

Allocution d'ouverture des représentants de l'ATILF

Au nom du comité d'organisation et de l'ensemble des membres du laboratoire ATILF, nous souhaitons remercier tous les congressistes qui sont venus à Nancy participer au XXVII^e Congrès international de linguistique et de philologie romanes¹. Nous voudrions aussi remercier la Société de linguistique romane, son président et son bureau pour la confiance qu'ils nous ont témoignée en nous confiant l'organisation de ce prestigieux congrès.

În numele comitetului de organizare și al întregului colectiv al institutului ATILF, vrem să mulțumim tuturor celor care au venit la Nancy pentru a participa la al douăzeci și șaptelea Congres internațional de lingvistică și filologie romanică. Mulțumirile noastre se îndreaptă de asemenea către Societatea de lingvistică romanică, către președintele și biroul său pentru încrederea pe care ne-au acordat-o, oferindu-ne posibilitatea de a organiza acest prestigios congres.

Le laboratoire ATILF (Analyse et Traitement Informatique de la Langue Française) est issu de trois anciennes unités : l'INaLF (Institut National de la Langue Française), une unité du CNRS, et deux équipes universitaires : le LanDisCo (Langue, Discours, Cognition) et le CRAPEL (Centre de Recherches et d'Applications Pédagogiques en Langues). Ses deux tutelles sont le CNRS et l'Université de Lorraine, et il réunit actuellement 138 membres : des enseignants-chercheurs et des chercheurs, des ingénieurs et des techniciens, du personnel administratif et des doctorants et post-doctorants. Le laboratoire est structuré en cinq équipes scientifiques : « Lexique », « Linguistique historique française et romane », « Discours, langue et cognition », « Didactique des langues et sociolinguistique » et « Ressources : normalisation, annotation et exploitation ».

Il laboratorio ATILF (Analyse et Traitement Informatique de la Langue Française) nasce dalla confluenza di tre precedenti unità: l'INaLF (Institut National de la Langue Française), un'unità del CNRS, e due gruppi di lavoro universitari, LanDisCo e CRAPEL. Le sue due istituzioni di riferimento sono il CNRS e l'Università della Lorena, e attualmente riunisce 138 membri tra docenti-ricercatori, ricercatori,

¹ Nos remerciements les plus vifs s'adressent à Xosé Afonso Álvarez Pérez (Lisbonne), Maria Reina Bastardas i Rufat (Barcelone), Myriam Benarroch (Paris), Giorgio Cadorini (Opava), Marco Maggiore (Florence), Mihaela-Mariana Morcov (Nancy), Simone Pisano (Sassari), Fernando Sánchez Miret (Salamanque) et Paul Videsott (Bolzano) pour les traductions en roumain, en italien, en sarde, en frioulan, en ladin, en catalan, en espagnol et en portugais qu'ils nous ont procurées.

ingegneri, tecnici, personale amministrativo, dottorandi e post-dottorandi. Il laboratorio è strutturato in cinque unità scientifiche: «Lessico», «Linguistica storica francese e romanza», «Discorso, linguaggio e cognizione», «Didattica delle lingue e sociolinguistica» e «Risorse: normalizzazione, annotazione e utilizzo».

Pour l'organisation de ce congrès, l'ATILF a non seulement pu compter sur l'infrastructure du laboratoire et l'engagement de ses membres, mais a pu bénéficier aussi du soutien de ses tutelles, le CNRS et l'Université de Lorraine, d'un ensemble de partenaires institutionnels : le Ministère de l'Enseignement Supérieur et de la Recherche, la Délégation Générale à la Langue Française et aux Langues de France, l'Agence Universitaire de la Francophonie, l'Institut de Linguistique Française, la Maison des Sciences de l'Homme Lorraine, et des collectivités territoriales avec lesquelles le laboratoire entretient depuis plusieurs années des partenariats importants : le Conseil Régional de Lorraine, le Conseil général de Meurthe-et-Moselle, la Communauté urbaine du Grand Nancy et la Ville de Nancy. Qu'eux tous reçoivent nos remerciements les plus chaleureux.

Pro s'ammàniu de su congressu, s'ATILF at àpiu sa fortuna manna de poder impreare sos mèdios de su laboratòriu e at retziu una manu manna de azudu dae sos chi bi trabballant ma at fintzas pòtiu aer s'apozu de sos entes suos amparadores: Su CNRS e s'Universidade de Lorena; una cumarca de assòtzios istitutzionales: Su Ministèriu de Imparonzu Superiore e de sa Chirca, Sa Delegatzione Zenerale de sa limba frantzesa e de sas limbas de Franza, s'Agentzia de s'Universidade pro sa Francofonia, Istitutu de linguistica frantzesa, sa Dommo de sas iscèntzias de s'òmine de Lorena; sas istitutziones de su territòriu chi traballant a costazu de su laboratòriu dae medas annos: Su Cossizu Regionale de Lorena, su Cossizu zenerale de Meurthe-et-Moselle, sa Comunidade de totu sas biddas a fùrriu de Nancy e sa Dommo Comunale de Nancy. A cada unu bos torramus milli gràssias sas prus sintzeras.

Nous voudrions aussi rendre hommage à l'action du comité de programme, présidé par Jean-Pierre Chambon en sa qualité de président de la Société de linguistique romane, lequel a sollicité différentes personnalités pour présenter des conférences plénières ou grand public ou pour présider des tables rondes, mais qui a surtout arrêté la thématique des seize principales sections du congrès, tout en en confiant la responsabilité à des présidents et vice-présidents. Une dix-septième section, restreinte à une seule journée, est par la suite venue compléter la gamme.

O volaressin ancje fâi onôr ae ativitât dal comitât programatic, che i à stât da cjâf Jean-Pierre Chambon inte sô cualitât di president da la Societât di linguistiche romane; cun di plui che al à interpeladis diuersis personalitâts par la presentazion des conferencis plenaris o di grant public e po ancje par la presidence des taulis tarondis, ma che so redut al à inmaneât la tematiche des sedis sezions primaris dal congrès, tal timp istès confidantjur la responsabilitât ai presidents e vice-presidents. Tal ultin e je rivade inmò une seziun disesiet, limitade a nome une zornade, che e complete la sielte.

Ce sont les présidents et vice-présidents qui ont eu à assumer la difficile charge de la sélection des communications à retenir parmi les 850 propositions qui nous sont parvenues. Ils ont eu à cœur d'assumer cette tâche dans un réel esprit d'ouverture, en faisant se côtoyer au sein de leur section, voire au sein d'une même session, un large éventail de langues analysées, de thématiques scientifiques et d'approches théoriques.

I presidenc y i vizepresidenc dles seziions à abù l'enciaria nia saurida da chirì fora i referac da porté dant anter les 850 propostes che é vegnudes manedes ite. Chest laour ài fat con n spirit davert, a na maniera che vigni seziion pò mostré su n gran numer de lingac adorés, de tematches analisedes y de teories metudes a descusion.

Afin de donner la possibilité à un plus grand nombre de romanistes de présenter leurs travaux, nous avons pris l'initiative, pour ce vingt-septième congrès de la Société, de mettre en place, au sein des différentes sections, des sessions de posters, qui, grâce au temps ensoleillé dont nous avons pu bénéficier durant toute la semaine, se sont déroulées en plein air, sous les arcades du beau bâtiment que la Faculté de droit, sciences économiques et gestion de l'Université de Lorraine a mis à notre disposition.

A fi de permetre que un nombre més gran de romanistes pogués presentar els seus treballs, hem pres la iniciativa d'organitzar, per aquest vint-i-setè congrés de la Societat, i en el sí de les diferents seccions, unes sessions en format pòster. Aquestes comunicacions s'han presentat a l'aire lliure, gràcies al bon temps i al sol que ens han acompanyat durant tota la setmana, sota els porxos del bell edifici que ens ha cedit la Facultat de Dret, Ciències Econòmiques i Gestió de la Universitat de la Lorena.

En matière d'édition d'actes, nous avons choisi d'innover, avec l'objectif de garantir une lecture partagée et donc une appropriation collective des apports diversifiés du congrès. Les actes sont ainsi publiés dans leur intégralité sur internet et accessibles à partir des sites de l'ATILF et de la Société de linguistique romane et parallèlement, sous une forme partielle, sur papier. La publication internet répond aux meilleures exigences actuelles de cet outil de diffusion avec en particulier, pour chaque article, un DOI, identifiant d'objet numérique, équivalent pour les documents numériques du numéro ISBN des publications papier, ainsi qu'une pagination continue, permettant des citations précises. Par ailleurs, l'intégralité des actes sera versée dans les collections électroniques de la Bibliothèque nationale de France pour garantir leur longévité.

En cuanto a la edición de las actas, hemos decidido innovar, con el objetivo de garantizar una lectura compartida que facilite una apropiación colectiva de las variadas aportaciones del congreso. Por este motivo las actas se han publicado íntegramente en internet y, paralelamente, una versión reducida ha aparecido en papel. La publicación en internet responde a las máximas exigencias actuales de este medio de difusión: los artículos se han etiquetado con un formato XML, se han paginado de manera que se puedan hacer referencias bibliográficas precisas y cuentan con un número DOI (identificador de objeto digital), que equivale al ISBN de las publicaciones en papel. Los textos de cada sección se pueden descargar desde los sitios del

ATILF y de la Société de linguistique romane. Además, a fin de garantizar su longevidad, las actas completas se volcarán en las colecciones electrónicas de la Biblioteca Nacional de Francia.

Pour ce qui est de la publication sur papier, elle regroupe le texte des événements communs – les allocutions prononcées lors de la cérémonie d'ouverture, les conférences plénières et les tables rondes – ainsi qu'un aperçu synthétique du travail de chacune des sections, assorti d'un choix représentatif de communications.

*Quanto à publicação em papel, reúne o texto dos acontecimentos comuns – as alo-
cuções pronunciadas na cerimónia de abertura, as conferências plenárias e as mesas
redondas – assim como uma amostra sintética do trabalho de cada uma das seções,
acompanhada por uma seleção representativa de comunicações.*

Secrétaire général du CILPR 2013
Secrétaire adjointe du CILPR 2013

Jean-Marie PIERREL
Éva BUCHI

Allocution d'ouverture du représentant de l'Université de Lorraine

Au nom de Pierre Mutzenhardt, Président de l'Université de Lorraine, je vous souhaite la bienvenue en Lorraine, à Nancy et dans les locaux de l'UL.

Comme vous le savez sans doute, l'Université de Lorraine est une jeune et en même temps une ancienne université, puisque créée en janvier 2012 par la fusion de quatre universités quadragénaires, et compte près de 55 000 étudiants dont environ 2 000 doctorants. Mais l'histoire remonte en fait encore plus loin, à 1572, année de création de la première université de Lorraine, soit 440 ans d'âge...

Le site lorrain associe l'Université de Lorraine à quatre Établissements Publics à caractère Scientifique et Technologique (CNRS, Inra, Inria et Inserm) et à un Centre Hospitalier Universitaire. Ces six acteurs sont engagés dans un partenariat académique fort depuis de nombreuses années et partagent des volontés politiques et stratégiques en recherche déclinées en six axes thématiques prioritaires sur des enjeux sociétaux et résolument orientées vers la pluridisciplinarité dans une articulation affirmée entre recherche fondamentale, finalisée et appliquée, tout ceci en partenariat étroit avec le monde socio-économique.

Ce site lorrain comprend plus de 4 000 personnels permanents dans les structures de recherche, soit 61 laboratoires répartis en dix pôles scientifiques qui couvrent presque toutes les disciplines scientifiques. Parmi ces dix pôles, trois couvrent les sciences humaines et sociales, juridiques et de gestion et l'un d'entre eux (dénommé *Connaissance, Langage, Communication, Sociétés*) abrite le laboratoire ATILF, coorganisateur de l'évènement d'aujourd'hui avec la Société de linguistique romane.

L'ATILF est l'un des laboratoires SHS les plus conséquents de l'Université de Lorraine et est parfaitement en phase avec la politique de l'établissement. En effet, il participe en particulier à l'axe prioritaire « Ingénierie des langues et des connaissances », qui représente une thématique fortement ancrée en Lorraine depuis l'installation du *Trésor de la Langue Française* (TLF) en 1962.

Aujourd'hui l'ATILF est acteur des programmes d'investissement d'avenir au travers de l'Equipex ORTOLANG (Outils et Ressources pour un Traitement Optimisé de la LANGue) et est fortement impliqué dans des projets de recherche internationaux d'envergure.

D'ailleurs, si ce XXVII^e Congrès international de linguistique et de philologie romanes a choisi la Lorraine pour se réunir, après une absence de 30 ans en France,

c'est sans aucun doute significatif de la place des travaux menés à l'ATILF dans ce domaine, et ses deux tutelles, l'Université de Lorraine et le CNRS, ne peuvent que s'en féliciter (tout en félicitant une fois de plus ses directeurs successifs).

Voilà, au nom de l'Université de Lorraine et de son Président, je vous renouvelle nos mots de bienvenue et, 1 200 ans après le Concile de Tours, en cette année anniversaire pour les langues romanes, je vous souhaite que vos travaux d'aujourd'hui aient un même retentissement suite à ce qui pourrait devenir le Congrès de Lorraine, ne voyant point dans cet amphithéâtre les marques élémentaires d'un concile.

Chargé de mission « Partenariat académique avec les EPST »
de l'Université de Lorraine

Pierre ARCHAMBAULT

Allocution d'ouverture du représentant du CNRS

C'est un honneur de participer à l'ouverture du XXVII^e Congrès international de linguistique et de philologie romanes, fort de près de 900 participants venus d'une quarantaine de pays différents. Depuis sa création, c'est la première fois que ce congrès est accueilli en Lorraine et la sixième fois en France, où il ne s'est pas tenu depuis trente ans. La linguistique, le lexique, les corpus, la littérature, la philologie, le multilinguisme, le traitement automatique de données textuelles, la création de nouvelles ressources informatisées et de nouveaux outils pour l'étude et l'utilisation des langues seront au cœur de vos échanges.

Ce congrès est un événement majeur de la Société de linguistique romane, qui a comme vocation de fédérer au niveau international des chercheurs et des institutions qui souhaitent développer les études de linguistique et philologie romanes. Fondée en 1924, elle compte aujourd'hui un millier de membres de 45 pays différents.

Le congrès est organisé sur place par l'Unité Mixte de Recherche ATILF (Analyse et Traitement Informatique de la Langue Française), qui associe l'Université de Lorraine et le CNRS.

Le Centre National de la Recherche Scientifique est un organisme public de recherche. Avec près de 34 000 personnes (dont 25 000 statutaires), un budget pour 2013 autour de 3 milliards et demi d'euros, le CNRS exerce son activité dans tous les champs de la connaissance, en s'appuyant sur plus de 1 100 unités de recherche et de service. Des chercheurs éminents ont travaillé et travaillent dans des laboratoires du CNRS, qui, avec dix-huit lauréats du prix Nobel et onze de la Médaille Fields¹, est fort d'une longue tradition d'excellence.

L'ATILF est une unité importante du dispositif de recherche du CNRS en linguistique, avec un nombre d'effectifs CNRS (ingénieurs, techniciens et chercheurs) qui dépasse la trentaine. Créé en 2001, ce laboratoire demeure une référence sur les plans national et international en linguistique historique, en didactique des langues, en analyse du discours ainsi que dans le domaine du lexique, une thématique à caractère transversal.

L'ATILF est porteur de l'EQUIPEX ORTOLANG (Open Resources and TOols for LANGUAGE), membre de la fédération CNRS Institut de Linguistique Française (ILF) et du consortium international Text Encoding Initiative (TEI). Il est aussi labo-

¹ Depuis la date de la conférence, le CNRS a vu augmenter en 2014 la liste de ses prix par attribution du prix Nobel d'économie et d'une médaille Fields de mathématiques.

ratoire support du Centre National de Ressources Textuelles et Lexicales (CNRTL), mis en place par le CNRS et impliqué au sein de l'infrastructure européenne Common Language Resources and Technology Infrastructure (CLARIN).

L'architecture de l'ATILF est articulée en cinq équipes : Lexique ; Linguistique historique française et romane ; Discours, langue et cognition ; Didactique des langues et sociolinguistique ; Ressources : normalisation, annotation et exploitation.

L'ATILF est une unité très dynamique sur le front des coopérations nationales et internationales et sur celui de la valorisation, notamment en ce qui concerne l'analyse de la langue française et son traitement informatique.

Pour conclure, je tiens à remercier les organisateurs d'un congrès qui demeure la référence internationale des études de linguistique romane. En particulier, ma gratitude va à Jean-Marie Pierrel, l'ancien directeur de l'ATILF, et Éva Buchi, l'actuelle directrice, pour avoir su relever un défi si important et mobiliser les effectifs de l'unité. Un remerciement spécial va au Président de la Société de linguistique romane, Jean-Pierre Chambon, qui a su impulser une dynamique remarquable à la Société et fédérer des forces internationales considérables.

Dedico infine una parola di riguardo alla comunità linguistica italiana così fortemente rappresentata per quantità e qualità di partecipanti a questo congresso, ricordando ai miei connazionali quanto fu generosa in tempi recenti la terra di Lorena nei confronti dei nostri emigranti che partirono d'Italia per ragioni prima politiche e poi economiche.

Je souhaite ainsi à tous les participants des échanges fructueux sur le plan scientifique et un agréable séjour dans cette ville culturellement très riche.

Directeur adjoint scientifique, en charge de l'économie, de la gestion et de la linguistique, de l'Institut des sciences humaines et sociales
du Centre National de la Recherche Scientifique

Stefano BOSI

Allocution d'ouverture du président de la Société de linguistique romane

Mesdames et Messieurs les vice-présidents, Monsieur le doyen, Monsieur le directeur adjoint, Monsieur le chargé de mission, chers consœurs et confrères, chers amis,

Quatre-vingt-cinq ans après le premier congrès de la Société de linguistique romane tenu à Dijon en 1928, les « pèlerins de la latinité » (la formule est d'Albert Henry) et/ou de la romanité que nous sommes reviennent en France d'oïl. Après de longues pérégrinations, nous devons ce retour à Jean-Marie Pierrel et Éva Buchi, que je voudrais remercier tout d'abord, ainsi que toutes les personnalités et institutions qui ont permis la réalisation de leur projet.

La ville lorraine qui nous accueille est devenue depuis quelques décennies, grâce à l'INaLF, puis à l'ATILF (CNRS/Université de Lorraine), à leurs chercheurs et aux projets que vous connaissez tous – le *Trésor de la langue française*, le *Dictionnaire du moyen français*, le *Französiches Etymologisches Wörterbuch* et le *Dictionnaire Étymologique Roman* –, l'une des capitales de la linguistique et de la lexicographie françaises, puis romanes. Si bien que beaucoup d'entre nous se sentent déjà un peu chez eux à Nancy, au plan intellectuel en tout cas.

Chacune de nos réunions triennales est pour les organisateurs l'occasion d'apporter des retouches à l'architecture de nos congrès et d'esquisser une certaine idée de nos deux disciplines, la linguistique romane et la philologie romane : des disciplines bien distinctes, mais non irrémédiablement opposées et dont chacun de nos congrès proclame au contraire la nécessaire complémentarité.

À son tour, le congrès de Nancy apporte certaines modifications à ce qui est devenu, au fil des éditions, notre schéma de base.

Une première innovation réside dans l'organisation de quelques conférences « en ville ». Cette initiative vise à faire partager à un plus grand nombre certains aspects de nos travaux. Une première conférence a déjà eu lieu, permettant d'annoncer le congrès ; d'autres suivront, consacrées à l'histoire du français, aux apports des savants allemands à la linguistique/philologie romane et aux noms de famille en Lorraine. Je remercie les confrères qui ont bien voulu accepter de prononcer ces conférences « grand public », qui seront suivies aussi, j'en suis sûr, par beaucoup d'entre nous.

Le congrès de Nancy met d'autre part l'accent, dans ses deux premières sections, sur une articulation plus nette et plus ferme de la linguistique romane avec les branches de la linguistique super-ordonnées : la linguistique latine et la linguistique générale.

On peut souhaiter, je crois, que cette articulation soit pérennisée dans nos prochains congrès. C'est là, d'ailleurs, un retour aux sources : en 1925, l'organigramme de la Société commençait par la linguistique générale, dont le responsable était Jules Vendryes, et par le latin, représenté notamment par Alfred Ernout.

Par ailleurs, les littératures romanes médiévales, qui ont déjà fait des apparitions disons épisodiques dans nos congrès, sous des intitulés divers, reviennent à Nancy avec une section qui leur est explicitement consacrée. Il me semble désirable que cette expérience puisse être prolongée.

La répartition des communications en sections tente, autant que possible, de mettre de l'ordre dans la richesse proliférante des propositions, grâce à la collaboration – essentielle – des présidents et des vice-présidents. À cet égard, la section « Syntaxe » a atteint à une belle ordonnance thématique de ses sessions.

Les conférences plénières et les table rondes nous permettront de lutter contre les tendances centrifuges et de faire le point sur quelques questions centrales : philologie textuelle, philologie et syntaxe historique, géolinguistique, sémantique, unité de la romanistique. Max Pfister nous rappellera, *in fine*, que la Lorraine est de longue date une terre de contact entre le monde roman et le monde germanique.

La très large majorité des grands ensembles de variétés romanes feront l'objet de communications, y compris le macédonien, le dalmate, l'istroroman, le latin dolomitique, le francoprovençal, l'occitan, l'asturien, le galicien ou le français exporté. En feuilletant le programme, il m'a semblé que l'italien méridional et le portugais du Brésil étaient particulièrement bien représentés. Les influences romanes sur diverses langues non romanes (le polonais, le basque, le guarani et le chinois) seront aussi envisagées.

Sept langues romanes, sauf erreur, seront employées comme métalangues : le roumain, l'italien, le français, le catalan, l'espagnol, le portugais, mais aussi le frioulan.

Cette double diversité des langues et des métalangues est un bien extrêmement précieux, qui est constitutif à tous égards de l'identité de notre Société.

La linguistique romane n'est pas l'addition de la linguistique française, de la linguistique espagnole, de la linguistique italienne etc., mais une discipline *sui generis* ayant constamment pour fondement et comme horizon la totalité des variétés romanes, dans leur diversité et dans leur unité, dans le présent comme dans le passé. Pour la linguistique romane, toutes les variétés romanes, standardisées ou non standardisées, standard ou sub-standard, possèdent un poids épistémique identique.

Ne serait-ce qu'en raison de leur nombre, les variétés primaires, c'est-à-dire les variétés non standardisées (dont les variétés anciennes pré-standard), revêtent une importance fondamentale pour la linguistique romane. Ces variétés sont partout menacées, si bien que la linguistique romane se trouve confrontée à la disparition progressive d'une partie essentielle de son objet. Cette disparition est une catastrophe scientifique et une raison supplémentaire pour souhaiter que la dialectologie romane conserve le rôle central qu'elle a tenu dans notre discipline. Certes, la notion de dia-

lectologie est notoirement mal formée : faire de la dialectologie, ce n'est pas autre chose que faire de la linguistique (et même au sens éminent de ce terme !), mais l'existence d'une dialectologie romane vigoureuse constitue un rempart protecteur contre le poids social écrasant des grandes langues standardisées et contre le *standardocentrisme* dans nos études.

