BRÈVE INTRODUCTION AUX PROBLÈMES ET MÉTHODES DE LA DIALECTOMÉTRIE

HANS GOEBL

Abstract. This article provides a brief introduction to the problems and methods of dialectometry (fr. "dialectométrie", ger. "Dialektometrie"), such as it is practiced in Salzburg. Introduced in 1973 by Jean Séguy, we use the term "dialectométrie" to refer to the global analysis of the geolinguistic deep structures of the data from a given linguistic atlas, using mathematical und visual methods borrowed from statistics, numerical classification and modern computer assisted quantitative cartography. Taking the results of a dialectometric analysis ("dialectometrization") of the French linguistic atlas ALF ("Atlas linguistique de la France", published by Jules Gilliéron and Edmond Edmont between 1902 and 1910) as an example, our article examines the different steps involved in a complete dialectometrization – according to the Salzburg standards – of the original maps of the ALF, from the establishment of the data matrix and the subsequent quantitative measuring of the linguistic similarities (and distances) between the 641 inquiry points on the ALF grid to the computer-aided visualization of the respective numerical results. We present the following dialectometric analyses in brief: similarity maps, parameter maps, interpoint maps (honeycomb and beam maps), correlation maps and dendrographic schemes (trees). Our article is accompanied by two explicative figures in black and white and 12 coloured choropleth and isarithmic maps which have been produced using the computer program VDM ("Visual DialectoMetry") created and tested in Salzburg by our research assistant Edgar Haimerl between 1997 and 2000.

1. REMARQUES PRÉLIMINAIRES

La dialectométrie (DM) est une méthode quantitative pour l'analyse et l'étude des structures de profondeur contenues (mieux: cachées) dans les données d'un atlas linguistique (AL) classique. Elle est partant un prolongement direct de la géographie linguistique traditionnelle. Le terme de « dialectométrie » a été forgé par le dialectologue toulousain Jean Séguy en 1973 (cf. Séguy 1973, 1). Les méthodes dialectométriques à proprement parler se sont rapidement développées et diversifiées par la suite, d'abord au sein de la philologie romane[1] et ensuite aussi bien au-delà[2] si tant est que, de nos jours, l'on est en droit de parler de l'existence de plusieurs « écoles » de DM. Mis à part nos propres apports en matière de DM –

[1] Voir, entre autres, les contributions de Melis (1984), Verlinde (1988), Álvarez/Dubert García/Sousa Fernández (2006) et de Viaplana (1999).

[2] Citons, à ce propos, les travaux des germanistes Hummel (1993) et Schiltz (1996).

RRL, **LIII**, *1–2*, p. 87–106, Bucureşti, 2008

BDD-A310 © 2008 Editura Academiei
Provided by Diacronia.ro for IP 216.73.216.25 (2026-11-07 16:04:16 UTC)

et qui sont étroitement liés aux noms des Universités de Ratisbonne/Regensburg (1973–1982) et de Salzbourg (1982 s.) – et ceux de mes disciples[3] de Salzbourg et d'Innsbruck, citons à ce propos l'école-DM néerlandaise de Groningen (avec John Nerbonne et Wilbert Heeringa)[4] et l'école-DM américaine d'Athens (Georgia) (avec William Kretzschmar)[5]. Cette diversification internationale des approches heuristiques et aussi des méthodes géostatistiques s'est avérée très féconde. Il est donc tout à fait légitime de parler aussi de l'existence d'une « École-DM de Salzbourg » (EDMS). Dans les lignes qui suivent, il sera uniquement question des conceptions théoriques et des méthodes statistiques et cartographiques utilisées par l'EDMS.

1.1. Quelques prérequis théoriques de l'EDMS

Le fondement théorique du travail pratique de l'EDMS consiste dans une vision particulière des données d'un AL. L'EDMS envisage la structuration dialectale des grands domaines linguistiques de notre planète comme le résultat de l'activité ordonnatrice de l'homme moyennant la création de ressemblances et de dissemblances géolinguistiques de toute sorte, lui permettant de s'approprier ainsi le territoire qu'il habite par voie symbolique[6]. C'est ainsi qu'est né le concept de « l'aménagement basilectal de l'espace par l'HOMO LOQUENS » qui peut être considéré comme le théorème de base de L'EDMS. Le concept de « l'aménagement basilectal de l'espace par l'HOMO LOQUENS » présuppose en outre que l'activité aménageuse de l'homme obéit à certaines régularités – pour ne pas dire « lois » – systémiques majeures[7] dont le dégagement constitue une des finalités premières de la DM.

1.2. La position méthodique de l'EDMS

Evidemment, la DM constitue, à l'instar de beaucoup d'autres sciences quantitatives métrisées, une discipline hybride, située au point de rencontre entre la géolinguistique, la statistique (avec la taxonomie ou classification numérique) et la cartographie quantitative numérisée. Ceci dit, il faut néanmoins souligner que les

[3] Voir à ce sujet les travaux de Roland Bauer de Salzbourg (dialectométrisation des données de l'atlas ladin ALD-I) et de Paul Videsott d'Innsbruck (analyse-DM de documents médiévaux de l'Italie septentrionale et création du mot et de la chose de l'« onomatométrie »).

[4] Voir, entre autres, Heeringa 2004, Heeringa/Nerbonne 2001 et Nerbonne/Heeringa 2001.

[5] Voir Nerbonne/Kretzschmar 2003 et 2006.

[6] L'EDMS admet en outre que l'instrument de l'AL constitue un excellent outil pour la représentation empirique de l'aménagement basilectal de l'espace par l'HOMO LOQUENS.

[7] Les lois spatiales découvertes par la DM constituent le pendant des « lois phonétiques » (« Lautgesetze ») découvertes par les néogrammairiens de la fin du XIX^{e} siècle. Alors que les « lois phonétiques » s'inscrivent dans la dimension du temps, les lois spatiales dégagées par la DM s'inscrivent dans celle de l'espace.

BDD-A310 © 2008 Editura Academiei
Provided by Diacronia.ro for IP 216.73.216.25 (2026-11-07 16:04:16 UTC)

finalités épistémologiques de la DM s'alignent avant tout sur celles de la linguistique diatopique classique et que, de ce fait, la statistique et la cartographie quantitative détiennent toujours – au sein de l'EDMS – une fonction carrément auxiliaire.

L'EDMS se sert donc des apports de ces deux disciplines quantitatives uniquement dans le but de faire avancer nos connaissances linguistiques sur le fonctionnement systémique (et global) de réseaux diatopiques définis par un nombre N aussi grand que possible de points d'enquête et un nombre p non moins important d'attributs linguistiques, relevés, si possible, sur le terrain au cours d'enquêtes dialectologiques directes. Dans l'optique de l'EDMS, la DM est une discipline linguistique résolument orientée vers l'étude de l'espace géographique dont l'utilité pour la linguistique générale consiste, entre autres, à ne pas être confondue avec (ou contaminée par) la sociolinguistique tout court ou certaines de ses sous-disciplines.

2. DE L'ATLAS LINGUISTIQUE À LA MATRICE DE DONNÉES

Pour une représentation globale des méthodes utilisées par l'EDMS voir la figure 1. Le premier maillon de la chaîne-DM se réfère à l'analyse (ou « taxation ») des données brutes de l'AL destiné à être dialectométrisé. Pour ceci nous avons choisi l' «Atlas linguistique de la France» (ALF) dont la dialectométrisation a été effectuée à Salzbourg entre 1997 et 2000[8]. C'est au cours de ce travail méticuleux qu'a été également mis sur pied le logiciel VDM (« Visual DialectoMetry ») par notre collaborateur de recherche Edgar Haimerl, qui permet d'effectuer confortablement tous les calculs statistiques et les visualisations respectives[9]. C'est d'ailleurs à l'aide de ce logiciel qu'ont été confectionnées les 12 cartes de cet article.

L'établissement de la matrice de données qui est un schéma à double entrée avec les dimensions N (points d'enquête) fois p (attributs linguistiques ou « cartes de travail ») présuppose tout d'abord l'intervention du dialectologue averti[10]. C'est à lui qu'incombe l'analyse linguistique détaillée du contenu des planches originales de l'ALF suivant les critères habituels de la phonétique, morphologie et lexicologie historiques. Les cartes 1 et 2 en constituent des échantillons représentatifs.

