Le raisonnement s'articulera principalement autour du concept sociolinguistique de *magasin dialectal*, introduit dans mes recherches

---

<sup>1</sup> L'expression « dialectologie française », certes courante faute de mieux, a toujours été de la plus grande ambiguïté, car elle repose sur l'acception variable de l'adjectif « française ». Parle-t-on de France politique, de France ethnique, etc. ? Dans cet ouvrage par exemple, si l'on est à peu près assuré que Jean Le Dû, Xavier Ravier ou Fabrice Jejcic s'intéressent bien à la géographie linguistique, font-ils tous (et à quel degré) de la dialectologie « française » ?

La question de la dénomination est en effet importante, mais très discutable. Elle demanderait à elle seule un développement spécifique. Déjà, Albert Dauzat (*La géographie linguistique*, Paris, Flammarion, 1922 : 5) aurait préféré *linguistique géographique*, « mais il est trop tard pour revenir sur une appellation consacrée par de nombreux travaux ».

sur le gallo lorsque je me trouvais à Rennes, repris depuis à plusieurs reprises. Ce concept peut, à mon sens, permettre de mieux comprendre structurellement l'ascension historique de la langue française, par une procédure de « mise à part » prise en charge par des élites que l'on connaît bien. Pour que cette ascension prenne au mieux sa force, il fallait aussi une dynamique d'ensemble qui ségrège pour longtemps les élites et le peuple. Le monde des laboureurs fut donc laissé à ses campagnes et garda notamment ses typologies sociales, culturelles et langagières originelles, l'ensemble étant globalement stabilisé jusqu'à l'arrivée des dialectologues (fin XIX<sup>e</sup>). Chaque ordre avait ainsi ses prérogatives et la France des élites, fondée sur le peuple dont elle se séparait pourtant de plus en plus, pouvait ainsi mettre en place adroitement sa propre mise en scène, présentée au reste du monde comme une nouvelle norme culturelle et langagière pour l'avenir, d'essence presque surnaturelle.

Aux élites la qualité, la modernité et le bon goût. Aux provinces paysannes le fatras, le maquis dialectal et la rusticité.

## 1. Vaugelas et l'Académie Française

Pour comprendre les principes essentiels des actions sur la langue aux XVII<sup>e</sup> et XVIII<sup>e</sup> siècles, on doit toujours repartir de ce contraste saisissant entre la faible représentation du français sur le terrain des pratiques et la très haute position sociolinguistique recherchée, affichée pour cette langue. On pourrait naïvement penser que cet écart pouvait affaiblir le projet. Il n'en est rien et c'est même le contraire. Les actions sur la langue, durant la période classique, supposent à peu près toutes un tri géo-social préalable. Sur ce point, Vaugelas, dans la célèbre Préface à ses *Remarques sur la langue françoise*, est on ne peut plus clair<sup>2</sup>. Le français qui servira de norme ne peut être que celui d'une élite sociale, le postulat social qui guide Vaugelas et les premiers académiciens y étant clairement exprimé, quand il écrit à propos du bon usage :

« [...] Il y a sans doute deux sortes d'usages, *un bon et un mauvais*. Le mauvais se forme du plus grand nombre de personnes, qui presque en toute chose n'est

---

<sup>2</sup>Les citations ou renvois qui suivent sont tous relatifs au fac-similé des *Remarques sur la langue françoise* édité par Jeanne Streicher. Paris, Droz, 1934.

pas le meilleur, et le bon au contraire est composé non pas de la pluralité, mais de l'élite des voix [...] »

Ce postulat non discuté et non discutable, la mission de l'Académie sera donc très nettement dessinée : ce sera la constitution d'une base de données représentatives du bon usage (la littérature), la codification grammaticale de ce bon usage (et non d'un autre), le rejet des formes et structures jugées périmées (lutte contre les archaïsmes), la chasse aux mots du vieux français et des provinces (*infra*, thème du « magasin dialectal ») et enfin la proposition d'un dictionnaire sélectif. On peut voir clairement à ce stade que la démarche d'ensemble consiste bien à produire l'abstraction langagière et normativisante d'une élite sociale qui cherche à se séparer symboliquement du terrain d'origine. Cette opération se place d'autant plus sur un plan symbolique que, par ailleurs, sur un plan concret, ce terrain dont on entend se séparer est la condition même de l'existence physique de l'élite : récoltes, élevage, bref, le Royaume de France profond, le tiers-État sans lequel l'élite ne pourrait ainsi se mettre à part en termes de langue, de culture, et de bon goût.

Vaugelas faisant partie du premier train d'académiciens, on ne sera pas étonné d'observer que trois articles des statuts fondateurs, approuvés par le Parlement en 1637, déterminent clairement ces objectifs<sup>3</sup>.

#### Article 24

La principale fonction de l'Académie sera de travailler avec tout le soin et toute la diligence possible à donner des règles certaines à notre langue et à la rendre pure, éloquente et capable de traiter les arts et les sciences<sup>4</sup>.

#### Article 25

Les meilleurs auteurs<sup>5</sup> de la langue françoise seront distribués aux Académiciens pour observer tant les dictionnaires que les phrases qui peuvent servir

---

<sup>3</sup>Depuis les années 1620, un groupe d'intellectuels se réunissait de manière informelle. Appuyés par Richelieu, ceux-ci débouchèrent donc sur une forme institutionnelle (fondation de l'Académie), dont on voit bien l'intérêt pour l'État à ce moment.

<sup>4</sup>Il faut ainsi comprendre que le français était donc auparavant (Moyen Âge, Renaissance) dépourvu de règles certaines (thème du maquis langagier), impur (= contaminé par le vieux, le provincial, le sale ou trivial), incapable de traiter les arts et les sciences (il faut donc le rapprocher du latin). Mais il est à remarquer qu'il n'est pas réputé incapable de traiter les lettres, puisque celles-ci sont une pierre angulaire de l'action (cf. article 25).

<sup>5</sup>L'Académie française ne disposant pas d'une machine toute prête et infaillible permettant de reconnaître les meilleurs auteurs, ceux-ci se recruteraient donc nécessairement dans cette « élite des voix » qui se trouve dans l'entourage de la Cour et de l'Académie. Toute littérature suppose on

de règles générales et en faire le rapport à la Compagnie qui jugera de leur travail et s'en servira aux occasions<sup>6</sup>.

Article 26

Il sera composé un Dictionnaire, une Grammaire, une Rhétorique et une poétique sur les observations de l'Académie<sup>7</sup>.

## 2. Le purisme

On constate que l'objectif de pureté figure explicitement dans cette première charte. Sans doute était-il en germe durant les époques précédentes (Renaissance, Malherbe)<sup>8</sup>, à cette réserve près qu'on ne concevait pas alors la pureté du langage comme apanage exclusif des élites sociales, d'Île de France notamment<sup>9</sup>. Du moins le système

---

le sait bien des réseaux d'influence sociale, à toute époque et en tous lieux. Mais dans ce cas précis, c'est assez nettement de contrôle, d'autorité et de clientélisme qu'il s'agit. Plus d'un écrivain déviant de l'époque classique fera les frais de cet état d'esprit. L'un des plus grands assurément, Corneille, subira directement les attaques très vives de l'Académie lors de la création du Cid (polémique de 1637).

Dans un courrier remarquable à Chapelain (l'un des premiers académiciens), Guez de Balzac écrivait (26 mars 1634) : « Vous me mandez que vous avez été reçu par grâce dans l'Académie des beaux esprits. Et moi je voudrais vous demander qui a reçu les beaux esprits qui vous ont reçu ? D'où vient le principe de l'autorité et la source de la mission ? ».

<sup>6</sup>Il faut souligner que la méthode ainsi tracée avait toute chance de ne déboucher sur rien de valable. Des collections de remarques ne font pas une norme.

<sup>7</sup>L'Académie, au XVII<sup>e</sup> siècle, ne produira qu'un Dictionnaire (1694), généralement jugé inférieur à deux concurrents antérieurs : Richelet, *Dictionnaire français etc.* (1680) et Furetière, *Dictionnaire universel* (1690). Furetière, préalablement académicien, fut chassé de la Compagnie en 1687, notamment accusé d'avoir dérobé des épreuves du Dictionnaire de l'Académie. L'Académie possédant depuis 1674 le privilège royal exclusif des dictionnaires, Richelet et Furetière durent publier leurs travaux à l'étranger.

À aucun moment, durant l'époque classique, l'Académie ne sera en mesure de produire un livre de grammaire, malgré cette mission fondamentale (article 26). Sans entrer dans le détail, on peut donner deux explications fondamentales à cette incapacité : surtout la méthode de travail (voir article 25 de l'Académie) qui ne pouvait fatalement déboucher que sur des collections infinies de remarques ; mais également la mobilité du français lui-même (notamment dans les années 1610-1660), qui entravait sérieusement toute vue systémique d'ensemble, cette vue eût-elle existé.

<sup>8</sup>Par principe général la plupart des centralisations politiques et administratives (c'est particulièrement net pour le Royaume de France) tendent vers une ou des unifications progressive(s) de codes langagiers. Ce même phénomène implique un sous-principe de tri, d'élimination, de choix, dont l'une des conséquences est la recherche (souvent chimérique) de pureté. Pour finir, cette déviation peut revenir sur la fragmentation sociale, en favorisant l'identification et la reproduction de certains groupes qui la mettent à profit. C'est ce que l'on voit dans cette partie.

<sup>9</sup>On connaît bien la position de Malherbe (fin XVI<sup>e</sup>/début XVII<sup>e</sup>) qui, en cas de doute (car nombre d'hommes de lettres le consultaient en matière de bon langage), renvoyait aux Crocheteurs du Port au Foin (quais de Paris). On peut le considérer comme héritier en cela d'une lointaine tradition médiévale plus unanimiste et bien moins tranchée sociologiquement dans laquelle l'écart entre



sociolinguistique général de cotation restait-il relativement souple. Les Provinces ou les provinciaux, le peuple, pouvaient à des degrés divers hériter de bonnes traditions en la matière.

Ce point de vue n'a plus guère de place à partir de la seconde moitié du XVII<sup>e</sup> siècle, compte tenu de ce qu'on vient d'observer. Au contraire, le purisme en matière de langage aura désormais comme corollaire l'éloignement par rapport au peuple et aux provinces de langue française<sup>10</sup>, à plus forte raison par rapport aux provinces ethniquement non françaises, romanes (langues d'oc, catalan) ou non romanes (Bretagne occidentale, Pays basque, Flandre etc.)<sup>11</sup>. Aussi l'objectif suivant (*infra*, nettoyage) deviendra-t-il fortement caractéristique du français et de l'esprit français. On peut tenter une métaphore : tout se passe comme si le français, durant cette période, était atteint de ce que les psychiatres qualifient de trouble obsessionnel-compulsif<sup>12</sup>. On le voit, de manière récurrente, se débarrasser de tout ce qu'il juge être souillure : tournures et mots populaires, régionaux, techniques etc., soit le programme à peu près complet du dictionnaire initial de l'Académie, tel qu'il est énoncé dans la préface. Ce qu'affirmait aussi Vaugelas, dans ce conseil donné à ceux qui recherchent la pureté du langage :

---

manants et élites n'a pas encore été porté à l'excès. Ce genre de positionnement d'un membre de l'élite sera désormais saugrenu, voire indigne. Vaugelas en particulier s'indigne explicitement des positions de Malherbe. Par exemple, quelle part de vérité des paysans illettrés, engoncés dans leurs pratiques et leurs affreux patois, pourraient-ils porter ?

<sup>10</sup>C'est-à-dire en fait de langue d'oïl.

<sup>11</sup> Toutefois, dès le XVII<sup>e</sup> (et on le verra mieux encore durant la Révolution : rapport Grégoire et divers autres), une forme de compensation, de retournement, est possible. Elle consiste à voir certaines langues de France géographiquement et typologiquement marginales (germaniques, italo-romanes) comme des avatars ou formes locales de langues européennes elles-mêmes reconnues dans la cour des grands (allemand, italien). Non sans ambiguïté, car on évoque alors une forme de trahison en faveur de l'étranger dans des provinces de ce genre. Mais, trahison ou pas, le cas de telles langues régionales marginales n'est-il pas au fond relativement moins grave que la trahison « culturelle », cette corruption multiforme du français par le peuple et les paysans de la France d'oïl qui, dès qu'ils parlent, insultent le génie même du français. C'est évidemment l'opinion probable de certaines élites, non la mienne. Pour penser autrement, il faudra encore d'autres écrivains, eux-mêmes marginalisés, comme Molière ou La Fontaine, pour n'en citer que deux.

<sup>12</sup>Dans l'éventail des troubles mentaux, l'obsessionnel-compulsif développe à l'excès des rituels de nettoyage (corps, domicile etc.) à caractère anxiogène. Tel se lave les mains plusieurs fois par jour. Tel autre accumule les douches. Tel autre en vient à refuser de serrer les mains. Tel autre est obnubilé par les saletés qui se collent à la semelle de ses chaussures et ne les laisse pas pénétrer dans la maison etc.

« [...] Il faut estre assidu dans la Cour et dans la frequentation de ces sortes de personnes, pour se prevaloir de l'un et de l'autre, et il ne faut pas se laisser insensiblement corrompre par la contagion des Provinces, en y faisant un trop long séjour ».