En ce qui concerne à présent les métalangues, vous savez que dans nos congrès, dans la vie de notre Société et dans la *Revue de linguistique romane*, toutes les langues romanes sont admises, mais rien que les langues romanes. Ce dispositif de défense de la *latinità* est dû aux refondateurs de notre Société. Nous savons gré à nos confrères non romanophones natifs qui veulent bien s'exprimer dans un idiome roman lors de nos congrès et dans notre *Revue*.

Si, dans le grand processus glottophagique en cours, les variétés romanes primaires sont tout simplement menacées de mort, les variétés romanes standardisées sont, quant à elles, menacées d'un début de déstandardisation, à travers la 'démaximalisation' de leurs fonctions, en particulier dans la communication scientifique, et ce en faveur de la langue de la puissance mondiale dominante. Certes, disait Georges Straka, ce ne sont pas les linguistes et les philologues qui font l'histoire des langues. Ils ont cependant, comme tout un chacun, leur petite part de responsabilité individuelle et collective en tant que parleurs et que scripteurs. On peut toujours choisir de résister à l'hégémonisme de la langue – et de la pensée – uniques.

Cette question est clairement liée aux bouleversements qui affectent les universités européennes. Sur ce point, je me borne à redire ce que j'ai eu l'occasion de déclarer au congrès de Valence : « la Société de linguistique romane est une compagnie de savants, dotée de sa logique propre, distincte de la logique strictement universitaire ». J'espère qu'en dépit des inévitables pressions, notre Société saura maintenir ce cap d'autonomie et de liberté.

Je signale ici que la Société de linguistique romane et le congrès se réjouissent de donner l'hospitalité à la réunion qui jettera les fondations d'un Conseil international d'études romanes regroupant les associations scientifiques actives dans notre champ de travail.

Vous savez sans doute que les romanistes aiment à caresser l'idée qu'ils forment une grande « famille ». Je préfère néanmoins reprendre, en terminant, une formule d'Albert Henry, formule que je mets à l'impératif : efforçons-nous donc de faire vivre, pour la durée de ce congrès, « une démocratie savante, représentant tout ce qui est Rome et ce qui est amour de la latinité » !

Je déclare ouvert le XXVII^e Congrès international de linguistique et de philologie romanes.

Professeur de linguistique romane et
langue et littérature d'oc à l'Université de Paris-Sorbonne,
Président de la Société de Linguistique Romane

Jean-Pierre CHAMBON

Lachmann et Bédier. La guerre est finie

Opposer Joseph Bédier et Karl Lachmann, en parlant de l'édition des textes, c'est comme inventer un duel entre une personne réelle et un fantôme. Sebastiano Timpanaro¹ et surtout Giovanni Fiesoli² ont montré que Lachmann n'est pas le responsable de la méthode dite lachmannienne. Il n'empêche que cette méthode, largement adoptée par les spécialistes, est un point de repère, grâce aussi à l'importance de la filière Lachmann-Paris-Bédier.

Comme dans le cas de la provenance des fabliaux ou celui des origines de l'épopée, en ce qui concerne l'édition de textes Bédier renverse totalement la position de son maître, Gaston Paris. Bédier a saisi avant les autres le changement épistémique entre la culture romantique et les nouvelles tendances historiographiques. Par exemple, au lieu de partir à la recherche des origines reculées et mystérieuses de l'épopée, Bédier focalise son attention sur les chansons de geste telles que nous les connaissons, et évite de s'égarer dans le haut Moyen Âge, sur des pistes trop problématiques. La critique philologique, élaborée par les philologues anglais et allemands sur des textes d'abord grecs et latins et ensuite allemands et romans aussi, avait l'ambition de restituer et reconstituer les œuvres qui sont arrivées jusqu'à nous, généralement grâce à des manuscrits. Une tâche d'envergure, si l'on remarque que ces philologues, pour la plupart protestants, ont affronté également le texte de la Bible. Depuis le milieu du XIX^e siècle, la philologie allemande a assuré un travail gigantesque de récupération et reconstitution. Son idéal était de trouver des principes logiques de validation et de restitution ; la perspective de la méthode des sciences naturelles était dominante. Et dans ces mêmes années, Eugène Viollet-Le-Duc refaisait les principaux monuments médiévaux selon la doctrine exposée dans son très savant Dictionnaire raisonné de l'architecture française (1854-1868).

En 1872, Gaston Paris met au point une édition du célèbre poème hagiographique sur *Saint Alexis* selon les principes de l'école allemande, qu'il avait acquis en tant qu'élève de Friedrich Diez, considéré comme le fondateur de la philologie romane³. L'édition de Paris, qui est la première édition lachmannienne d'un texte roman et un véritable monument de maîtrise philologique, reconstituait le texte de ce poème en utilisant tous les manuscrits, y compris ceux qui transmettaient des remaniements du texte original, datable de la fin du XI^e siècle. Paris résumait les résultats de sa recher-

¹ Timpanaro (1963).

² Fiesoli (2000).

³ Paris / Pannier (1872).

che sur la tradition textuelle dans un arbre généalogique, ou *stemma codicum*. Cet arbre était fondé sur le principe des fautes communes. En pratique, si deux ou plusieurs manuscrits contiennent les mêmes fautes, on considère qu'ils remontent sans aucun doute à un seul ancêtre commun, puisqu'il est invraisemblable que des manuscrits qui ne sont pas apparentés présentent les mêmes fautes aux mêmes endroits.

Une fois l'arbre reconstruit, l'accord de deux ou plusieurs traditions indique les leçons originaires, qu'il faudra considérer comme correctes, sur la base d'un raisonnement inverse : comme il est impossible que deux manuscrits contiennent les mêmes innovations au même endroit, l'accord de deux contre un indique la leçon correcte. Pour donner un exemple, dans le cas d'une œuvre transmise par trois manuscrits, ou trois sous-archétypes, ABC, on peut considérer probablement comme correctes les leçons communes à A et B, ou à A et C, ou encore à B et C. Étant donné, en revanche, un arbre AB ; C, l'accord de A avec C ou de B avec C indique encore une fois la leçon correcte, tandis que dans le cas d'une opposition AB contre C, il n'existe aucun critère logique pour donner raison à l'un ou à l'autre. C'est un contre un.

C'est justement sur cela que Bédier prend appui dans son article révolutionnaire de 1928⁴. Mais il faut préciser qu'entre 1872 et 1928, la France, comme d'ailleurs l'Italie, avait fait son entrée dans le concert ecdotique en produisant des éditions critiques lachmanniennes d'un niveau excellent, par exemple celles du *Roman de Troie* par Léopold Constans (1904-1912) ou du *Couronnement de Louis* (1888) et du *Roman de la Rose* par Ernest Langlois (1914-1924), ou encore les œuvres de Machaut éditées par Hoepffner (1908-21), que l'on peut comparer aux produits du lachmannisme allemand, comme Chrétien de Troyes par Wendelin Foerster (1884, 1887, 1890, 1899, etc.), ou les *Fables* de Marie de France par K. Warnke (1898) ou les magnifiques éditions, toujours allemandes, des troubadours occitans, parmi lesquelles il suffira de citer celle de Bernart de Ventadorn (1925), par Carl Appel.

À vrai dire, le premier travail lachmannien sur des textes romans est l'étude, qui pourtant n'aboutira pas à une édition, de la tradition du *Fierabras* (1869) menée par Gustav Gröber, qui par la suite, en 1877, nous a donné sa recherche magistrale sur les *Liedersammlungen der Troubadours*. En Italie, après l'édition très précoce d'Arnaut Daniel par Ugo Angelo Canello (1883), pour trouver des éditions significatives il faut arriver aux *Realì di Francia* de Andrea da Barberino, édités par Giuseppe Vandelli (1892-1900), ou à la *Divina Commedia* et aux autres ouvrages de Dante (*Vita nuova* et *Rime* édités par Michele Barbi, en 1907 et 1932) : nous sommes là autour du sixième centenaire de la mort du grand poète (1921).

L'importance attribuée à l'établissement critique du texte est telle que Léon Gautier, après avoir publié son commentaire classique de la *Chanson de Roland*, en 1872, sur la base de l'édition lachmannienne de Theodor Müller (1863), modifie radicalement son commentaire à partir de 1880 (vii^e édition), parce qu'il apporte des retouches au texte de Müller en le conformant aux résultats des premiers travaux d'Edmund

⁴ Bédier (1928). Voir à ce propos Leonardi (2009).

Stengel (1877), et en anticipant ainsi l'édition que Stengel ne publiera qu'en 1900. Une décision, faut-il ajouter, déplorable, car Stengel avait conçu son texte sur la base d'un arbre qui était lachmannien, bien sûr, mais décidément erroné, tandis que l'arbre précédent de Müller est encore valable, même selon Bédier.

À cette activité d'édition d'inspiration rigoureusement lachmannienne participe aussi, et en première ligne, Joseph Bédier, notamment avec ses deux éditions du *Lai de l'ombre* (1890 et 1913), dont la seconde est conçue comme une réponse à un compte-rendu de Gaston Paris, qui à son tour avait proposé son propre arbre.

Dans son article de 1928, Bédier change sa perspective, et semble se poser pour la première fois le problème de la facilité avec laquelle on peut remplacer un arbre par un autre. Il s'amuse même à multiplier les combinaisons possibles, en proposant au moins 14 arbres. C'est un grand jeu de l'intelligence, mais pas si surprenant, si l'on envisage posément les faits. Évidemment, les manuscrits du *Lai de l'ombre* contiennent très peu de fautes indiscutables, et les philologues sont obligés de chercher des fautes là où il n'y en a pas : la combinatoire devient infinie. En général, on a remarqué que les copistes consacrent beaucoup plus d'attention et de soin aux textes d'un niveau stylistique plus élevé : dans les poèmes troubadouresques aussi, il est très difficile de trouver des fautes certaines. Je signale, en tout cas, que quelqu'un s'est chargé de contrôler tous les arbres conçus par Bédier, et a conclu que celui qui demeure valable est le premier⁵.

Mais Bédier avait fait une autre remarque, qui nous laisse franchement rêveurs. En recueillant tous les arbres proposés par les romanistes pour leurs textes, Bédier remarquait que sur 110 arbres, 105 sont bipartis. Arrigo Castellani, qui a essayé de refaire les comptes de Bédier⁶, est arrivé à établir une liste de 120 arbres, dont 82-83% à deux branches. Bref, cette loi admet peu d'exceptions, même si dans le calcul de Castellani leur nombre est un peu plus élevé. Mais pourquoi tous les textes conservés auraient-ils donné lieu à deux copies, d'où descendraient toutes les autres? Pourquoi faut-il que les arbres à trois, quatre ou plusieurs branches soient si rares? Les philologues, à partir de Sebastiano Timpanaro, ont réfléchi là dessus, et ont trouvé bon nombre de raisons qui justifient cette prédominance des arbres bipartis. Leurs arguments sont pour la plupart statistiques. Et chacun peut les contrôler. Ensuite, Castellani a expliqué la prédominance des arbres à deux branches par les modalités du processus de copie, qui doit être exécuté le plus vite possible. D'où sa notion ingénieuse de « production maximum ».

Le résultat de la recherche de Bédier était en tout cas intéressant, même s'il a peut-être exagéré son importance. Mais il faut souligner que, pour Bédier, la prédominance des arbres bipartis dépendait, plutôt que de la réalité de la tradition manuscrite, de raisons qui nous échappent, mais qui pourtant jettent l'ombre du doute sur

⁵ Whitehead / Pickford (1973). Pour une nouvelle proposition stématique (toujours bífide) au sujet du *lai* de Jean Renart voir Trovato (2013).

⁶ Castellani (1957).

le mécanisme d'élaboration des arbres mêmes. Parmi les explications avancées par Bédier, en effet, il y a le soupçon que l'éditeur vise spontanément un arbre biparti parce que celui-ci lui laisse cette liberté de choix que l'arbre à trois branches anéantirait. Le travail de collation des manuscrits étant de nature essentiellement binaire (leçon considérée comme correcte *vs* leçon considérée comme fautive), on peut penser que le système de reconstruction de la recension amène lui aussi inévitablement à une solution binaire.

Pour comprendre le succès de sa thèse, sans aucun doute révolutionnaire, il faut tenir compte du fait que Bédier était désormais le chef de file de la romanistique française, qu'il était célèbre aussi pour sa traduction-remaniement très élégante du *Roman de Tristan* (1900), et presque un héros national pour sa contribution à la documentation des crimes allemands pendant la première guerre mondiale⁷. Incontestablement, dans l'histoire que je suis en train d'évoquer, l'antagonisme franco-allemand a joué un rôle considérable. Rabaisser la gloire philologique des Allemands, voilà ce qui pouvait représenter une belle satisfaction. Et Bédier était doué d'une capacité rhétorique exceptionnelle, quitte à frôler le raisonnement captieux.

Certes, l'opération de Bédier est d'une envergure considérable. Si son article consacré au *Lai de l'ombre* date de 1928, l'année précédente il avait publié les *Commentaires* à son édition de la *Chanson de Roland*, dont la première parution date de 1922⁸. Dans ses *Commentaires*, Bédier expose les critères qu'il a suivis pour l'édition de la *Chanson*, des critères qui à première vue apparaissent déjà, pour ainsi dire, bédieriens.

Le texte de base (le « bon manuscrit ») est, comme dans la quasi-totalité des éditions de la *Chanson de Roland*, celui d'Oxford, qu'il faut suivre presque toujours parce que, selon Bédier, « le texte d'Oxford a autant d'autorité à lui seul que tous les autres textes réunis » (p. 93). Et pourtant, on pourrait affirmer – au grand scandale des bédieriens – que ces *Commentaires* laissent apercevoir un établissement du texte de type lachmannien, pour les raisons suivantes. Premièrement : Bédier souligne, avec une grande évidence, sa fidélité à l'arbre de Müller, établi grâce à des critères lachmanniens. Deuxièmement : celui de Müller est l'un des ces arbres bipartis que Bédier condamne dans son article de 1928. Lorsqu'il s'agit de la *Chanson de Roland*, Bédier n'insiste jamais sur la fiabilité douteuse de ce type d'arbre. Troisièmement : comme on vient de le dire, Bédier affirme, à raison, que « le texte d'Oxford a autant d'autorité à lui seul que tous les autres textes réunis ». Mais dans l'enthousiasme de son plaidoyer, il ne s'aperçoit pas que cette affirmation signifie également, en bonne logique : « [le témoignage de] tous les autres textes réunis a autant d'autorité que celui d'Oxford ». Quoi qu'il en soit, dans la *Section V* (p. 178-195), Bédier examine tous les cas où les versions du groupe b⁹ « peuvent être de bon secours », et sa réponse est souvent favo-

⁷ Voir sur ce sujet Corbellari (1997, chap. IX).

⁸ Bédier (1927).

⁹ Que Bédier dénomme a.

nable. De plus, il reviendra encore sur son texte, avec de nouvelles retouches à O¹⁰. Que Bédier souhaite, de toute façon, réduire au maximum le recours aux manuscrits b est un choix personnel, que tout lachmannien trouverait acceptable. Remarquons, en outre, que l'étude d'ensemble de la recension de la *Chanson de Roland* est extrêmement approfondie, ce qui serait même trop dans une édition de l'école bédérienne.

Que l'on puisse vouer au mépris public la fréquence des arbres avec une conformation bipartite et, en même temps, baser l'édition de la plus grande *chanson de geste* française du Moyen Âge sur un arbre à deux branches, est une question psychologique, ou peut-être rhétorique. Chaque fois, les lecteurs sont emportés par la feu de la démonstration de Bédier : certes, dans ses *Commentaires* il écrit : « on ne doit pas s'étonner, ni s'inquiéter, si l'on aboutit, dans le cas de la *Chanson de Roland*, à un classement à deux branches. En faveur de celui-ci, j'ai pu invoquer, j'invoquerai encore, comme on a vu et comme on verra, nombre de faits » (*Com.* 134). Il n'y avait donc pas de raison de s'inquiéter face à la majorité des arbres bipartis.

Je parlais de logique. Eh bien, toujours dans cette année fatale 1928, et toujours dans son article consacré au *Lai de l'ombre*, Bédier à l'occasion de discuter une tentative révolutionnaire de bâtir des arbres, celle de dom Henri Quentin; une tentative aujourd'hui oubliée, sauf sa ressemblance avec quelques expériences modernes de construction mécanique des arbres. À la fin de sa revue des suggestions critiques de dom Quentin, Bédier, en les repoussant, accuse le savant bénédictin d'avoir renoncé à « tracer les lignes de faite du schéma, celles qui doivent relier les manuscrits réels à l'archétype », de ne s'être pas soucié de « peser les variantes » pour bâtir des groupes et des sous-groupes. C'est un résumé parfait du travail d'un éditeur lachmannien, même si, dans ce cas aussi, Bédier refuse la notion lachmannienne de « faute ». Et alors ? Alors, il faudrait penser que Bédier ne s'est jamais débarrassé tout à fait des principes de Lachmann.

Quoi qu'il en soit, après 1928 s'est affirmée une sorte de vulgate bédérienne, selon laquelle il ne serait pas indispensable d'étudier systématiquement la tradition manuscrite d'un texte : il suffit de choisir « le bon manuscrit », ou « le meilleur manuscrit », et l'éditer de la façon la plus fidèle possible. Les éditions se sont succédé, toujours selon le même schéma, toutes pareilles, même si les meilleurs disciples de Bédier condamnaient le choix de la facilité et de la paresse : en effet, des philologues tels qu'Albert Henry (Adenet le Roi, 1951-1971) ou Jean Rychner (Marie de France, 1966) ont basé leurs éditions sur des recherches approfondies et systématiques, en discutant et en résolvant au fur et à mesure des problèmes que les lachmanniens résolvent préliminairement dans la discussion de l'arbre.

La notion de « bon manuscrit » est l'une des plus vagues et contradictoires dans le domaine de l'édition des textes. Le grand Maurice Delbouille dénonçait le fait qu'avant comme après Bédier « le “ meilleur manuscrit ” a continué d'imposer ses

¹⁰ Bédier (1937-1938).

leçons les moins heureuses »¹¹. Si l'on considère dans l'ensemble les usages des divers philologues, le « bon manuscrit » est le plus complet, ou celui qui offre la leçon la plus correcte ; ou encore, il s'agit du manuscrit avec les traits linguistiques les plus proches de ceux de la région d'appartenance de l'auteur. En outre, le « bon manuscrit » est, préférablement, l'un des plus anciens, parce que de cette façon il a moins subi la détérioration du temps et le cumul des altérations dues au travail de copie.

Je me permets de remplacer ces définitions, qui souvent se contredisent mutuellement, par une seule : le « bon manuscrit » est un manuscrit qui nous transmet le texte dans les *meilleures conditions* : moins de fautes et moins d'altérations. Mais comment le reconnaître ? Le seul moyen est d'approfondir la recension, ce qui nous ramène en pleine méthode lachmanienne.

J'ai eu l'occasion de réfléchir sur cette question lors de la découverte du manuscrit T du *Bestiaire d'Amours*¹². Un manuscrit que Xenia Muratova, qui l'a découvert, aurait tendance à définir « bon », car il est complet, apparemment correct et relativement ancien. Malheureusement, l'examen de la tradition, que je peux faire aisément comme auteur de l'édition critique de ce texte¹³, dont les apparats offrent toutes les données utiles pour une telle appréciation, m'a révélé immédiatement que T remonte à une mauvaise copie, dont il a hérité forcément les fautes et les banalisations¹⁴. Malgré son aspect apparemment correct, il ne s'agit donc pas d'un « bon manuscrit ». Lorsque Bédier célèbre le manuscrit O de la *Chanson de Roland* comme le « bon manuscrit », il a beau jeu, parce que la supériorité de O est universellement reconnue. Mais la qualité de T n'a rien à voir à celle de O. J'ajouterai que le manuscrit T du *Bestiaire* s'insère dans l'arbre d'une façon tellement précise que l'on peut prévoir presque constamment quelles seront ses leçons.

En tout cas, le succès de la théorie de Bédier a eu également des conséquences sur la géographie philologique. Car cette théorie a été accueillie presque sans exception en France, tandis que dans les autres pays elle a été moins suivie. Et il faut reconnaître que, souvent, la multiplication d'éditions de textes français basées sur un « bon manuscrit », sans une étude systématique de la tradition, n'assure pas un véritable progrès de nos connaissances. Un grand philologue bédien, Jacques Monfrin, le reconnaît, lorsqu'il déclare : « on ne peut s'étonner du discrédit où est tombée cette méthode qui a justifié tous les abandons, mais il est étrangement simplificateur d'en attribuer, surtout sous les formes caricaturales qu'elle a parfois prises, la paternité à Joseph Bédier »¹⁵.

Je passe, sans pour autant l'absoudre, sur le niveau souvent avilissant qu'ont atteint les débats entre bédiériens et lachmanniens. Mais je souligne qu'il est utile,

¹¹ Delbouille (1976).

¹² Muratova (2005).

¹³ Segre (1957).

¹⁴ Pour la démonstration, voir Segre (2010).

¹⁵ Monfrin (1984-1986, 354).

et fécond, de prendre en compte, d'un côté, les changements d'avis des bédieriens, de l'autre, plus généralement, la conscience, présente depuis longtemps chez les néolachmanniens les plus avertis, que l'édition critique n'est pas une idole, mais l'un des traitements du texte suggérés par la pratique philologique.

Car il y a des éditions qui excluent à juste titre toute retouche trop lourde, comme les éditions diplomatiques et interprétatives ; il y a des approches dialectologiques des textes, qui seraient faussées par les conjectures. Et il y a des éditions synoptiques du même texte selon différents manuscrits. Ni les répertoires des rimes d'un auteur ni les concordances n'admettent aucune intervention du philologue.

L'édition critique n'est fondamentale que quand on cherche à récupérer quelque chose qui puisse représenter vraisemblablement et au niveau le plus élevé de probabilité l'aspect original du texte¹⁶. J'ajoute que les hypertextes, s'ils sont réalisés avec intelligence, nous permettent de parcourir l'ensemble des variantes beaucoup plus aisément qu'une édition critique, et de multiplier et comparer les différentes possibilités, en les mettant à l'épreuve dans des hypothèses de reconstitution.

L'expérience de la philologie romane italienne montre, mieux que d'autres expériences nationales, comment l'ecdotique a abouti à une heureuse intégration des principes les plus valables du bédierisme et du lachmannisme : pour cette raison (et je m'excuse de cette simplification forcée), presque tous les exemples que je vais citer seront tirés de cette expérience.

Après la seconde guerre mondiale, il y a eu en Italie une reprise très particulière de la méthode lachmanienne, qui par ailleurs avait déjà obtenu des résultats excellents avec Michele Barbi et Pio Rajna. Le protagoniste de cette reprise a été Gianfranco Contini. Contini a donné naissance à une grande école, qui s'affirme d'abord avec l'édition (1960) des poètes italiens du XIII^e siècle¹⁷, établie sous sa direction par un groupe de jeunes et même très jeunes chercheurs, parmi lesquels Franca Ageno, D'Arco S. Avalle, Romano Broggin, Domenico De Robertis, Ezio Raimondi et Cesare Segre. Quelques années plus tard, Contini crée, avec Alfredo Schiaffini, une collection de textes critiques, les « Documenti di Filologia », qui ont renouvelé le panorama philologique et littéraire de la Romania.