[8] Pour une ample information cf. nos contributions de 2000, 2002 et 2003.

[9] Le logiciel VDM est à la libre disposition de tous les chercheurs qui en font la demande. L'initiation au maniement pratique de VDM peut être faite au cours d'un stage de 2 à 3 jours à effectuer à Salzbourg.

[10] Il y a là une différence foncière face à l'école dialectométrique de Groningen qui, en dernière analyse, écarte pour l'établissement de la matrice de similarité l'expertise humaine (c.-à-d. du linguiste) tout en se servant du calcul automatisé des distances interdialectales à l'aide de la « distance de Levenshtein ». Pour ce faire, les spécialistes de Groningen recourent à une matrice des données brutes de l'AL à dialectométriser entièrement informatisée : cf. à ce sujet Nerbonne/ Kretzschmar 2003 et surtout Heeringa 2004.

BDD-A310 © 2008 Editura Academiei
Provided by Diacronia.ro for IP 216.73.216.25 (2026-11-07 16:04:16 UTC)

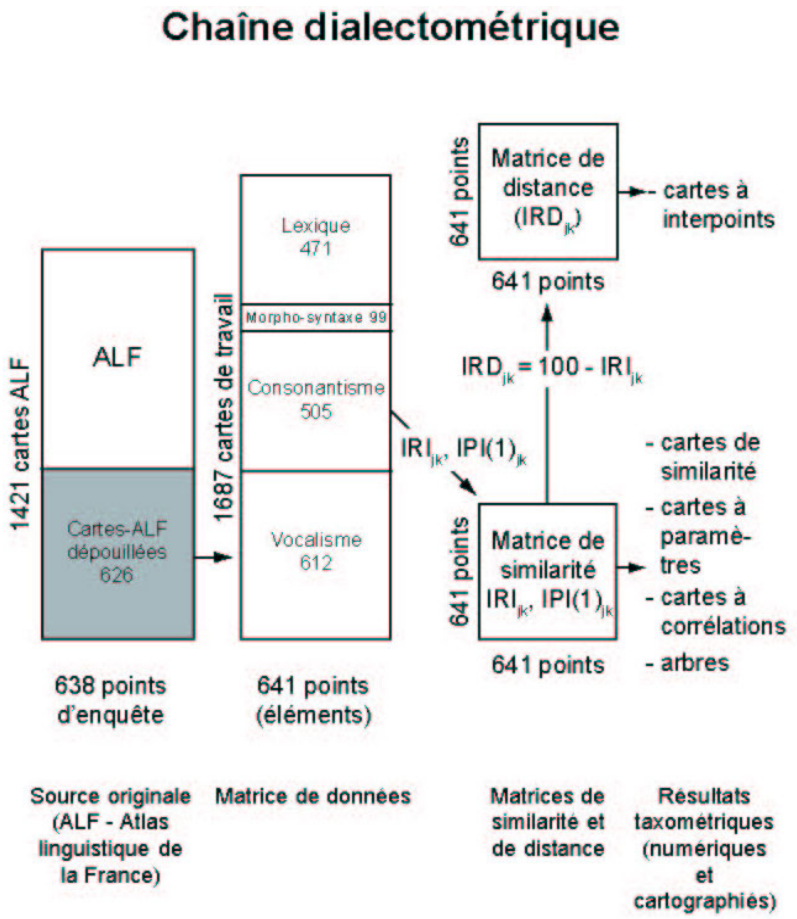


Fig. 1 – Schéma de l'enchaînement des méthodes dialectométriques présentées dans cet article.

La carte 1 renseigne sur la répartition spatiale des succédanés de -T- latin intervocalique (dans l'étymon latin MATÚRU) au sein du réseau de l'ALF. Le profil choroplèthe en montre très clairement l'opposition entre le sud de la France, linguistiquement conservateur (par le biais du maintien du -d- roman intervocalique), et le nord et l'est, linguistiquement plus innovateurs (moyennant l'amuïssement complet du -T- latin intervocalique)[11].

Evidemment, les données de la carte-ALF 891 (*mûr, mûre*) permettent aussi des analyses similaires pour les autres nexus de l'étymon latin MATÚRU, en particulier pour les succédanés de l'A prétonique, de l'U long latin accentué en position libre, de l'R posttonique en position intervocalique et de l'U final. De cet

[11] Il va de soi que le caractère de la carte 1 est double : a) linguistique: par la saisie de différents résultats linguistiques, b) extralinguistique: par la saisie de zones plus ou moins dynamiques.

BDD-A310 © 2008 Editura Academiei
Provided by Diacronia.ro for IP 216.73.216.25 (2026-11-07 16:04:16 UTC)

exemple, il appert qu'il est possible de tirer, au cours d'une analyse-DM taxatoire, plusieurs cartes de travail *phonétiques* d'une seule planche-ALF originale[12].

Il va de soi que la taxation multiple ne peut pas être appliquée aux planches-ALF *lexicales*: voir à ce sujet la carte 2 qui montre la répartition spatiale des dénominations galloromanes de la jument (selon la carte-ALF 736 *la jument*). Soit dit en passant, de telles analyses ont une longue tradition au sein de la philologie romane, en commençant par les apports de l'onomasiologie classique du type « Wörter und Sachen », en passant par les cartes linguistiques présentées dans tous les grands manuels de linguistique ou de dialectologie romane pour aboutir à une myriade d'articles spécialisés accompagnés d'une ou de plusieurs cartes géolinguistiques effectuées à l'aide d'une « carte muette » de l'AL examiné. Par le biais de la taxation manuelle exécutée par un expert, l'EDMS se sert donc de la plénitude d'innombrables expériences analytiques en matière de géographie linguistique[13].

Voici le bilan chiffré[14] de la dialectométrisation de l'ALF effectuée à Salzbourg entre 1997 et 2000 :

Planches originales-ALF dépouillées	626 (= 44% de l'effectif total de la série A de l'ALF)
Cartes de travail analysées :	
– pour la phonétique	1117
– pour la morphologie	99
– pour le lexique	471
Nombre total des cartes de travail analysées	1687

Du côté p (nombre des attributs), notre matrice de données comprend donc 1687 lignes (ou vecteurs-attributs) alors que, du côté N (nombre des points d'enquête), elle consiste en 641 colonnes (ou vecteurs-points d'enquête). Le chiffre 641 s'explique par le fait que le réseau sur lequel reposent nos recherches, comprend, outre les 638 points-ALF originaux, aussi 3 points factices, à savoir le français standard (P. 999, localisé en pleine Ile-de-France), l'italien standard (P. 998, localisé à l'extrême pointe orientale de la Provence) et le catalan standard

[12] Une présupposition élémentaire de cette espèce de taxation est que la planche-ALF en question soit « mononyme ». Ceci signifie qu'elle doit disposer, théoriquement, de 638 formes romanes issues d'un seul étymon latin (ou autre).

[13] Ajoutons que ce genre de travail correspond parfaitement à ce qui avait été fait par tous les grands maîtres de la géographie linguistique classique au cours de l'établissement de cartes typisées d'un AL donné. Pour un renouveau de cette tradition quelque peu tombée en désuétude, voir l'excellent traité de Brun-Trigaud/Le Berre/Le Dû (2005) qui détient plus de 500 cartes-ALF typisées (en couleurs).

[14] Voir aussi les figures 2 dans Goebl 2002 (11) et 2003 (67).

BDD-A310 © 2008 Editura Academiei
Provided by Diacronia.ro for IP 216.73.216.25 (2026-11-07 16:04:16 UTC)

(P. 997, localisé au sud du Roussillon). L'insertion du français comme dialecte artificiel s'est avérée particulièrement révélatrice parce que nous avons pu étudier ainsi les multiples aspects de la francisation capillaire de la Galloromania.

La matrice de données constitue donc, en tant que représentation taxatoire de pratiquement la moitié des données brutes de l'ALF, une source empirique préalablement ordonnée très vaste et non moins complexe qui, malgré sa nature *qualitative*, se prête facilement à un dépouillement ultérieur par voie *quantitative*.