Si l'on doute encore du processus, voici donc un mot-clef, celui de *contagion*, qui indique bien l'ampleur de l'obsession et de la coupure. La langue saine et pure se trouve dans une zone bien cernée de la France, celle du « bon usage » (*supra*). Aussi doit-on organiser une véritable quarantaine langagière, qui seule mettra le français à l'abri pour longtemps.

Confrère de Vaugelas, Faret livre directement deux autres mots-clefs dans un discours très connu sur les tâches de l'Académie :

« [...] *nettoyer* la langue des *ordures* qu'elle avoit contractées, ou dans la bouche du peuple ou dans la foule du Palais et dans les impuretés de la chicane, ou par les mauvais usages des courtisans ignorants [...] »<sup>13</sup>

Contagion, ordures, nettoyage. Par le purisme, la centralisation française s'engageait nettement dans un programme de nettoyage, de mise en ordre qui allait se dérouler par la suite, au-delà des constitutions et régimes politiques, fondant jusqu'à nous une programmation subliminale des élites de tout poil.

Un tel programme amènerait à rejeter toujours plus les usages archaïques, témoins de périodes confuses et arriérées (Moyen Âge, Renaissance), les usages régionaux d'oïl (versions rurales plus attardées de l'ancienne langue), et, évidemment plus éloignées typologiquement du français, les langues non françaises du pays ou les langues romanes autres que le français.

C'est, en somme, en le travaillant, en le nettoyant et en l'amputant sur ces différents axes complémentaires, que l'on ferait du français une langue vraiment « classique », égale ou supérieure au grec et au latin dans le panthéon des langues-phares de l'humanité<sup>14</sup>.

---

<sup>13</sup>Notamment cité par Caput, Jean-Pol (1972 : 203). *La langue française, histoire d'une institution*. Paris, Larousse.

<sup>14</sup>Tout au long de la période classique, et par la suite encore, existera toutefois une tradition de refus de cette centralisation, avec ouverture vers les variétés archaïques, provinciales ou populaires. Un des meilleurs exemples de cette attitude en matière de choix de la langue, reste celui de La Fontaine au XVII<sup>e</sup> siècle, type « marotique » selon Voltaire. Cette tradition sera maintenue ensuite, de Marivaux à Céline ou Giono plus près de nous, avec des facettes différentes. Mais il est également visible que de tels écrivains, décalés par rapport à la centralisation

### 3. Nettoyage, mise à part et magasin dialectal

En soi, le nettoyage de la langue n'a rien d'anormal, on peut le prendre comme une preuve de bonne santé et de dynamisme. Il peut correspondre alors à l'adaptation nécessaire des langues aux évolutions des environnements dans lesquels elles sont utilisées. C'est là une contrainte socio- et éco-linguistique qui pèse notamment et de manière éclatante sur le lexique. Les mots vieillissent, de nouveaux arrivent (plus adaptés à l'époque et aux besoins), déterminant de nouveaux parcours sémantiques pour les unités conservées etc. Il serait d'ailleurs plus approprié de parler de *toiletage* nécessaire, effectué automatiquement par les langues naturelles à la surface du globe, le plus souvent sans le secours des grammairiens et lexicographes.

Ainsi, depuis l'ancien français, des pans entiers du lexique français, fonctionnels à l'époque de Béroul ou de Chrétien de Troyes se sont-ils normalement dilués. Par exemple, le vocabulaire de la féodalité, fondé, on le sait, sur des systèmes étymologiques germaniques, s'est-il transformé en vaste coquille vide durant le moyen français et à l'approche de la Renaissance. Et les *harnois*, *quintaines*, *adoubements* etc. sont bien pour nous aujourd'hui des coquilles vides à des degrés différents, que l'on qualifie d'ailleurs de mots d'époque. À cet égard, une véritable « dé-germanisation » du lexique français est en cours depuis l'ancien français, *dégermanisation* ayant d'ailleurs suivi une *déceltisation* commencée bien plus tôt. On peut même dire qu'elle a commencé avec l'existence du français lui-même, roman et contradictoire par principe de la pression du germanique ancien. Voici pour le « naturel » des langues.

Le problème se pose autrement à l'époque classique, car ce sont maintenant des hommes qui se donnent l'autorité pour procéder à ce *toiletage*. Mais, comme on l'a vu, l'autorité est construite sur la base d'une restriction géo-sociale, aussi le *toiletage* devient-il nettoyage, et le nettoyage devient-il presque automatiquement expurgation<sup>15</sup>.

Le plus remarquable pour notre réflexion est que cette expurgation se fait à peu près toujours de la même manière, parce que le plan d'action était impliqué par les primats de départ. Ce plan très visible

---

recherchée du français, ont toujours eu des difficultés, au moins provisoires, à être admis, reconnus par leur époque.

<sup>15</sup>Vaugelas notamment illustre cette litanie, qui égrène dans ses *Remarques sur la langue françoise* des mots « vieux », « guère bons », des « fautes familières à toutes les provinces », des mots « que l'on ne dit plus » etc.

dans l'ordre de la phonétique et du lexique, un peu moins sans doute pour celui de la morphosyntaxe, consiste le plus souvent à rejeter les souches lexicales ou les prononciations qui ancreraient le français dans la langue d'oïl ancienne, et, indirectement, dans la Romania. Pour le lexique notamment, deux grandes solutions se dégagent alors, soit l'abandon pur et simple de la souche, consacré par les dictionnaires de la période classique, soit la réorientation sémantique de l'unité. Dans les deux cas, on parvient à peu près au même résultat : le français ne peut plus être immédiatement rattaché à son terrain d'origine, les liens sont brisés et il faut beaucoup de soin pour les reconstituer.

Mais il y a plus. Ce qu'il rejette pour lui, le français le laisse et le confirme pour les provinces. Car cette confirmation présente l'intérêt d'alimenter l'auto-reconnaissance du français dans sa mise à part, qui par cette rupture se confirmera lui-même toujours plus dans son statut de langue élue ou d'exception. Les provinces d'oïl (et d'oc dans une large mesure) deviennent alors un vaste magasin dans lequel le français va déposer les chaussures sales qu'il ne veut plus voir entrer dans la maison. Là, dans ce magasin, on quitte le temps historique réel et les ruraux y entretiennent par ignorance, paresse et manque d'énergie<sup>16</sup>, tout le bric-à-brac dont le français s'est séparé à diverses époques. Comme on le rappellera un peu plus bas, les premiers dialectologues eux-mêmes semblent convaincus de cette « bassesse » des patois face au français.

Bric-à-brac, disparité, données hétéroclites, archaïques, périmées, surprenantes, amusantes voire offensantes pour une oreille éduquée, urbaine. Retenons ce cortège de qualifications.

Ces dernières années, j'ai choisi de dénommer *magasin dialectal* ce vaste ensemble construit négativement par l'émergence française.

Derrière *magasin*, il y a bien entendu l'image de la vieille échoppe fourre-tout des provinces d'autrefois : ce que vous ne trouvez plus en ville depuis bien longtemps, vous avez de bonnes chances de le trouver là. C'est aussi un magasin *dialectal* pour deux raisons conjointes. Tout d'abord parce que seuls savent vraiment y fouiner les diachroniciens/dialectologues dont c'est le métier, et qui y voient là une aubaine<sup>17</sup>. Ensuite parce que la qualification le plus souvent

---

<sup>16</sup> Encore une fois, ces termes ne sont pas de mon fait, mais apparaissent invariablement depuis le XVII<sup>e</sup> siècle.

<sup>17</sup>La dialectologie (XIX<sup>e</sup> et XX<sup>e</sup> siècles) se présente historiquement comme une technique exploratoire des campagnes concédée par la grande romanistique et l'ancien français qui tiennent le haut du pavé universitaire à la fin du XIX<sup>e</sup> siècle. Il faut dire en effet que, dans une large

négative de « dialecte » en français commun rend difficile tout contenu « positif » de ce magasin<sup>18</sup>. Comment notamment un « dialecte » de paysans jugés rétrogrades<sup>19</sup> pourrait-il donner une quelconque leçon d'authenticité au français ?<sup>20</sup>

C'est surtout un magasin au sens le plus historique ou étymologique du terme. Cet endroit où l'on entrepose de manière surprenante, souvent anarchique, des marchandises d'origine et d'intérêt divers. Quelque chose qui ressemble aussi au bazar<sup>21</sup>. Le point le plus important est sans doute cette absence d'ordre, ou cet ordre d'un autre type, non conforme au rationalisme qui se développe et s'installe précisément à partir du XVII<sup>e</sup> siècle. Une masse où l'on ne voit pas le principe d'arrangement, s'il en existe un. Tout se passe donc comme si le regard central, urbain, français, était inapte à comprendre (c'est-à-dire à saisir dans son ensemble, à organiser) cet univers, pour la bonne raison qu'il ne possède pas le code d'agencement, et que son propre code de lecture ne convient pas à ces campagnes constamment décalées et parfois rebelles.

\*

Un détour rapide par l'ethnographie permettra de mieux articuler le raisonnement. Bronislaw Malinowski<sup>22</sup>, après Lucien Lévy-Bruhl<sup>23</sup>,

---

mesure, les romanistes en vue de l'époque envoient sur le terrain des enquêteurs dont ils attendent qu'ils leur permettent d'enrichir leurs théorisations de la fragmentation et de la taxinomie des langues romanes dans leur ensemble. De cette « concession » universitaire aux gens de terrain est sans doute restée une composante de la dialectologie, à propos de laquelle William Labov écrit ceci : « la dialectologie en tant que discipline paraît avoir perdu tout contact avec la linguistique théorique, et, désormais, les tenants de la géographie linguistique se contentent le plus souvent de recueillir leurs matériaux et de les publier » (*Sociolinguistique*, 1976 [Les Editions de Minuit] : 362).

<sup>18</sup> It. *magazzino*, arabe *māḡẓān* (désignant principalement un entrepôt).

<sup>19</sup> C'est ainsi qu'ils sont le plus souvent vus et qu'ils sont « dits », du XVII<sup>e</sup> au milieu du XX<sup>e</sup> siècle. Les mentions plus positives, objectives ou euphorisantes, viendront plus tard (deuxième moitié du XX<sup>e</sup> siècle essentiellement), avec l'avancée de l'écologie générale.

<sup>20</sup> Cette argumentation a été notamment retournée par certains courants militants depuis quelques dizaines d'années. Parmi les plus récents, le mouvement gallésant fait usage régulièrement du procédé, listant et réhabilitant des séries de mots (plus ou moins fonctionnels) généralement vus comme « paysans » ou à tout le moins bizarres (type *parlement* « acte de parler » [qui avait cette signification en ancien français, et ne l'a plus], *demanderie* « demande en mariage », *netée* [pour nuit] etc.).

<sup>21</sup> Cette fois tiré du persan, passé par le portugais.

<sup>22</sup> Ethnologue ou ethnographe anglais, d'origine polonaise (1884-1942). Il a travaillé dans la zone pacifique et notamment aux Îles Tobriand (Mélanésie), laissant l'un des ouvrages les plus innovateurs et l'un des plus beaux titres du champ avec *Argonauts of the Western Pacific* (Londres, 1922, en français : *Les Argonautes du Pacifique occidental*). Différents aspects de son œuvre sont aujourd'hui critiqués ou même rejetés, mais son apport est doublement important pour nous. D'une part B. M. est un réel acteur et théoricien de l'observation participante, principe suivi par différents dialectologues et sociolinguistes à la même époque et depuis (souvent automatiquement et sans

s'est beaucoup préoccupé de la mentalité des « primitifs »<sup>24</sup>, puisque l'on s'exprimait ordinairement ainsi à la fin du XIX<sup>e</sup> et durant les premières décennies du XX<sup>e</sup>. Pour le premier comme pour le second, l'ethnographe, l'ethnologue, doivent se déplacer d'une culture dans l'autre, car les requis pour fonctionner, s'intégrer, comprendre, ne sont pas les mêmes.

Comme l'écrit Lévy-Bruhl<sup>25</sup> :

Il faut essayer de nous mettre dans leur attitude mentale, au lieu de leur prêter la nôtre<sup>26</sup>.

Plusieurs remarques puissantes de l'un et l'autre, feront couler beaucoup d'encre dans les cercles de l'anthropologie et indirectement des sciences du langage.

---

réelle théorisation). Pour Malinowski, le chercheur doit entrer préalablement dans l'univers des enquêtés, s'appropriiser en somme, pour espérer reconstruire avec les mots et les concepts des enquêtés la trame et les linéaments de la société prise comme un ensemble fonctionnel. Ce qui amène au deuxième apport : l'interprétation. L'univers social des Tobriandais (pour ne retenir que ceux-ci) est un système équilibré qui ne se comprend que de l'intérieur et résiste aux interprétations rationaliste et historique. C'est un autre monde humain, avec d'autres références que celles de l'observateur, d'autres codes. Et l'on débouche alors sur une conclusion très ambiguë si on la décrypte seulement au premier degré : les sociétés primitives n'ont pas d'histoire et de chronologie. C'est pourquoi j'ajouterai, dans la même veine, que les terrains dialectaux n'ont pas d'histoire et de chronologie, sauf celles qu'on tente de leur donner de l'extérieur (notamment par la dialectologie).