Il faut remarquer que tous les auteurs choisis par les chercheurs italiens appartiennent en fait au monde roman: France du Nord et France du Sud, Italie, Espagne. Cela élargissait le domaine de l'expérience de l'édition des textes, et en même temps renforçait l'aspect comparatiste de la tradition italienne dans le domaine de la philologie romane, qui dans les autres pays ne travaille le plus souvent que sur des textes dans la langue nationale correspondante.

¹⁶ Voir, par exemple, le débat constructif entre Roncaglia / McMillan / Segre / Lecoy (1978).

¹⁷ Contini (1960).

Le jeune Contini, après avoir étudié la méthode lachmannienne à l'école de Santorre Debenedetti et avoir réfléchi sur l'ouvrage capital de Giorgio Pasquali¹⁸, avait fait un séjour de deux ans à Paris (1935-1936), à l'école de Bédier, auquel il a consacré un superbe portrait¹⁹. Bédier a été constamment un point de repère pour Contini, qui n'a pas cessé pour autant d'être lachmannien ou plutôt néo-lachmannien²⁰. En effet, il a aisément accueilli plusieurs points décisifs de la philologie de Bédier : éviter les reconstitutions en mosaïque des textes, et même suivre, pour l'aspect linguistique, un seul texte (qui n'est pas toujours le « bon manuscrit ») ; éviter des restaurations traumatiques des textes ; éviter d'être influencé par les grammaires historiques, mais partir de données documentaires et localisables.

Contini soumet à une analyse méthodologique, je dirais presque scientifique, le travail de Bédier. Son principe fondamental est la conception de l'édition en tant qu'« hypothèse de travail » : une formule où ceux qui ne sont pas experts dans la méthodologie scientifique voient un aveu de faiblesse (hypothèse), en oubliant que, dans le domaine scientifique, cette expression indique l'un des rapprochements progressifs de la vérité. « Hypothèse de travail » signifie : la meilleure solution que n'importe quelle science a pu fournir pour un problème. Si nous faisons confiance au philologue, nous pouvons considérer comme valable son travail, jusqu'au moment où un autre philologue, sur la base de données plus nombreuses et plus sûres, propose une solution plus satisfaisante. C'est ainsi que la science avance dans son chemin. Nous sommes là aux antipodes du texte critique conçu comme une vérité ultime.

L'un des résultats principaux de la réflexion de Contini est la notion de « diffraction des leçons », une forme de la *lectio difficilior* pourvue d'une considérable rentabilité opérationnelle. Mais l'opération la plus importante de Contini, du point de vue symbolique aussi, est son retour au problème des arbres bipartis, en prenant en considération justement l'élaboration la plus ancienne du domaine lachmannien : l'arbre du *Saint Alexis* proposé par Gaston Paris. En effet, selon Contini, l'arbre correct du *Saint Alexis* serait un arbre à trois branches, et non à deux branches, comme le croyait Paris²¹. Le poème hagiographique le plus illustre aurait échappé à la force mystérieuse des arbres bipartis, tandis que Gaston Paris serait l'une des nombreuses victimes du binarisme²².

¹⁸ Pasquali (1934).

¹⁹ Contini (1939).

²⁰ Tout en prônant une reprise de la méthode lachmannienne, Contini antéposait le préfixe néo (néo-lachmannienne) pour souligner d'emblée que sa propre méthode était bien différente de celle utilisée jusqu'à 1928. Dans des pages malheureusement très difficiles, Contini a discuté chaque point de son opération d'édition des textes, voir Contini (1986).

²¹ Contini (1970).

²² Il est significatif qu'après la prise de position de Bédier nombre de philologues soient allés à la recherche d'arbres tripartis. Contini (1978) avait déjà signalé la probabilité d'un arbre triparti, mais voir aussi Crespo (2012).

Le nombre des expériences des néo-lachmanniens dans l'édition de textes est tel que le problème des arbres bipartis se situe dans des lointains brumeux. Les arbres à deux ou trois branches sont les cas les plus simples d'une forêt où l'on peut trouver des descendants de deux archétypes distincts, des rédactions doubles et plurielles, et où on se voit obligé d'avoir recours à des représentations horizontales de la tradition, ou à des représentations stéréoscopiques, ou encore à la notion linguistique de diasystème²³.

Il faut accepter également que dans plusieurs cas, la méthode lachmannienne n'est pas utilisable. Il s'agit d'une phénoménologie très vaste, dont il faut se rendre compte et comprendre les raisons. Quant à moi, si la méthode néo-lachmannienne a été fondamentale pour mon édition de la *Chanson de Roland* et pour le *Bestiaire d'Amour*, ainsi que pour plusieurs textes italiens, je la trouve inutilisable pour les *lais* de Marie de France, comme Jean Rychner et moi-même l'avons constaté lors d'un débat amical²⁴.

Je ne chercherai pas à décrire la situation d'ensemble des études actuelles d'ecdotique, mais je peux affirmer, en synthèse, qu'à mon avis, la guerre entre lachmanniens et bédériens est désormais un vestige archéologique. On a compris qu'entre les deux méthodes il n'y a pas d'opposition, mais une différence de finalité.

La reconstitution de l'archétype n'est pas le seul but du travail d'édition ; les éditions interprétatives sont souvent indispensables, et peuvent représenter la base de nombreuses recherches. Les recherches sur la tradition des textes ne peuvent pas toujours utiliser la méthode des fautes, qui pourtant reste fondamentale dans les cas où elle est applicable.

Il faudrait en tout cas une certaine prudence lorsqu'on définit comme éditions critiques celles qui n'utilisent pas tous les procédés de la critique textuelle. Certes, pour moi, souvent désigné comme un dangereux lachmannien, il a été surprenant de constater que, dans le manuel pour l'édition des textes médiévaux publié en 2002 par deux illustres chartistes tels que Pascale Bourgain et Françoise Vieliard, on affirme que la méthode lachmannienne est

²³ L'existence de deux archétypes du *Bestiaire d'Amours* est démontrée dans mon édition, Segre (1957), qui fournit également une méthode pour la recherche des contaminations ; Chiarini (1964) a résolu le problème de la prétendue double rédaction du *Libro de Buen Amor* ; Contini (1984) cherche à démontrer la paternité dantesque du *Fiore* grâce à la notion de « mémoire interne » ; Avallé (1965) démontre l'existence d'une koinè poitevine en partant du *Sponsus* provençal ; les chansons d'Arnaut Daniel, selon Perugi (1978), révèlent, à travers la « théorie du hiatus », certains traits conservateurs des textes des *canzonieri*. Pour une réflexion sur le poids stématique des variantes adiaophores voir Squillaciotti (1999). Mon édition italienne de la *Chanson de Roland* est une tentative d'édition stéréoscopique, car elle aide le lecteur, grâce à des caractères spéciaux, à suivre le texte de α ou celui, plus récent, de β . Malheureusement, je n'ai pas réussi à me faire comprendre par mes collègues, et dans l'édition française, Segre (1989), j'ai dû me rabattre sur une présentation plus traditionnelle. J'ai proposé aussi une nouvelle notion d'apparat, en considérant la stratification dialectale des textes comme un cas de « diasystème ».

²⁴ Segre (1959) ; Rychner (1966) ; Segre (1968). Ce qui est ressorti de cet échange d'opinions, c'est que l'arbre que je proposais n'était pas plus sûr que celui proposé par Jean Rychner. Les deux étaient équivalents.

« un merveilleux outil de logique, très rigoureux, extrêmement satisfaisant pour l'esprit, dans des conditions généalogiques normales. On ne peut échapper à ces principes, qui découlent du fait même de la reproduction par copie. Les éléments essentiels de cette méthode restent en place jusqu'à nos jours, avec les différentes étapes de la recension des témoins, de l'examen de la tradition et de l'établissement du texte »²⁵.

Un autre signe intéressant est le retour à l'étude des grandes éditions lachmanniennes des Langlois et des Foerster : non seulement pour leur rendre le prestige qu'ils méritent, mais pour approfondir le précieux travail déjà accompli, qu'il ne faut pas gaspiller mais poursuivre.

Nous savons bien que le domaine de la critique textuelle n'est jamais inerte. L'un des plus grands lachmanniens italiens, D'Arco Silvio Avalle, auteur d'une édition exemplaire de Peire Vidal²⁶, a produit un changement considérable. Avalle a préparé des *Concordanze della lingua poetica italiana delle origini* (CLPIO), dont le premier volume, qui pour l'instant est le seul publié, a paru en 1992 chez la maison d'édition Ricciardi. Ce volume monumental recueille tous les textes poétiques italiens attestés par des manuscrits d'avant l'an 1300. Bien entendu, ces poèmes ne sont pas transcrits d'après leurs éditions critiques, remplies de leçons conjecturales qui représenteraient des entrées étrangères aux textes, mais d'après les manuscrits, en respectant fidèlement leurs particularités graphiques.

D'autre part, dans son étude la plus ancienne, consacrée à la *Letteratura medievale in lingua d'oc nella sua tradizione manoscritta* (1961)²⁷, Avalle renouait inévitablement avec les vieilles recherches de Gröber. En étudiant les *canzonieri* provençaux, Gröber les avait classés, et avait étudié leur généalogie, non pas sur la base des fautes, trop rares pour offrir au philologue un fondement solide, mais par rapport à la succession des poèmes, qui peut être considérée comme un indice du rapport entre les manuscrits (*Kritik der Sammlungen*). C'était un type inhabituel de recherche, qui a été adopté essentiellement par ceux qui ont étudié les grandes anthologies lyriques médiévales, provençales, françaises, italiennes. Tout d'abord, Napoleone Caix. En outre, Gröber cherchait à reconstruire les phases fondamentales de la diffusion des poèmes : feuilles volantes, recueils d'un seul auteur, recueils anthologiques, etc. Son but était de retracer les différentes phases de la tradition de chaque texte.

Après sa récupération par Avalle, la recherche fondée sur les critères de Gröber a eu une impulsion nouvelle, se répandant très rapidement. Aujourd'hui, par exemple, l'attention est centrée sur les recueils d'auteur ou leurs traces²⁸, sur la présence éventuelle de variantes d'auteur, sur la reconstitution des grands collecteurs des recueils, sur la définition des différentes couches linguistiques, sur la reconstruction de l'ordre originel éventuel des poèmes²⁹. Il s'agit d'une série d'opérations génétiques, qui, au

²⁵ Bourgain / Vieillard (2002, 16).

²⁶ Avalle (1960).

²⁷ Nouvelle édition augmentée : Avalle (1993).

²⁸ Voir la synthèse de Meliga (2006).

²⁹ Voir, par exemple, les recherches rassemblées par Leonardi (2001), avec des contributions

lieu de viser le modèle, nous amènent à reconstruire la micro-histoire des textes. Il faut tenir compte du fait que, surtout dans les usages de la poésie italienne des origines, la distance chronologique entre la rédaction et la transcription n'est souvent que de quelques années, au plus de quelques décennies : notre regard se pose presque sur des copies de manuscrits autographes. Dans ce cas, on est tenté – et Avalle lui-même l'a fait – d'opposer deux méthodes, de mettre face à face deux vérités, dans le sillage d'Averroès³⁰. Toutefois, je voudrais le préciser, il ne s'agit pas de comparer deux entités ontologiques (dans notre cas, la « vérité » du texte et la « vérité » du témoin)³¹, mais de choisir entre deux approches différentes, dont chacune apporte un enrichissement de connaissance.

Università di Pavia / Accademia Nazionale dei Lincei

Cesare SEGRE †

Références bibliographiques

- Antonelli, Roberto, 1985. « Interpretazione e critica del testo », in: Asor Rosa, Alberto (ed.), *Letteratura italiana*, IV. *L'interpretazione*, Turin, Einaudi, 141-243.
- Avalle, D'Arco Silvio, 1960. Peire Vidal, *Poesie*, édition établie et commentée par D'A.S.A., Milan-Naples, R. Ricciardi, 1960, 2 vol. (Documenti di Filologia).
- Avalle, D'Arco Silvio, 1965. *Sponsus. Drama delle vergini prudenti e delle vergini stolte*. Testo letterario a cura di D'A.S.A., testo musicale a cura di Raffaello Monterosso, Milan-Naples, R. Ricciardi.
- Avalle, D'Arco Silvio, 1972. « La critica testuale », in: Jauss, Hans Robert / Köhler Erich (ed.), *Grundriss der romanischen Literaturen des Mittelalters*, I, *Généralités*, Heidelberg, Winter, 538-558 ; pubblicato in seguito con il titolo « Fenomenologia ecdotica del Medioevo romanzo », in : Id., *La doppia verità. Fenomenologia, ecdotica e lingua letteraria del Medioevo romanzo*, Florence, Edizioni del Galluzzo, 2002, 125-153.
- Avalle, D'Arco Silvio, 1985. « I canzonieri : definizione di genere e problemi di edizione », in: *La critica del testo, Problemi di metodo ed esperienze di lavoro*, Atti del Convegno di Lecce (22-26 ottobre, 1984), Rome, Salerno, 363-382, Id. *La doppia verità. Fenomenologia, ecdotica e lingua letteraria del Medioevo romanzo*, Florence, Edizioni del Galluzzo, 2002, 155-173.
- Avalle, D'Arco Silvio, 1993. *I manoscritti della letteratura in lingua d'oc*. Nuova edizione a cura di L. Leonardi, Turin, Einaudi, 1993.
- Bédier, Joseph, 1900. *Le Roman de Tristan reconstitué d'après les poèmes français du XII^e siècle*, Paris, Piazza.

de R. Antonelli, A. Petrucci, M. Palma, P. Larson, C. Bologna, S. Zamponi, G. Frosini, G. Savino, T. De Robertis, Valentina Pollidori, M. L. Meneghetti.

³⁰ Avalle (1985), où l'auteur parle d'une double vérité : « la verità dei protagonisti, che è quella cui aspirano le edizioni critiche di singoli autori, e la verità dei testimoni per il momento affidata [...] alle edizioni diplomatiche o semi-diplomatiche » (p. 275). Avalle (1972) avait déjà exprimé un certain penchant pour Bédier.

³¹ La ligne de travail qui privilégie l'analyse de la matérialité des témoins manuscrits ne concerne pas que les chercheurs italiens ; mais il faut attribuer aux Italiens une priorité, à partir même de l'étiquette de 'filologia materiale', avec laquelle on désigne d'habitude cette approche, et qui a été formulée par Antonelli (1985, 207) : voir, par exemple, Tyssens (1991) et Cadioli / Meneghetti (2008).

- Bédier, Joseph, 1927. *La Chanson de Roland (Commentaires)*, Paris, Piazza.
- Bédier, Joseph, 1928. «La tradition manuscrite du *Lai de l'Ombre* : réflexions sur l'art d'éditer les anciens textes», R 54, 161-196, 321-356.
- Bédier, Joseph, 1937-1938. «De l'édition princeps de la *Chanson de Roland* aux éditions les plus récentes. Nouvelles remarques sur l'art d'établir les anciens textes», R 63, 433-469; R 64, 145-244, 489-521.
- Bourgain, Pascale / Viellard, Françoise, 2002. *Conseils pour l'édition des textes médiévaux*, fascicule III : *Textes Littéraires*, Paris, École Nationale des Chartes, 2002.
- Cadioli, Alberto / Meneghetti, Maria Luisa, 2008. *La materialità nella filologia*, a cura di A.C. e M.L.M. = «Moderna» 10.
- Castellani, Arrigo, 1957. *Bédier avait-il raison? La méthode de Lachmann dans les textes du Moyen Âge*, Fribourg, Editions universitaires.
- Chiarini, Giorgio, 1964. Juan Ruiz, Arcipreste de Hita, *Libro de buen amor*. Edizione critica a cura di G.C., Milan-Naples, R. Ricciardi.
- Contini, Gianfranco, 1939. «Ricordo di Joseph Bédier», *Letteratura* 3, 145-152.
- Contini, Gianfranco, 1960. *Poeti del Duecento*, a cura di G.C., Milan-Naples, Ricciardi.
- Contini, Gianfranco, 1970. «La Vita francese di Sant'Alessio e l'arte di pubblicare i testi antichi» (1970), Id., *Frammenti di filologia romanza. Scritti di ecdotica e linguistica (1932-1989)*, textes recueillis par Giancarlo Breschi, Florence, Edizioni del Galluzzo, 2007, II, 957-986.
- Contini, Gianfranco, 1978. «Fragments inconnus d'un ancien chansonnier français à Einsiedeln», in Güntert, Georges / Jung, Marc-René / Ringger, Kurt (ed.), *Orbis Mediaevalis. Mélanges de langue et de littérature médiévales offerts à Reto Raduolf Bezzola à l'occasion de son quatre-vingtième anniversaire*, Bern, Francke, 29-59, Id., *Frammenti di filologia romanza*, Florence, Edizioni del Galluzzo, 2007, II, p. 1025-60.
- Contini, Gianfranco, 1984. *Il Fiore e il Detto d'amore attribuibili a Dante Alighieri*, a cura di G.C., Milan, A. Mondadori.
- Contini, Gianfranco, 1986. *Breviario di ecdotica*, Milan-Naples, R. Ricciardi.
- Corbellari, Alain, 1997. *Joseph Bédier, écrivain et philologue*, Genève, Droz.
- Crespo, Roberto, 2012. *Il primo episodio del Couronnement Louis*, Modène, Mucchi.
- Delbouille, Maurice, 1976. «La philologie médiévale et la critique textuelle», in Boudreault, Marcel / Möhren, Frankwalt (ed.), *Actes du XIII^e Congrès International de Linguistique et Philologie Romanes tenu à l'Université Laval (Québec, Canada) du 29 août au 5 septembre 1971*, Québec, Université Laval, 57-73.
- Fiesoli, Giovanni, 2000. *La genesi del lachmannismo*, Florence, Sismel.
- Gröber, Gustav, 1869. *Die handschriftlichen Gestaltungen der chanson de geste «Fierabras» und ihre Vorstufen*, Leipzig.
- Gröber, Gustav, 1877. *Die Liedersammlungen der Troubadours*, RSt 2, 337-360.
- Leonardi, Lino, 2001. *Canzonieri della lirica italiana delle origini. Saggi critici*, Florence, Edizioni del Galluzzo.
- Leonardi, Lino, 2009. «L'art d'éditer les anciens textes (1872-1928). Les stratégies d'un débat aux origines de la philologie romane», R 127, 273-302.
- Meliga, Walter, 2006. «Le raccolte d'autore nella tradizione trobadorica», in: Lo Monaco, Francesco / Rossi, Luca Carlo / Saffai, Niccolò, «*Liber*», «*fragmenta*», «*libellus*» prima e dopo Petrarca. In ricordo di D'Arco Silvio Avalle. Seminario Internazionale di Studi, Bergamo 23-25 ottobre 2003, Florence, SISMEL – Edizioni del Galluzzo, 81-91.

- Monfrin, Jaques, 1984-1986. «Problèmes d'éditions de textes», in Actes du XVII^e Congrès International de Linguistique et Philologie Romanes (Aix-en-Provence, 29 août-3 septembre 1983), Aix-en-Provence, Université de Provence, vol. IX, 351-364.
- Muratova, Xenia, 2005. «Un nouveau manuscrit du Bestiaire d'Amours de Richard de Fournival », in: van den Abeele / Baudouin (ed.), *Bestiaires médiévaux. Nouvelles perspectives sur les manuscrits et les traditions textuelles. Communications présentées au XV^e Colloque de la Société Internationale Renardienne*, Louvain-la-Neuve, Institut d'études médiévales de l'Université Catholique de Louvain, 265-281.
- Paris, Gaston / Pannier, Léopold, 1872. *La Vie de Saint Alexis. Poème du XI^e siècle et renouvellements du XII^e, XIII^e et XIV^e siècles*, Paris, Frank.
- Pasquali, Giorgio, 1934. *Storia della tradizione e critica del testo*, Florence, Le Monnier.
- Perugi, Maurizio, 1978. *Le canzoni di Arnaut Daniel. Edizione critica*, a cura di M.P., Milan-Naples, R. Ricciardi, 2 voll.
- Roncaglia, Aurelio / McMillan, Duncan / Segre, Cesare / Lecoy, Félix, 1978. «Premesse ideologiche della critica testuale», in: Vàrvaro, Alberto (ed.), *XIV Congresso internazionale di Linguistica e Filologia Romanza. Atti. Napoli, 15-20 aprile 1974*, Naples-Amsterdam, Macchiaroli-Benamins, I, 479-514.
- Rychner, Jean, 1966. *Les lais de Marie de France*, publiés par J.R., Paris, Champion.
- Segre, Cesare, 1957. *Li bestiaires d'Amours de Maistre Richart de Fornival e Li Response du Bestiaire*, a cura di C.S., Milan-Naples, R. Ricciardi.
- Segre, Cesare, 1959. «Per l'edizione critica dei lai di Maria di Francia», *CN* 19, 215-237.
- Segre, Cesare, 1968. Compte-rendu de Rychner, Jean, 1966, *CCM* 11, 243-246.
- Segre, Cesare, 1971. *La Chanson de Roland*. Edizione critica a cura di C.S., Milan-Naples, R. Ricciardi.
- Segre, Cesare, 1989. *La Chanson de Roland*. Édition critique par C.S.. Nouvelle édition revue traduite de l'italien par Madeleine Tyssens, Genève, Droz, 2 voll.
- Segre, Cesare, 2010. «Due interventi sul *Bestiaire d'Amours* di Richard de Fournival», *MedRom* 34, 396-405.
- Squillacioti, Paolo, 1999. *Le poesie di Folchetto di Marsiglia*. Edizione critica a cura di P.S., Ospedaletto, Pacini.
- Timpanaro, Sebastiano, 1963. *La genesi del metodo del Lachmann*, Florence, Le Monnier (édition augmentée : Padoue, Liviana, 1981).
- Trovato, Paolo, 2013. «La tradizione manoscritta del *Lai de l'ombre*. Riflessioni sulla tecnica d'edizione protonovecentesca», *R* 131, 338-380.
- Tyssens, Madeleine, 1991. *Lyrique romane médiévale. La tradition des chansonniers*, Actes du Colloque de Liège, 1989, édités par M.T., Liège, Université de Liège.
- Whitehead, Frederick / Pickford, Cedric E., 1973. «The Introduction of the *Lai de l'ombre*. Sixty Years Later», *R* 94, 145-156.

Du chemin parcouru entre Coquebert de Montbret père et fils et la plus récente dialectométrie : une reconsidération critique

1. De Coquebert de Montbret père et fils à Jules Gilliéron : deux points de repère de la genèse de l'analyse empirique de l'espace de la France

Depuis un certain temps je défends l'idée que l'ALF et l'œuvre entière de Jules Gilliéron (1854-1926) dépendent d'un substrat intellectuel « typiquement français » et qu'ils doivent donc être considérés dans une perspective historique particulière. La même remarque vaut pour la genèse de la dialectométrie qui, de par son souci enveloppant de généralisation spatiale, a navigué dans le même sillage¹.