3. DE LA MATRICE DE DONNÉES AUX MATRICES DE SIMILARITÉ ET DE DISTANCE

Le dépouillement quantitatif de la matrice données obéit, en ligne générale, aux préceptes de la classification (ou taxonomie) numérique[15] selon lesquels il s'agit de mesurer la similarité (ou ressemblance) entre deux vecteurs-attributs à l'aide d'un indice de similarité approprié. Evidemment, il faut s'entendre, à cet effet, sur le concept de similarité, pour la saisie de laquelle la classification numérique internationale offre de nombreuses possibilités. Il est donc nécessaire que le géolinguiste-dialectométricien s'en serve tout en respectant sa propre acception de la similarité interdialectale. C'est ainsi que nous avons défini, il y a belle lurette, l'« Indice Relatif d'Identité » (IRI_{jk})[16]. L'IRI_{jk} qui mesure le pourcentage des co-identités taxatoires par paires, s'est révélé comme étant très utile en matière de DM. Ceci n'empêche que l'EDMS utilise, de nos jours, aussi d'autres indices de similarité dont chacun dispose d'un champ d'application particulier[17].

Toutes les cartes quantitatives (no. 3-12) de cet article reposent sur des calculs de similarité effectués à l'aide de l'IRI_{jk}.

Le résultat de l'application de l'IRI_{jk} aux N vecteurs-points d'enquête de la matrice de données, est une matrice de similarité aux dimensions N^2 qui contient le total de toutes les relations de similarités calculables entre les N points d'enquête[18]. Une des particularités saillantes de l'EDMS consiste dans le dépouillement cartographique multiple de la matrice de similarité, réalisé uniquement selon les besoins et les nécessités de la géographie linguistique à proprement parler.

[15] Voir à ce sujet les grands manuels de Bock 1974 (allemand), Chandon/Pinson 1981 (français) et Sneath/Sokal 1973 (anglais).

[16] Pour plus ample information voir les paragraphes respectifs dans nos contributions de 1981 (en français : 357s.), 1982 (en allemand : 23s.) et 1984 (en allemand : I, 74s.).

[17] Voir à ce sujet les discussions (étayées par des cartes appropriées) présentées dans Goebl 1987, 2002 et 2003.

[18] Par le biais d'une transformation très simple (IRI_{jk} = 100 + IRD_{jk} [« Indice Relatif de Distance »]) l'on passe d'ailleurs facilement de la matrice de similarité à la matrice de distance dont on a besoin pour l'étude des distances interponctuelles: voir le paragraphe 7 et la carte 9.

BDD-A310 © 2008 Editura Academiei
Provided by Diacronia.ro for IP 216.73.216.25 (2026-11-07 16:04:16 UTC)

4. NORMES ET CONVENTIONS TECHNIQUES DE LA MISE EN CARTE DES CALCULS-DM

Le réseau de l'ALF a été polygonisé suivant les préceptes de la géométrie de Delaunay-Voronoï[19]. Il s'agit là d'une procédure standard de la cartographie quantitative internationale dont l'utilité a été décrite maintes fois: cf. Okabe/ Boots/Sugihara 1992, *passim*. Il en résulte deux sortes de fond de carte: a) un fond de carte à plages polygonales, dit « carte choroplèthe », b) un fond de carte à grille triangulaire, dit « carte isarithmique ». C'est sur le principe des cartes choroplèthes (coloriées) que reposent les cartes quantitatives 3-8 et 11-12. La carte 9 en revanche, utilise le fond de carte polygonal d'une façon linéaire, alors que la syntaxe iconique de la carte 10 se base directement sur la triangulation du réseau-ALF. Le nombre des côtés de polygone (et de triangle) s'élève à 1791: voir les légendes numériques des cartes 9 et 10.

Le grand défi visualisateur que l'EDMS a relevé de bonne heure, consiste à représenter la variation numérique calculée par une variation iconique analogue et facilement saisissable par l'œil de l'observateur. Pour atteindre ce but, l'EDMS se sert de deux moyens: a) de la visualisation par les couleurs de l'arc-en-ciel (ou du spectre solaire), b) d'algorithmes de visualisation appropriés qui répartissent les scores numériques calculés en un nombre adéquat de classes iconiques (ou chromatiques). Là encore, l'EDMS s'est toujours servi d'expériences internationales et interdisciplinaires, en l'occurrence de celles de la cartographie statistique[20].

Les cartes présentées dans cet article reposent sur 6, 8 et 10 intervalles chromatiques et sur deux algorithmes d'intervallisation: MINMWMAX et MEDMW[21]. Le but des algorithmes MINMWMAX et MEDMW est de répartir les scores numériques à visualiser des deux côtés de la moyenne arithmétique respective si bien que les couleurs « chaudes » (étalées entre le rouge et le jaune) correspondent aux scores situés au-dessus de la moyenne arithmétique alors qu'aux couleurs « froides » (allant du bleu foncé au vert) font écho les scores situés au-dessous de la moyenne arithmétique.

Chacune des cartes quantitatives dispose en outre d'une légende numérique et d'un histogramme[22] qui renseignent, respectivement, sur les détails statistiques de l'intervallisation choisie et sur la nature stochastique de la distribution de similarité (ou de distance) mise en carte.

[19] Pour plus de détails cf. Goebl 1981, 363s. et 1984, I, 90s.

[20] Cf. à ce sujet le manuel classique de Dickinson 1973.

[21] Pour une description détaillée des algorithmes MINMWMAX et MEDMW cf. Goebl 1984, I, 93s. et 1987, passim.

[22] Pour le mode de construction des histogrammes cf. Goebl 1981, 363 ainsi que 1984, I, 97s. La courbe gaussienne superposée aux histogrammes n'a qu'une valeur indicative. Elle a été calculée à partir des paramètres suivants de la distribution de fréquence empirique respective : minimum, maximum, moyenne arithmétique, déviation-standard (écart-type).

BDD-A310 © 2008 Editura Academiei
Provided by Diacronia.ro for IP 216.73.216.25 (2026-11-07 16:04:16 UTC)

5. LA MENSURATION DE LA SIMILARITÉ INTERDIALECTALE

Toute matrice de similarité consiste en N vecteurs ou distributions de similarité dont chacune comprend N scores de similarité. Pour la dialectométrisation de l'ALF, il en résulte 641 distributions de similarité dont chacune se réfère à un des 641 points-ALF pris en considération. Il est donc indiqué d'en tirer 641 cartes choroplèthes dont chacune dispose d'un point de référence (ou de repère) et de 640 scores numériques oscillant entre 100 et 0 %. La leçon linguistique que le géolinguiste peut tirer des cartes de similarité, est très riche: c'est pourquoi les cartes de similarité constituent le pivot heuristique central de l'EDMS.

5.1. Présentation et interprétation des cartes 3-5

5.1.1. Présentation et interprétation de la carte 3

Il s'agit d'une carte de similarité « typiquement » gasconne[23]. Le pourtour des zones rouges et orange circonscrit grosso modo le domaine linguistique gascon. L'essentiel du message iconique qui émane du profil choroplèthe de la carte 3, réside cependant moins dans la détection de domaines dialectaux particuliers que dans la découverte et l'analyse de l'agencement foncièrement régulier du gradient spatial. Cette constatation vaut d'ailleurs pour toute carte de similarité.

L'agencement hautement régulier des différents gradients spatiaux représente l'émanation directe de la composante *diatopique* du changement linguistique général dont on ne connaît d'habitude que la composante *diachronique* sous la forme des fameuses « lois phonétiques » (« Lautgesetze »), découvertes et décrites par les Néogrammairiens de Leipzig à la fin du XIX^{e} siècle.

Le changement linguistique *diatopique* s'étend donc harmonieusement le long des deux dimensions de l'espace, exactement à l'instar de ce qui arrive, pour le changement linguistique *diachronique*, le long de la seule dimension de l'axe du temps. C'est là une des découvertes les plus lourdes de conséquences qui ont été faites par la DM en général et l'EDMS en particulier.