<sup>23</sup>Lucien Lévy-Bruhl (1857-1939), plus encore que Bronislaw Malinowski, est aujourd'hui victime d'oubli ou de franc rejet. Son approche des « sociétés primitives », développée à la même époque que l'*Atlas Linguistique de la France* ou les travaux dialectologiques parents (cela doit être souligné), donne à comprendre que ces sociétés fonctionnent sur des bases mystiques ou pré-logiques, le modèle interprétatif occidental ne s'y appliquant de toute façon pas.

Les critiques *a posteriori* sont faciles, d'autant que L. L.-B. a tout mis de son côté, sa position sociale de bourgeois d'une part, ses titres d'ouvrages d'autre part, qu'on en juge : *Les fonctions mentales dans les sociétés inférieures* (1910), *La mentalité primitive* (1927), *L'âme primitive* (1931). Malinowski lui-même a critiqué Lévy-Bruhl, comme beaucoup d'autres par la suite. Mais au-delà de ces critiques certainement nécessaires, un clivage puissant persiste entre les logiques rurales et urbaines, entre « primitifs » et « civilisés », sociétés « froides » et sociétés « chaudes » (Claude Lévi-Strauss). Ce principe peut être habillé différemment, il nous interpelle directement ici. C'est ce que nous poursuivons.

<sup>24</sup> Bien entendu, en théorie, les primitifs sont ailleurs et à des degrés divers (nombreuses colonies, outremer), car l'ensemble du raisonnement doit être replacé dans l'époque et ses conditionnements. Mais bien qu'on ne le dise guère (et surtout qu'on ne l'écrive guère, ou avec des pincettes), est également là à ce moment la grande frontière entre d'une part l'homme occidental, urbain, dynamique et en devenir, et d'autre part l'homme « d'avant », des bois et des campagnes, rural, statique, condamné au retard culturel et technologique.

<sup>25</sup> Dans *Carnets*, PUF, Quadrige, 1998 (original de 1949).

<sup>26</sup> Quel dialectologue, quel sociolinguiste de terrain ne s'est pas aperçu de cette évidence ? François Laplantine utilise à ce propos le terme d'*acculturation à l'envers* (*L'anthropologie*, Paris, Payot, 2001 : 158). La qualification vaut sans doute pour l'ethnographie comme pour la dialectologie.

Elles nous intéressent directement et concernent l’histoire, plus justement la perception du temps. Pour les « primitifs », le passé se présenterait comme un tout non hiérarchisé, précisément un « vaste magasin d’événements » :

Toute notion de succession d’époque est absente de leur esprit. Le passé est un vaste *magasin* d’événements [...] tous les événements passés sont mis par eux sur un seul et même plan, sans qu’il y ait de gradation (Malinowski, *Argonautes*).

Claude Lévi-Strauss<sup>27</sup> qualifiera plus tard (avec réserves) ces sociétés de *froides*, pour les différencier des *sociétés chaudes*, ou sociétés à histoire. Dans un entretien (*Le Nouvel Observateur*, samedi 5 juillet 1980), il refuse d’ailleurs la simplification binaire de cette opposition présumée et donne quelques précisions utiles pour nous :

D’abord, entendons-nous sur la notion de société « froide ». Comme dans celle de société « chaude », il s’agit dans ma pensée de limites théoriques, mais qui ne sont jamais réalisées dans les faits ; je répète qu’aucune société n’a jamais existé dont on puisse dire qu’elle soit absolument froide ou chaude.

Second point : ces notions ne caractérisent pas un état intrinsèque de telle ou telle société mais l’attitude subjective que chacune d’elle adopte à l’égard de la dimension historique.

Les plus simples des sociétés qu’étudient les ethnologues ont, à la limite, une attitude négative envers l’histoire ; elles se voudraient telles que les dieux ou les premiers ancêtres les créèrent, sans y parvenir, bien entendu, car elles sont elles-aussi dans l’histoire, mais elles font ce qu’elles peuvent pour l’ignorer. Au contraire, une société comme la nôtre se veut résolument dans l’histoire, et elle cherche dans l’image qu’elle se fait de son mouvement (il serait plus exact de dire : dans les images conflictuelles que les groupes qui la composent se font de ce mouvement) à trouver le moteur de son changement.

\*

Voici donc une bonne base pour repasser au terrain dialectal. L’histoire ne serait pas en question, mais bien plutôt la perception du temps, qui n’a rien d’universel. Le temps est sans doute ce que l’on en

---

<sup>27</sup> Il est moins nécessaire sans doute de présenter Claude Lévi-Strauss (1908-2009). Le thème que nous abordons ici fait ordinairement partie de son corpus de réflexions, mais une mention particulière doit être faite de *La pensée sauvage* (Paris, Plon, 1962), la qualification anglophone existant aussi : « Savage Mind » (par exemple chez Jack Goody, *The domestication of the Savage Mind*, 1977). Comme dans l’extrait observé, C. L.-S. ne défend sans doute pas une bipolarisation tranchée entre pensée « sauvage » et pensée « occidentale » ou pensée « scientifique ». Plutôt deux attitudes mentales en principe présentes en toutes sociétés. Le problème est donc celui du dosage, qui, pour nous, rejoint en bonne partie la question de la francisation et de la résistance des langues régionales.

fait, sa conception est toujours relative, produite par un groupe, une société donnés. Passez au groupe suivant et vous aurez probablement un temps perçu différemment. Si l'on suit Malinowski, pour l'aborigène des Tobriand, ce qui s'est passé il y a quelques mois ne se distingue pas de ce qui s'est passé des siècles auparavant. Perception d'une société dans laquelle les siècles n'ont pas de sens, en tout cas pas le sens que nous leur donnons, où l'on se trouve dans la logique étiologique de l'ancêtre ou des dieux (voir l'entretien de C. Lévi-Strauss), précisément ce magasin hétéroclite qui ne prend sa justification, sa logique et son organisation qu'à partir du mythe et d'un réseau *autre* de correspondances.

Face à ce système, certaines sociétés, dont « la nôtre » (c'est-à-dire l'occidentale) se caractériseraient par leur aptitude à classer, organiser les événements, à les dérouler chronologiquement et à fonder globalement la dynamique sociale sur l'aptitude conjointe :

a) à la reconnaissance des cadres chronologiques/historiques dont nous héritons et qui nous permettent d'avancer en préparant notre futur,

b) au dépassement contrôlé de la coutume et de l'ancien, qui nous permet (nous) semble-t-il d'« aller de l'avant » en prétendant savoir d'où l'on vient.

Ce système de valeurs et de références se trouve aussi en lien direct avec le rationalisme et avec une séparation croissante entre mythes et terrains concrets. Voilà aussi pourquoi la pensée « occidentale », fondamentalement « urbaine », tend constamment à séparer les phénomènes religieux, métaphysiques, du projet social. Elle y voit des ordres différents alors qu'à des degrés divers, « primitifs » et campagnards sont prioritairement sensibles, voire convaincus, de la complémentarité, pour ne pas dire de l'inséparabilité du tout.

Il faudrait bien plus de temps pour espérer éclaircir les problèmes de ce genre, du moins peut-on proposer quelques directions. De la même manière que Claude Lévi-Strauss relativise ou rejette la polarisation *sociétés froides* vs. *sociétés chaudes*, les sociétés du magasin dialectal que nous évoquons échappent-elles à l'« occidentalité », au rationalisme, au temps historique etc. ?

On voit bien que non, car la pression du modèle français est extrêmement puissante -et de longue date-, par la gamme des relais dont elle dispose, surtout depuis le début du XX<sup>e</sup> siècle : École, administration, politique, télévision. Vu ainsi, le magasin dialectal



cède sans arrêt des éléments tout en s'arc-boutant sur d'autres. Cette aptitude à résister, à lâcher du lest, ne pourrait se comprendre sans la référence au macro-conflit évoqué par le biais de deux modèles culturels en interaction.

Quand la dialectologie s'est mise en place à la fin du XIX<sup>e</sup> siècle, nombre d'observateurs, à l'instar de Karl Jaberg (voir la citation liminaire), constataient la vitalité des patois et langues régionales. *La réalité* était bien effectivement autour d'eux. Mais en même temps, nombre d'entre eux voyaient cet ensemble rural comme promis à destruction rapide, d'autant qu'il se présentait comme dégénéré, disparate, bourré d'incohérences et d'emprunts, incapable de résister à la bienfaitrice ou « normale » pression française. Une réflexion commune soulignait en effet la puissance culturelle du français (modèle de langue claire, organisée, urbaine, libératrice des esprits etc.)<sup>28</sup> dont la pénétration avait commencé, mais sans doute pas assez vite et assez profond aux yeux de la plupart. L'École moderne devait donner le coup de grâce en élevant les esprits et la langue de tous. Les patois, les langues régionales allaient donc disparaître, plus justement s'effacer devant le français en quelques décennies.

Sans doute nos prédécesseurs seraient-ils surpris de constater le maintien, un siècle plus tard, des langues régionales. Certes avec beaucoup moins de locuteurs, nombre de trous géographiques et sociaux, mais toujours là. Il y a donc bien dans ce domaine, un mirage collectif qui m'a souvent intrigué ces dernières années, et que je compléterai maintenant. Comme j'ai eu l'occasion de l'écrire<sup>29</sup>, le recul résulterait « plutôt (d') un déplacement à vitesse variable, et de la langue nationale, et des langues régionales. On pense toujours la fin proche, mais elle ne se produit pas quand on l'attend. Un peu comme Achille qui n'atteindra jamais la tortue »<sup>30</sup>.

---

<sup>28</sup> Vulgate conceptuelle dont on trouvera trace directement chez Gilliéron lui-même : « une langue claire qui [satisfait] nos besoins modernes et dont la supériorité a été reconnue par les étrangers » (*Généalogies des mots qui désignent l'abeille*, 1918 : 258). Plus intéressante encore, sa qualification du terrain : « dans leur pauvreté génétique, dans leur misère actuelle, les parlers -nous n'en exceptons pas la langue illustrée par Mistral- ont recours à la grande pourvoyeuse qu'est la langue littéraire et leurs emprunts sont moins souvent peut-être les effets d'une servilité à son égard que ceux d'une nécessité causée par l'engrenage des lois phonétiques qui provoquent un désarroi lexical, auquel ils ne savent plus remédier autrement que par l'emprunt. » (*Pathologie et thérapeutique verbale*, I, 1915 : 13-14). Citations extraites de Sever Pop & Rodica Doina Pop, *Jules Gilliéron : vie, enseignement, élèves, œuvres, souvenirs*. Louvain, 1959.

<sup>29</sup>F. Manzano (2000), « De la dynamique du français, langue d'État et de pouvoir. Quelques rappels sur la diffusion entre le XV<sup>e</sup> et le XX<sup>e</sup> siècle », dans Dubois, Kasbarian, Queffélec, *L'expansion du français dans les Suds*. Publications de l'Université de Provence.

<sup>30</sup> Référence au paradoxe philosophique de Zénon d'Elée.

Beaucoup de choses démontrent ce lent déplacement, comme par exemple cette tendance pour la dialectologie de la fin du XX<sup>e</sup> siècle, en fait bien cohérente, à se déplacer des patois et langues régionales vers les français régionaux<sup>31</sup>. Tendance ici-même évoquée par Robert Chaudenson (: 51).

Pour autant, les cartes de distribution des langues et identités ethnolinguistiques n'ont pas fondamentalement bougé. Certes, en des centaines d'endroits, le patoisant n'existe plus, ou il est déclaré disparu. Il n'y en a vraiment plus. Certains l'ont vu, il existe peut-être, c'est surtout là-bas, ici non. Mon grand-père, ma grand-mère, untel qui est mort il y a quelques années. Rideau de fumée, recul en trompe-l'œil, comme si la vraie programmation de tout acteur du magasin était au fond de survivre, résister pour reproduire au mieux son identité globale en entrant le moins possible dans le modèle d'interprétation, de découpage qu'apportent avec eux dialectologues et sociolinguistes, main armée scientifique de ce modèle culturel auquel on résiste. Garder notre temps, avec nos mots, nos expressions, ne pas entrer dans ces placages théoriques (patois, dialecte, langue etc.) qui ne sont d'aucun réel secours pour vivre ici. On pourrait même dire l'inverse.

Toujours céder pour reproduire, mais céder le moins possible dans les coutumes, les perceptions, les faire-valoir, la langue, les mots.

\*

Le magasin dialectal est donc ce dépôt d'un autre temps auquel dialectologues et géolinguistes s'intéressent directement depuis plus d'un siècle. Quels que soient ses aspects amusants, surprenants, décalés, il cadre de manière optimale avec l'aménagement local, couple insécable de la langue et de la société ; et si le français local parvient à le prendre en charge, à le perpétuer au mieux, c'est autant de gagné, plus justement de non perdu<sup>32</sup>. En fin de compte, c'est donc l'écologie globale, l'osmose entre langage et aménagement local qui

---

<sup>31</sup> Il y aurait pour certains un parfum de trahison. Pour d'autres un simple déplacement des objectifs, une tentative d'adaptation à l'évolution même du terrain. Peut-être considérerait-on à ce moment l'entreprise atlantographique comme achevée et risquant de tourner en rond. Enfin et surtout, dans les années 70, le stock des vrais patoisants semblait fondre comme neige au soleil, la plus grande masse des matériaux dialectaux se retrouvant versée *de facto* dans les français locaux/régionaux. Pour beaucoup d'entre nous s'esquissait ainsi l'idée que ce passage des langues régionales aux français régionaux était l'un des derniers avatars du magasin dialectal (sans qu'ils utilisent ce terme, bien évidemment).