Rappelons brièvement les jalons les plus importants de ces implications. L'ALF est caractérisé, entre autres, par trois propriétés qui ont grandement favorisé sa réussite fulgurante :

- un maillage d'enquête « économique » et « rationnel », basé sur le réseau départemental de la France et partageant avec lui, de ce fait, tous les avantages y ayant trait, comme l'équidistance entre les mailles de la grille et le nombre « opérationnel » (= limité) des points d'enquête².
- un questionnaire d'accessibilité « pan-galloromane », c'est-à-dire contenant des items universellement compréhensibles et partant – métrologiquement parlant – pourvus d'un haut degré de fiabilité et de validité.
- les consignes très strictes données à l'enquêteur Edmond Edmont qui consistaient avant tout à n'enregistrer que la première réponse des sujets interviewés et d'éviter toute sorte de mensuration.

Avec la curiosité de l'historien, l'on peut se demander deux choses :

Tout ceci a-t-il été inventé de toutes pièces par Gilliéron lui-même,

ou l'a-t-il conçu sous le coup d'une tradition spirituelle ou intellectuelle antérieure dans laquelle il avait baigné dès 1876, l'année de son arrivée en France?

¹ Voir aussi nos contributions de 2002(b) et 2006.

² Rappelons que la limitation programmatique des travaux de sondage à quelque 600 points d'enquête par rapport au total des 36 000 communes de la France équivaut, en dernière analyse, au tirage d'un échantillon aléatoire et représentatif.

Une autre certitude est son indépendance intellectuelle complète de l'entreprise géolinguistique allemande « Deutscher Sprachatlas » de Georg Wenker (1852-1911), pourtant antérieure à la sienne de quelque 20 ans³.

Mais d'où avait-il pris l'idée heureuse de construire le réseau d'enquête de l'ALF suivant le compartimentage départemental de la France? Toujours est-il que cette décision s'est révélée, après coup, comme étant vraiment *heureuse*, vu son excellente maniabilité, qui se détache considérablement de la contrepartie allemande lourde – et non pas riche, en fin de compte – de plus 50 000 points d'enquête.

Or, je pense que les idées linguistiques de Gilliéron se nourrissaient – au cours de la période de gestation de l'ALF – d'un substrat intellectuel préexistant dont on parle peu en matière de linguistique ou de géolinguistique.

L'esprit de la départementalisation de la France – qui, elle, était une des premières mesures prises par la Révolution française, devenue par la suite très efficace – a favorisé, entre autres, une nouvelle considération de l'espace de la Nation, tout en la définissant non pas comme un continuum amorphe, mais bien plutôt comme la somme d'un nombre relativement petit de vecteurs, entre eux égaux et équidistants⁴. En bref, c'était un acte de mensuration visant à une discrétisation rationnelle de l'espace de la France, lourd de conséquences. Précisons que les préambules intellectuels de cette réorganisation de l'espace de la France se situaient déjà aux XVII^e et XVIII^e siècles.

Mais, ce qui compte le plus en l'affaire, c'est la dissémination à tous azimuts de l'usage de la nouvelle grille départementale de la France pour toutes sortes de relevés et d'études empiriques qui ont été réalisés au cours du XIX^e siècle. Il en est né ce qu'on appelle la « Statistique Départementale ».

Dans ce contexte, marqué par une curiosité empirique toujours portant sur l'ensemble de la France, brillent – en ce qui concerne nos propos de linguistes – les figures et l'œuvre de Coquebert de Montbret père⁵ et fils⁶, devenus les auteurs, entre autres, de la première enquête linguistique du territoire entier de la France (et aussi bien au-delà), réalisée entre 1806 et 1812⁷.

Une autre des émanations les plus connues de cette nouvelle veine d'investigation est la « Carte figurative de l'instruction populaire de la France », dressée en 1827 par le mathématicien Charles Dupin (1784-1873) : voir la figure 1⁸.

³ Cf. Herrgen (2001).

⁴ Cf. Ozouf-Marignier (1989).

⁵ Il s'agit de Charles-Etienne Coquebert de Montbret (1755-1831).

⁶ Il est question d'Eugène-Barthélémy Coquebert de Montbret (1785-1847), un des cinq enfants de Charles-Etienne.

⁷ Au sujet de Coquebert de Montbret père et fils voir surtout la somme magistrale d'Isabelle Laboulais-Lesage de 1999, la thèse récente de Sven Ködel soutenue en 2013 aux Universités de Bamberg et Paris 7 (Diderot) et les indications fournies par S. Pop (1950, 19-26).

⁸ Pour la personne et l'œuvre de Ch. Dupin, cf. Palsky (1996, 62-67).

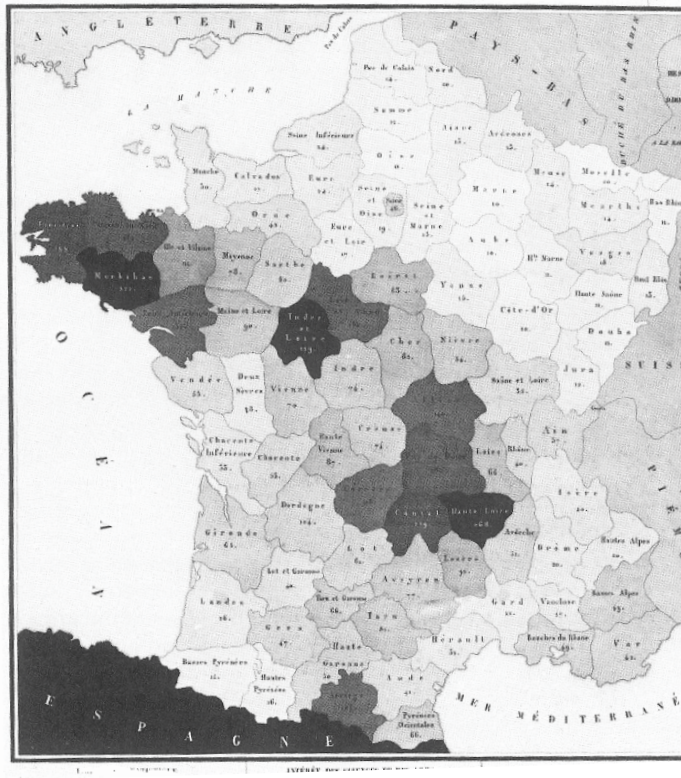


Figure 1 : Charles Dupin (1784-1873) : Carte figurative de l'instruction populaire de la France en 1827.

Cette carte présente déjà quelques particularités empiriques et techniques qui caractérisent également la production cartographique de la dialectométrie. Elle repose sur des données judicieusement collectées, visualise un sujet thématique bien défini (= le degré d'alphabétisation) et a été dressée moyennant la technique visuelle de la carte en plages (= carte choroplèthe). Son message cartographique repose sur une dizaine de teintes grises, réparties entre le noir complet et le blanc, dont le principe sémiotique est simple : plus c'est blanc, meilleure est l'alphabétisation des habitants des différents départements de la France.

Une considération sommaire du profil choroplèthe de la carte fait ressortir une ligne séparatrice située approximativement entre St-Malo et Genève. En 1827 et après, le message de cette carte a provoqué d'énormes conflits, voire des scandales, sur la scène intellectuelle, culturelle et politique de la France, et ceci à cause de l'inégalité évidente entre le sud et le nord de la Nation⁹.

⁹ Cf. à ce sujet la contribution de Roger Chartier (1992).

La « Statistique Départementale » reposait sur des piliers méthodiques très solides : il y avait déjà une théorie de départ – une sorte d'idée d'une « Belle au bois dormant », dont il fallait trouver la cachette ou les traces –, des techniques bien rodées pour la collecte des données et aussi une maîtrise cartographique toujours en essor pour mettre au point la meilleure des visualisations des recherches entreprises¹⁰.

Évidemment, la galerie de telles cartes pourrait être facilement élargie. La production française respective en est très riche, pour ne pas dire exubérante¹¹.

À partir de 1860, il y avait même une cartographie statistique institutionnalisée, et pratiquée sous les auspices du Ministère des Travaux Publics, dont l'émanation la plus connue était l'« Album des statistiques graphiques¹² ». C'était une collection de cartes thématiques à couleurs, toutes en format in folio, qui avait le privilège d'être omniprésente dans tous les organismes et bureaux de l'État. En bref, la vie intellectuelle et politique de la France était imprégnée – au moment de la gestation de l'ALF – de cartes, d'idées cartographiques et de tous les prérequis méthodiques, théoriques et pratiques y ayant trait.

Et Jules Gilliéron : combien de tout ceci connaissait-il? Lui, qui pourtant était le fils d'un géologue? Évidemment, il est possible qu'il n'en sût rien, mais est-ce réellement vraisemblable?

Ma conviction est que l'entreprise de l'ALF représente, en dernière analyse, une des dernières émanations de la « Statistique départementale » telle qu'elle a été pratiquée tout au long du XIX^e siècle – depuis la Révolution française et le Premier Empire – pour beaucoup d'autres variables, qu'elles soient d'origine humaine ou non.

Évidemment, depuis la période de l'ALF les méthodes de la géographie linguistique ont changé. Le grand dilemme devant lequel se trouvaient les auteurs d'atlas linguistiques du XX^e siècle était la nature foncièrement bipolaire du travail d'atlas en tant que tel : ou bien collecter les données dialectales à l'aune de *l'espace* tout en mettant l'accent – dans le binôme *géographie linguistique* – sur le mot *géographie*, ou bien collecter les données à l'aune de la *faculté langagière* des locuteurs (dont on connaît le caractère hétéroclite) tout en mettant l'accent sur l'épithète *linguistique*.

Pour Gilliéron, la collecte de données exclusivement *basilectales* – issues d'un acte conscient de traduction de la part de témoins bilingues – était un prérequis de fer. Cette procédure lui semblait être la garante de la bonne *intercomparabilité* des données collectées. Cette idée – absolument normale au sein de toutes les sciences empiriques telles que la sociologie ou la psychologie – fut encore partagée et défendue par Albert Dauzat (1877-1955) au moment de la planification du « Nouvel Atlas linguistique de la France » (NALF)¹³. Or, nous savons bien que les premières réalisations

¹⁰ Voir à ce sujet l'excellente analyse de Marie-Noëlle Bourguet de 1988.

¹¹ Pour une description très détaillée de ce contexte, je renvoie expressément à l'excellent livre du géographe français Gilles Palsky de 1996.

¹² Cf. Palsky (1996, 141 sq.).

¹³ Pour le programme du NALF cf. Tuaillon (1983).

concrètes du NALF – toutes conçues et réalisées par d'éminents dialectologues de renom tels que Pierre Gardette (1906-1973), Jean Séguy (1914-1973) ou Pierre Nauton (1917-1970) – n'ont pas été faites sous le signe de l'intercomparabilité des données recueillies : ni entre elles, ni avec celles de l'ALF ou d'un atlas voisin, mais bien plutôt sous le signe de *l'authenticité*, tant des enquêtes – faites souvent en *patois* et partant sans le biais de la traduction bilingue consciente – que des réponses, recueillies non pas en forme standardisée, mais au cours d'une *conversation dirigée*. D'où l'inévitable augmentation, chaleureusement saluée par les dialectologues, de réponses doubles, triples et même au-delà, toutes hautement idiosyncrasiques, mais dépourvues de toute intercomparabilité.

Au début, la géographie linguistique romane n'a nullement été entravée par cette bipolarité. A preuve de cela, je cite la double dédicace, formulée par Walther von Wartburg (1888-1971) au début du premier volume du FEW, à l'intention de Wilhelm Meyer-Lübke (1861-1936), le traditionnaliste néogrammairien, et à Jules Gilliéron (1854-1926), l'innovateur empirique et théorique.

À la différence de Pierre Gardette, Jean Séguy, en tant qu'auteur de l'« Atlas linguistique de la Gascogne » (ALG), voulut longtemps respecter les principes prônés par Albert Dauzat et produire, de ce fait, « de l'intercomparable ». Il a fini par succomber d'abord à la tentation de *l'authenticisme* empirique, pour se convertir ensuite – surtout après l'expérience faite avec le quatrième volume de l'ALG – en *dialectométricien*.

Par le biais de l'analyse *quantitative* des données *qualitatives* débordantes contenues dans son ALG, il voulait en extraire une sorte de sève vitale cachée, comparable à la « Belle au bois dormant », dont on devine l'existence, mais ignore le domicile.

Aux dires de Xavier Ravier, son fidèle compagnon de travail, Jean Séguy voulait faire tenir (ou réduire) le gascon – entendu comme une masse de « 47 millions de chiures de mouche » – « dans une formule ou un schéma¹⁴ ».

Ce qu'il est devenu de la *dialectométrie* depuis 1973, année du lancement du mot et aussi de la chose, j'en parlerai tout de suite. D'entrée de jeu, j'aimerais bien souligner encore une chose. Les études de dialectométrie, dont il existe de nos jours beaucoup de variantes, ont démontré que sous les enchevêtrements souvent inextricables ou carrément chaotiques d'un nombre inouï d'aires linguistiques appartenant à toutes les catégories linguistiques imaginables, il se cache d'étonnantes régularités spatiales qui peuvent être considérées comme autant d'émanations de lois géolinguistiques très claires et simples qui, elles, constituent la contrepartie des « Lautgesetze » (lois linguistiques) d'ordre diachronique découvertes par les néogrammairiens de Leipzig dans le dernier quart du XIX^e siècle.

¹⁴ Cf. Ravier (1976, 390).

2. La dialectométrie : une méthode inductive et exploratoire

Au début des années 70 du XX^e siècle, la dialectométrie (DM) est née sous l'impulsion de la curiosité de quelques géolinguistes¹⁵ qui voulaient savoir si, derrière (ou : sous) l'apparent désordre des structures spatiales des cartes de n'importe quel atlas linguistique (AL), il existe un ordre enveloppant et, partant, ou bien « majeur » ou bien « caché ». Ce but ne pouvait être rejoint que par la synthèse, faite à la main¹⁶ ou à l'aide de l'informatique alors en état de gestation¹⁷, de beaucoup de cartes d'atlas. Par une vieille tradition de la philosophie occidentale, une telle procédure – qui part du plan du « particulier » pour arriver à celui du « général » – est appelée « induction ». Comme elle se lance aussi à la recherche de choses jusqu'alors inconnues, elle a par définition aussi une vocation « exploratoire ».

Après plus de quatre décennies d'intenses recherches-DM appliquées à des atlas linguistiques romans¹⁸, allemands¹⁹ et anglais²⁰, il est clair que l'ordre sous-jacent (c'est-à-dire « caché » ou « de profondeur ») assumé d'entrée de jeu, existe réellement et qu'il est l'émanation (ou la création) d'une activité particulière des humains. Au sein de l'« École dialectométrique de Salzbourg » (EDMS), il est question, dans ce contexte, depuis belle lurette de la « gestion basilectale de l'espace par l'homo loquens », considérée comme une des multiples facultés innées des humains et à laquelle est attribué le rang d'un « universal » du langage à *plein titre*.

C'est ainsi qu'en face d'un atlas linguistique (AL) à dialectométrer, nous nous sommes toujours posé les questions suivantes :

- Quelle est la variation géolinguistique *globale* répertoriée sur toutes les planches de l'AL en question, en l'occurrence de l'ALF ?
- Comment se reflète, sur toutes les planches de l'ALF, la *gestion basilectale de la Galloromania* de la part des Français, des Suisses romands et des Wallons ? Pour préciser : nos recherches-DM relatives à l'ALF englobaient, certes, le réseau entier de l'ALF, mais ont dû se limiter, pour des raisons d'ordre économique, à l'analyse de quelque 40% des 1421 planches originales de la série A de l'ALF²¹. Ce pourcentage est pourtant suffisamment grand pour pouvoir en tirer des conclusions représentatives et, de ce fait, fiables.

¹⁵ Je renvoie, à ce sujet, à deux articles de Jean Séguy (1971 et 1973 : avec la mention du néologisme *dialectométrie*) ainsi qu'à mes propres contributions de 1971 (avec N. Winterleitner) et 1976.

¹⁶ C'était le cas pour Jean Séguy, décédé déjà en 1973.

¹⁷ J'avais le privilège de pouvoir me servir, dès le début, de l'informatique.

¹⁸ L'essentiel de nos recherches-DM romanes porte sur les grands atlas nationaux ALF, AIS (Italo-romania), ALPI (Ibéro-romania) et ALDC (domaine catalan).

¹⁹ Cf. Goebel/Scherrer/Smečka (2013).

²⁰ Cf. Goebel/Schiltz (1997).

²¹ Pour plus de détails sur la dialectométrisation des données de l'ALF, cf. Goebel (2002a et 2003).

3. Préparation cartographique du réseau d'enquête de l'ALF

Rappelons d'abord trois caractéristiques élémentaires de l'ALF :

- Le réseau de l'ALF comprend 638 localités²² où Edmond Edmont²³ a mené 639 enquêtes.
- Jules Gilliéron a inauguré, par les modalités typographiques de la publication des données de l'ALF, la tradition – devenue plus tard caractéristique de la géographie linguistique *romane* en général – de présenter, sur chaque planche d'atlas, une synopse de toutes les transcriptions recueillies sur le terrain, relatives à un item particulier du questionnaire utilisé²⁴. C'est le principe du « full text map » qui s'oppose à celui du « symbol map » utilisé par les *germanistes* et les *anglicistes*.
- Dès le début, la présentation des données transcrites originales sur la superficie d'une planche d'atlas a rendu nécessaire – pour les besoins de leur dépouillement rationnel – le recours à un instrument heuristique spécial, à savoir aux « cartes muettes ».

Ces faits expliquent pourquoi la « géographie linguistique » créée par Jules Gilliéron ne pouvait se faire sans un recours continu et massif à la pensée et aux méthodes d'une discipline située en dehors de la linguistique, à savoir de la *cartographie*. Le maniement et le remplissage des *cartes muettes* demandaient en outre, de la part des adeptes de la nouvelle méthode, le déploiement de *réflexions géo-classificatoires* visant à subdiviser les données originales d'une planche d'atlas en groupement spatiaux utiles suivant la logique d'un principe classificateur préalablement choisi. Les annales des études romanes fournissent une quantité énorme d'applications très réussies de cette démarche géo-classificatoire²⁵.

De même qu'aux temps de la publication de l'ALF quand l'imprimeur Protat Frères (de Mâcon) a mis en vente des cartes muettes du réseau de l'ALF pour permettre aux lecteurs de ce dernier d'y reporter le rendement de leurs réflexions géo-classificatoires, il est également de nos jours indispensable de créer, pour le réseau d'un AL en train d'être dialectométrisé, un support cartographique capable de rece-

²² Le nombre réel des *points d'enquête* de l'ALF est 638 et non pas 639 comme il est indiqué très souvent par erreur. L'ALF comporte cependant, sur chacune des 1421 planches de la série A, les résultats transcrits de 639 *enquêtes* faites par Edmond Edmont. La solution de cette « énigme » est qu'Edmont avait réalisé deux enquêtes au point-ALF 284, Saint-Pol-sur-Ternoise, sa ville natale.

²³ Rappelons qu'Edmond Edmont (1849-1926) était l'enquêteur de l'ALF.

²⁴ Du temps de la publication des données de l'ALF, le principe combien heureux du « full text map » était bien loin d'être hors conteste. Un des adversaires les plus farouches de ce principe était Paul Meyer qui s'employait énergiquement pour une publication des données de l'ALF sous la forme de tableaux : cf. Storost (1994, 77). Quelle catastrophe pour la géographie linguistique si P. Meyer avait réussi à imposer ses visées saugrenues!!

²⁵ Ceci concerne surtout les apports de l'onomasiologie romane qui, elle, se basait avant tout sur les atlas linguistiques.

voir les informations spatiales tant *qualitatives* que *quantitatives* utilisées au cours de l'analyse-DM. Un tel support présuppose l'établissement d'un compartimentage polygonal du réseau en question, dit aussi « tessellation de Voronoï ».

La polygonisation de Voronoï se fait à travers les étapes suivantes²⁶ :

- triangulation du réseau des points d'enquête
- construction successive des médiatrices sur les côtés de triangle du réseau
- De la fusion de toutes les médiatrices ainsi construites naît le réseau polygonal dont nous avons besoin.

4. La taxation des données d'un atlas linguistique

Voir les cartes 1 et 2.

Au sein de l'EDMS, l'on entend par la « taxation » (des planches originales d'un AL) l'analyse d'un secteur déterminé des informations originales contenues dans une carte originale de l'AL examiné, et l'établissement successif d'une « image » des données originales respectives sous la forme d'une « carte de travail » (CT). Précisons que ce genre de travail classificatoire s'aligne parfaitement sur une tradition déjà ancienne en matière de linguistique romane qui avait été inaugurée par Jules Gilliéron lui-même²⁷.

Les CT issues d'une campagne de taxation disposent de structururations internes très variables qui, elles, peuvent osciller entre 2 et n « aires taxatoires ». L'élément linguistique de base qui régit chaque aire taxatoire, est appelée « taxat ». Au cours de la dialectométrisation de plus de 600 des 1421 planches originales de l'ALF entreprise peu avant le tournant du millénaire à Salzbourg²⁸, nous en avons extrait 1681 CT appartenant aux catégories de la *phonétique*, du *lexique* et de la *morpho-syntaxe*, dont la figure 2 montre la distribution quantitative.

²⁶ Pour plus amples détails cf. Goebel (1981, 363 sq.) et Goebel (1984 I, 90 sq.).

²⁷ Cf. Gilliéron (1918) : à la fin de cet ouvrage se trouve une typisation lexicale relative à la carte 1 (abeille) de l'ALF. Voir également Jaberg (1908), où l'on trouve toute une gamme de typisations colorisées de différentes planches de l'ALF.

²⁸ Voir nos contributions de 2000, 2002a et 2003 qui en présentent les principes et quelques résultats.