Ajoutons que la bipartition de la Galloromania en un secteur teinté de couleurs « chaudes » et un autre teinté de couleurs « froides », se présenterait de la même manière aussi à partir d'autres points de repère « sudistes ». Signalons encore deux autres particularités de la carte 3 : a) que toute carte de similarité peut être interprétée aussi par voie *analogique*. C'est ainsi que les polygones rouges sont censés marquer la présence des « meilleurs amis » de la « personne de repère » 676, et les polygones en bleu foncé celle des « pires ennemis » (ou des « antipodes ») de la « personne de référence »; b) l'asymétrie accusée de

[23] Pour tous les aspects de l'interprétation intra- et extralinguistique des cartes de similarité cf. Goebl 1981, 368 s. ainsi que 1984, I, 100s.

BDD-A310 © 2008 Editura Academiei
Provided by Diacronia.ro for IP 216.73.216.25 (2026-11-07 16:04:16 UTC)

l'histogramme qui caractérise la position excentrique (ou isolée) du dialecte gascon de Riscle (P.-ALF 676) par rapport au reste du réseau de l'ALF. Nous allons voir, par la suite, que l'étude comparative du taux de symétrie des distributions de similarité peut assumer une signification capitale en matière de géolinguistique.

5.1.2. Présentation et interprétation de la carte 4

Le point de référence correspond au français standard (P. 999) si bien que le profil choroplèthe de cette carte reflète l'état de la francisation de la Galloromania tout court. L'agencement spatial des polygones rouges et orange – qui correspondent aux scores de similarité les plus grands – est typique de tout processus de diffusion spatiale, qu'elle soit continue ou discontinue (c'est-à-dire par « parachutage »). Evidemment, la répartition des couleurs « chaudes » et « froides » est complémentaire à celle de la carte 3. À noter en outre l'amoncellement cunéiforme de plusieurs polygones verts le long du couloir rhodanien qui s'explique par l'infiltration pluriséculaire de traits linguistiques oïliques dans le domaine occitan.

5.1.3. Présentation et interprétation de la carte 5

Alors que les cartes 3 et 4 reposent sur le total de notre matrice de données et représentent, de ce fait, des profils d'identité d'ordre « général », la carte 5 ne s'appuie que sur les cartes de travail de nature lexicale dont le nombre (471) est d'ailleurs suffisamment grand pour permettre des jugements comparatifs fiables. Comme les normes-DM de la carte 5 correspondent exactement à celles de la carte 4, une comparaison directe entre ces deux cartes de similarité est possible. Or, l'on constate que l'agencement du gradient spatial des deux cartes est légèrement différent tout en obéissant aux mêmes principes: valeurs de similarité maximales aux alentours du point de référence, valeurs de similarité minimales à la périphérie (i. e. aux antipodes) de la carte, étalement spatial graduel des similarités intermédiaires (intervalles 2-5). L'on peut donc dire que l'aménagement de l'espace par les lexèmes ne s'écarte pas considérablement de celui effectué par les autres unités linguistiques prises en considération. Nous reviendrons sur ce sujet plus tard.

5.1.4. Présentation et interprétation de la carte 6

Du point de vue cartographique, la carte 6 est une véritable carte de similarité ; du point de vue linguistique, il n'en est rien. C'est qu'elle visualise non pas des similarités linguistiques relatives au point de repère 999 (français standard), mais des proximités géographiques ou mieux encore : euclidiennes. Les proximités géographiques ont été calculées à l'aide du logiciel VDM (épaulé par le logiciel MapInfo) – à partir des coordonnées cartésiennes (en *x* et *y*) des 641 points-ALF et mises en carte, par la suite, à l'instar des scores de similarité.

BDD-A310 © 2008 Editura Academiei
Provided by Diacronia.ro for IP 216.73.216.25 (2026-11-07 16:04:16 UTC)

Par une comparaison visuelle d'une carte de similarité et de la carte de proximité respective, il est possible de juger, d'une manière rapide mais tout à fait superficielle voire subjective, de la plus ou moins grande corrélation spatiale qui existe entre les aménagements dialectal et euclidien du même espace. Ajoutons que le traitement statistique (et non plus visuel) de ce problème constitue la base de la DM corrélative (cf. le paragraphe 8).

6. DE LA MENSURATION DE LA SIMILARITÉ AUX CARTES À PARAMÈTRES : DEUX SYNOPSES DU « COEFFICIENT D'ASYMÉTRIE DE FISHER » (CAF)

Dans le paragraphe 5.1.1., nous avons déjà évoqué le rôle classificateur que pourrait jouer la considération systématique du taux d'asymétrie de différentes distributions de similarité. La même remarque vaut d'ailleurs aussi pour les autres paramètres statistiques des mêmes distributions de similarité, telles que les valeurs minimales et maximales, la moyenne arithmétique, la médiane, la déviation-standard (ou écart-type), etc.

Pour N (= 641) distributions de similarité calculées à partir d'une matrice de données examinée, il est donc possible d'en extraire N (= 641) paramètres caractéristiques préalablement définis, de les réunir en synopse sous la forme d'une nouvelle distribution de fréquence et de la visualiser par la suite suivant les normes habituelles. Les cartes 7 et 8 en constituent des exemples éloquents.

En ce qui concerne la valeur linguistique à proprement parler des paramètres statistiques mis en synthèse, il importe, d'entrée de jeu, que le dialectométricien en ait arrêté ses idées d'une façon aussi précise que possible. Quant à la synthèse des scores du CAF, son importance linguistique peut être décrite comme suit : le CAF saisit le « brassage (ou compromis) linguistique » (all. « Sprachausgleich »). On entend par là la présence concomitante, à un endroit donné du réseau examiné, de traits linguistiques de petite, moyenne et grande diffusion spatiale[24]. Evidemment, le mélange entre ces trois sortes de traits linguistiques peut varier considérablement d'une région à l'autre.

Précisons qu'un pourcentage très élevé de traits linguistiques à *grande* diffusion spatiale est la conséquence directe d'une forte irradiation linguistique et que l'inverse est vrai pour un pourcentage considérable de traits linguistiques à *petite* diffusion spatiale: il s'agit là d'une zone linguistiquement plutôt isolée ou conservatrice.

Or, les valeurs *négatives* du CAF (représentées par les couleurs « froides ») renvoient à des zones où, comme conséquence d'un brassage linguistique très intense, le pourcentage de traits linguistiques à grande diffusion spatiale est très

[24] Pour une description détaillée de la nature statistique et de l'importance linguistique du CAF cf. Goebl 1981, 394s. et 1984, I, 150s.

BDD-A310 © 2008 Editura Academiei
Provided by Diacronia.ro for IP 216.73.216.25 (2026-11-07 16:04:16 UTC)

haut, alors que les valeurs *positives* du CAF (symbolisées par les couleurs « chaudes ») se réfèrent à des régions restées plutôt à l'écart du brassage linguistique pan-galloroman.

La logique spatiale de la carte 7 est très claire: dans le nord, il existe une grande configuration circulaire en bleu foncé qui entoure le domaine d'Oïl en entier. La même remarque vaut pour le domaine francoprovençal dans le sud-est. Le domaine d'Oc, par contre, est caractérisé par l'implantation ponctuelle de plusieurs zones en rouge et orange (Gascogne, Languedoc, Provence) entre lesquelles s'étendent des couloirs de transition en ocre et jaune. Le message linguistique de cette architecture polygonale bipolaire est claire : la configuration circulaire dans le nord est le résultat de la diffusion radiale des parlers directeurs du centre du domaine d'Oïl. Les zones périphériques en bleu foncé accusent l'action d'un brassage linguistique très intense, alors que les zones centrales en vert constituent des régions « pacifiées », caractérisées par un dynamisme irradiateur ralenti (ou « saturé »).