<sup>32</sup> Du point de vue de la plupart des acteurs du magasin, l'intérêt typologique et diachronique pour la connaissance de la Romania et de la langue d'oïl est obscur. C'est typiquement un point de vue de « spécialiste ».

détermine l'appropriation des langues régionales ou, à défaut, la récupération de cette adaptation par le français local. Si des mouvements trop rapides et violents se produisent dans l'espace global, la crise, voire à terme la mort du magasin, peuvent survenir<sup>33</sup>.

Mots et structures de ce magasin sont des segments du patrimoine que l'on se transmet, parfois spécifiques et incontournables<sup>34</sup>, parfois seulement là depuis toujours [argument en soi suffisant], fragments de la coutume et pour cette raison même légitimes et tenaces le plus longtemps possible.

\*

Parmi les traits du magasin dialectal, on peut mettre en avant celui de la *vénéralité* du gallo<sup>35</sup>. Déjà évoqué (note 20), il oppose à une vision synchronique péjorative du patois, l'idée que la forme relevée peut renvoyer aux origines de la langue. Le patois reçoit alors une forme de noblesse (en se rapprochant du français, que l'on appelle souvent « vieux français »), censée rattraper la violence sociolinguistique que la langue locale subit sous nos yeux. Bien entendu, l'argument est à double tranchant car il peut aussi bien servir que desservir dans la défense de la langue régionale. Notamment par le fait que cette conservation tardive elle-même peut aussi prouver l'immutabilité fondamentale de la langue régionale et du magasin, quand la « vraie » langue (entendez le français) est précisément capable de se réformer, de se débarrasser des vieilleries qui l'encombrent, bref d'être moderne et dynamique. Ce qui nous ramène encore à la dynamique anthropologique du magasin : ce qui se trouve là est précisément ce que la noblesse et la bourgeoisie (qui a pris sa place) ont laissé, concédé aux paysans, organisant elles-mêmes la vraie langue de la modernité et de l'avenir.

---

<sup>33</sup> Tel fut le cas de nombres de sociétés aborigènes étudiées au début du XX<sup>e</sup> siècle, celle des Tobriand par exemple, totalement disparue (et combien d'autres, pulvérisées). Les sociétés paysannes de France, à des degrés divers, semblent finalement mieux résister sous nos yeux, c'est-à-dire mieux protéger leur « magasin » par différentes stratégies de rupture et de coexistence avec le modèle central, occidental, français (ce que les sociétés dites primitives n'ont pu faire). Pour une approche plus générale des langues de France, on peut se référer à F. Manzano "Les langues régionales sont-elles égales dans le recul ?" (2005, disponible sur : [hal.archives-ouvertes.fr](http://hal.archives-ouvertes.fr)).

<sup>34</sup> C'est tout le volant technique des vocabulaires locaux, qui fait qu'on ne dit pas la même chose avec le patois ou avec le français. On reconnaît là une grande part des matériaux de l'ALF ou du NALF.

<sup>35</sup> Néologisme proposé dès "Sur le statut sociolinguistique du gallo" (1995), reformulé ensuite dans "Le gallo à la fin du XX<sup>e</sup> siècle" (1997), mais qui peut s'appliquer à bien d'autres régions. Voir à ce sujet les deux volumes des Cahiers de Sociolinguistique, dir. F. Manzano (Presses Universitaires de Rennes) : *Langues et parlers de l'Ouest* (1995-1996), *Vitalités des parlers de l'Ouest et du canada francophone à la fin du XX<sup>e</sup> siècle* (1997).

Cet archaïsme du gallo est réel et se révèle principalement à l'observation des structures phonétiques et du matériel lexical<sup>36</sup>. Le linguiste n'est pas seul à pouvoir effectuer ce constat. Les locuteurs le font aussi assez souvent, en faisant cadrer archaïsme et adaptation, déclarant que le mot est plus « *précis* » en gallo, ou « *ne peut se dire qu'en gallo* » etc<sup>37</sup>. De proche en proche se bâtit un discours sur la légitimité du gallo, d'autant plus délicat que ce discours se positionne par rapport au français langue réputée ici comme ailleurs langue de la *clarté*, de la *précision* etc<sup>38</sup>.

Pour aborder concrètement cette problématique du magasin dialectal, de la *vénéralité* et de la typologie en quelque sorte intemporelle du « magasin », je repartirai d'un extrait de texte déjà étudié, publié par Paul Sébillot<sup>39</sup> (*La goule es fées*), lequel a recueilli à son époque différents contes de Haute-Bretagne. Sans doute les traits du gallo ont-ils évolué en un siècle environ, mais relativement peu<sup>40</sup>. Nous négligerons donc ces variations éventuelles, d'autant qu'elles demanderaient d'abord à être vérifiées dans la zone de production du texte<sup>41</sup>. Soulignons enfin qu'il s'agit d'un texte écrit et par là passablement francisé. Un exemple parmi d'autres : il y a très peu de chances que l'on ait prononcé *femme* comme en français (ce que pourrait suggérer le texte écrit), car hier comme aujourd'hui on prononce [fôm]<sup>42</sup> ou [fòm] en gallo.

La version originale est dénommée Gallo-1. Dans la version Gallo-2, on met en italique les mots dont l'identification passe

---

<sup>36</sup> Ce n'est toutefois qu'un cas particulier du conservatisme plus général des langues et parlers de l'ouest.

<sup>37</sup> D'après différentes enquêtes menées dans le cadre des programmes poursuivis à l'université de Rennes. Tous les dialectologues ont entendu des propos de ce genre, toutes langues régionales confondues. En Haute-Bretagne, tout le monde sait, pense ou dit que *clencher* [une porte] (régional, local) n'est pas la même chose que *fermer* [une porte].

<sup>38</sup> Voir un peu plus haut la position de Gilliéron (note 28).

<sup>39</sup> Assurément l'un des plus célèbres folkloristes de la fin du XIX<sup>e</sup> siècle, originaire de Haute-Bretagne (1843-1918), contemporain de Lévy-Bruhl et Malinowski. Il fut d'ailleurs président de la Société d'anthropologie vers la fin de sa vie. Beaucoup de ses travaux sont consacrés à la Bretagne, mais il s'est intéressé aussi à d'autres régions françaises ou d'autres pays européens.

<sup>40</sup> C'est donc plutôt leur application, leur portée, leur masse, qui s'est modifiée. L'ensemble des locuteurs et des actes de parole en gallo n'étant pas le même en 1912 ou en 2012.

<sup>41</sup> En outre, comme nous l'avons souligné plus haut, l'une des caractéristiques du « magasin dialectal » est son immobilité relative par rapport au français, sa densité et la difficulté pour la langue véhiculaire de le pénétrer efficacement. La plupart des lexies que nous allons lister sont toujours connues en Haute-Bretagne, mais pas forcément utilisées, ou utilisées en concurrence avec le français mais dans un jeu du « chat et de la souris » pas très simple à aborder dans ce cadre.

<sup>42</sup> Avec un phonème articulé très à l'arrière.

principalement par un transcodage, phonétique ou morphologique mais sur lesquels nous ne nous attarderons pas ici<sup>43</sup>. On souligne d'autre part ceux qui peuvent appeler un commentaire-type de vénérabilité.

Gallo 1 :

Un saï qu'la mère Milie, qu'était saïge-femme de son état, était assise [.] dans l'coin d'son fouyer, o ouït queuqu'un qui cognait à l'hu de son hôte.

O débarrit la porte, et o vit entrer sez ielle eune veille femme qui li dit comme héla de venir do ielle tot cont' Saint-Leunaire, à cette fin d'assister une créature qu'était en ma d'éfant. Milie répondit qu'o voulait ben : o prit ses solées, mint sur son dos eune petite devantière, rapport à la fret ; o cutit son feu et sieuvit la veille qui cheminait devant ielle et marchait par les sentes comme s'il avait été jòù.

Gallo 2 :

Un saï qu'la mère Milie, qu'était saïge-femme de son état, était assise [.] dans l'coin d'son fouyer, o ouït queuqu'un qui cognait à l'hu de son hôte.

O débarrit la porte, et o vit entrer *sez ielle eune veille* femme qui li dit comme *héla* de venir do ielle tot cont' Saint-Leunaire, à cette fin d'assister une créature qu'était en ma d'éfant. Milie répondit qu'o voulait *ben* : o *print* ses *solées*, mint sur son dos *eune* petite devantière, rapport à la *fret* ; o cutit son feu et *sieuvit* la *veille* qui cheminait devant *ielle* et marchait par les sentes comme s'il avait été *jòù*.

Français :

Un soir qu'la mère Milie, qu'était sage-femme de son état, était assise [.] dans l'coin de son foyer, elle entendit quelqu'un qui cognait à la porte de sa maison.

Elle ouvrit la porte, et elle vit entrer chez elle une vieille femme qui lui dit comme çà de venir avec elle tout près de Saint-Lunaire<sup>44</sup>, à cette fin d'assister une créature qu'était en mal d'enfant. Milie

---

<sup>43</sup>Par exemple fr. /ü/ = gallo /œ/ (cf. *eune, Saint-Leunaire*), *o* = pronom sujet 3ème pers. fém., *ielle* = id. complément.

<sup>44</sup> Commune à l'ouest de Saint-Malo et Dinard, Ille-et-Vilaine.

répondit qu'elle voulait bien : elle prit ses chaussures, mit sur son dos un petit tablier (= fichu ?), rapport au froid ; elle couvrit son feu et suivit la vieille qui cheminait par les sentiers comme s'il avait fait jour.

En plus de la problématique ici abordée, ce court texte aura l'avantage de nous replacer dans le thème général de la qualification du gallo en tant que langue. On voit bien qu'une traduction globale n'est pas totalement nécessaire. Pourtant, on bute régulièrement sur des formes grammaticales, des mots, qu'il faut ou adapter ou déduire et traduire, parfois avec difficulté : *sai*, *fouyer*, *o* etc.

Le paradoxe est donc que l'intelligibilité du gallo ne se fait pas rapidement, bien que l'on ait conscience d'évoluer immédiatement dans un schéma macro-syntaxique qui semble le même qu'en français (en fait le schéma général de la langue d'oïl). Ceci donne automatiquement une priorité à la différenciation par le lexique (et la phonétique) :

a) soit que les mots ne soient différents du français que par leur structure phonétique : parfois assez éloignés comme *sai* pour *soir*, parfois plus facilement repérables (particulièrement en contexte) comme *fouyer* [fuyè] pour *foyer*, *ma* pour *mal* etc<sup>45</sup>.

b) soit que les mots soient inconnus du français, en totalité (mots propres) ou en partie (emplois divergents, partiellement continués en français, voir ci-après).

Ce faisceau de variation lexicale et sémantique peut temporairement faire barrage à la compréhension du francophone moyen. Il est pourtant assez vite surmonté par un système de transcodage « par équivalence »<sup>46</sup> qui s'effectue d'autant plus facilement que les patrons morpho-syntaxiques fondamentaux sont voisins de ceux du français. Mais du fait de la grande proximité entre morpho-syntaxe du gallo et morpho-syntaxe française, la qualification du gallo en tant que *langue* s'avère impossible. Dans la tradition linguistique, diachronique et

---

<sup>45</sup>Ces variations s'expliquent par la typologie phonétique dialectale du gallo au sein de la langue d'oïl, typologie conservatrice, archaïque, par rapport à celle du français contemporain. Sur ces points, voir notamment J.-P. Chauveau, *Evolutions phonétiques en gallo*, 1989.

<sup>46</sup>On apprend très vite à convertir les traits et à les additionner. Ainsi convertit-on [ë] en /é/ (type [abominë] = *abominer* « détester »), les [ay] issus de /é/ fermé roman en [wa] (type *mai* = fr. *moi*). Même chose (mais en additionnant le fait que le -r final n'est pas prononcé en gallo) pour *sai* [say] qui, par l'intermédiaire d'un \*saïr, renvoie au fr. *soir* [swar]. C'est cette chaîne d'opérations que je qualifie de *transcodage par équivalence*.

dialectologique, le gallo ne varie (parfois considérablement) que par le biais de matériaux au fond *périphériques* : lexicale, sémantique, structures phonétiques. On considère que cela ne suffit pas à faire une langue.

Voyons à présent quelques uns de ces cas de variation pour les expliciter et montrer comment ils peuvent être versés au dossier de la défense du gallo par le biais de la *vénéralité* et des valeurs voisines de la *légitimité* et de la *cohérence* du magasin dialectal.