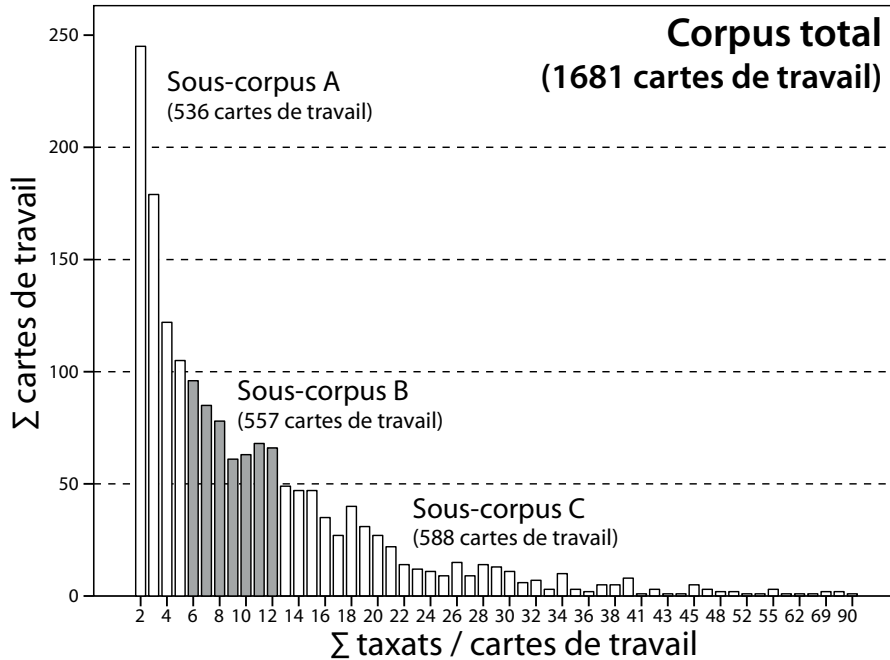


Figure 2 : Histogramme – pourvu de 55 colonne

s – montrant, pour notre corpus-DM tiré de l'ALF, la relation entre le nombre de taxats par carte de travail (en abscisse) et le nombre absolu des cartes de travail (en ordonnée)

Les colonnes de l'histogramme en question représentent une courbe décroissante de gauche à droite²⁹. Il s'agit de 55 colonnes de hauteurs inégales qui couvrent un espace « polynymique » allant de 2 taxats/CT jusqu'à 90 taxats/CT : voir l'abscisse sur la figure 2.

Les hauteurs de 55 colonnes signalent les nombres respectifs des CT analysées. Il y a donc 245 CT « *bi*-nymes » (= avec *deux* types linguistiques ou taxats), 179 CT « *tri*-nymes » (= avec *trois* taxats), 122 CT « *tétra*-nymes » (= avec quatre taxats) et seulement une CT de nature 90-nyme. Elle se réfère d'ailleurs à la carte 545 de l'ALF, relative aux dénominations de la *fauvette*.

²⁹ Il s'agit d'une courbe exponentielle. Nous avons d'ailleurs découvert les mêmes régularités dans toutes nos analyses dialectométriques qui portent sur tous les grands atlas romans ainsi que sur ceux des études anglaises. Le caractère exponentiel de ces courbes a été utilisé, en 1985, par le linguiste quantitatif allemand Gabriel Altmann pour la définition de ce qu'il a appelé alors la « loi de Goebel ».

Le bilan taxatoire de notre analyse-DM de l'ALF se présente donc comme suit :

corpus	Σ cartes de travail (CT)	Σ aires taxatoires / taxats
total	1.681	18.047
phonétique	1.117	10.642
lexique	465	6.514
morpho-syntaxe	99	891

Dans l'ensemble, la richesse structurale des quelque 18 mille aires taxatoires analysées est très grande. N'oublions pas que la structure spatiale de chaque carte de travail (CT) peut varier dans quatre directions :

1. du nombre des taxats
2. de la taille des aires taxatoires
3. de la configuration des aires taxatoires et
4. de la localisation des aires taxatoires à l'intérieur du réseau d'atlas en question.

Il en résulte une variabilité vraiment déroutante. Une des conséquences de cette grande variabilité est le fait que les différentes aires taxatoires ne se recouvrent presque jamais. Il en résulte un effet d'*enchevêtrement* très complexe entre les aires taxatoires qui semble être typique de beaucoup de phénomènes spatialement répartis.

La carte 1 montre une des 1117 CT *phonétiques* analysées. Avec ses 17 taxats *phonétiques* – qui symbolisent autant de succédanés du nexus vocalique latin -ĀTU (dans MERC-ĀTU « marché ») – elle représente une CT déjà relativement complexe.

La carte 2, en revanche, appartient à notre corpus *lexical*, riche de seulement 465 CT : elle est franchement moins complexe³⁰.

Une des premières expériences que nous ayons faites au cours de nos travaux-DM concernait l'impact du nombre absolu des CT synthétisées par voie-DM. Nous avons constaté alors que les structures-DM de profondeur obtenues à l'aide de la synthèse de 1681 CT (= corpus total) ressemblent de très près à celles obtenues avec 50%, 25% et 15% de cet effectif, à condition que l'extraction de l'urne pleine en ait été faite par voie aléatoire³¹. La même expérience a pu être réitérée avec d'autres atlas linguistiques romans et non romans.

De cette réalité, étonnante au premier coup d'œil, nous avons conclu que les structures-DM de profondeur mises au jour, sont imprégnées d'une grande *redondance* interne qui, elle, serait comparable à celle dont disposent tous les systèmes ou signaux communicatifs, qu'ils soient de nature linguistique ou non.

³⁰ La distribution spatiale et l'origine des noms de jours de la semaine de la Galloromania ont été analysées à plusieurs reprises : cf. Gilliéron/Roques (1912, 85-108), Rohlf (1949), Wartburg (1949) et Baehr (1952).

³¹ Cf. Goebel (1984 I, 206-219).

Une autre réflexion concernait les conditions de la genèse de cette redondance dont l'origine était indubitablement humaine. Depuis un certain temps, nous sommes convaincus que toutes ces régularités quantitatives sont l'émanation directe d'un comportement particulier des dialectophones par rapport aux défis communicatifs de l'espace qu'ils habitent. Nous lui avons conféré l'appellation de « gestion (ou : aménagement) basilectale de l'espace par l'homo loquens »³².

5. Brève présentation de la « chaîne dialectométrique » de l'EDMS

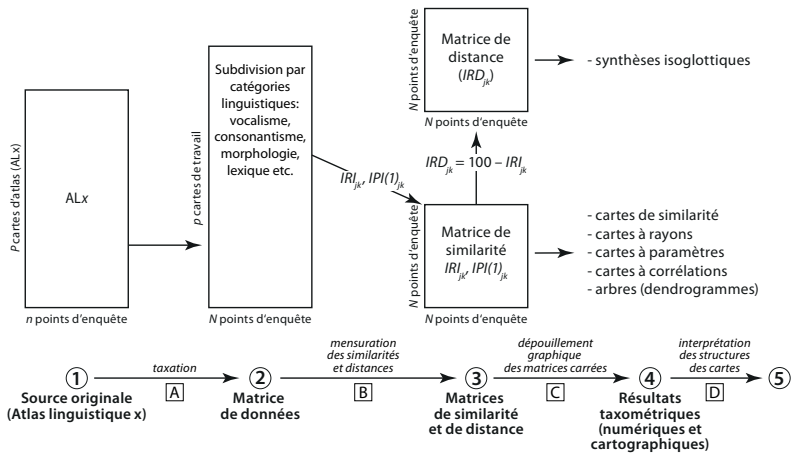


Figure 3 : Diagramme de la chaîne méthodique utilisée par l'« Ecole Dialectométrique de Salzbourg »

Les méthodes de l'EDMS sont alignées en forme de « chaîne » : voir la figure 3. Le tableau suivant renseigne sur les fonctions des différents « chaînons » ou étapes de cet alignement³³.

³² Comparer à cela le titre de notre contribution de 2005 et la définition de ce concept dans Goebel (2002a, 5).

³³ Pour une introduction détaillée aux buts et méthodes de la DM de Salzbourg nous renvoyons à Goebel (1984), vol. I-III (en allemand) et à nos contributions de 2000, 2002a, 2003 et 2012 (toutes en français). Voir en outre les adresses-web suivantes :

- bibliographie complète de nos travaux-DM : https://www.sbg.ac.at/rom/people/prof/goebl/dm_publi.htm
- trois sites-DM de travail et de présentation :

<http://www.dialectometry.com/>

<http://www.dialectometry.com/dmdocs/index.html>

<http://ald.sbg.ac.at/dm/>

³⁴ VDM : „Visual DialectoMetry“, logiciel dialectométrique créé par Edgar Haimerl (Seattle, EUA) entre 1997 et 1999, mis à jour continuellement depuis cette date.

³⁵ L'indice de similarité standard de la DM de Salzbourg est l'« Indice Relatif d'Identité » (IRI_{jk}) auquel fait écho l'« Indice Relatif de Distance » (IRD_{jk}) selon la formule $IRI_{jk} + IRD_{jk} = 100$.

Étapes méthodiques	États des données en voie de traitement-DM	Commentaires
A	1-2	Choix d'un atlas linguistique [1] pour la dialectométrisation ; définition et application des principes taxatoires respectifs ; taxation [A] d'un nombre aussi grand que possible de cartes originales de l'atlas linguistique retenu ; saisie électronique successive des cartes de travail (CT) analysées et établissement de la matrice de données [2] (avec N points d'enquête et p cartes de travail) à l'aide du logiciel VDM ³⁴ . Les données taxées et la matrice de données se situent au niveau de l'échelle métrologique nominale.
B	2-3	Choix et application d'une mesure <i>quantitative</i> pour déterminer les similarités (quantitatives) entre les N points d'enquête : ici : choix de l'IRI _{jk} (« Indice Relatif d'Identité ») ³⁵ ; établissement consécutif de la matrice de similarité (carrée : N × N) [3, en bas]. Celle-ci contient la variation <i>qualitative</i> de la matrice de données après sa conversion dans une forme <i>quantitative</i> .
C	3-4	Transformation (partielle ou intégrale) du contenu <i>quantitatif</i> de la matrice de similarité en représentations et schémas visuels. Le choix des outils de visualisation dépend uniquement des besoins heuristiques du géolinguiste. En utilisant le logiciel VDM : choix de l'algorithme de visualisation approprié et du nombre des paliers chromatiques (intervalles).
D	4-5	Surtout par voie optique : interprétation et évaluation théoriques et pratiques [D] des structures spatiales des mises en cartes effectuées. Genèse de nouvelles représentations conceptuelles et théoriques [5] des données de départ [1].

Dans la panoplie salzbourgeoise des indices de similarité, il y a, entre autres, aussi l'indice IPI_{jk} (« Indice Pondéré d'Identité ») qui dispose de certains avantages taxométriques et visualisateurs auxquels le dialectométricien peut avoir recours le cas échéant.

Précisons encore que dans cette démarche pluridisciplinaire, le savoir nécessaire pour l'établissement de la matrice de données et pour la définition de l'indice de similarité est à rechercher du côté de la « classification numérique » (dite aussi « automatique »)³⁶ alors que les méthodes pour la confection et la visualisation appropriée des cartes en couleurs sont fournies par la cartographie numérisée³⁷, appelée de nos jours aussi « visualistique » (en anglais : « visualistics »). Ces facilités n'existaient d'ailleurs pas encore en tant que telles au commencement de nos activités-DM.

Une des particularités de la procédure pratique de l'EDMS est de toujours privilégier la dimension heuristique traditionnelle de la géographie linguistique : à savoir le plan des cartes. Tous les résultats quantitatifs des calculs-DM effectués sont directement convertis en schémas visuels appropriés et offrant un accès aussi suggestif (et facile !!) que possible aux géolinguistes. Soulignons en outre que l'EDMS pratique également une sélection judicieuse des procédures statistiques utilisées tout en évitant l'application de celles dont la *complexité mathématique* n'est pas à la hauteur de leur *utilité géolinguistique*.

À partir de 1997, tous les calculs et toutes les visualisations-DM s'effectuent à l'aide d'un logiciel-DM appelé « Visual DialectoMetry » (VDM), qui avait été créé par notre ami Edgar Haimerl et qui, depuis cette date, a connu plusieurs remaniements et élargissements³⁸.

6. La mesure de similarité entre deux vecteurs de la matrice des données Voir la figure 4.

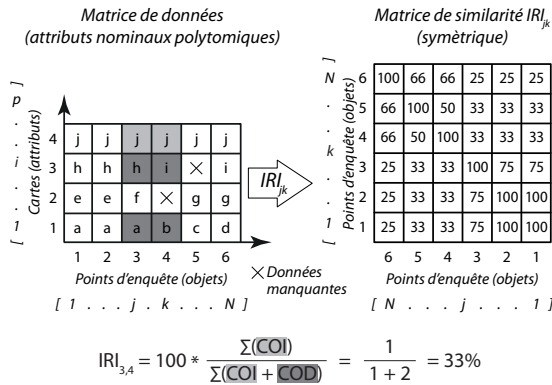


Figure 4 : Schéma de la mensuration de la similarité linguistique entre les différents points d'enquête d'un atlas linguistique

³⁶ Cf., entre autres, les grands manuels taxométriques de Hans Hermann Bock (1974) ainsi que de Peter H. A. Sneath et Robert R. Sokal (1973).

³⁷ Voir les manuels cartographiques de Roger Brunet (1987) et de Jean-Paul Bord (2012).

³⁸ Le logiciel VDM est à la libre disposition de qui en fait la demande.

IRI_{3,4} : Indice Relatif d'Identité entre les points d'enquête 3 et 4

COI : co-identité (entre les paires de taxat j et j)

COD : co-différences (entre les paires de taxat a et b ainsi que h et i)

L' « Indice Relatif d'Identité » (IRI_{jk}) mesure la proportion de taxats identiques entre deux vecteurs-points d'enquête (j et k) de la matrice de données. Théoriquement, les scores de l'IRI_{jk} peuvent varier entre 1 et 100 (%). Suivant la formule combinatoire $N/2 (N - 1)$ et pour l'effectif $N = 641$, le bilan total des calculs à effectuer s'élève à 205'120 scores qui, par la suite, seront stockés dans la matrice de similarité carrée (aux dimensions $N \times N$).

Signalons encore que nous avons ajouté, aux 638 points d'enquête *originaux* de l'ALF, trois points *artificiels* relatifs aux langues standard du *français*, de l'*italien* et du *catalan*³⁹ qui permettent d'étudier l'impact de ces trois registres linguistiques sur le tissu des dialectes de l'ALF. La taille de N s'élève donc à 641 unités.

La matrice de similarité, outre d'être carrée, est aussi parfaitement symétrique. La valeur-IRI de 33% calculée pour les localités factices 3 et 4 se retrouve donc des deux côtés de la diagonale de la matrice de similarité : voir à nouveau la figure 4.

Une fois la matrice de similarité établie, il faut que le dialectométricien-classificateur se mette à la dépouiller judicieusement : voir à ce propos le moment C sur la figure 3. En plus, c'est le premier des nombreux pas *exploratoires* qui mènent des structures « de surface » aux structures de « profondeur ».

7. La mise en carte (visualisation) de scores quantitatifs : à l'exemple de plusieurs cartes de similarité

Voir les cartes 3-8.

Le défi heuristique à relever en la matière, est double⁴⁰ :

- du point de vue *statistique* :
Il s'agit d'extraire un des N vecteurs de la matrice de similarité, de le dépouiller de la valeur relative à la diagonale qui est toujours égale de 1 (ou 100), et de visualiser les N-1 valeurs (ou scores) restantes.
- du point de vue *cartographique* :
Il s'agit de *convertir la variabilité numérique* des N-1 scores de similarité extraits en une *variabilité visuelle analogue* pour fournir au géolinguiste-dialectométricien des informations aptes à être rapidement saisies par voie mentale et interprétées après du point de vue linguistique.

La « machine visualisatrice » employée à cet effet opère avec un certain nombre d'algorithmes de visualisation dont le plus utilisé à Salzbourg est l'algorithme MINM-WMAX. Il répartit les N-1 valeurs à visualiser équitablement des deux côtés de la

³⁹ Cf. à ce propos Goebel 2000, 212.

⁴⁰ Voir à ce propos Goebel 1981 et 1984 (vol. I, passim).

moyenne arithmétique (MA) de la distribution de similarité respective, tout en définissant des intervalles numériquement toujours égaux des deux côtés de la MA⁴¹.

Il en résulte des profils iconiques remarquablement stables et bien ordonnés auxquels le logiciel VDM permet de conférer des silhouettes plus ou moins fines, allant de 2 à 18 paliers chromatiques. Les couleurs que nous utilisons obéissent à l'agencement chromatique du spectre solaire qui est d'ailleurs celui de l'arc-en-ciel.

En général, le bleu est associé aux valeurs numériques petites alors que le rouge renvoie aux scores maximaux. Chaque carte de similarité – comme d'ailleurs toutes nos visualisations – dispose, en outre, d'une légende numérique et d'un histogramme qui fournit des informations qui intéressent plutôt le *statisticien* que le *linguiste*.

8. Structure de profondeur A : l'importance et le rôle heuristiques des cartes de similarité

Voir les cartes 3-8.

Les cartes 3-8 constituent six spécimens de cartes de similarité relatives à différents sous-corpus (équivalents à autant de matrices de données) tirés de notre matrice de données-ALF intégrale.

Les cartes 3 et 4 montrent deux des 641 cartes de similarité calculées à partir de notre corpus intégral (contenant 1681 CT). Elles disposent, comme chaque carte de similarité, d'un point de référence dont le polygone est toujours en blanc. Par la structuration – toujours très régulière – de leur profil choroplèthe, elles indiquent le décalage spatial des similarités linguistiques des N-1 points d'atlas restants par rapport aux points de référence préétablis : 394 (Auderville, Manche) et 778 (Agde, Hérault).

Une des grandes découvertes de la DM de Salzbourg est d'ailleurs l'étonnante régularité avec laquelle ce décalage s'inscrit dans l'espace. Et ceci malgré l'indéniable « chaotisme » qui se dégage de la considération particularisante de quelque 18 000 aires taxatoires réparties sur les 1681 CT extraites de l'ALF⁴².

En ligne générale et dans les conditions de visualisation déjà mentionnées, l'entourage immédiat du point de référence prend les couleurs *rouge* (= intervalle 6) et *orange* (= intervalle 5), qui constituent toujours une aire (quantitative) cohérente. Par contre, les couleurs *bleu foncé* (= intervalle 1) et *bleu clair* (= intervalle 2) caractérisent les « antipodes » de la carte de similarité. Si la visualisation a été effectuée à l'aide de six paliers chromatiques, les polygones en *jaune* (= intervalle 4) et *vert* (= intervalle 3) représentent des *zones-tampon* (ou de *transition*) entre – métaphoriquement parlant – les « meilleurs amis », d'un côté, et, de l'autre, les « pires ennemis » du basilecte du point de référence.

⁴¹ Voir à ce sujet surtout Goebel (1981, 1983 et 1984).

⁴² Cf. Goebel (2003, 65).

Du point de vue linguistique, toute carte de similarité donne une réponse quantitative et finement nuancée à la question de savoir quelle est la position de tel dialecte à l'intérieur d'un espace dialectal examiné. J'ajoute que de telles questions ont été posées très souvent (et depuis longtemps) au sein de la romanistique et d'autres philologies modernes.

Les cartes 3, 4 et 5 montrent des profils géotypologiques pour les domaines du normand, du languedocien et du berrichon. Les profils choroplèthes respectifs sont très différents, surtout en ce qui concerne les cartes 3 (de type d'oïl) et 4 (de type d'oc).

Évidemment, les trois profils de similarité que nous venons de voir ont démontré que le décalage de la similarité linguistique dans l'espace peut être très variable. Toujours est-il qu'il s'agit là d'un phénomène *humain* ou, pour être plus précis, d'une émanation directe de la « gestion basilectale de l'espace par l'homo loquens », évoquée au début de cet article.

Or, il serait intéressant de comparer le décalage quantitatif produit dans l'espace par une activité d'origine *humaine* avec celui issu d'un modèle *géométrique* relatif à la même superficie. Un tel modèle consiste à définir les proximités aériennes réciproques entre les points de l'ALF à l'aide de la formule bien connue de Pythagore ($a^2 + b^2 = c^2$). La visualisation des 641 vecteurs de la matrice des proximités euclidiennes crée autant de profils choroplèthes absolument réguliers qui, eux, peuvent servir d'étalon pour une comparaison avec les calculs de similarité.

La carte 6 en constitue un spécimen relatif au point de référence 103 (Bué, Cher). En la comparant avec la carte 5, l'on comprend aisément que la gestion *basilectale* (= humaine) du réseau de l'ALF s'écarte considérablement de la gestion *géométrique* (= euclidienne) du même espace.

Or, la comparaison *visuelle* des deux profils choroplèthes multicolores peut être complétée par une évaluation *quantitative* des deux séries de 641 scores, relatives, d'un côté, aux *valeurs de similarité* (selon l'IRI) et, de l'autre, aux *scores de proximités* (selon la formule de Pythagore). Un tel calcul de corrélation peut être fait moyennant le coefficient de corrélation de [Auguste] Bravais et [Karl] Pearson $r(BP)$. Pour les 2×641 scores qui sont à la base des cartes 5 et 6, il en résulte la valeur +0,697. Prise comme telle, cette valeur ne nous dit absolument rien. Ce n'est que dans le cadre de la DM corrélatrice que de telles valeurs peuvent revêtir une signification *linguistique* à proprement parler : voir le paragraphe 11.

Les cartes 7 et 8 reposent sur deux sous-corpus linguistiquement différents : alors que la carte 7 se réfère à notre sous-corpus *phonétique* (avec 1 117 CT), la carte 8 a été construite à l'aide de notre sous-corpus *lexical* (avec 465 CT).

L'allure de leurs profils choroplèthes est – malgré certaines similarités de base – quelque peu différente. On a l'impression que la stabilité interne du profil choroplèthe de la carte 7 (phonétique) dépasse celle du profil choroplèthe de la carte 8 (lexique).

9. Structure de profondeur B : l'analyse « interponctuelle » et la carte à cloisons

Voir les cartes 9 et 10.

L'analyse dite « interponctuelle » connaît deux versants : celui de la visualisation de *distances* et celui de la visualisation de *similarités* interponctuelles, c'est-à-dire relevées entre deux points d'enquête contigus. La première des deux visualisations, appelée « carte à cloisons », correspond exactement à la synthèse traditionnelle d'isoglosses destinée à dégager des bourrelets de frontière plus ou moins épais⁴³. Il s'agit donc de la modélisation, faite avec les moyens de la DM, d'une technique déjà ancienne et, par conséquent, bien connue des géolinguistes.

La seconde visualisation, dite « carte à rayons », est, en revanche, de date récente. Du point de vue taxométrique, elle constitue le complément de la carte à cloisons, puisqu'elle visualise des *similarités* interponctuelles (s) dont le rapport avec les *distances* interponctuelles (d) est régi par la formule suivante : $s + d = 100$.

Le message linguistique des deux types de visualisation est également complémentaire. Alors que la carte à *cloisons* visualise, à travers l'épaisseur et le coloriage froid des côtés de *polygone*, le degré d'« aversion » qui existe entre deux lococetcs contigus, la carte à *rayons* représente, par l'épaisseur et le coloriage chaud des côtés de *triangle*, des relations d'« amitié » à courte distance⁴⁴.

La structure iconique des cartes 9 et 10 ressort de l'interaction visuelle de 1791 éléments graphiques (i. e. de côtés de polygone) et fournit une vision très nuancée des enchevêtrements interponctuels qui existent entre le Nord et le Sud de la France. Il n'offre cependant aucune trace d'une confrontation *linéiforme* – c'est-à-dire revêtant les apparences d'une « véritable frontière » – entre les domaines d'oïl et d'oc. Sur les deux cartes, l'œil perçoit, par contre, une vaste ceinture de passage, sillonnée d'effets de cloisonnements finement gradués.

Une fois de plus, la traversée du Nord au Sud semble être spatialement plus « ordonnée » sur la carte 9 (basée sur le sous-corpus *phonétique*) que sur la carte 10 (basée sur le sous-corpus *lexical*).