La configuration circulaire mineure en bleu foncé qui entoure le domaine francoprovençal, renvoie aux multiples conflits et contacts linguistiques que devait subir la vieille latinité de Lugdunum/Lyon avec les parlers d'Oïl (de type franc-comtois) dans le nord, et les parlers d'Oc (de type provençal) dans le sud. La répartition spatiale antagoniste issue de configurations *linéaires* (en bleu foncé) dans le nord et de configurations *ponctuelles* (en rouge) dans le sud, s'explique par la grande différence qui existe, entre les domaines d'Oc et d'Oïl, du point de vue des dynamismes diachroniques respectifs. Le domaine d'Oïl expansionniste s'oppose (et le faisait depuis longtemps) au domaine d'Oc conservateur alors que le domaine francoprovençal semble s'étioler entre les feux croisés tant du nord que du sud.

Soulignons encore que la carte 7 repose uniquement sur des données-ALF d'ordre *phonétique* alors que la carte 8 a été calculée à partir de données *lexicales*. En comparant les deux profils choroplèthes l'on constate que les dynamismes géolinguistiques les plus intenses (voir les polygones en bleu foncé) se concentrent, en ce qui concerne l'expansion des *phonèmes* (voir la carte 7), dans l'est, alors que le même phénomène se manifeste, pour l'expansion des *lexèmes* (voir la carte 8), au beau milieu de la Galloromania.

7. LA DIALECTOMÉTRIE INTERPONCTUELLE : CARTE À CLOISONS ET CARTE À RAYONS

La méthode des interpoints constitue, en dernière analyse, la continuation-DM des fameuses synthèses isoglottiques qui représentaient, pour un laps de temps relativement long, le seul moyen cartographique pour le dépouillement synthétique des données d'un AL[25]. Il va de soi que la confection de synthèses isoglottiques se

[25] Voir à ce sujet les synthèses d'isoglosses qui avaient été faites à partir des données de l'ALF par A. Rosenqvist (1919) et K. v. Ettmayer (1924).

BDD-A310 © 2008 Editura Academiei
Provided by Diacronia.ro for IP 216.73.216.25 (2026-11-07 16:04:16 UTC)

heurtait, jadis, à beaucoup d'obstacles d'ordre surtout cartographique qui, à l'ère de la cartographie numérisée, ont complètement disparu. Le traitement-DM du problème isoglottique[26] repose sur les étapes suivantes :

- triangulation du réseau d'enquête: pour l'ALF il en résulte une grille de 1791 côtés de triangle,
- polygonisation consécutive de la grille triangulaire d'où résultent 641 polygones qui sont entourés de 1791 côtés de polygone[27],
- extraction des 1791 scores interponctuels de *similarité* (selon l'IRI_{jk}) de la matrice de *similarité* respective et leur visualisation consécutive le long de 1791 côtés de *triangle*: voir la carte 10,
- extraction des 1791 scores interponctuels de *distance* (selon l'IRD_{jk}) de la matrice de *distance* respective et leur visualisation consécutive le long de 1791 côtés de *polygone*: voir la carte 9.

Nous avons utilisé, pour les deux processus de visualisation, les conventions cartographiques suivantes:

- Pour la carte 9 (carte à *cloisons*: visualisation de *distances* interponctuelles): variation de l'épaisseur et du coloriage des côtés de *polygone*; à l'augmentation des scores de distances correspondent un accroissement de l'épaisseur des côtés de polygone et la transition du coloriage vers le bleu[28] foncé.
- Pour la carte 10 (carte à *rayons*: visualisation de *similarités* interponctuelles): variation de l'épaisseur et du coloriage des côtés de *triangle*; à l'augmentation des scores de similarité correspondent un accroissement de l'épaisseur des côtés de triangle et le déplacement du coloriage vers le rouge[29].

Evidemment, le message iconique des cartes 9 et 10 est complémentaire: alors que la carte 9 représente l'implantation spatiale de grands cloisonnements interponctuels (c'est-à-dire concernant les interstices entre deux points-ALF voisins), la carte 10, en revanche, montre la répartition spatiale de grandes plages interactives d'un côté et de zones de passage plus ou moins déconnectées de l'autre.

Soulignons encore, sur la carte 9, l'absence complète de phénomènes linéaires continus susceptibles d'être qualifiés de « frontières linguistiques ». Il en appert le caractère quelque peu trompeur de la métaphore de la frontière linguistique. C'est qu'à travers la syntaxe iconique de la carte 9, nous ne voyons que des zones ou bandes, situées dans le centre, l'est et le nord de la Galloromania, où s'amoncellent, d'une façon d'ailleurs très inégale, les valeurs maximales du cloisonnement interponctuel.

[26] Pour plus ample information cf. Goebl 1982, 49s., 1984, I, 183s. et surtout 1983, *passim*.

[27] Rappelons que l'utilité de la polygonisation d'un réseau d'atlas pour le tracement combiné d'isoglosses a déjà été démontrée par le germaniste Carl Haag en 1898 (!!).

[28] Association symbolique: la couleur bleue renvoie à la « froideur » des relations interpersonnelles.

[29] Association symbolique: la couleur rouge renvoie à la « chaleur » des relations interpersonnelles.

BDD-A310 © 2008 Editura Academiei
Provided by Diacronia.ro for IP 216.73.216.25 (2026-11-07 16:04:16 UTC)

Signalons en outre que la complémentarité des cartes interponctuelles 9 et 10 ressort également de la symétrie parfaite des deux histogrammes respectifs et aussi du fait que de l'addition de la valeur minimale 5,40 de la légende numérique de la carte 9, et de la valeur maximale 94,60 de la légende numérique de la carte 10, résulte la valeur globale 100. C'est la conséquence directe de la formule IRI_{jk} + IRD_{jk} = 100 qui régit la relation entre les indices de similarité et de distance.

8. LA DM CORRÉLATIVE

La DM corrélative n'a été ajoutée à la panoplie méthodique de l'EDMS qu'en 2004[30]. C'est donc un outil-DM relativement jeune. Il permet l'étude comparative – par l'application de calculs de corrélation et la visualisation consécutive des résultats y ayant trait – de plusieurs sortes d'aménagement d'un territoire donné: aménagements linguistique et euclidien, phonétique et consonantique, vocalique et consonantique, etc. La DM corrélative est un outil très souple dont la vertu exploratoire dépasse encore celle des autres méthodes de l'EDMS.

Regardons encore une fois – pour une bonne intelligence de la carte 11 – les cartes 5 et 6. La perception visuelle nous dit que les deux profils choroplèthes sont « moyennement » corrélés sans que nous soyons à même de conférer à notre jugement une forme plus précise. Evidemment, le même effort de comparaison peut être accompli aussi avec les moyens de la statistique descriptive, laquelle, pour ce faire, nous fournit, entre autres, le « coefficient de corrélation de Bravais et Pearson » [r(BP)]. Les scores du r(BP) oscillent dans un espace numérique situé entre -1 (corrélation négative), 0 (corrélation inexistante) et +1 (corrélation positive ou parfaite).

En appliquant le r(BP) aux 641 paires de valeurs (IRI_{jk} et ProxGéo) des cartes 5 et 6, l'on obtient le score de corrélation de +0,799 qui, en effet, est relativement proche de la valeur maximale +1. En effectuant les mêmes calculs pour les 640 paires de valeurs qui restent, tout en se basant sur la matrice de similarité contenant toutes les cartes de travail disponibles (1687), l'on obtient 641 scores du r(BP) qui oscillent entre -0,15 et +0,93: voir la légende numérique de la carte 11. De la mise en carte consécutive de ces 641 scores de corrélation résulte un profil choroplèthe dont l'allure harmonieuse voire esthétique est vraiment surprenante au premier abord.

Comment faut-il interpréter – toujours du point de vue du (géo)linguiste – l'agencement finement articulé des plages choroplèthes rouges et bleues ? Les zones rouges correspondent à des régions où l'étalement des similarités linguistiques dans l'espace obéit dans une grande mesure aux impératifs euclidiens de ce dernier. Ceci signifie que, sur une carte de similarité quelconque, le taux de la

[30] Voir à ce sujet notamment notre article de 2005(a) où ce nouvel outil-DM a été présenté pour la première fois. Il en est question également dans Goebl 2004 et 2005b et dans nos contributions publiées en 2006 et 2007.