*l.1-o ouit*, passé-simple de *oir* « entendre », toujours fonctionnel à la fin du XX<sup>e</sup> siècle en gallo<sup>47</sup>. Cas typique d'enracinement roman, puisque le gallo rejoint par ce biais les grandes langues méridionales conservatrices comme l'espagnol, le catalan et l'occitan notamment (esp. cat. *oir* « entendre, percevoir », oc. *ausir*). L'étymon est le latin AUDIRE<sup>48</sup> auquel s'approvisionne toujours le français, mais par le biais de ses formations savantes uniquement : *auditeur*, *audition*, *audible* etc. On peut observer que le français a connu ce verbe comme usuel<sup>49</sup>, mais dans les périodes les plus anciennes de la langue : ancien et moyen français (*oir* puis *ouïr*). Remplacé par *entendre* (< lat. INTENDERE), le verbe originel est pratiquement perdu aujourd'hui puisqu'il n'apparaît que dans des locutions figées, des archaïsmes, des usages techniques (type : *par ouï dire*, *j'ai ouï dire*, *ouïr les témoins*). Mais alors que ce type est chassé du système normatif du français classique, il peut subsister chez les campagnards d'oïl, dans l'ouest en particulier, et s'intégrer de facto à un modèle d'opposition binaire *archaïque*, *paysan* (gallo) vs. *actuel*, *urbain* (français). Ce qui était normal en français jusque là devient donc rapidement vieux ou même ridicule, dans la bouche des provinciaux.

*l.2-l'hu*, même souche que le français *huis* (< lat. ŪSTIUM « porte »). Schéma évolutif à peu près comparable au précédent. Déjà réputé vieilli en français au XVIII<sup>e</sup> siècle. N'existe plus

---

<sup>47</sup>Le recensement des formes contemporaines du gallo de Haute-Bretagne a été entrepris et publié, notamment depuis deux décennies. Voir par exemple Bertrand Obrée, dir. & Bretagne Gallèse, *Motier de galo/Dictionnaire de gallo*, Bretagne Gallèse, 1995. Régulièrement paraissent de nouveaux dictionnaires régionaux.

<sup>48</sup> Le trait souscrit marque la voyelle tonique.

<sup>49</sup>Au sens de verbe banal, c'est-à-dire conjugué. En ce sens, *ouïr* n'est plus aujourd'hui qu'une relique en français. Les dictionnaires de la fin du XVIII<sup>e</sup> le disaient déjà vieilli et déféctif.

aujourd'hui que dans la série relique très limitée *huis clos*, *huissier*, *huisserie*<sup>50</sup>.

*l.3-son hôté*. En ancien français et jusqu'au XVII<sup>e</sup> siècle, l'*hôtel* (< lat. HOSPITALE) désigne bien la demeure particulière, la maison, comme aujourd'hui en gallo, mais aussi en occitan (*ostal*, *ostau*). Le français a produit des réductions successives vers « *hôtel (particulier)* »<sup>51</sup> (XVIII<sup>e</sup> siècle), et « *hôtellerie* », XIX<sup>e</sup>. Voici donc une lexie qui paraît à nouveau très archaïque, dialectale et marginale du point de vue français.

*l.4-do*. Issu de D(E) + APUD. On trouve également la version monovocalique originelle *o* (< lat. APUD), normale et usuelle en gallo.

Cette même conjonction *o* a existé en ancien français. Le français l'a remplacée par la forme construite APUD HOC qui conduit à *avec*<sup>52</sup>. Notons encore que le lat. APUD est toujours directement représenté dans la Romania méridionale (comme en gallo donc), par exemple en occitan et catalan *amb* (diverses variantes : *ob*, *ab*, *am*, *an*, etc.)<sup>53</sup>.

*l.5-créature*. Si cette formation évoque le fr. *créature* (< lat. CREATURA), le signifié n'est pas le même. Comme souvent, le français moderne oscille entre l'abstraction (*créature* = être vivant, être humain, type *les créatures de Dieu*)<sup>54</sup> et les péjorations que l'on

---

<sup>50</sup>Commentaire-type des dictionnaires à la fin du XVIII<sup>e</sup> et au XIX<sup>e</sup> : « Vieux mot qui signifie, porte, et qui n'est plus guère usité qu'au palais, dans la loc., À huis clos, À portes fermées et sans que le public soit admis » (Lorain, *Abrégé du Dictionnaire de l'Académie Française*. Paris, Desrez-Didot, 1838).

<sup>51</sup>C'est ce que dit Furetière, à propos de ce mot : « En son ancienne signification signifie logis, maison où on demeure. Se dit plus communément du logis des princes et grands seigneurs ». Et au XVIII<sup>e</sup> : « Grande maison, demeure somptueuse d'une personne de qualité, d'un personnage éminent, d'un riche particulier » (Lorain, *Abrégé...*).

<sup>52</sup>La forme construite apparaît au Moyen Âge déjà. On la rencontre alors sous les formes *avoc*, *avuec*, *avoec*.

<sup>53</sup>Dernier à l'avoir semble-t-il défendue en français, Ronsard. Voir chez Jean Lemaire de Belges, début XVI<sup>e</sup> : *Et faux espoir que j'avoye d'user Mes jours o toy* (cité par Greimas et Keane, *Dictionnaire du Moyen-Français*. Larousse, 1992). Avec cet exemple et d'autres, on peut voir que c'est en fait le français qui s'est mis à part, non seulement de sa sphère originelle (langue d'oïl), mais plus généralement de l'ensemble de la Romania.

<sup>54</sup>Au XVII<sup>e</sup> on identifie aussi le signifié dérivé de « protégé, obligé » : *Faites que je sois votre créature, et me donnez quelque charge* (Sorel, cité par Dubois-Lagane-Lerond, *Dictionnaire du français classique*. Larousse, 1971).



connaît (*une créature* = celui qui n'existe que par celui qui l'a créé, ou *une créature* = femme de mauvaise vie, depuis le XVII<sup>e</sup> siècle). Ce mot, en réalité, a bien disparu du français, du moins dans cet emploi. En revanche, comme beaucoup de ceux que nous relevons ici il est très normalement passé outre-atlantique. Un exemple, Pascal Poirier, dans son Glossaire acadien<sup>55</sup>, mentionne *CRÉATURE, CRIATURE, CREYATURE. Personne du sexe féminin : Il y avait beaucoup de criatures à l'assemblée ; il n'y avait là que des creyatures...*<sup>56</sup>

*l.6-devantière* est une désignation du « tablier ». Une désignation très motivée (= *devant* < DE ABANTE + suffixe), dont on retrouve à nouveau le principe plus au sud dans des mots comme l'occitan et le catalan *davantal* « tablier ». L'espagnol adopte le même principe de dénomination avec *delantal* (cf. esp. *delante* « devant »). On trouve ceci chez Furetière : *Devanteau. Vieux mot qui signifiait autrefois tablier, et qui n'est plus en usage que parmi le petit peuple.*

*l.8-o cutit*<sup>57</sup>, comme *l.4-o débarrit*, illustre la subsistance du paradigme unique d'un parfait en *-i* en gallo, y compris pour les verbes dits du 1<sup>er</sup> groupe. Ce temps du passé reste fréquent en gallo à la fin du XX<sup>e</sup> siècle. On en arrive à nouveau à une sorte de « paradoxe » typique du *magasin dialectal* : les paysans de Haute-Bretagne font sur ce point comme la meilleure société le faisait aux XVII<sup>e</sup> et XVIII<sup>e</sup> siècles, en utilisant un registre qui depuis cette époque s'est spécialisé en français comme temps littéraire et mondain<sup>58</sup>. Vieilli et stigmatisé en français contemporain, le passé-

---

<sup>55</sup> Pascal Poirier, *Le Glossaire acadien* (début du XX<sup>e</sup> siècle, réédité par Pierre M. Gérin (Centre d'Etudes Acadiennes). Document non daté, années 1990. Cet ouvrage a été mis en ligne par l'université de Moncton, c'est une mine de renseignements : <http://www2.umoncton.ca/cfdocs/cea/livres/doc.cfm?livre=glossaire>.

<sup>56</sup> P. Poirier cite un extrait du *Médecin malgré lui* : *Et v'la que la pauvre creyature est devenu jaune comme un coing*. Ou [velà que la pauvre creiature en est devenue jaune comme un coing].

<sup>57</sup> Le verbe *cuter* « cacher », si l'on ose dire, est « encore plus » du magasin dialectal. D'après Antoine Thomas (1913, CR des séances de l'Académie des Belles Lettres), ce verbe serait issu d'une base celtique (CUD-, var. CUZ-) représentée dans plusieurs rameaux, dont le gaulois et le breton plus tard. Recensé en ancien français sous les formes *cuter, cutter, se cuter, cute* « cache, cachette » (postulant un étymon secondaire latinisé \*CUDITARE), c'est un type de distribution préférentielle occidentale. Une fois encore, le français l'a perdu ou éliminé.

<sup>58</sup>Dans une optique voisine, on notera que des formes comme *o print, mint* sont tout à fait proches des occurrences normales en français du XVI<sup>e</sup> siècle.

simple, très coté jusqu'en français classique, devient en même temps suranné, voire comique, chez les provinciaux<sup>59</sup>.

*l.9-sente* (< lat. SEMITA) a existé comme mot simple en ancien français pour être progressivement remplacé par le dérivé *sentier* (< \*SEMTARIU). Notons qu'une langue comme le catalan dispose encore conjointement du simple (*senda*) et de dérivés (*sender*, *sendera*). Cette forme *sente* était encore recensée par le *Dictionnaire de l'Académie* de 1778, à côté de *sentier*.<sup>60</sup>

#### 4. Pour conclure

Ce qui précède n'a effectivement rien de propre au domaine gallo. Des observations parallèles à celles-ci sont chose normale dans différentes régions de langue d'oïl, d'oc, ailleurs encore, sans compter l'Amérique du Nord et les créoles où une grande partie de ces matériaux ont été mécaniquement transplantés par les émigrés (de l'ouest notamment). Chemin faisant, on remarquera que de telles spécificités sont à chaque fois vues comme particularisantes, définitoires de la langue et de l'identité locales, non comme les résurgences d'un fonds plus ancien, plus large, partagé avec d'autres. Après les cas examinés précédemment, nous en ajouterons quelques autres avant de conclure.

Car en réalité la liste serait longue et concorderait bien souvent avec les corpus atlantographiques, les glossaires, où ont été patiemment entreposées une bonne partie de ces données ; mais celui qui utilise *asteure* « en ce moment, maintenant » (var. *astoure*, *asture*, *achteure* etc.), *anuit* « aujourd'hui » (var. *aneut* etc.)<sup>61</sup>, *bailler* « donner » (oc.

---

<sup>59</sup>Chez le boucher, à Saint-Brice en Coglès (35), une cliente de la campagne entre, rencontre une connaissance. Les deux évoquent une troisième personne : « Hier, je fus chez lui... ».

<sup>60</sup> Nos dictionnaires le comportent encore, mais comme mot exclusivement littéraire.

<sup>61</sup> De : AD NOCTEM, « cette nuit, la nuit d'aujourd'hui », qu'on retrouve toujours en catalan par exemple : *anit* « cette nuit » (standard et régions). Un glissement s'est produit en différentes régions, qui sélectionne comme on le voit *anuit* = « aujourd'hui » en perdant une part du sémantisme originel. Pour ce qui concerne les variantes graphiques, aux variations dues aux faisceaux d'évolutions diachroniques locales se sont surajoutées les complications du militantisme orthographique récent. Par exemple, pour le domaine gallo, suivant l'école ou le(s) militant(s) on aura ici : *aneit*, *anoet*, *anet*, *anë*, *ane*. Ce sont là des complications supplémentaires pour aujourd'hui et pour l'avenir, mais là n'est pas notre propos.

*balhar*, var. *bailar*)<sup>62</sup>, *linceu(l)* « drap »<sup>63</sup> etc. perçoit avant tout ces différentes lexies courantes, suivant les lieux, comme ordinairement gallèses, picardes, acadiennes, occitanes, roussillonnaises etc. Et ce n'est nullement une question de mauvaise foi. En Haute-Bretagne par exemple le mot *asteure*, est un de ces mots rituels et emblématiques qu'on avance le plus souvent comme caractéristique du gallo ou patois (suivant le degré de conscience). Mais il n'a rien de spécifiquement gallo (sauf sans doute ses prononciations particulières ici ou là). Aux variantes phonétiques près, d'autres le verront aussi bien et avec la même certitude comme typique du picard, de la Louisiane ou de plusieurs régions de l'Amérique francophone<sup>64</sup>. Pour chacun, mais parallèlement en somme, c'est bien un élément du tissu identitaire qui fait qu'on est vraiment d'ici.

Ces mots sont donc de véritables mots balises, des repères anthropologiques et identitaires faisant eux-mêmes partie d'une chaîne, d'un tissu écolinguistique qui cimente l'identité locale. Être de Haute-Bretagne, c'est ponctuer son discours d'*asteure*, d'*apouyettes*, c'est *abominer* son prochain, *avoir une poganée* et *battre de la goule* comme une telle qu'on connaît bien au bourg, sans oublier les *bérouettes* et les *guernettes* que les non Bretons reconnaîtront facilement. Mais c'est aussi quitter le *village* pour aller au *bourg*, aux courses, à la messe ou au café-épicerie, acheter et *mangeu* une galette-saucisse au *marcheu*.<sup>65</sup> Tout cela peut se dire dans une variété très proche formellement du gallo « originel », comme en français parlé du lieu. Mais au bout du compte, c'est toujours la présence de l'ensemble (sorti du magasin dialectal) qui fait la reconnaissance et entretient le magasin.