Encore une petite remarque d'ordre quantitatif : l'effectif total des scores de distance (selon l'IRD_{jk}) pris en compte sur les cartes à cloisons 9 et 10, comprend seulement 0,873% du total des scores disponibles dans les matrices de distance respectives⁴⁵. La clarté très suggestive de ces deux cartes repose donc sur des effectifs très réduits de la matrice de distance sous-jacente. En termes de *taxométrie*, la *valeur classificatoire* des deux cartes est donc relativement *faible*.

⁴³ Voir les tableaux en appendice du travail séminal de Karl von Ettmayer (1924) où se trouve un grand nombre de schémas isoglottiques exécutés à la main et contenant des bourrelets de frontière de toute sorte. Voir aussi les cartes isoglottiques reproduites dans Jochnowitz 1973.

⁴⁴ Pour des spécimens de la carte à rayons, cf. Goebel (1983 et 2012, 137-138).

⁴⁵ Cette relation se calcule comme suit : $1791 : 205\,120 \times 100$.

Heureusement, il y a des méthodes taxométriques beaucoup plus « enveloppantes » et, de ce fait, plus sûres : voir les paragraphes suivants.

10. Structure de profondeur C : au-delà des cartes de similarité : les cartes à paramètre

Voir les cartes 11-14.

Par « paramètre » l'on entend, en matière de *statistique descriptive*, une des nombreuses valeurs « caractéristiques » qui dénotent une quelconque distribution de fréquence, tel que le *minimum*, le *maximum*, la *MA*, l'*écart-type*, la *symétrie*, etc. En considérant les histogrammes au bas des cartes 3 et 4, l'on constate aisément que la symétrie des distributions de similarité peut être très variable.

À y regarder de près, cette propriété *statistique* revêt aussi une grande importance *linguistique*. Sur la carte 4, l'histogramme de droite, relatif à un parler du languedocien maritime, est fortement biaisé vers la gauche. Ceci signifie que des 640 scores de similarité 432 (= 53 + 245 + 134) se trouvent au-dessous de la MA⁴⁶ et seulement 208 (= 100 + 76 + 32) au-dessus du même seuil⁴⁷.

Le bilan analogue pour l'histogramme de la carte 3 se présente comme suit :

- valeurs situées au-dessous de la MA : 272 (= 9 + 168 + 95)
- valeurs situées au-dessus de la MA : 368 (= 13 + 198 + 157)

Quand on admet qu'une carte de similarité reflète – au moins en théorie – la position communicative du dialecte d'un point d'atlas, les valeurs mentionnées ci-dessus suggèrent que la position communicative *générale*⁴⁸ du dialecte du point-ALF 778 (situé en Languedoc méridional) est moins bonne que celle du point-ALF 394 (situé dans le nord de la Manche) et que, par conséquent, la plus ou moins grande symétrie des histogrammes respectifs renvoie, en dernière analyse, à cette potentialité interactive.

Or, il s'est avéré que le meilleur indice pour saisir la *symétrie* d'une distribution de similarité et, avec elle, sa vertu *communicative*, est le « coefficient d'asymétrie de [Ronald A.] Fisher » (CAF). En calculant le CAF de chacune des N distributions de similarité et en visualisant par la suite les N scores ainsi dégagés, l'on obtient une carte choroplèthe finement articulée (voir les cartes 11-14) dont la valeur linguistique est considérable. En établissant la synthèse des CAF nous poussons – métaphoriquement parlant – encore plus avant notre « plongée exploratoire dans les profondeurs » des données de l'ALF.

⁴⁶ Il s'agit des valeurs appartenant aux intervalles 1-3 : sur la carte 4, elles sont toutes marquées par des couleurs « froides » (bleu foncé, bleu clair et vert).

⁴⁷ Il s'agit des valeurs appartenant aux intervalles 4-6 : sur la carte 4, elles sont toutes marquées par des couleurs « chaudes » (jaune, orange et rouge).

⁴⁸ Il s'agit de toujours mettre en rapport les différents lococlectes de l'ALF avec l'ensemble de son réseau.

Du point de vue statistique, les scores du CAF oscillent des deux côtés de la valeur 0 : ils prennent des valeurs *négatives* quand les scores de similarité (selon IRI_{jk}) s'accumulent entre le maximum et la MA, et des valeurs *positives* quand le cumul intervient entre le minimum et la MA.

Sur les cartes 11-14, les polygones en *bleu foncé* – équivalant à des scores négatifs du CAF – renvoient à des dialecticités *bien connectées* avec le reste du réseau, alors que les polygones *rouges* – correspondant toujours à des scores *positifs* du CAF – se réfèrent à des variétés *détachées*, *conservatrices* et *peu enclines* à s'ouvrir à des interactions et échanges linguistiques.

Depuis un certain temps, j'utilise le concept de *compromis linguistique* (dérivé du terme allemand *Sprachausgleich*) pour désigner le degré d'interaction qu'un lococlecte entretient avec le reste de son réseau. Au vu des cartes 9 et 10, l'on peut donc dire que les lococlectes marqués en bleu (foncé et clair) sont très interactifs, alors que le contraire est vrai pour les lococlectes marqués en rouge.

L'agencement spatial des deux tonalités est d'ailleurs très éloquent. Alors que les polygones en bleu foncé forment, d'un côté, un grand cercle fermé (autour du domaine d'oïl) et, de l'autre, un petit cercle à demi ouvert (relatif au francoprovençal) et évoquent par là l'idée d'une *irradiation spatiale* de grande envergure, les polygones rouges constituent des agglomérations *ponctuelles* suggérant l'idée d'immobilisme et de défense « tous azimuts ».

En effet, nous sommes en présence d'une analyse-DM très efficace dont l'utilité est avant tout d'ordre *diachronique*. C'est ainsi que le grand agencement circulaire bleu dans le nord évoque l'action irradiatrice continue des dialectes du type francilien, alors que la pince entrouverte bleue autour du domaine francoprovençal renvoie aux contacts et conflits linguistiques que ce géotype entretenait tant vers le nord que vers le sud.

Que le coloriage en bleu foncé – équivalent à des scores négatifs du CAF – soit en relation directe avec l'interaction linguistique tout court, est corroboré par le comportement du seul îlot linguistique compris dans le réseau de l'ALF, à savoir du P. 635 (Andraut, Gironde). L'on y trouve un dialecte oïlique de type saintongeais qui avait été transféré en territoire gascon au cours du XV^e siècle. Comme il avait été exposé à toute sorte d'interactions au cours des derniers six siècles, il va de soi qu'il figure en bleu foncé.

Cette analyse-DM permet donc de repérer, au sein d'un réseau donné, des zones linguistiquement « mouvementées » et des régions relativement calmes ou même « inertes ».

Ajoutons encore que les quatre cartes en question reposent sur autant de sous-corpus différenciés du point de vue *quantitatif* et *qualitatif*.

- Pour les cartes 11 et 12 :

Les deux sous-corpus utilisés se distinguent par leur degré de *polynymie*. Alors que la carte 11 repose sur 536 CT dont les structures internes sont relativement simples, la carte 12 est basée sur 588 CT dont les structures sont très complexes. Pour mieux comprendre cet effet, il suffit de jeter un regard sur la figure 2. L'on y perçoit l'emplacement de trois sous-corpus (A, B et C) dont la structuration interne varie entre 2-5 taxats par CT (= sous-corpus *oligo-nyme*⁴⁹ A) et 13-90 taxats/CT (= sous-corpus *poikilo-nyme*⁵⁰ C).

Entre parenthèses : la carte 1, pourvue de 17 taxats et, partant, d'une structure plutôt complexe, appartient au sous-corpus C, alors que la carte 2, disposant de seulement trois taxats et, de ce fait, d'une structuration très simple, fait partie du sous-corpus A.

Or, il s'est avéré, au cours de nos analyses-DM, que de la synthèse de beaucoup de CT *oligo-nymes* (= simples ou « claires ») ressortent pratiquement les mêmes structures de profondeur que l'on obtient moyennant la synthèse de beaucoup de CT *poikilo-nymes* (= complexes ou « chaotiques »). Cet effet quelque peu surprenant montre que les structures-DM de profondeur disposent d'une grande redondance interne, comparables en cela au comportement d'autres signaux anthropiques.

Il n'est donc pas étonnant que les cartes 11 et 12 montrent pratiquement la même structuration choroplèthe caractérisée par une circularité « oïlique » et une pince entrouverte « francoprovençale » en bleu foncé dans le nord, et par trois « balouards de résistance » en rouge situés dans le sud.

Pour les cartes 13 et 14 :

Alors que le profil choroplèthe de la carte 13 – dressé à l'aide de 1 117 CT *phonétiques* – s'aligne grosso modo sur celui des cartes 11 et 12, le profil choroplèthe de la carte 14, reposant sur 465 CT *lexicales*, s'en écarte un peu. L'effet oïlique de circularité y est moins saillant : ce qui ressort avec beaucoup de clarté, c'est, d'un côté, une large bande d'*interaction* et de *compromis linguistique* qui se creuse, d'une façon très compacte, au beau milieu de la Galloromania, et, de l'autre, la réduction de la zone colorée en rouge, orange et jaune, causée par le détachement du Limousin et de l'Auvergne du bloc occitano-catalan.

11. Structure de profondeur D : la dialectométrie corrélative

Voir les cartes 15-18.

La DM corrélative, qui est une des pièces fortes de l'arsenal méthodique de l'EDMS, existe depuis 2005⁵¹ et a fourni depuis beaucoup de preuves de sa grande utilité heuristique. Le module de VDM relatif à la DM corrélative permet de calculer la corrélation entre deux matrices de similarité (etc.) à dimensions extérieures égales

⁴⁹ Grec *olígos* „petit, peu“.

⁵⁰ Grec *poikilos* „bigarré, polychrome“.

⁵¹ Cf. Goebel (2005, *passim*).

(= deux fois $N \times N$), et de visualiser rapidement le rendement de ces calculs sous la forme d'une carte choroplèthe à couleurs avec N points (ou polygones).

La *nature phénoménologique* des deux variables à corrélérer peut varier à loisir : l'on peut corrélérer, entre elles, différentes *catégories linguistiques*, les mettre en relation, l'une après l'autre, avec l'*espace euclidien*, ou comparer, par voie de corrélation, les résultats numériques de différentes *mesurations de similarité*. Ce qui importe au cours de ces opérations statistico-visualisatrices, c'est de ne jamais perdre de vue leur finalité et signification *linguistiques* à proprement parler. Ajoutons encore que la profondeur des structures synthétiques rejointes par la DM corrélatrice, est très grande.

Nous commençons la description de l'efficacité de la DM corrélatrice par la présentation d'une corrélation cartographique entre les *similarités linguistiques* d'un côté et les *proximités euclidiennes* de l'autre : voir la carte 15. Le but en est d'étudier de plus près les multiples facettes du décalage spatial que subissent les scores de similarité linguistique des différentes cartes de similarité.

En reprenant la leçon donnée par les cartes 5 (= de *similarité*) et 6 (= de *proximité*) il est possible de calculer 641 scores de corrélation – toujours selon le coefficient de corrélation $r(\text{BP})$ – et de les visualiser par la suite suivant les normes habituelles de l'EDMS. Il en résulte la carte 15 qui, d'un côté, repose sur le corpus-ALF intégral (= 1681 CT) et, de l'autre, sur les proximités euclidiennes calculées selon le théorème de Pythagore.

Le profil choroplèthe de la carte 15 est d'une clarté et régularité exemplaires. Pour bien pouvoir le comprendre en tant que *linguiste*, il faut avant tout connaître la signification *linguistique* de la variabilité *numérique* calculée, et du symbolisme chromatique y ayant trait.

Or, les teintes *chaudes* renvoient à une *harmonie intrinsèque* du déploiement des similarités linguistiques et des proximités euclidiennes dans l'espace, alors que les teintes *froides* signalent de fortes *frictions* et *divergences* entre ces deux « énergies » spatiales. Evidemment, de telles disharmonies sont dues à tout ce qui peut influencer de l'extérieur sur l'évolution « naturelle » de la langue dans l'espace, soit en l'accélérant, soit en la retardant. Les facteurs en question sont bien entendu d'ordre *historique*, *politique* et *social*.

Dans cette perspective, la Galloromania se présente comme le théâtre d'une grande confrontation entre le Nord et le Sud où le premier a fini par prendre le dessus tout en déstabilisant graduellement le Sud, en passant par le Massif Central pour arriver aux Pyrénées. La ceinture médiane de la Galloromania, allant de la Gironde à l'ouest jusqu'à l'éventail du francoprovençal à l'est, constitue la zone des affrontements les plus violents.

Une fois de plus, la position de l'îlot linguistique du P.-ALF 635 (Andraut) – où l'on trouve le score $r(\text{BP})$ minimal – nous fournit la preuve du bien fondé de notre supposition selon laquelle la divergence de fond entre les gestions *basilectale* et *euclidienne* de l'espace crée les scores de corrélation les plus bas et, partant, des coloriages

en bleu foncé. Rappelons que la « Petite Gavacherie » dont Andraut (= P.-ALF 635) fait partie, avait été fondée au XV^e siècle en territoire d'oc comme colonie linguistique d'oïl d'origine saintongeaise. Il va de soi que la fondation d'un îlot linguistique en terre linguistiquement étrangère constitue une violation éclatante de la relation « naturelle » entre les gestions *basilectale* et *euclidienne* de l'espace.

Évidemment, il est possible de répéter le même exploit-DM pour la *phonétique* et l'*espace* d'un côté (voir la carte 16), ainsi que pour le *lexique* et l'*espace* de l'autre (voir la carte 17). Les plus grandes divergences entre les profils choroplèthes des deux cartes à corrélations se situent en Provence, alors que pour le reste de la Galloromania, la distribution spatiale des corrélations entre les gestions linguistique et euclidienne de l'espace reste sensiblement la même.

Selon toute vraisemblance, l'allure éminemment régulière des cartes 15-17 est due, entre autres, à la grande ancienneté des divergences qui se sont manifestées, au cours des siècles, entre les gestions *linguistique* et *euclidienne* de la Galloromania.

La dernière des quatre cartes corrélatives que nous venons de présenter, se réfère à un problème éminemment *grammatical* : voir la carte 18. Il s'agit de savoir dans quelle mesure la gestion de l'espace moyennant deux *catégories linguistiques différentes* s'annonce comme convergente ou divergente.

Les mensurations et les visualisations y ayant trait fournissent un profil choroplèthe très clair : les gestions *lexicale* et *phonétique* divergent considérablement le long d'un couloir « delphino-forme » allant de la Gironde aux Alpes alors qu'elles convergent tant au nord qu'au sud de la Galloromania. À remarquer, une fois de plus, l'allure très régulière du profil choroplèthe dans son ensemble.

Ajoutons encore que le couloir *delphino-forme* mentionné ci-dessus a fait son apparition dans beaucoup d'autres analyses-DM relatives à l'ALF. Il est cependant bon de ne pas le confondre avec le tracé du « croissant » de mémoire ronjatiennne. Ce dernier se situe plus au nord et dispose d'un contour beaucoup plus mince⁵².

Précisons en outre que l'application de la DM corrélative à d'autres domaines linguistiques – romans ou autres – a fourni également des résultats très clairs et bien structurés : cf. à ce propos Goebl (2005, passim, et 2012, 155 sq.).

12. Structure de profondeur E : l'analyse dendrographique

Voir les cartes 19-20.

Le niveau E de l'analyse-DM des structures de profondeur concerne la DM dendrographique ou – dans la terminologie technique de la classification numérique (ou taxométrie) – la « classification ascendante hiérarchique » (CAH). La CAH est largement employée au sein de beaucoup de sciences naturelles et humaines. Elle sert au

⁵² Pour l'histoire du *croissant*, voir le travail séminal de Guylaine Brun-Trigaud de 1990 : pour les contours géographiques traditionnels du « croissant » tel qu'il avait été défini en 1913 par Jules Ronjat (1864-1925), voir la carte à la page 274 (ibidem).

groupement algorithmique de N objets – définis, chacun, par p attributs – en structures hiérarchiques pourvues de ramifications binaires. Du point de vue statistique, il s'agit d'une analyse dite « multivariée ». Dans la chaîne-DM, la CAH ne représente cependant qu'une seule des nombreuses possibilités de dépouillement de la matrice de similarité (voir la figure 3). Vu la longue tradition de la *pensée arborescente* au sein de pratiquement toutes les branches de la linguistique, ce genre d'analyse offre un certain attrait « naturel » aussi pour les besoins de la géolinguistique.

Les algorithmes à utiliser (ou utilisables) dans le cadre de la CAH sont très variés. Ils opèrent, tout en exécutant des impératifs mathématiques préalablement définis, des fusions (ou agglomérations) binaires entre les N objets de la matrice de similarité en question. En général, leur action agglomératrice débute par les différences *minimales* qui existent entre les N^2 valeurs de la matrice de similarité pour en rejoindre les différences *maximales* : d'où le qualificatif « ascendant » de la CAH. Quant au choix de l'algorithme-CAH le plus approprié aux besoins de la DM, il n'y a pas de solution fixe ou « passe-partout ». C'est pourquoi le dialectométricien doit sélectionner, dans l'arsenal algorithmique de la CAH, la procédure dont les résultats correspondent le plus à ses propres besoins et visées.

Un des algorithmes-CAH qui répond le mieux à ces impératifs est celui qui a été proposé, en 1963, par le statisticien américain Joe Ward Jr. Il fournit, en ligne générale, d'excellents résultats quand il est appliqué à des données de nature géolinguistique. Un des critères pour déterminer l'utilité *linguistique* d'un algorithme-CAH est la cohérence spatiale des différents clusters (= « chorèmes ») extraits de l'arbre calculé.

Évidemment, l'arbre en tant que configuration ramifiée ne saurait jamais suffire aux besoins d'un *géolinguiste* épris de schémas *cartographiques* et, par conséquent, *spatiaux*. Il faut donc projeter le rendement classificatoire de l'arbre dans l'espace. J'appelle ce transfert « spatialisation ». C'est ainsi que chacun des branchages de l'arbre (appelés par nous « dendrèmes ») retenus comme linguistiquement « significatifs », est transposé dans le secteur approprié du réseau en question qui, lui, est appelé à Salzbourg « chorème ». Inutile d'ajouter que le coloriage des dendrèmes correspond parfaitement à celui des chorèmes.

Le jeu alternatif des différents chorèmes disséminés dans l'espace établit donc une excellente classification des locolestes examinés, alors que le message hiérarchique concomitant de l'arbre permet d'y ajouter des nuances quantitatives – et aussi diachroniques – supplémentaires.

Ceci est dû au fait que tout arbre issu d'une CAH peut être utilisé également à une *reconstruction diachronique* telle qu'elle est conçue et pratiquée depuis belle lurette par la *lexicostatistique* et la *glottochronologie*⁵³. La logique de cette perspective veut que la ramification croissante à partir de la racine de l'arbre corresponde au déploiement diachronique de la fragmentation linguistique du territoire en question.

⁵³ Cf. à ce sujet Goebel (2003, 84-89).

Jetons à ce propos un coup d'œil sur la carte 19 : les deux embranchements de la première bifurcation correspondent, en effet, aux domaines d'oc (= dendrèmes/chorèmes 1-4) et d'oïl (= dendrèmes/chorèmes 5-8). La bifurcation suivante – toujours selon la logique de l'algorithme de Ward – engendre la séparation entre le francoprovençal (= dendrème/chorème 5) et le reste du domaine d'oïl (= dendrèmes/chorèmes 6-8). Evidemment, il est possible de continuer ce genre d'analyse jusqu'au niveau des feuilles. VDM en offre les possibilités techniques nécessaires.

Une autre particularité – d'ordre statistique – mérite également d'être mentionnée. C'est que les différents dendrèmes (et chorèmes) doivent leur genèse à la considération serrée de leur *variabilité quantitative interne*. C'est ainsi que naît une certaine hiérarchisation des dendrèmes/chorèmes en fonction de leur hétérogénéité quantitative interne (et aussi réciproque) : plus un dendrème est proche de la racine de l'arbre, plus grande en est l'hétérogénéité interne. Et vice versa.

Ceci signifie que, sur la carte 19, la variabilité numérique interne du dendrème/chorème relatif au domaine d'oïl (= dendrèmes/chorèmes 5-8) est moins grande que celle du dendrème/chorème relatif au domaine d'oc (= dendrèmes/chorèmes 1-4)⁵⁴.

Soulignons encore que les cartes 19 et 20 se distinguent par leurs corpus de base : la carte 19 repose sur notre sous-corpus *phonétique* (avec 1 117 CT) alors que la carte 20 émane d'une matrice de données *lexicale* (avec 465 CT). Il va de soi que les deux arborescences et leurs spatialisations ne sont pas complètement identiques. Ceci vaut surtout pour l'agencement (relationnel) des huit dendrèmes à l'intérieur des arbres correspondants.

Pour en faciliter la comparaison, nous avons conféré, sur les cartes 19 et 20, le même coloriage (et numérotation) aux chorèmes placés dans une situation géographique comparable, d'où il résulte cependant une numérotation non plus continue (c'est-à-dire allant de 1 à 8) mais discontinue des différents dendrèmes de la carte 20. L'on peut en déduire que la première bifurcation aboutit, sur les deux cartes, aux mêmes résultats (= nette séparation entre Oc et Oïl), certes, mais aussi que la hiérarchisation des huit dendrèmes/chorèmes colorés à l'intérieur de l'arbre de la carte 20 se distingue quelque peu de la situation analogue sur la carte 19. Toujours est-il que la cohérence spatiale des chorèmes telle qu'elle se manifeste près de la racine, persiste aussi au *milieu* de l'arbre et près des *feuilles*. Ceci prouve le haut degré de « classifiabilité » de nos données.

13. Epilogue

Le fondement théorique de la DM de Salzbourg est l'hypothèse de l'existence de la gestion basilectale de l'espace par l'homo loquens dont les manifestations réelles – toutes saisissables par le biais d'un atlas linguistique de type ALF, AIS, ALPI ou

⁵⁴ La raison en est que la racine du méga-dendrème 5-8 est plus proche du tronc de l'arbre que celle du méga-dendrème 1-4.

ALDC – obéissent à certaines régularités voire lois générales qui ne peuvent être découvertes qu'à travers un certain nombre d'opérations *inductives* ou de *synthèse*. Il importe que les résultats de ces opérations de synthèse soient dûment visualisés afin qu'ils puissent déclencher les opérations mentales voulues.

Il est en outre très important de ne pas oublier qu'il y a, en dehors de la géolinguistique à proprement parler, de nombreuses sciences étroitement liées avec l'espace dont les méthodes géo-classificatoires sont très proches de celle de la DM, et avec lesquelles une coopération interdisciplinaire s'impose. Nous n'en citons ici que la *géographie humaine* et la *génétique des populations*⁵⁵. Dans la rétrospective, il est bien évident que de telles disponibilités – ou faudrait-il dire « velléités » – interdisciplinaires existaient déjà du temps de Coquebert de Montbret père et fils.