BDD-A310 © 2008 Editura Academiei
Provided by Diacronia.ro for IP 216.73.216.25 (2026-11-07 16:04:16 UTC)

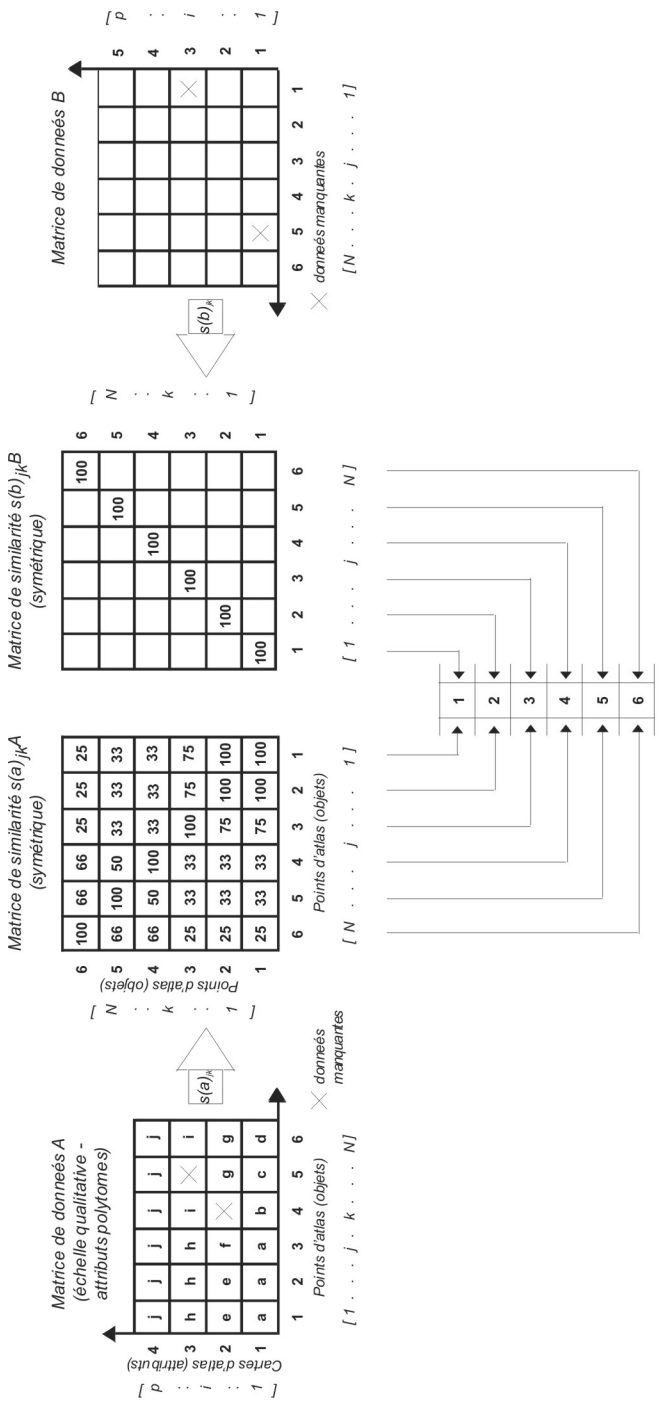


Fig. 2 – Schéma du calcul de la carte à corrélations à l'aide de deux matrices (carrées) de similarité (ou de proximité, de distance, etc.) disposant de N vecteurs d'objets.

BDD-A310 © 2008 Editura Academiei
Provided by Diacronia.ro for IP 216.73.216.25 (2026-11-07 16:04:16 UTC)

similarité linguistique décroît, avec l'éloignement du point de référence respectif, d'une façon plus ou moins régulière (c'est-à-dire statistiquement linéaire). L'inverse est vrai pour les zones teintées de bleu. C'est ici que l'étalement des similarités linguistiques dans l'espace a été libéré complètement des contraintes euclidiennes de l'espace, évidemment pour des raisons sociales et politiques de toute sorte. Ceci signifie que l'échelonnement spatial des similarités linguistiques peut ou bien dépasser les impératifs de la géométrie euclidienne ou bien rester en retrait par rapport à ces derniers. Ces écarts peuvent survenir à la suite d'un rayonnement culturel et social dans le premier des deux cas, ou représenter la conséquence d'un isolement du même ordre dans l'autre.

Le profil choroplèthe de la carte 11 doit être lu et interprété surtout dans une perspective diachronique. C'est une carte qui témoigne de l'affrontement pluriséculaire de deux forces linguistiques antagonistes, dont la plus vigoureuse s'élance du nord vers le sud alors que l'autre bat successivement en retraite vers les Pyrénées: voir l'agencement semi-circulaire des plages choroplèthes en rouge, orange, mauve et ocre au pied des Pyrénées.

La bande hyperbolique en bleu par contre – et qui traverse la Galloromania de l'Atlantique jusqu'aux confins orientaux de la Suisse Romande – constitue le théâtre d'innombrables contacts et de conflits linguistiques de toute sorte survenus depuis 2000 ans. C'est là qu'une éventuelle « harmonie primitive » entre les aménagements linguistique et euclidien de l'espace a été complètement désarçonnée, évidemment pour des raisons d'ordre extra-linguistique.

Pour cette supposition, la carte 11 nous fournit une excellente butte-témoin. Il s'agit de la position excentrique (marquée par un polygone isolé en bleu foncé) du P.-ALF 635 (Andraut) qui se trouve au sud de Bordeaux et qui est entouré, sur la carte 11, de polygones en jaune et ocre. La localité d'Andraut appartient à l'îlot linguistique de la Petite Gavacherie qui, elle, avait été fondée à la fin du XV^{e} siècle par des colons saintongeais en territoire gascon. Or, il est évident que les circonstances de la fondation d'un îlot linguistique sont absolument étrangères à l'idée d'une éventuelle « harmonie primitive » entre les aménagements linguistique et euclidien de l'espace.

9. LA DM DENDROGRAPHIQUE

La DM dendrographique (ou arborescente) est un instrument très utile tant du logiciel VDM que de l'EDMS. Il s'agit d'ailleurs d'un outil classificatoire très répandu qui est utilisé aussi par beaucoup d'autres sciences humaines, biologiques et naturelles. La particularité salzbourgeoise dans l'application de dendrogrammes de classification de toute sorte consiste dans la transposition automatique voire « obligatoire » de la structure ramifiée des arbres dans l'espace: voir la carte 12. C'est que la DM salzbourgeoise s'intéresse beaucoup plus aux schémas spatiaux qu'à leurs pendants arborescents.

BDD-A310 © 2008 Editura Academiei
Provided by Diacronia.ro for IP 216.73.216.25 (2026-11-07 16:04:16 UTC)

La construction de l'arbre présenté sur la carte 12 repose sur les prérequis suivants :

- établissement d'une matrice de similarité (calculée à l'aide de l'IRI_{jk} et 1687 cartes de travail,
- choix de l'algorithme dendrographique de Joe Ward, Jr.,
- mise en vedette (arbitraire) de neuf embranchements (« dendrèmes ») de l'arbre et leur projection (sous la forme de « chorèmes ») dans le réseau polygonal de l'ALF.

L'algorithme de Ward appartient, en termes de statistique, à la « classification ascendante hiérarchique » (CAH)[31]. Les spécialistes de la CAH ont défini, au cours des 60 à 70 dernières années, un nombre très élevé d'algorithmes dendrographiques de toute sorte dont l'utilité pratique, pour une discipline empirique donnée, n'est pas garantie d'entrée de jeu. Cette réserve vaut également pour la géolinguistique. C'est pourquoi le logiciel VDM permet l'application et le calcul de plusieurs algorithmes-CAH à la fois et, partant, l'expérimentation comparée des arborescences respectives. Fort de ces expériences nous pouvons assurer que l'utilité classificatoire et/ou exploratoire de l'algorithme-CAH de Ward est, en matière de géolinguistique, très grande.