J'insiste un peu car cette approche ouvre une perspective intéressante sur les terrains dialectaux. Elle nous permet de

---

<sup>62</sup> Par exemple, oc. *Balha-me una cigarreta* (fr. *Donne-moi une cigarette*), gallo *Bay ma ça !* (fr. *Donne-moi ça !*). Du lat. BAJULARE.

<sup>63</sup> Issu du lat. LINTEOLUM (donc tissage/drap de lin), devenu nom général du drap en français et ailleurs (ancien français *lincel*, *linçuel*), avant d'être sémantiquement réduit puis évacué en français (= drap mortuaire).

<sup>64</sup> Parmi bien d'autres exemples, un groupe de musique cajun se dénomme actuellement *Bon Temps Asteur*. Des associations locales, groupes divers, captent régulièrement cette lexie en France, au Québec etc. Voir ici même F. Carton : 106.

<sup>65</sup> On ne peut développer davantage ici, mais la cohérence du magasin dépasse évidemment la langue, impliquant une chaîne très complexe qui va du paysage physique aux vêtements, aux postures etc. Cet ensemble d'articulations est très important car il permet de mieux comprendre pourquoi les sociétés rurales de France résistent ou au contraire décrochent.

comprendre un certain nombre de choses à partir de principes simples mais étayés.

Tout se passe donc comme si des alluvions d'époques différentes coexistaient au sein des terrains régionaux, dans une dimension « autre » que celle du français central : le français, lui, arrange, sépare, organise les alluvions et les étalonne diachroniquement, produisant et mettant en scène sa propre mise à part. Il n'y a pas que les linguistes qui séparent ancien, moyen, moderne, classique, contemporain, haut, bas etc. Les locuteurs usuellement francophones sont généralement très habiles à situer leurs pratiques et celles des autres, à les référer, les hiérarchiser, les classer, les ordonner. C'est une seconde nature sur laquelle il n'est pas trop nécessaire d'insister.

Le terrain dialectal, quant à lui, n'a pas de diachronie, c'est d'ailleurs pourquoi il essaye de s'en donner. Par exemple, quand l'informateur vous dit que *ça c'est un mot des vieux, des anciens, nous on ne le dit plus*. Ou quand un autre vous dit une *x<sup>ième</sup>* fois que lui-même ou son grand-père était puni pour avoir parlé patois à l'école (il se place donc dans le « temps » de la société française ou principe urbain). Dans un cas comme dans l'autre, on remet plus justement un peu de relativité, de chronologie, précisément dans un monde où en principe la chronologie est tout à fait secondaire. On a toujours dit *hucher* « crier, appeler », *affutiaux* ou *hannes* « vêtements », peu importe finalement de savoir d'où viennent ces mots et depuis quand ils sont là. C'est bien une société que l'on peut dire « froide » et travaillée en surface par la « chaude », laquelle tente constamment dans ce contact, mais sans y parvenir intégralement, de lui imprimer ses rouages mentaux et culturels. Et c'est bien finalement parce que les deux ensembles (culturels, temporels, identitaires) ne sont jamais parfaitement miscibles que l'abandon du magasin comme la mort des idiomes locaux reculent sans cesse, ainsi qu'on l'a remarqué plus haut.

\*

Pour les locuteurs des différents pôles « régionaux » (ou langues régionales), ces alluvions ont été parfaitement intégrées et parviennent donc à constituer des systèmes périphériques, mais cohérents, fonctionnels, de la langue d'oïl notamment<sup>66</sup>. Pour qui l'observe avec

---

<sup>66</sup> C'est dans les zones de langue d'oïl que la dynamique exposée est sans doute la plus criante. C'est en somme le premier cercle, là se trouve le magasin le plus franchement décalé, distancié,

attention, cet ensemble est bien un « magasin dialectal », la mémoire romane du français archaïque et de la langue d'oïl. En réalité, cela ne s'analyse pas, cela se dit, se vit, se transmet en dehors des contraintes de rationalité ou de bienséance qui caractérisent le pôle français. Et, surtout des contraintes de tri, car cela se retouche le moins possible.

En revanche, pour les locuteurs du français, depuis l'époque classique, le sens de l'actualité, de l'appropriation, de la nouveauté, sont constitutifs de la personnalité collective. L'insécurité grammaticale/sociolinguistique, l'inquiétude normative, traversent nos propos et charpentent nos attitudes sociales. La machine à éliminer, à épurer, tourne à plein depuis trois bons siècles, et si la recherche de modernité n'est pas le premier souci des campagnards de France<sup>67</sup>, elle est au contraire déterminante de ce côté et tire toujours le français vers plus d'urbanité. Nous sommes en cela, collectivement et quelque effort qu'on fasse, les dignes héritiers de Vaugelas.

---

péjoré. Mais on a vu que, plus globalement, le raisonnement concerne de proche en proche l'ensemble de la Romania, au premier chef les langues romanes de France : oc, catalan, francoprovençal etc. Autant de cercles supplémentaires fondant la niche éco- et socio-linguistique globale du français.

<sup>67</sup> L'évitement de la modernité est même une nécessité structurelle du « magasin ».

*Francis Manzano*  
**Postface**

Les 12 et 13 mars 2009, une vingtaine de dialectologues se retrouvaient à Lyon, à l'initiative du Centre d'Études Linguistiques (Équipe d'accueil 1663). Presque tous avaient travaillé directement à l'une de pièces du *Nouvel Atlas Linguistique de la France*, d'abord dans le cadre des opérations soutenues financièrement par le CNRS, puis au sein de plusieurs groupements ultérieurs, RCP, GDR, GRECO<sup>1</sup>. Plusieurs des participants travaillaient en outre à différents projets en cours, souvent complémentaires (THESOC, ALiR etc.).

La présente publication reprend les réflexions et travaux présentés durant ces deux journées. Chaque texte est de la responsabilité de son auteur, aussi existe-t-il quelques variations dans les transcriptions, suivant les alphabets utilisés par les participants : API ou système des romanistes.

Le choix d'engager une réflexion autour de la thématique « Mémoires du terrain : enquêtes, matériaux, traitement des données » visait assez clairement à ce que la réunion soit un temps de réflexion collective sur le bilan de la dialectologie en un siècle environ, sur la manière dont les enquêtes mais aussi les traitements de données avaient évolué à travers ce siècle, sur la valeur même du bilan et des données remontées du terrain, sur l'avenir de la géolinguistique et de la dialectologie, pour reprendre les termes qui font à peu près consensus<sup>2</sup>.

Quelle vision les dialectologues ont-ils de leur propre famille scientifique ? de leurs origines, de leur fonction sociale, de leur place relative dans les sciences du langage et de la société, de leur rôle dans les sociétés globales du début du XXI<sup>e</sup> siècle ?

---

<sup>1</sup> Cet historique des faits est en partie rappelé ici par Robert Chaudenson ou Jean Le Dù, par d'autres aussi plus partiellement et ponctuellement au cours des différentes communications.

<sup>2</sup> Les opinions des intervenants n'ont nullement été modifiées, modalisées par l'éditeur. Chacun reste ici responsable de ses propos. J'observerai simplement que les chocs, voire les contradictions sont souvent nécessaires pour qu'une évolution heuristique du champ soit vraiment possible.

Va-t-on vers la fin de la dialectologie ? Ou cette fin s'est-elle en vérité déjà produite ? Ou bien cette fin s'éloigne-t-elle au fur et à mesure qu'on croît l'atteindre ?

Y a-t-il encore un avenir ? Mais dans ce cas, les dialectologues seront-ils prêts à le construire, le configurer, à démontrer l'intérêt des recherches qu'ils poursuivent ? Autant de questions et de débats soulevés en mars 2009 et à d'autres occasions, sur lesquels je reviendrai rapidement.

Certains de nos lecteurs jugeront peut-être que les lignes qui suivent sont un peu réductrices. J'en suis conscient, mais un espace de conclusion doit rester bref. Il faut aussi et surtout forcer quelque peu les raisonnements et la communauté des géolinguistes, sans quoi c'est la communauté universitaire et sociale qui le fera. Probablement les conséquences seront-elles alors sans doute plus violentes, en termes de pérennité, de recrutements, de crédits, de reconnaissance sociétale tout simplement.

Ces remarques, non exhaustives, seront faites en trois temps.

## **1. Origines et missions de la dialectologie**

Ce n'est un mystère pour personne. En 2009 comme déjà depuis plusieurs années, la dialectologie, les dialectologues, s'interrogeaient sur leur place, leur rôle. En réalité, la dialectologie s'interroge depuis ses débuts. Derrière l'euphorie des premières entreprises, planait déjà la menace. À quoi servaient ces dialectologues suisses, français, européens, qui parcouraient les campagnes de la fin du XIX<sup>e</sup> à la seconde guerre mondiale ? Étaient-ils, comme on le voyait souvent, de simples collecteurs, vérificateurs, à travers la métaphore floristique qui revient souvent, par exemple sous la plume de Jaberg<sup>3</sup> (à propos de l'avis de Gaston Paris sur les travaux de Gilliéron) :

Lui, qui dominait admirablement toutes les méthodes de la philologie romane, reconnut du premier coup d'œil la fécondité de l'idée nouvelle qui consistait à présenter sur des cartes géographiques, sortes de tableaux synoptiques, toute la flore dialectale d'un pays.

Au fond, dès lors, un pli est pris. Toute entreprise de dialectologie implique un dédoublement relativement clair. D'un côté, celui qui

---

<sup>3</sup> Karl Jaberg. *Aspects géographiques du langage*. Paris, Droz (1936 : 12).

conçoit, de l'autre celui qui enquête. Il peut s'agir de la même personne, mais le plus souvent la séparation est physique, comme on le voit dès l'ALF et comme le confirmera plus d'une fois le NALF. D'où cette idée réductrice, par la suite amplifiée, que les dialectologues font les enquêtes, savent les faire, mais peut-être ne savent faire que cela. C'est probablement injuste, mais l'évolution d'un paysage universitaire et scientifique se fait malheureusement avec de telles réductions<sup>4</sup>.

Il y aurait donc quelque part des universitaires ayant un regard plus ample et savant. Des Paris, des Gilliéron, des Dauzat, puis des Gardette, des Séguy etc. qui ont une vue globale de ce qu'il faut/faudrait faire et de l'interprétation des données au sein d'un plan plus général et plus ambitieux<sup>5</sup>.

Une autre réflexion taraude également le champ. Si herbier il y a, recueil floristique comme le suggérait Jaberg, jusqu'à quand fera-t-on des herbiers ? De plus en plus précis, de plus en plus détaillés, diversifiés, numérisés ... À quoi bon pourrait-on dire si le programme d'ensemble est de moins en moins clair ?

\*

Souvenons-nous. La mission initiale était de fournir à la « grande » linguistique comparative des langues romanes des matériaux réels issus du terrain et des langues régionales orales. Ce que l'on recherchait était au fond la photographie la plus exacte possible de l'ensemble des ramifications de la famille romane, réputée exemplaire tant par la connaissance diachronique qu'on en avait que par sa vitalité sur le terrain à travers notamment ces trois grands pays qu'étaient la France, l'Espagne et l'Italie. L'étude des patois allait compléter en un siècle la connaissance précise qu'on avait déjà des grandes langues de culture et langues nationales de la Romania<sup>6</sup>.

On peut donc considérer que la mission était dans les grandes lignes réalisée vers la fin du NALF (pour la France du moins), par exemple

---

<sup>4</sup> Voir plus haut, Chaudenson : 50,51.

<sup>5</sup> Ce clivage n'est certainement pas exclusif de la dialectologie. Plusieurs sciences humaines de terrain le partagent (le rapport au terrain est donc bien déterminant), par exemple l'archéologie qui a, parallèlement, ses concepteurs/analystes, ses fouilleurs, parfois réunis, souvent séparés. La distribution du pouvoir chez les dialectologues, les archéologues, les géographes etc. entre en ligne de compte, bien évidemment.

<sup>6</sup> Ces pays sont en effet exemplaires par la présence de grandes langues de culture, historiquement bien établies, nationales et/ou internationales, précisément des langues dont la connaissance allait être complétée par l'étude des « patois » romans. Dans ce lot, le Portugal joua un rôle moindre, car c'est connu, le Portugal est probablement le seul pays à peu près monolingue de la Romania.



et de mon point de vue, avec l'atlas remarquable des Pyrénées-Orientales (Henri Guiter), qui donnait une vision très épurée, quasi phonologique, de la variation intégrale à travers toutes les communes d'un département, plus deux bandes également exhaustives côté languedocien et côté espagnol.

## 2. De nouvelles directions ?

Malgré tout, on peut voir que la dialectologie est au fond repartie depuis dans plusieurs directions. J'en indiquerai trois essentielles, mais sans doute en trouverait-on davantage dans le détail.