Il est cependant non moins évident que – du point de vue *empirique* et *méthodique* – un fil conducteur rouge s'étend de leurs activités à la DM actuelle.

Université de Salzbourg

Hans GOEBL

Remerciements

Préparation et confection de tous les clichés de cet article : Werner Goebel (Vienne)

Contrôle linguistique de mon texte français : Stéphanie Lescure (Marbourg, Hesse)

Que mon fils Werner et Mme Lescure reçoivent ici l'expression de ma sincère et profonde gratitude.

Abréviations souvent utilisées

AL	atlas linguistique
CAF	coefficient d'asymétrie de [Ronald A.] Fisher
CAH	classification ascendante hiérarchique
CT	carte de travail
DM	dialectométrie, relative à la dialectométrie, dialectométrique
EDMS	Ecole dialectométrique de Salzbourg
IRD _{jk}	Indice Relatif de Distance (entre les points d'atlas j et k)
IRI _{jk}	Indice Relatif d'Identité (entre les points d'atlas j et k)
MA	moyenne arithmétique

⁵⁵ Citons, à titre d'exemple, la contribution interdisciplinaire de Scapoli / Goebel / Sobota / Mamolini / Barraï de 2005.

MEDMW	algorithme d'intervallisation (centré sur la MA, avec la formation, des deux côtés de la MA, de n classes égales en nombre de points d'atlas)
MINMWMAX	algorithme d'intervallisation (centré sur la MA, avec la formation, des deux côtés de la MA, de n classes dont la largeur numérique est égale)
r(BP)	coefficient de corrélation de [Auguste] Bravais et [Karl] Pearson
VDM	Visual DialectoMetry (logiciel dialectométrique utilisé à Salzbourg ; librement disponible sur demande)

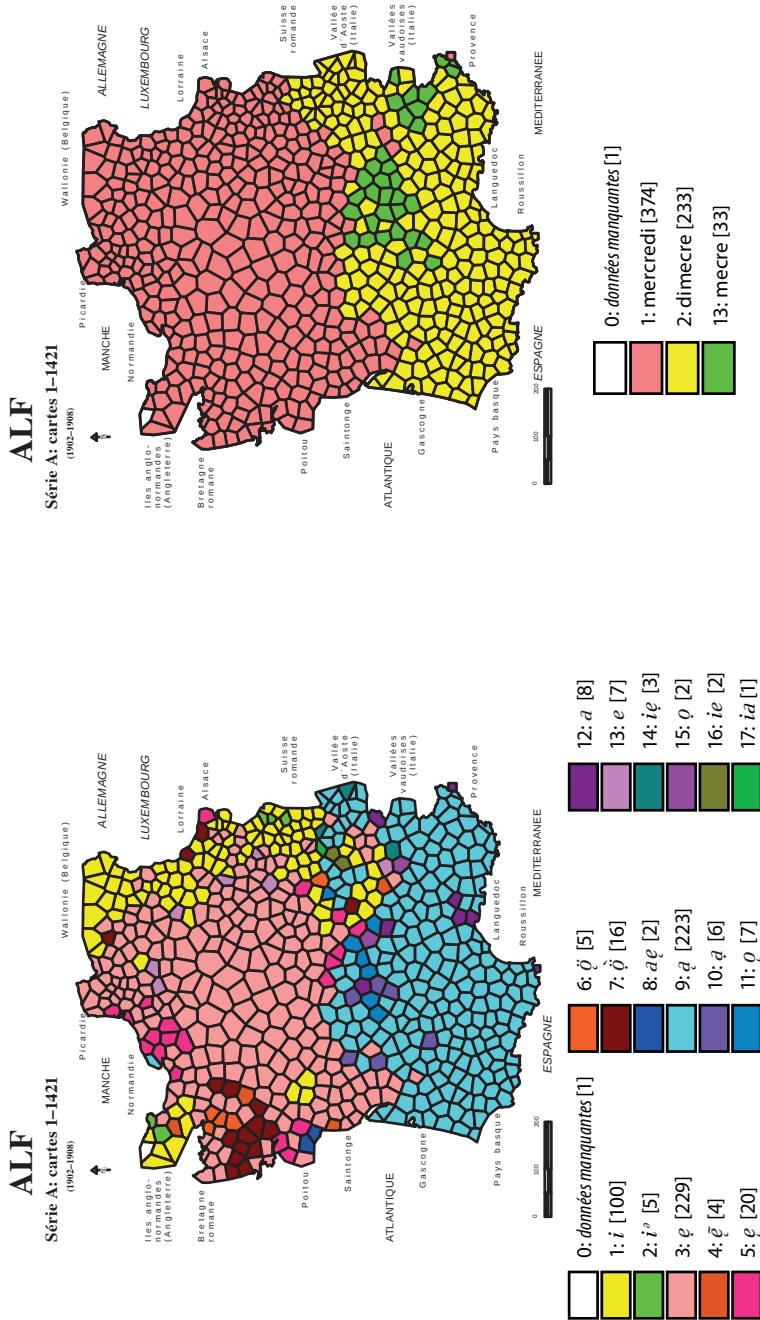
Références bibliographiques

- AIS = Jaberg, Karl / Jud, Jakob, 1928-1940. *Sprach- und Sachatlas Italiens und der Südschweiz*, 8 vol., Zofingen, Ringier (réimpression : Nendeln, Kraus 1971).
- ALDC = Veny, Joan / Pons i Griera, Lúdia, 2001-2013. *Atles Lingüístic del Domini Català*, 6 vol., Barcelona, Institut d'Estudis Catalans.
- ALF = Gilliéron, Jules / Edmont, Edmond, 1902-1910. *Atlas linguistique de la France*, 10 vol., Paris, Champion (réimpression : Bologne, Forni 1968).
- ALPI = *Atlas lingüístico de la Península Ibérica*, 1962, vol. I : Fonética, Madrid, CSIC.
- Altmann, Gabriel, 1985. « Die Entstehung diatopischer Varianten. Ein stochastisches Modell ». *Zeitschrift für Sprachwissenschaft* 4, 139-155.
- Baehr, Rudolf, 1958. « Zu den romanischen Wochentagsnamen », in : Lausberg, Heinrich / Weinrich, Harald (ed.), *Romanica. Festschrift für Gerhard Rohlfs*, Halle, Niemeyer, 26-56.
- Berschlin, Helmut / Felixberger, Josef / Goebel, Hans, 2008². *Französische Sprachgeschichte. Lateinische Basis, interne und externe Geschichte, sprachliche Gliederung Frankreichs. Mit einer Einführung in die historische Sprachwissenschaft*, Hildesheim, Olms.
- Bock, Hans Hermann, 1974. *Automatische Klassifikation. Theoretische und praktische Methoden zur Gruppierung und Strukturierung von Daten (Cluster-Analyse)*, Göttingen, Vandenhoeck & Ruprecht.
- Bord, Jean-Paul, 2012. *L'univers des cartes. La carte et le cartographe*, Paris, Belin.
- Bourguet, Marie-Noëlle, 1988. *Déchiffrer la France. La statistique départementale à l'époque napoléonienne*, Paris, Editions des Archives Contemporaines.
- Brunet, Roger, 1987. *La carte. Mode d'emploi*, Paris, Fayard.
- Brun-Trigaud, Guylaine, 1990. *Le croissant : le concept et le mot. Contribution à l'histoire de la dialectologie française au XIX^e siècle*, Lyon, Centre d'études linguistiques Jacques Goudet.
- Brun-Trigaud, Guylaine / Le Berre, Yves / Le Dù, Jean, 2005. *Lectures de l'Atlas linguistique de la France de Gilliéron et Edmont. Du temps dans l'espace. Essai d'interprétation des cartes de l'Atlas linguistique de la France de Jules Gilliéron et Edmond Edmont augmenté de quelques cartes de l'Atlas linguistique de la Basse Bretagne de Pierre Le Roux*, Paris, CTHS.
- Chartier, Roger, 1992. « La ligne Saint-Malo – Genève », in : Nora, Pierre (ed.), *Les lieux de mémoire*, vol. 3 : *Les France, tome 1 : Conflits et partages*, Paris, Gallimard, 738-775.

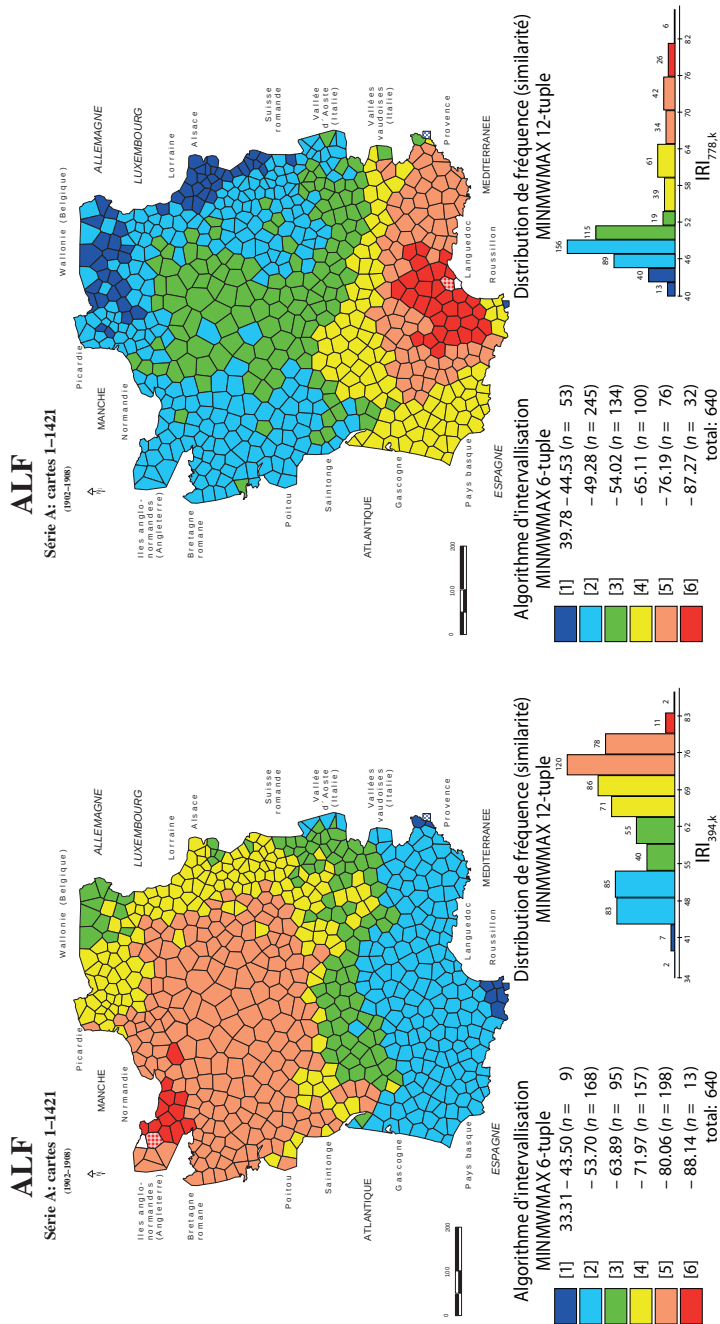
- Ettmayer, Karl von, 1924. « Das Wesen der Dialektbildung, erläutert an den Dialekten Frankreichs », *Denkschriften* [= Mémoires] *der Akademie der Wissenschaften in Wien*, philosophisch-historische Klasse, 66/3, 1-56, 7 Tafeln [= tableaux].
- FEW = Wartburg, Walther von (1922/1928 sq.). *Französisches etymologisches Wörterbuch*, Bonn, Bouvier.
- Gilliéron, Jules, 1918. *Généalogie des mots qui désignent l'abeille d'après l'Atlas linguistique de la France*, Paris, Champion.
- Gilliéron, Jules / Roques, Mario, 1912. *Etudes de géographie linguistique d'après l'Atlas linguistique de la France*, Paris, Champion.
- Goebel, Hans, 1976. « La dialectométrie appliquée à L'ALF (Normandie) », in : Vârvaro, Alberto (ed.), *Atti del XVI Congresso Internazionale di Linguistica e Filologia Romanza* (Naples 1974), Naples, Marchiaroli/Amsterdam, Benjamins, vol. 2, 165-195.
- Goebel, Hans, 1981. « Eléments d'analyse dialectométrique (avec application à l'AIS) », *RLiR* 45, 349-420.
- Goebel, Hans, 1983. « Parquet polygonal et treillis triangulaire. Les deux versants de la dialectométrie interponctuelle », *RLiR* 47, 353-412.
- Goebel, Hans, 1984. *Dialektometrische Studien. Anhand italo-romanischer, rätoromanischer und galloromanischer Sprachmaterialien aus AIS und ALF*, Tübingen, Niemeyer, 3 vol.
- Goebel, Hans, 2000. « La dialectométrisation de l'ALF : présentation des premiers résultats », *Linguistica* 40, 209-236.
- Goebel, Hans, 2002a. « Analyse dialectométrique des structures de profondeur de l'ALF » *RLiR* 66, 5-63.
- Goebel, Hans, 2002b. « Sprachatlanten : woher? womit? wozu? Einige buntgemischte Reflexionen am Gartenzaun zwischen Romanistik und Germanistik », in : Anreiter, Peter / Ernst, Peter / Hausner, Isolde / Kalb, Helmut (ed.), *Namen, Sprachen und Kulturen. Imena, Jeziki in Kulture. Festschrift für Heinz Dieter Pohl zum 60. Geburtstag*, Vienne, Edition Präsens, 257-274.
- Goebel, Hans, 2003. « Regards dialectométriques sur les données de l'Atlas linguistique de la France (ALF) : relations quantitatives et structures de profondeur », *Estudis Romànics* 25, 59-96.
- Goebel, Hans, 2004. « VDM – Visual DialectoMetry. Vorstellung eines dialektometrischen Software-Pakets auf CD-ROM (mit Beispielen zu ALF und Dees 1980) », in : Dahmen, Wolfgang / Holtus, Günter / Kramer, Johannes / Metzeltin, Michael / Schweickard, Wolfgang / Winkelmann, Otto (ed.), *Romanistik und neue Medien. Romanistisches Kolloquium XVI* (Rauischholzhausen, 2000), Tübingen, Narr, 209-241.
- Goebel, Hans, 2005. « La dialectométrie corrélative : un nouvel outil pour l'étude de l'aménagement dialectal de l'espace par l'homme », *RLiR* 69, 321-367.
- Goebel, Hans, 2006. « Warum die Dialektometrie nur in einem roman(ist)ischen Forschungskontext entstehen konnte », in : Dahmen, Wolfgang / Holtus, Günter / Kramer, Johannes / Metzeltin, Michael / Schweickard, Wolfgang / Winkelmann, Otto (ed.), *Was kann eine vergleichende romanische Sprachwissenschaft heute (noch) leisten. Romanistisches Kolloquium XX* (Göttingen, 2004), Tübingen, Narr, 291-317.
- Goebel, Hans, 2008. « La dialettomettrizzazione integrale dell'AIS. Presentazione dei primi risultati », *RLiR* 72, 25-113.
- Goebel, Hans, 2011. « Quo vadis, atlas linguistique? Einige wissenschaftshistorische und zeitgeistkritische Reflexionen zur atlasgestützten Geolinguistik », in : Schlaak, Claudia / Busse, Lena

- (ed.), *Sprachkontakte, Sprachvariation und Sprachwandel. Festschrift für Thomas Stehl zum 60. Geburtstag*, Tübingen, Narr, 5-27.
- Goebel, Hans, 2012. « Introduction aux problèmes et méthodes de l' «Ecole dialectométrique de Salzbourg» (avec des exemples gallo-, italo- et ibéroromans) », in : Álvarez Pérez, Xosé Afonso / Carrilho, Ernestina / Magro, Catarina (ed.), *Proceedings of the International Symposium on Limits and Areas in Dialectology (LimiAr)*, Lisbonne, Centro de Linguística da Universidade de Lisboa, 117-166 (dans le web : http://limiar.clul.ul.pt/proceedings/general_session/7_goebel.pdf).
- Goebel, Hans / Scherrer, Yves / Smečka, Pavel, 2013. « Kurzbericht über die Dialektometrisierung des Gesamtnetzes des «Sprachatlases der deutschen Schweiz» (SDS) », in : Schneider-Wiejowski, Karina / Kellermeier-Rehbein, Birte / Haselhuber, Jakob (ed.), *Vielfalt, Variation und Stellung der deutschen Sprache* [Mélanges pour le 70^e anniversaire de Ulrich Ammon], Berlin/Boston, Walter de Gruyter, 153-175.
- Goebel, Hans / Schiltz, Guillaume, 1997. « A Dialectometrical Compilation of CLAE 1 and CLAE 2. Isoglosses and Dialect Integration », in : Viereck, Wolfgang / Ramisch, Heinrich (ed.), *The Computer Developed Linguistic Atlas of England*, vol. 2, Tübingen, Niemeyer, 13-21.
- Goebel, Hans / Winterleitner, Norbert, 1971. « Projekt einer sprachstatistischen Auswertung von in Sprachatlanten gespeicherter linguistischer Information mit Hilfe elektronischer Rechenanlagen », *Linguistische Berichte* 14, 60-61.
- Herrgen, Joachim, 2001. « Die Dialektologie des Deutschen », in : Auroux, Sylvain / Koerner, E. F. Konrad / Nederehe, Hans-Josef / Versteegh, Kees (ed.), *History of the Language Sciences / Geschichte der Sprachwissenschaften / Histoire des sciences du langage. An International Handbook on the Evolution of the Study of Language from the Beginning to the Present / Ein internationales Handbuch zur Entwicklung der Sprachforschung von den Anfängen bis zur Gegenwart / Manuel international sur l'évolution de l'étude du langage des origines à nos jours*, Berlin/New York, de Gruyter, vol. 2, 1513-1535.
- Jaberg, Karl, 1908. *Sprachgeographie. Beitrag zum Verständnis des Atlas linguistique de la France Aarau, Sauerländer* (version espagnole : *Geografía lingüística. Ensayo de interpretación del « Atlas lingüístico de Francia »*). Traducción de Antonio Llorente y Manuel Alvar, Granada, Universidad de Granada, 1959).
- Jochnowitz, George, 1973. *Dialect Boundaries and the Question of Franco-Provençal*, La Haye, Mouton.
- Ködel, Sven, 2013. *L'enquête Coquebert de Montbret (1806-1812) sur les langues et dialectes de France et la représentation de l'espace linguistique français sous le Premier Empire*, Bamberg, Paris, thèse doctorale.
- Laboulais-Lesage, Isabelle, 1999. *Lectures et pratiques de l'espace. L'Itinéraire de Coquebert de Montbret, savant et grand commis d'Etat (1755-1831)*, Paris, Champion.
- Ozouf-Marignier, Marie-Vic, 1989. *La formation des départements : la représentation du territoire français à la fin du 18^{ème} siècle*, Paris, Editions de l'Ecole des Hautes Etudes en Sciences Sociales.
- Palsky, Gilles, 1996. *Des chiffres et des cartes. Naissance et développement de la cartographie quantitative française au XIX^e siècle*, Paris, CTHS.
- Pop, Sever, 1950. *La dialectologie. Aperçu historique et méthodes d'enquêtes linguistiques*, Louvain, Chez l'auteur/Gembloux, Duculot.
- Ravier, Xavier, 1976. « Jean Ségué et la traversée du langage gascon. Réflexions sur une topogénèse géolinguistique », *RLiR* 40, 389-402.
- Rohlf, Gerhard, 1949. « Les noms des jours de la semaine dans les langues romanes », in : *Mis-*

- celânea de filologia, literatura e história cultural, à memória de Francisco Adolfo Coelho (1847-1919)*, Lisbonne, Centro de Estudos Filológicos, 88-94.
- Rosenqvist, Arvid, 1919. « Limites administratives et division dialectale de la France », *Neuphilologische Mitteilungen* 20, 87-119 (carte publiée à nouveau in : Berschin/Felixberger/Goebel 2008², 261).
- Scapoli, Chiara / Goebel, Hans / Sobota, Slawomir / Mamolini, Elisabetta / Rodriguez-Larralde, Alvaro / Barrai, Italo, 2005. « Surnames and Dialects in France : Population structure and cultural evolution », *Journal of Theoretical Biology* 237, 75-86.
- Séguy, Jean, 1971. « La relation entre la distance spatiale et la distance lexicale ». *RLiR* 35, 335-357.
- Séguy, Jean, 1973. « La dialectométrie dans l'Atlas linguistique de la Gascogne », *RLiR* 37, 1-24.
- Sneath, Peter H. A. / Sokal, Robert R., 1973. *Numerical Taxonomy. The Principles and Practice of Numerical Classification*, San Francisco, W. H. Freeman.
- Storost, Jürgen, 1994. « Noch einmal zur Diez-Stiftung », *RJb* 45, 74-84.
- Tuaillon, Gaston, 1983. « Les atlas linguistiques régionaux de France », *BALI* III Serie, Dispensa n. 7, 68-89.
- Wartburg, Walther von, 1949. « Los nombres de los días de la semana », *RFE* 33, 1-14.