La lecture et l'interprétation de la carte 12 peuvent se faire dans une perspective ou bien synchronique ou bien diachronique. Du point de vue *synchronique*, l'agencement des dendrèmes dans l'arbre et des chorèmes sur la carte, nous renseigne sur la structuration spatiale de la Galloromania dans une visée hiérarchique. Au vu de la carte 12, il appert que les macro-chorèmes 1-5 d'un côté et 6-9 de l'autre, constituent les composantes les plus importantes de la Galloromania. Il en appert en outre que le francoprovençal représente une branche latérale du domaine d'Oïl, et que le domaine d'Oc consiste en deux groupements majeurs qui comprennent, d'un côté, le gascon et le languedocien (chorèmes-dendrèmes 6-7) et, de l'autre, le provençal maritime, (chorème-dendrème 9), le limousin, l'auvergnat et le provençal alpin (chorème-dendrème 8).

Du point de vue *diachronique* par contre, l'analyse progressive, à partir de la racine de l'arbre, des ramifications binaires permet de reconstruire – évidemment par voie hypothétique – le déroulement de la fragmentation linguistique de la Galloromania. Ceci présuppose d'entrée de jeu l'existence d'une Galloromania linguistiquement homogène (voir la racine de l'arbre) et, par la suite, la mise en place de sa dialectalisation ultérieure. La première étape en est la désagrégation en Oc et Oïl (voir la première ramification après la racine), suivie par des scissions analogues en Oïl et francoprovençal dans le nord, et en Oc du Sud et Oc du Nord

[31] Les algorithmes-CAH opèrent par des agglomérations binaires qui débutent au niveau des objets à classifier (« feuilles de l'arbre ») pour arriver au tronc ou à la racine de l'arbre. On en trouve des descriptions détaillées dans tous les manuels modernes de classification numérique : cf. Bock 1974, 407s. (en allemand), Chandon/Pinson 1981, 123s. (en français) et Sneath/Sokal 1973, 204s. (en anglais).

BDD-A310 © 2008 Editura Academiei
Provided by Diacronia.ro for IP 216.73.216.25 (2026-11-07 16:04:16 UTC)

dans le sud (voir les ramifications respectives dans l'arbre de la carte 12). Précisons en outre qu'il s'agit là d'une méthode chère aux tenants de la lexicostatistique[32].

10. REMARQUES FINALES

Dans cet article, nous n'avons pu effleurer que très sommairement les multiples problèmes théoriques et pratiques de la DM en général et de l'EDMS en particulier. Soulignons encore une fois les points forts de la DM : c'est un outil diagnostique très efficace pour la détection de structures de profondeur contenues dans les données d'un AL. Or, le nombre des AL mal ou peu explorés est très grand. La DM peut être également appliquée à tout autre stock de données organisé sous la forme bidimensionnelle d'un AL classique[33].

N'oublions pas non plus que l'avènement des AL à la fin du XIX^{e} siècle a quelque peu bouleversé la linguistique contemporaine, surtout à cause de la découverte de la complexité déroutante de la distribution spatiale des différents traits linguistiques, même de ceux ayant la même origine étymologique.

Malheureusement, les possibilités cartographiques et aussi statistiques de l'époque n'ont permis qu'un traitement très superficiel de cette situation empirique quelque peu décourageante. À la distance d'un siècle, l'on doit donc constater, hélas, que l'écrasante majorité des collections atlantographiques tant nationales que régionales sont encore loin d'avoir fourni tous leurs secrets. Il est donc grand temps de combler cette lacune.

N'oublions pas non plus que l'analyse empirique de la gestion d'un espace donné est largement pratiquée aussi par d'autres sciences tant humaines que naturelles si bien que l'étude comparée de deux sortes de gestion du même espace – dont l'une évidemment devrait être de nature linguistique – constitue un défi interdisciplinaire très intéressant. À ce propos, nous ne citons que la génétique des populations qui, depuis plusieurs décennies, coopère étroitement avec la géolinguistique[34].

Répétons encore une fois que le logiciel VDM est à la libre disposition de tous ceux qui s'y intéressent et que l'initiation à son usage courant demande l'accomplissement d'un stage de formation de deux ou trois jours à effectuer à Salzbourg.

Remerciements. – Financement des travaux de recherche: « Fonds zur Förderung der wissenschaftlichen Forschung in Österreich“ (FWF), Vienne : projets 12 414 et 13 349.

[32] Voir à ce sujet nos observations respectives publiées en 2002 (29-34) et 2003 (84-89).

[33] Nous renvoyons à ce sujet à notre analyse dialectométrique des données de l'atlas scripturaire d'Anthonij Dees de 1980 publiée en 2006 (b).

[34] Cf.à ce sujet Goebl (2005b).

BDD-A310 © 2008 Editura Academiei
Provided by Diacronia.ro for IP 216.73.216.25 (2026-11-07 16:04:16 UTC)

– Établissement des 2 figures et des 12 cartes de cet article: Slawomir Sobota, Université de Salzbourg.
– Création et mise à jour régulière du logiciel VDM: Edgar Haimerl, Seattle (États Unis d'Amérique).
– Toilette stylistique de mon texte français : Lily Ditz-Fuhrich, Université de Salzbourg.

Que Mme Ditz-Fuhrich, M. Sobota, M. Haimerl ainsi que les responsables de l'organisme de recherche autrichien FWF reçoivent ici l'expression de ma profonde reconnaissance.

ABRÉVIATIONS ET SYMBOLES SOUVENT UTILISÉS

AL	Atlas linguistique
ALF	Atlas linguistique de la France : cf. la bibliographie.
CAF	coefficient d'asymétrie de Fisher: cf. le paragraphe 6.
CAH	classification ascendante hiérarchique: cf. le paragraphe 9.
DM	dialectométrie
-DM	dialectométrique
EDMS	Ecole dialectométrique de Salzbourg: cf. le paragraphe 1.
IRD_{jk}	Indice Relatif de Distance (relevé entre les points d'enquête j et k: cf. le paragraphe 3.
IRI_{jk}	Indice Relatif d'Identité (relevé entre les points d'enquête j et k): cf. le paragraphe 3.
N	Nombre des points d'enquête d'un atlas linguistique ou des vecteurs-points d'enquête d'une matrice de données: cf. le paragraphe 2.
p	Nombre des planches d'un atlas linguistique ou des vecteurs-attributs d'une matrice de données: cf. le paragraphe 2.
P., PP.	point, points (d'enquête) d'un atlas linguistique
r(BP)	Coefficient de corrélation de Bravais et Pearson: cf. le paragraphe 8.
VDM	Visual DialectoMetry : logiciel dialectométrique: cf. le paragraphe 2.

BIBLIOGRAPHIE

ALD-I: H. Goebl, R. Bauer, E. Haimerl (eds.), 1998), *Atlant linguistich del ladin dolomitich y di dialec vejins*, 1ª pert, *Atlante linguistico del ladino dolomitico e dei dialetti limitrofi,* 1ª parte, *Sprachatlas des Dolomitenladinischen und angrenzender Dialekte*, 1. Teil, Wiesbaden, Dr. Ludwig Reichert-Verlag, 7 vol., 3 CD-ROM (Salzbourg, Fachbereich Romanistik, 1999) et 1 DVD (Salzbourg, Fachbereich Romanistik, 2001).

ALF: J. Gilliéron, E. Edmont, (eds.), 1902–1910, *Atlas linguistique de la France*, Paris, Champion, 10 vol. (réimpression : Bologne, Forni, 1968).

R. Álvarez, F. Dubert García, X. Sousa Fernández, 2006, « Aplicación da analise dialectométrica aos datos do atlas lingüístico galego », in: R. Álvarez, F. Dubert García, X. Sousa Fernández (eds.), *Lingua e Territorio*, Santiago de Compostela, ILG/CCG, 461–493.

Bauer, R., 2002/3, « Dolomitenladinische Ähnlichkeitsprofile aus dem Gadertal. Ein Werkstattbericht zur dialektometrischen Analyse des ALD-I », *Ladinia,* 26–27, 209–250.

Bauer, R., 2003, *Dialektometrische Analyse des Sprachatlas des Dolomitenladinischen und angrenzender Dialekte (ALD-I),* Salzbourg, Fachbereich Romanistik, 2 vols.