- a) Celle des mesures, des quantifications, direction issue des réflexions de Jean Séguy et Henri Guiter dans les années 70<sup>7</sup>. Ces réflexions, auxquelles se réfère ici en partie l'intervention de Jean-Louis Fossat, sont toutefois restées à la marge et le terme générique produit durant la décennie 1970-1980 (*dialectométrie*) ne semble pas avoir rayonné dans l'ensemble de la dialectologie romane. Surtout, il ne semble pas avoir relancé cette dernière de manière décisive<sup>8</sup>. De telles directions se trouvent pourtant dans la prolongation de l'un des axes fondamentaux de la dialectologie, celui de l'étude des taxinomies, des classifications typologiques et des frontières de langues (rôle scientifique ancien et primordial plus haut rappelé par Jean-Philippe Dalbera).
- b) Celle de l'extension romane. C'est le nouvel atlas roman, qui donne en fait l'impression que le poisson se mord la queue, comme si la dialectologie, désorientée à la fin du XX<sup>e</sup> siècle, était repartie, avec armes et bagages, vers le macro- après avoir épuisé ses loupes sur le micro-. Ceci n'est pas un jugement de valeur, car les méthodes ont changé assez profondément. On procède pourtant comme si le problème

---

<sup>7</sup> Je pense notamment à leurs contributions respectives dans la *Revue de Linguistique Romane*, tome 35, 1971 (J. S. "La relation entre la distance spatiale et la distance lexicale", H. G. "Fréquences verbales dans les langues romanes") ; ou dans l'ouvrage de référence par ailleurs publié au CNRS en 1973 : *Les dialectes romans de France à la lumière des atlas régionaux* (contributions de Jean Séguy, Xavier Ravier, Henri Guiter, et de plusieurs directeurs de sections du NALF). Henri Guiter est en outre l'auteur de plusieurs études métriques remarquables (mais de faible diffusion) sur la « frontière » occitano-catalane.

<sup>8</sup> Ce qui ne signifie pas pour autant que la piste soit erronée. On a vu ici-même que la question des distances est reposée par le projet AMPER (Lai & Contini).

était un problème de focale ou d'exhaustivité, alors que l'exhaustivité est à l'évidence impossible dans une telle matière.

- c) Celle de la sociolinguistique. Que la sociolinguistique prolonge la dialectologie et replace l'ensemble au sein d'une linguistique « sociale » globale est à peu près évident pour tous. Depuis les premiers jusqu'à nous (différents textes le montrent encore dans ce volume), la famille des dialectologues travaille directement sur des corrélations entre langage et société, langage et objets, langage et espace physique etc. Le legs est profond et William Labov le soulignait il y a plusieurs années (*Sociolinguistique*, 1972-1976), estimant en somme que le programme sociolinguistique était en place 60 ans avant lui, notamment chez Louis Gauchat dans ses études sur les patois romans de Suisse.

Mais cette reconnaissance de la dialectologie par la sociolinguistique a aussi clairement montré les limites naturelles de la première comme de la seconde. Par exemple, dès lors qu'on passe d'un informateur principal à plusieurs (système labovien, inspiré de Gauchat et d'autres), comment ne pas voir qu'une étude sociolinguistique plus fine supposerait une multiplication des informateurs, non seulement en âges, mais aussi en fonctions, métiers etc. à la ville comme à la campagne. À nouveau pointe la vis sans fin, de proche en proche n'irait-on pas vers la société entière, l'espace géographique entier ?

Il y aurait donc quelque part et quoi qu'on fasse, nécessité de réduire avant de pouvoir modéliser. Cela pèse constamment sur nos disciplines. D'où aussi un sentiment inépuisable de culpabilité par rapport au réel pour qui fait les enquêtes et qui les exploite. Cette belle carte fixée dans l'atlas ou maintenant mise en ligne sur un serveur, n'est-elle pas seulement l'expression de ce qu'une équipe, des chercheurs, ont cru comprendre/entendre du réel ? Chacun sait à ce sujet les choix qu'il a dû faire, les solutions qu'il a dû adopter face aux problèmes. On le dit généralement en avertissement de l'atlas, de l'article ou du livre, mais la frustration reste de règle<sup>9</sup>. En réalité, cette frustration doit même s'imposer si l'on veut rebondir.

---

<sup>9</sup> L'opération achevée, apparaît souvent plus clairement ce que l'on aurait dû faire, ce qu'il faudra reprendre, préciser, modifier, dépasser ultérieurement. On pourra certes dire que toute science

Une remarque supplémentaire peut-être faite quant à cette réunion naturelle entre sociolinguistique et dialectologie à la fin du XX<sup>e</sup> siècle. Quand la sociolinguistique (*lato sensu*) s'est répandue durant la décennie 1970-1980 et par la suite, alors que les derniers atlas étaient publiés, un mouvement parut clairement. L'université, la recherche, produisirent de nouvelles spécialisations, régionales, thématiques. Voilà qu'à nouveau, on envoyait sur le terrain des gens issus de tel ou tel village, de telle ou telle vallée, dont la famille était spécialisée dans telle ou telle activité etc. On refaisait donc, certes sous une autre forme et probablement avec d'autres intentions, à peu près ce que le champ avait produit 100 ans plus tôt quand les professeurs de linguistique française ou romane envoyaient sur le terrain étudiants ou enseignants de la région à étudier en vue de constituer de nouveaux glossaires patoisants.

Nos disciplines ne seraient-elles pas tout simplement piégées par ce terrain qui change, condamnées à se reproduire comme un Phénix qui, à la longue, ennuerait passablement les décideurs malgré ses belles couleurs et ses résurrections successives ?

### **3. Trois paradoxes pour conclure**

Ces différentes caractéristiques ou travers de la dialectologie, de la géographie linguistique, nous les connaissons en réalité parfaitement. Plusieurs communications lors de ce colloque les ont déplorées. Cependant, elles sont souvent passées sous silence ou minimisées. Une discrétion d'ailleurs normale, sans doute honorable, même s'il faut la secouer. Ce que nous faisons ou ce que nous faisons faire à d'autres, n'est rien d'autre qu'un travail sur la mémoire vive des sociétés qui se transforment. On aurait tort de se sentir coupable de cet acte, éminemment social, patrimonial, comme le souligne également Jean-Philippe Dalbera. Un acte qui, en outre, prend toute sa valeur quand les années ont passé, quand les pages de l'histoire se sont tournées. C'est toujours ce recul qui met le mieux en lumière le

---

procède ainsi par tâtonnements ou ajustements successifs. Mais c'est très grave pour une approche intimement liée au terrain, c'est-à-dire à ces informateurs qu'on ne retrouvera plus, ou qu'on retrouvera transformés dans des sociétés locales qui auront elles-mêmes changé. Imaginons une seconde que nous puissions aujourd'hui retrouver les campagnes arpentées par Edmont, Scheuermeier, Rohlf's et tant d'autres. Que de choses nous saurions à présent y faire pour en tirer le meilleur parti... Cela paraît donc la raison inévitable d'un travail sur des matériaux vivants, il faut toujours en tenir compte.

formidable apport de la dialectologie, pas le regard sur l'actualité. Sans cet immense corps de données réalisé en plus d'un siècle, nous aurions sans doute une vision erronée du XX<sup>e</sup> siècle avant l'invention de l'enregistrement, avant la multiplication des images. Ces herbiers jaunis que nous appelons des atlas linguistiques nous présentent donc des données disparues ou que nous ne reconnaissons plus. Les critiques que l'on peut en faire sont nombreuses, c'est vieillot selon certains, c'est d'un intérêt scientifique à discuter, mais la mémoire est là et c'est surtout grâce à cela que la famille romane (notamment) est plus que jamais exemplaire dans le monde global.

Pour cet ensemble de raisons, toute société intelligente devrait anticiper dans la perspective du recueil des patrimoines, car les sociétés dont les patrimoines ne sont pas recueillis, valorisés, sont dans une bonne mesure des sociétés en danger. Patrimoines langagiers certes, mais aussi paysagers, culturels, comme le démontrent les études onomastiques que plusieurs d'entre nous envisagent naturellement.

\*

Premier paradoxe. Un organisme international comme l'UNESCO donne de plus en plus de puissance à cette notion de patrimoine à protéger, à enregistrer pour le transmettre aux générations futures. Dans le même temps, non seulement la dialectologie recule partout en visibilité dans les pays européens qui l'ont portée, la France en particulier ; mais, pire encore, la linguistique diachronique, mamelle de la géographie linguistique, s'effondre dans les universités, là où en principe elle avait une place importante et constituait un vivier naturel pour des recherches linguistiques de terrain. L'avenir est assez sombre, le problème étant qu'arrivent et qu'arriveront ainsi de plus en plus de gens non formés à la typologie et à la diachronie des langues et pourtant sincèrement enthousiastes face à de telles recherches patrimoniales.

Il est probable que nous payons ainsi, plusieurs années après, le déport caricatural (mais compréhensible) du pendule scientifique du tout diachronique/externe vers le tout synchronique/interne, en dépit de l'irruption de la sociolinguistique à la fin du XX<sup>e</sup> siècle. Comme toujours, le risque est que l'on a de bonnes chances de se réveiller collectivement quand il sera trop tard pour corriger l'erreur de trajectoire.

\*

On en arrive au deuxième paradoxe. Alors que la dialectologie est globalement victime de l'évolution du paysage des sciences du langage en Europe et en France en particulier, des modes aussi, différents pays dits « émergents », privés d'une tradition d'études diachroniques et géolinguistiques, se mobilisent au contraire pour les projets de recueil des langues parlées dans leurs territoires, non sans difficultés<sup>10</sup>. Il se trouve que Melissa Barkat et Élisabeth Motte-Florac évoquent plus haut un projet d'atlas linguistique du monde arabe. Elles ne sont pas les seules. Depuis plusieurs années déjà, certains linguistes tunisiens militaient pour un atlas linguistique de leur pays. Les choses ont très bien avancé, on peut se référer aux états produits par Salah Mejri<sup>11</sup>. De tels développements sont à verser au dossier<sup>12</sup>.

Résumons un peu ce paradoxe. La dialectologie européenne, romane et française en particulier, est une lumière pour différents chercheurs dans les pays méditerranéens, ailleurs dans le monde aussi<sup>13</sup>. Beaucoup considèrent ce modèle comme particulièrement attractif, souhaitable, incontournable à terme, en partie reproductible<sup>14</sup>. Mais dans le même temps, Européens et Français voient très souvent les dialectologues et la dialectologie (quand ils les voient) comme des dinosaures, comme une époque révolue, une discipline archaïque qui n'aurait plus sa place dans le monde moderne. Qui se serait mise elle-même hors-circuit<sup>15</sup>.

---

<sup>10</sup> Où l'on voit que la dialectologie est une école du respect de soi et des autres, une école de liberté tout simplement aussi. Le quadrillage des territoires européens n'a pas toujours été facile. Une dialectologie sous l'Ancien Régime eût été sans doute inconcevable, comme il y a peu, toute recherche de terrain était à peu près interdite dans des portions entières de la Méditerranée. L'exemplarité de la dialectologie française et romane repose aussi sur cette remarque.

<sup>11</sup> Notamment (2011 : 35-51), "L'atlas linguistique de Tunisie". *International Journal of The Sociology of Language*, n°211, et 2005, également cité par Barkat & Florac : 331.

<sup>12</sup> Autre illustration de cette re-jonction. Après avoir joué un rôle déterminant dans l'affirmation de la sociolinguistique (avec l'enracinement méthodologique que l'on a rappelé), William Labov a réalisé en collaboration un *Atlas of North American English : Phonology and Phonetics*, Berlin, Mouton de Gruyter, 2006, 318 p.

<sup>13</sup> En témoigne par exemple le succès de la *Société internationale de dialectologie et de géolinguistique* (SIDG) dont certains d'entre nous sont membres.

<sup>14</sup> Du coup, les dialectologues français ont un rôle pédagogique, aujourd'hui encore. La demande de l'étranger devient en effet de plus en plus forte, ce qui est, on le comprend bien, signe positif et négatif en même temps. On rappellera qu'un processus du même genre fut lié à l'ALF et aux enseignements de Jules Gilliéron, déclencheur de formation de chercheurs étrangers à Paris, ces derniers porteurs de la dialectologie dans leurs pays respectifs. S. & R.-D. Pop en dressent une liste intéressante [pour 1883-1926], dans leur ouvrage déjà cité : *Gilliéron, vie, enseignement, élèves, œuvres, souvenirs* (1959 : 53-63).

<sup>15</sup> Un dernier pan, celui des militants. Depuis les années 1970-80 leur nombre s'accroît, en liaison avec la visibilité croissante des langues régionales au moment même où celles-ci paraissent perdre

\*

Terminons par le troisième paradoxe.

Admettons que la dialectologie soit périmée, engageons alors une petite fiction. Fermons les centres, n'assurons plus de formation. Ce n'est pas très difficile compte-tenu du fait que la plupart des dialectologues « historiques » (une majorité de nos participants en 2009) sont sortis du système universitaire depuis 10 ou 20 ans ; et compte-tenu du fait que ceux qui les ont remplacés (au passage de plus en plus sociolinguistes, plus que géolinguistes) partent à la retraite en ce moment et dans les années qui viennent.