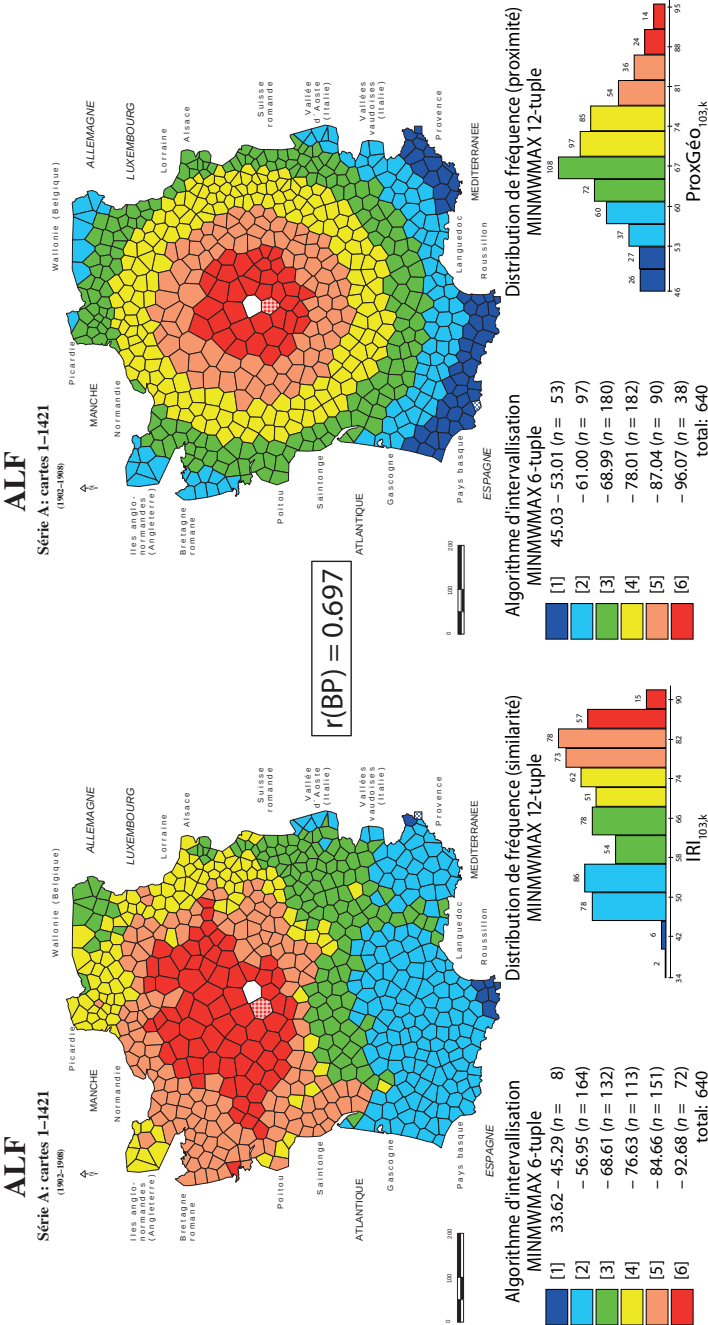


Carte 2: Exemple d'une carte de travail 3-nyme de nature *lexicale* (selon ALF 839 *mercredi*) : les réalisations lexicales du concept « mercredi »



Carte 3: Profil choroplèthe d'une carte de similarité relative au point-ALF 394 (Auderville, Manche)
Corpus: corpus intégral (1681 cartes de travail)
Indice de similarité: IRIjk
Algorithme d'intervallisation: MINMWMAX 6-tuple

Carte 4: Profil choroplèthe d'une carte de similarité relative au point-ALF 778 (Agde, Hérault)
Corpus: corpus intégral (1681 cartes de travail)
Indice de similarité: IRIjk
Algorithme d'intervallisation: MINMWMAX 6-tuple



Carte 5: Profil choroplèthe d'une carte de similarité relative au point-ALF 103 (Bué, Sancerre, Cher)

Corpus: corpus intégral (1681 cartes de travail)

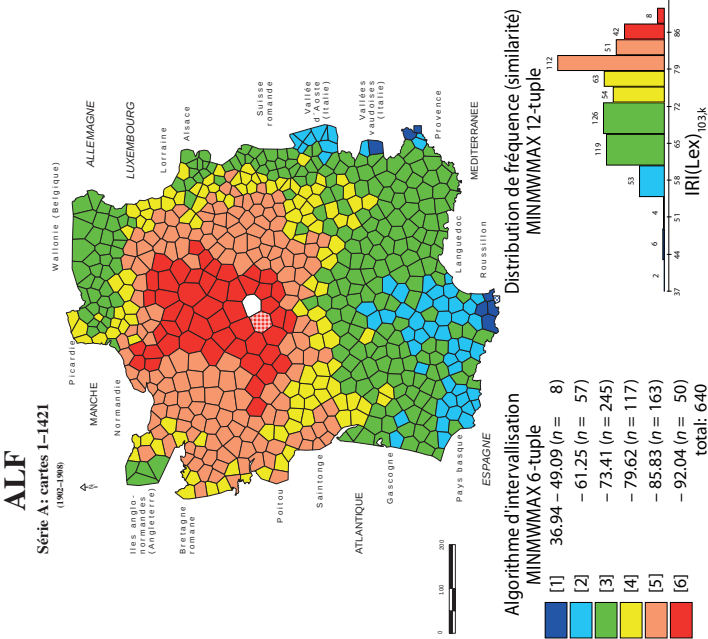
Indice de similarité: IRIjk

Algorithme d'intervallisation: MINMWMAX 6-tuple

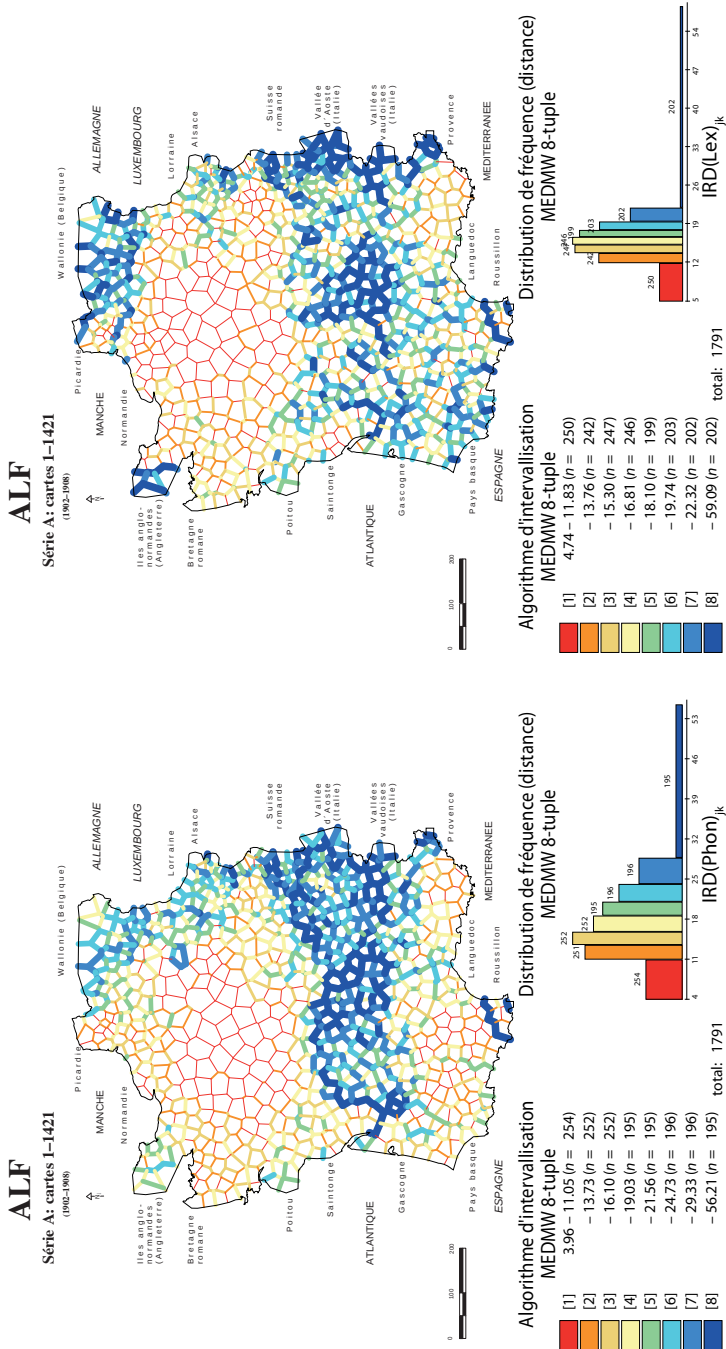
Carte 6: Profil choroplèthe d'une carte de proximité relative au point-ALF 103 (Bué, Sancerre, Cher)

Indice de proximité: théorème de Pythagore

Algorithme d'intervallisation: MINMWMAX 6-tuple

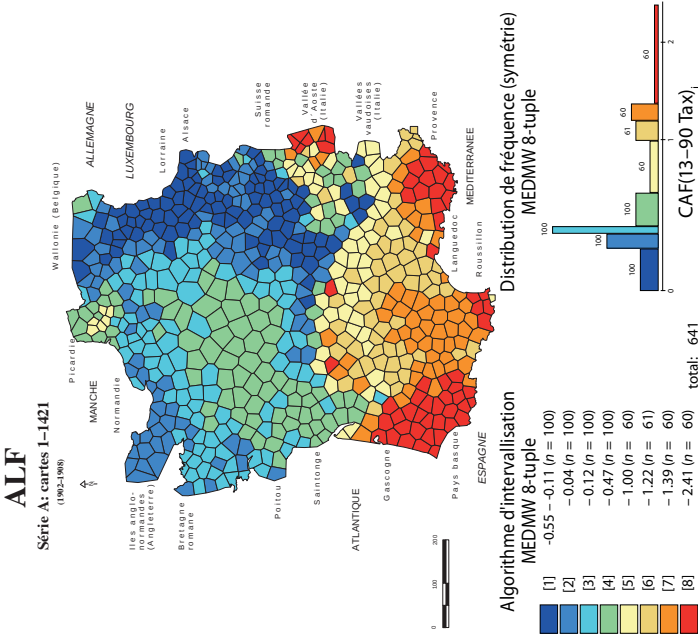


Carte 8: Profil choroplèthe d'une carte de similarité relative au point-ALF 103 (Bué, Sancerre, Cher)
Corpus: sous-corpus *lexical* (465 cartes de travail)
Indice de similarité: IR1jk
Algorithme d'intervallisation: MINMW/MAX 6-tuple

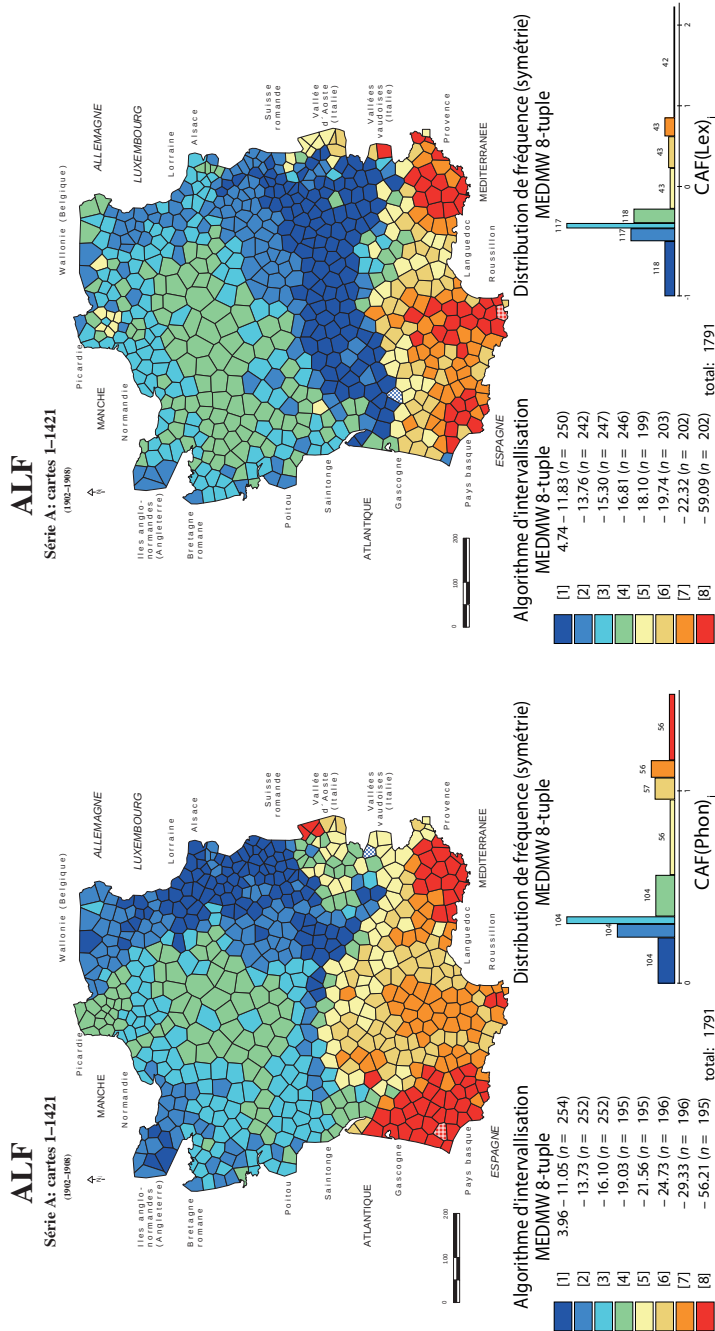


Carte 10: Analyse interponctuelle (carte à cloisons)
Corpus: sous-corpus *lexical* (465 cartes de travail)
Indice de distance: TRD_{jk}
Algorithme d'intervalisation: MEDMW 8-tuple

Carte 9: Analyse interponctuelle (carte à cloisons)
Corpus: sous-corpus *phonétique* (1117 cartes de travail)
Indice de distance: IRDpk
Algorithme d'intervalisation: MEDMW 8-tuple

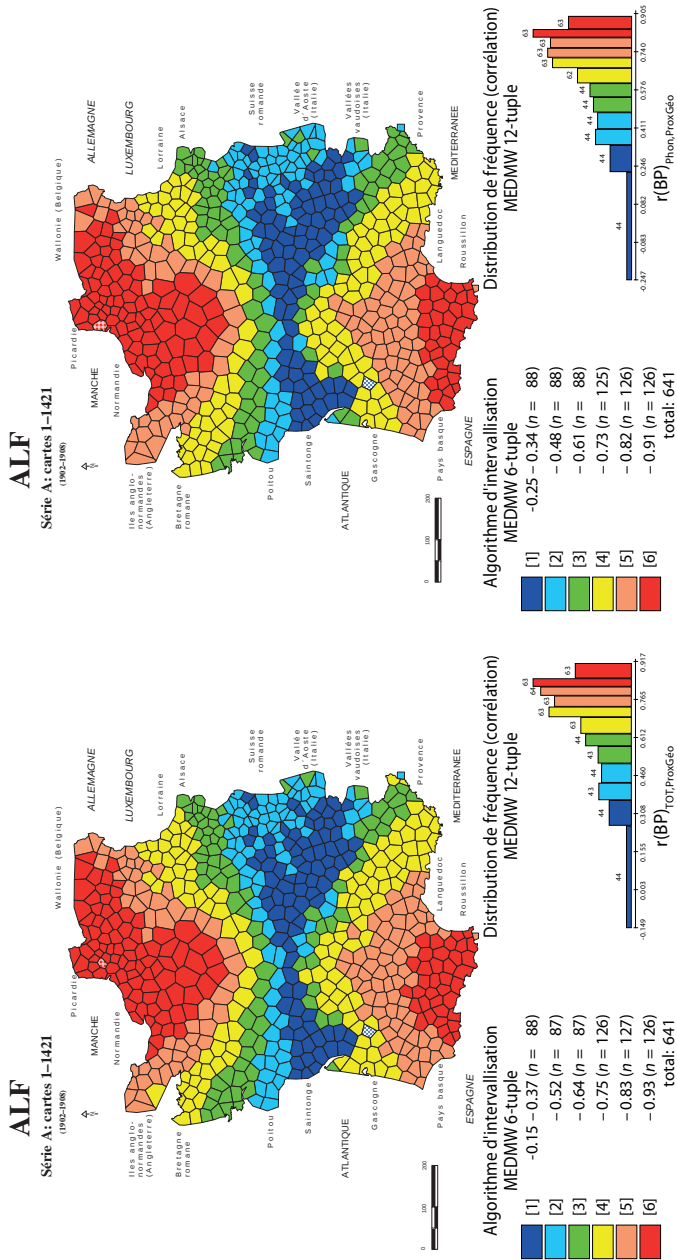


Carte 11 : Synopse des scores du « coefficient d'asymétrie de Fisher » (CAF)
Corpus: sous-corpus A (entre 2 et 5 taxats par carte de travail):
536 cartes de travail
Indices de similarité: IRIjk
Algorithme d'intervallisation: MEDMW 8-tuple



Carte 13: Synopse des scores du « coefficient d'asymétrie de Fisher » (CAF)
Corpus: sous-corpus *phonétique* (1117 cartes de travail)
Index de similarité: IRI_{ijk}
Algorithme d'intervallisation: MEDMW 8-tuple

Carte 14: Synopse des scores du « coefficient d'asymétrie de Fisher » (CAF)
Corpus: sous-corpus *lexical* (465 cartes de travail)
Indice de similarité: IRI_{jk}
Algorithme d'intervallisation: MEDMW 8-tuple



Carte 15: Profil choroplèthe d'une carte à corrélations

Matrices carrées corrélées:

- matrice de similarité (selon IRI_{jk})
- corpus *intégral* (1681 cartes de travail)
- matrice de proximité (selon le théorème de Pythagore)

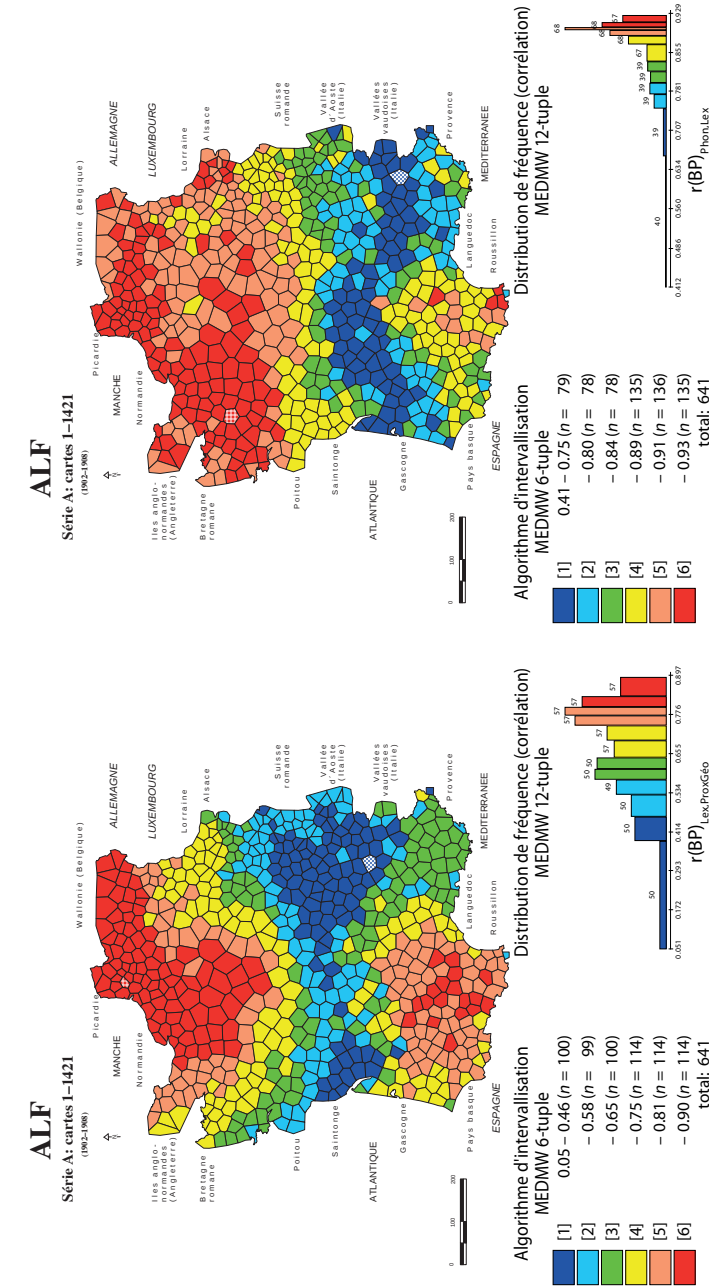
Algorithme d'intervallisation: MEDMW 6-tuple

Carte 16: Profil choroplèthe d'une carte à corrélations

Matrices carrées corrélées:

- matrice de similarité (selon IRI_{jk})
- sous-corpus *phonétique* (1117 cartes de travail)
- matrice de proximité (selon le théorème de Pythagore)

Algorithme d'intervallisation: MEDMW 6-tuple



Carte 17: Profil choroplèthe d'une carte à corrélations

Matrices cartées corrélées:

- matrice de similarité (selon IRI_{jk})
sous-corpus *lexical* (465 cartes de travail)
- matrice de proximité (selon le théorème de Pythagore)

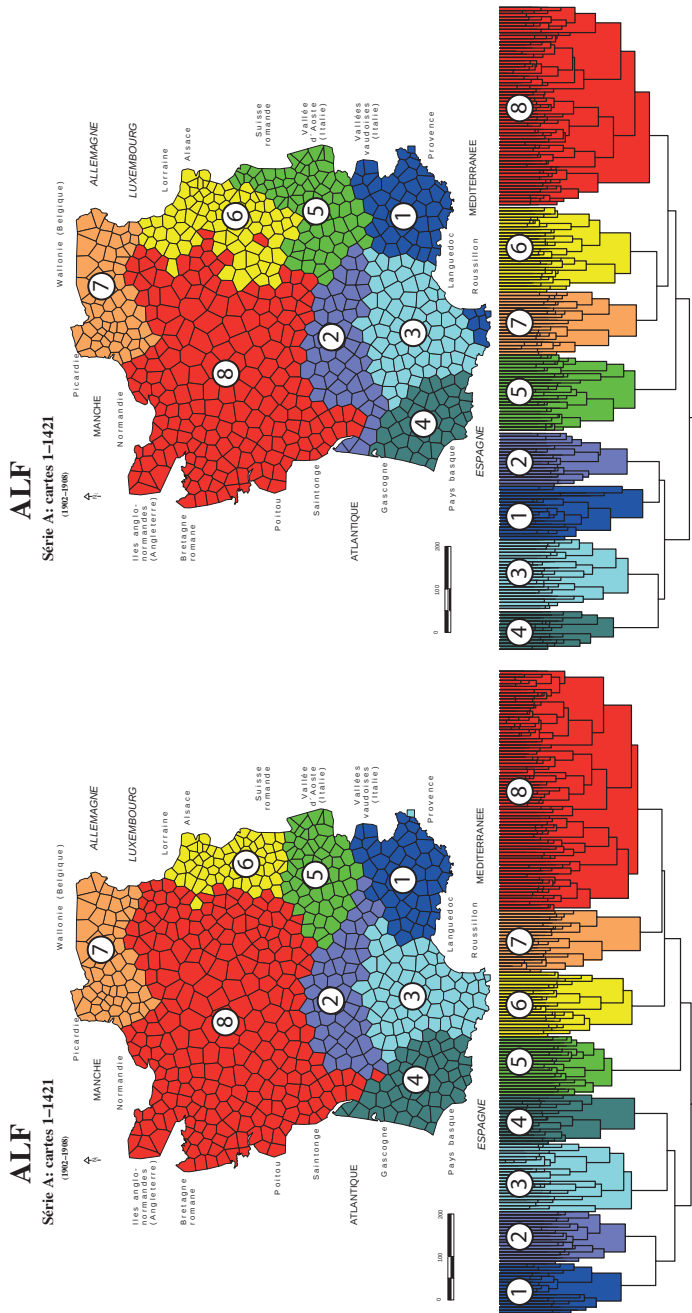
Algorithme d'intervallisation: MEDMW 6-tuple

Carte 18: Profil choroplèthe d'une carte à corrélations

Matrices de similarité (selon IRI_{jk}) corrélées:

- sous-corpus *phonétique* (1117 cartes de travail)
- sous-corpus *lexical* (465 cartes de travail)

Algorithme d'intervallisation: MEDMW 6-tuple



Carte 19: Analyse dendrographique (selon l'algorithme de Joe Ward Jr.)

- En bas: arbre (classification en 8 dendrèmes)
- En haut: spatialisé de l'arbre (classification en 8 chorèmes)

Indice de similarité: IRI_{jk}
Corpus: sous-corpus *phonétique* (1117 cartes de travail)

Carte 20: Analyse dendrographique (selon l'algorithme de Joe Ward Jr.)

- En bas: arbre (classification en 8 dendrèmes)
- En haut: spatialisé de l'arbre (classification en 8 chorèmes)

Indice de similarité: IRI_{jk}
Corpus: sous-corpus *lexical* (465 cartes de travail)