BDD-A310 © 2008 Editura Academiei
Provided by Diacronia.ro for IP 216.73.216.25 (2026-11-07 16:04:16 UTC)

Bauer, R., 2004, « Dialekte – Dialektmerkmale – dialektale Spannungen. Von „Cliquen“, „Störenfrieden“ und „Sündenböcken“ im Netz des dolomitenladinischen Sprachatlasses ALD-I », *Ladinia,* 28, 201–242.

Bock, H. H., 1974, *Automatische Klassifikation. Theoretische und praktische Methoden zur Gruppierung und Strukturierung von Daten (Cluster-Analyse),* Göttigen, Vandenhoeck & Ruprecht.

Brun-Trigaud, G., Y. Le Berre, J. Le Dû, 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, Editions du Comité des travaux historiques et scientifiques.

Chandon, J.-L., S. Pinson, 1981, *Analyse typologique. Théories et applications*, Paris, New York, Barcelone, Milan, Masson.

Dees, A., 1980, *Atlas des formes et des constructions des chartes françaises du 13^{e} siècle*, Tübingen, Niemeyer.

Dickinson, G. C., 1973, *Statistical Mapping and the Presentation of Statistics*, Londres, Arnold.

Ettmayer, K. von, 1924, « Über das Wesen der Dialektbildung, erläutert an den Dialekten Frankreichs », in *Denkschriften [Mémoires] der Akademie der Wissenschaften in Wien, philosophisch-historische Klasse*, vol. 66, 1–56, 7 planches.

Goebl, H., 1981, « Eléments d'analyse dialectométrique (avec application à l'AIS) », *Revue de linguistique romane,* 45, 349–420.

Goebl, H., 1982, « Dialektometrie. Prinzipien und Methoden des Einsatzes der Numerischen Taxonomie im Bereich der Dialektgeographie », in *Denkschriften [Mémoires] der Österreichischen Akademie der Wissenschaften, philosophisch-historische Klasse*, vol. 157, Vienne, 1–123.

Goebl, H., 1983, « Parquet polygonal et treillis triangulaire: les deux versants de la dialectométrie interponctuelle », *Revue de linguistique romane,* 47, 353–412.

Goebl, H., 1984, *Dialektometrische Studien. Anhand italoromanischer, rätoromanischer und galloromanischer Sprachmaterialien aus AIS und ALF*, Tübingen, Niemeyer, 3 vols.

Goebl, H., 1987, « Points chauds de l'analyse dialectométrique: pondération et visualisation », *Revue de linguistique romane,* 51, 63–118.

Goebl, H., 2000, « La dialectométrisation de l'ALF: présentation des premiers résultats », *Linguistica,* 40, 209–236.

Goebl, H., 2002, « Analyse dialectométrique des structures de profondeur de l'ALF », R*evue de linguistique romane,* 66, 5–63.

Goebl, H., 2003, « Regards dialectométriques sur les données de l'Atlas linguistique de la France (ALF): relations quantitatives et structures de profondeur », *Estudis Romànics,* 25, 59–120.

Goebl, H., 2004, « Sprache, Sprecher und Raum. Eine kurze Darstellung der Dialektometrie. Das Fallbeispiel Frankreich », *Mitteilungen der Österreichischen Geographischen Gesellschaft,* 146, 247–286.

Goebl, H, 2005a, « La dialectométrie corrélative. Un nouvel outil pour l'étude de l'aménagement dialectal de l'espace par l'homme », *Revue de linguistique romane,* 69, 321–367.

Goebl, H., 2005b, « Dialekte und Familiennamen in Frankreich. Ein interdisziplinärer Vergleich mit den Mitteln der Dialektometrie », in: G. Hauska (ed.), *Gene, Sprachen und ihre Evolution. Wie verwandt sind die Menschen – wie verwandt sind ihre Sprachen?,* Regensburg, Universitätsverlag Regensburg, 68–99.

Goebl, H., 2006a, « Recent Advances in Salzburg Dialectometry », *Literary and Linguistic Computing,* 21, 4, 411–435.

Goebl, H., 2006b, « Sur le changement macrolinguistique survenu entre 1300 et 1900 dans le domaine d'Oïl. Une étude diachronique d'inspiration dialectométrique », *Linguistica,* 46, 3–43.

Goebl, H., 2007, « A bunch of dialectometric flowers: a brief introduction to dialectometry », in: U. Smit *et al.* (eds.), *Tracing English through time. Explorations in language variation. In honour of Herbert Schendl on occasion of his 65th birthday*, Vienne, Braumüller, 133–171.

BDD-A310 © 2008 Editura Academiei
Provided by Diacronia.ro for IP 216.73.216.25 (2026-11-07 16:04:16 UTC)

Haag, C., 1898, *Die Mundarten des oberen Neckar- und Donautales (Schwäbisch-alemannisches Grenzgebiet: Baarmundarten),* Reutlingen, Hutzler.
Heeringa, W., 2004, *Measuring Dialect Pronunciation Differences using Levenshtein Distance*, Groningen, Groningen Dissertations, 46.
Heeringa, W., J. Nerbonne, 2001, « Dialect areas and dialect continua », *Language Variation and Change,* 13, 375–400.
Hummel, L., 1993, *Dialektometrische Analysen zum Kleinen Deutschen Sprachatlas (KDSA). Experimentelle Untersuchungen zu taxometrischen Ordnungsstrukturen als dialektaler Gliederung des deutschen Sprachraums*, Tübingen, Niemeyer, Teil 1 : Text, Teil 2 : Karten.
Melis, L., S. Verlinde, P. Derynck, 1984, « La notion de dialecte supralocal: essai de définition dialectométrique », *Orbis,* 33, [1984/1989], 70–132.
Nerbonne, J., W. Heeringa, 2001, « Computational Comparison and Classification of Dialects », *Dialectologia et Geolinguistica,* 9, 69–83.
Nerbonne, J., W. Kretzschmar, 2003, « Introducing Computational Techniques in Dialectometry », « Computers and the Humanities », 37, 245–255.
Nerbonne, J., W. Kretzschmar, 2006, « Progress in Dialectometry: toward Explanation », *Literary and Linguistic Computing,* 21, 387–397.
Okabe, A., B. Boots, K. Sugihara, 1992, *Spatial Tesselations. Concepts and Applications of Voronoi Diagrams*, Chichester, New York, Brisbane, Toronto, Singapour, Wiley.
Rosenqvist, A., 1919, « Limites administratives et division dialectale de la France », *Neuphilologische Mitteilungen,* 20, 87–119.
Schiltz, G., 1996, *Dialektometrischer Atlas von Südwest-Baden (DASB). Konzepte eines dialektometrischen Informationssystems*, Marburg, Elwert, 1996, 4 vols.
Séguy, J., 1973, « La dialectométrie dans l'Atlas linguistique de la Gascogne », *Revue de linguistique romane,* 37, 1–24.
Sneath, P. A., R. R. Sokal, 1973, *Numerical Taxonomy. The Principles and Practice of Numerical Classification*, San Francisco, Freeman.
Verlinde, S., 1988, « La dialectométrie et la détection des zones dialectales : l'architecture dialectale de l'Est de la Belgique romane », *Revue de linguistique romane,* 52, 151–181.
Viaplana, J., 1999, *Entre la dialectologia i la lingüística. La distància lingüística entre les varietats del català nord-occidental*, Barcelone, Publicacions de l'Abadia de Monteserrat.
Videsott, P., 2003, *Padana scriptologica. Skriptologische und skriptometrische Untersuchungen zu altnorditalienischen Texten von den Anfängen bis 1525*, Innsbruck, Leopold-Franzens-Universität, 2 vols.
Videsott, P., 2005, « Continuità e trasformazione. La strutturazione geolinguistica dell'Italia settentrionale medievale e moderna. Un confronto tra dati scrittologici e dialettali », *Rivista italiana di dialettologia,* 29, 181–236.
Videsott, P., 2006, « Onomatometria: i metodi della tassonomia numerica applicati all'onomastica », *Rivista italiana di onomastica*, 12, 423–467.

BDD-A310 © 2008 Editura Academiei
Provided by Diacronia.ro for IP 216.73.216.25 (2026-11-07 16:04:16 UTC)