2020-2030, extinction naturelle, il n'y a plus de dialectologues. Sinon quelques dinosaures, pour reprendre ce mot. Ou encore quelques dialectologues nouvelle version par la force des choses, moins ou pas formés à la diachronie, aux méthodes d'enquête. Ils font ce qu'ils peuvent, ils ont en général comblé leur handicap initial, mais la fibre semble vraiment cassée.

2030-2050 et après, des voix se font entendre, de plus en plus nettement. Comment et pourquoi la recherche scientifique a-t-elle abandonné ces terrains de la transmission des langues et des patrimoines, alors que tout le monde se bat autour de nous pour réorganiser le monde à partir de valeurs universelles d'identité, de patrimoine, de culture, plutôt que de valeurs liées à la globalisation-mondialisation économique et culturelle ?

\*\*

Malheureusement, c'est peut-être ce scénario qui se prépare. Il reste toutefois un espoir, celui d'une rapide prise de conscience collective, débordant de très loin dialectologues et sociolinguistes eux-mêmes. On ne manque pas de signes à ce sujet. Le fait notamment que la question des langues et identités en danger devienne sous nos yeux un sujet de préoccupation profond<sup>16</sup>, qu'une demande apparaisse dans des zones du monde où l'on n'aurait pas soupçonné cela quelques années plus tôt, tout cela donc peut permettre dans les années à venir des conditions globalement favorables à une revitalisation des études de type dialectologique.

---

beaucoup de terrain. C'est un public nouveau pour la dialectologie, demandeur, mais assez souvent en porte-à-faux (pour des raisons qu'on ne peut aborder ici) avec la recherche universitaire.

<sup>16</sup> L'UNESCO, déjà évoquée, milite de plus en plus clairement pour la valorisation du patrimoine culturel immatériel (2003, *Convention pour la sauvegarde du patrimoine culturel immatériel*). Elle tend à le protéger et recense également les langues en danger par le biais d'un *Atlas des langues en danger* : <http://www.unesco.org/new/fr/culture/themes/endangered-languages>.

Probablement ces études repartiront-elles en priorité dans des zones à peu près vierges de ce point de vue, permettant à nouveau la réactivation de liaisons scientifiques naturelles avec l'ethnologie et l'ethnolinguistique par exemple.

Alors sans doute cette dialectologie romane, française, autre, curieusement reconnue et en même temps minimisée, redeviendra l'une des références des sciences du langage et de la société.

\*\*\*

Il me reste à remercier l'ensemble des collègues, chercheurs et militants des langues régionales, communicants ou non, qui ont permis en 2009 une rencontre stimulante, gaie et parfois morose en même temps, durant laquelle de bonnes idées et propositions ont été engrangées. Notamment nos amis dialectologues venus de l'université de Turin. Je tiens à remercier tout spécialement William Labov et Gillian Sankoff qui nous ont honorés de leur présence et dont les communications ont été écoutées avec le plus grand intérêt par les linguistes de l'université de Lyon. Je tiens aussi à remercier mon épouse Marie-France pour son aide dans l'organisation du colloque ainsi que ses relectures attentives, ainsi que Brigitte Horiot pour ses conseils et informations toujours avisés.

Merci également au Conseil Scientifique et au service de la recherche de l'Université Lyon 3, alors animés par Grégory Lee, à la région Rhône-Alpes qui a soutenu cette manifestation et nous aide depuis en différentes occasions, une région dont l'action en faveur des langues régionales du Rhône-Alpes (francoprovençal, occitan) n'est plus à démontrer<sup>17</sup>.

Enfin, depuis le colloque de 2009, plusieurs dialectologues, géolinguistes, onomasticiens, ont disparu. J'évoquerai en particulier Jacques Chaurand (octobre 2009), Martina Pitz (mai 2010), Gaston Tuailon (juin 2011), René Lepelley (août 2011), Marianne Mulon (octobre 2011), Lothar Wolf (juin 2012).

Tous, et bien d'autres, ont patiemment contribué à constituer cette mémoire du terrain dont on peut se féliciter et qui nous réunissait en 2009.

Ces travaux leur sont dédiés.

Lyon, juin 2012

---

<sup>17</sup> *Assises des langues régionales*, 21 juin 2012.

**EXTRAIT DES PUBLICATIONS EN VENTE AU CENTRE  
D'ÉTUDES LINGUISTIQUES**

(CEL, 18 rue Chevreul, 69362 Lyon cedex 07)

Les publications suivantes (antérieures à 2011) sont toujours disponibles et peuvent être commandées par courriel :

[cel@univ-lyon3.fr](mailto:cel@univ-lyon3.fr)

[Prix port payé]

**Série « dialectologie »**

*Le croissant : le concept et le mot. Contribution à l'histoire de la dialectologie française au XIX<sup>e</sup> siècle*, par Guylaine Brun-Trigaud, 1990. (53,36 € TTC).

*Dialectologie et Littérature du domaine d'oïl occidental. Sur le thème des enquêtes dialectales*. Actes du IV<sup>e</sup> colloque tenu à Poitiers, du 6 au 8 mars 1990, publiés par Brigitte Horiot, 1995. (22,87 € TTC).

*Français de France et Français du Canada. Les parlars de l'Ouest de la France, du Québec et de l'Acadie*, sous la direction de Pierre Gauthier et Thomas Lavoie, 1995. (42, 69 € TTC).

*L'onomastique au carrefour des sciences humaines*. Actes du XI<sup>e</sup> colloque de la Société française d'Onomastique, Université Lyon 3 (10-13 octobre 2001), édités par Brigitte Horiot et Corina Veleanu, 2004. (21,10 € TTC).

*La dialectologie hier et aujourd'hui (1906-2006)*. Actes du colloque international tenu à l'Université Lyon 3 (7, 8 et 9 décembre 2006). Textes rassemblés et édités par Brigitte Horiot, 2009. (25 € TTC).

**Série « Langues en contact »**

*Le français, ailleurs et toujours : place et fonctions du français dans les autres langues*. Actes du colloque international tenu à l'Université Lyon 3 (9 et 10 mai 2005). Textes rassemblés et édités par Brigitte Horiot, 2008. (22 € TTC).



### **Série « démodialectologie »**

*Des données empiriques à l'objet d'étude : la constitution d'un corpus de démodialectologie (français populaire d'Abidjan),* par Mariette Meunier-Crespo, 2 tomes, 1994. (42,69 € TTC).

### **Série « histoire de la linguistique »**

*Introduction à l'étude du « Dictionnaire étymologique ou origines de la langue française » de Gilles Ménage (1694). Les étymologies de Ménage : science et fantaisie,* par Isabelle Leroy-Turcan, 1991. (21,95 € TTC).  
*Noëls en français et en dialectes du XVI<sup>e</sup> siècle,* édités par Marina Fey, 2008. (22 € TTC).

### **Série « didactique »**

*Il signor Bianchi in ospedale. Méthode audio-orale d'italien pour le personnel hospitalier,* par Jean-Louis Chauzit, 1991. (12,20 € TTC).  
*Le fonctionnement de la démocratie espagnole actuelle (textes, documents et terminologie),* par Mariette Meunier-Crespo, 1998. (22, 87 € TTC).

### **Série « germanique ancien »**

*Althochdeutsch Syntax und Semantik.* Akten des Lyonner Kolloquiums zur Syntax und Semantik des Althochdeutschen (1-3 März 1990), publiés par Yvon Desportes, 1992. (30,49 € TTC).  
*Les préverbes et la préverbation en allemand au IX<sup>e</sup> siècle,* par Yvon Desportes, 4 tomes, 1998. (213, 43 € TTC).

### **Série « langues de spécialité »**

*Le subjonctif futur dans la langue juridique espagnole actuelle,* par Mariette Meunier-Crespo, 1998. (21,34 € TTC).

### **Hors série, mélanges, collectifs**

*Mélanges offerts au professeur Lothar Wolf « Je parle, donc je suis... de quelque part ».* Textes rassemblés et édités par Brigitte Horiot, Elmar Schafroth, Marie-Rose Simoni-Aurembou, 2005. (36,92 € TTC).

### **Publications du CEL, Nouvelle série**

*Unité et diversité de la linguistique.* Sous la direction de Francis Manzano. 301 pages, 2011. Publications du CEL, Nouvelle série 2011-1. (15 € TTC).

*Maghreb. Une francophonie sur la brèche, une interface en Méditerranée,* par Francis Manzano. 306 pages, 2011. Publications du CEL, Nouvelle série 2011-2. (15 € TTC).

## **Centre d'Etudes Linguistiques** **[Linguistique, dialectologie]**

Le Centre d'Etudes Linguistiques (EA 1663) est une équipe d'accueil de l'Université de Lyon (Université Jean Moulin – Lyon 3, Faculté des langues), dont les locaux se trouvent au 18, rue Chevreul – 5<sup>ème</sup> étage (Lyon, 7<sup>ème</sup> arrondissement).

L'équipe regroupe des chercheurs romanistes, anglicistes ou germanistes autour des thématiques de la linguistique comparative, de la dialectologie, de la sociolinguistique et de la pragmatique.

Les zones géographiques couvertes par le CEL sont celles de l'Europe de l'Ouest et de la Méditerranée occidentale (langues en contact). Les domaines d'expertise du groupe concernent principalement les langues et cultures régionales de France et de l'Europe romane, l'anthropologie linguistique, la dialectologie et la sociolinguistique (73<sup>ème</sup> section du CNU et 7<sup>ème</sup> section du CNU).

Le Centre d'Etudes Linguistiques est à ce titre support d'un master recherche spécifique « Linguistique et dialectologie », orienté vers la formation de jeunes chercheurs en sociolinguistique et dialectologie – programme sur la page :

[http://facdeslangues.univ-lyon3.fr/medias/fichier/master-lce-linguistique-et-dialectologie-05-04-2011\\_1302260055589.pdf](http://facdeslangues.univ-lyon3.fr/medias/fichier/master-lce-linguistique-et-dialectologie-05-04-2011_1302260055589.pdf)

Il est également laboratoire d'accueil et de direction pour des recherches doctorales en « Sciences du langage » et « Cultures et langues régionales » :

<http://facdeslangues.univ-lyon3.fr/formation/diplomes-nationaux/doctorat/doctorat-linguistique-linguistique-generale-langues-slaves-langues-romanes-didactique-des-langues-64398.kjsp?RH=LAN-RECHcel>

Les correspondances doivent être adressées au professeur directeur de l'équipe, et directeur des publications du CEL :

cel@univ-lyon3.fr  
francis.manzano@univ-lyon3.fr

D'où viennent la dialectologie, la géographie linguistique, et où vont-elles ?

Les 12 et 13 mars 2009, une vingtaine de dialectologues se retrouvaient à Lyon, à l'initiative du Centre d'Études Linguistiques pour un colloque scientifique intitulé « Mémoires du terrain ». Presque tous avaient travaillé directement à l'une de pièces du *Nouvel Atlas Linguistique de la France* (NALF), ou travaillaient à différents projets en cours, en France ou en Italie notamment (THESOC, AliR, ALI, ATPM etc.).

Le regard porté sur la discipline et sur les langues régionales étudiées par les uns et les autres révèle l'actualité de certaines démarches, la bonne santé relative des enquêtes et des descriptions dialectologiques en cours. Il révèle aussi la qualité des matériaux recueillis et inventoriés par les dialectologues dans une véritable tradition d'études de terrain.

Mais, parallèlement, les motifs d'inquiétude ne manquent pas. La communauté sociale, la communauté scientifique (celle des linguistes notamment) comprennent-elles véritablement le rôle et l'actualité d'une discipline vue par certains comme marginale, ou même désuète. En un mot, la dialectologie est-elle en train d'arriver au bout de son propre processus de développement, ou peut-elle au contraire contribuer à fournir les références scientifiques et méthodologiques indispensables au profond mouvement de gestion et de réhabilitation des identités, patrimoines immatériels et langues en danger dans le monde ?

Contributions de :

Jean Le Dù, Jean-Philippe Dalbera, Robert Chaudenson, Gérard Bodé, Jean-Paul Chauveau, Fernand Carton, Fabrice Jecic, Fernande Krier, Guylaine Brun-Trigaud, Pierre-Aurélien Georges, Jean-Claude Bouvier, Jean-Louis Fossat, Xavier Ravier, Jean-Pierre Lai, Michel Contini, Sabina Canobbio, Matteo Rivoira, Melissa Barkat-Defradas, Élisabeth Motte-Florac, Gérard Taverdet, Francis Manzano.

\*

*Francis Manzano est issu du centre de dialectologie romane de l'université Toulouse 2. Il a travaillé sur les zones de contact entre occitan et catalan dans les départements de l'Aude et des Pyrénées-Orientales (Thèse de Doctorat d'État en toponymie, 1987), et a enseigné plusieurs années au Maroc et en Tunisie (coopération universitaire : Université de Fès, ENS de Meknès et de Sousse). De 1991 à 2007, à l'université Rennes 2, il a développé des analyses sur les langues régionales de France et de la Méditerranée occidentale, la géolinguistique du gallo, ainsi que la sociolinguistique du Maghreb.*

*Il dirige depuis 2007 l'EA 1663, CEL [Linguistique – dialectologie], le département de linguistique et le master « recherche » Linguistique et dialectologie, à l'université de Lyon.*